Essays in Behavioral Macroeconomics and Mechanism Design

by

Joel P. Flynn

B.A. (Hons), University of Cambridge (2017)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Economics
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
June 2023

Authored by: Joel P. Flynn
Department of Economics
April 10, 2023

Certified by: Daron Acemoglu
Institute Professor
Thesis Supervisor

Certified by: George-Marios Angeletos
Professor of Economics
Thesis Supervisor

Certified by: Stephen Morris
Peter A. Diamond Professor of Economics
Thesis Supervisor

Accepted by: Abhijit Banerjee
Chairman, Departmental Committee on Graduate Theses

Essays in Behavioral Macroeconomics and Mechanism Design

by

Joel P. Flynn

Submitted to the Department of Economics
on April 10, 2023, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

Abstract

This thesis is in two parts. The first part of the thesis, "Essays in Behavioral Macroeconomics," is motivated by the simple observation that the macroeconomy is complicated; many households and firms interact across myriad markets in ways that change over time. This part of the thesis studies, empirically and theoretically, the microeconomic foundations and macroeconomic implications of hypotheses inspired by these complications: that people adopt simplified and misspecified narratives to understand the world; and that people will only pay attention to the macroeconomy when it is important to them.

In the first chapter, "The Macroeconomics of Narratives" (coauthored with Karthik A. Sastry), we study the macroeconomic implications of *narratives*, or beliefs about the economy that affect decisions and spread contagiously. Empirically, we use natural-language-processing methods to measure textual proxies for narratives in US public firms' end-of-year reports (Forms 10-K). We find that: (i) firms' hiring decisions respond strongly to narratives, (ii) narratives spread contagiously among firms, and (iii) this spread is responsive to macroeconomic conditions. To understand the macroeconomic implications of these forces, we embed a contagious optimistic narrative in a business-cycle model. We characterize, in terms of the decision-relevance and contagiousness of narratives, when the unique equilibrium features: (i) non-fundamental business cycles, (ii) non-linear belief dynamics (narratives "going viral") that generate multiple stable steady states (hysteresis), and (iii) the coexistence of hump-shaped responses to small shocks with regime-shifting behavior in response to large shocks. Our empirical estimates discipline both the static, general equilibrium effect of narratives on output and their dynamics. In the calibrated model, we find that contagious optimism explains 32% and 18% of the output reductions over the early 2000s recession and Great Recession, respectively, as well as 19% of the unconditional variance in output. We find that overall optimism is not sufficiently contagious to generate hysteresis, but other, more granular narratives are.

In the second chapter, "Attention Cycles" (coauthored with Karthik A. Sastry), we document that, in aggregate downturns, US public firms' attention to macroeconomic conditions rises and the size of their input-choice mistakes falls. We explain these phenomena with a business-cycle model in which firms face a cognitive cost of making

precise decisions. Because firms are owned by risk-averse households, there are greater incentives to deliver profits by making smaller input-choice mistakes when aggregate consumption is low. In the data, consistent with our model, financial markets punish mistakes more in downturns and macroeconomically attentive firms make smaller mistakes. Quantitatively, attention cycles generate asymmetric, state-dependent shock propagation and stochastic volatility of output growth.

In the third chapter, "Strategic Mistakes" (coauthored with Karthik A. Sastry), to study the equilibrium implications of decision frictions, we introduce a new class of control costs in continuum-player, continuum-action games in which agents interact via an aggregate of the actions of others. The costs that we study accommodate a rich class of decision frictions, including *ex post* misoptimization, imperfect *ex ante* planning, cognitive constraints that depend endogenously on the behavior of others, and consideration sets. We provide primitive conditions such that equilibria exist, are unique, are efficient, and feature monotone comparative statics for action distributions, aggregates, and the size of agents' mistakes. We apply the model to make robust equilibrium predictions in a monetary business-cycle model of price-setting with planning frictions and a model of consumption and savings during a liquidity trap when endogenous stress worsens decisions.

The second part of this thesis, "Essays in Mechanism Design," studies two contentious issues in the allocation of resources in the modern economy: How should we account for diversity when we allocate resources in two-sided matching markets? How should digital goods and information be priced and regulated?

In the fourth chapter, "Priority Design in Centralized Matching Markets" (coauthored with Oğuzhan Çelebi), we observe that in many centralized matching markets, agents' property rights over objects are derived from a coarse transformation of an underlying score. Prominent examples include the distance-based system employed by Boston Public Schools, where students who lived within a certain radius of each school were prioritized over all others, and the income-based system used in New York public housing allocation, where eligibility is determined by a sharp income cutoff. Motivated by this, we study how to optimally coarsen an underlying score. Our main result is that, for any continuous objective function and under stable matching mechanisms, the optimal design can be attained by splitting agents into at most three indifference classes for each object. We provide insights into this design problem in three applications: distance-based scores in Boston Public Schools, test-based scores for Chicago exam schools, and income-based scores in New York public housing allocation.

In the fifth chapter, "Adaptive Priority Mechanisms" (coauthored with Oğuzhan Çelebi), we ask how authorities that care about match quality and diversity should allocate resources when they are uncertain of the market they face? Such a question appears in many contexts, including the allocation of school seats to students from various socioeconomic groups with differing exam scores. We propose a new class of *adaptive priority mechanisms* (APM) that prioritize agents as a function of both scores that reflect match quality and the number of assigned agents with the same socioeconomic characteristics. When there is a single authority and preferences over scores and diversity are separable, we derive an APM that is optimal, generates a unique outcome, and can be specified solely in terms of the preferences of the authority. By contrast, the ubiquitous priority and quota mechanisms are optimal if and only if the authority is risk-neutral or extremely risk-averse over diversity, respectively. When there are many authorities, it is dominant for each of them to use the optimal APM, and each so doing implements the unique stable matching. However, this is generally inefficient for the authorities. A centralized allocation mechanism that first uses an aggregate APM and then implements authority-specific quotas restores efficiency. Using data from Chicago Public Schools, we estimate that the gains from adopting APM are considerable.

In the sixth and final chapter, "Nonlinear Pricing with Under-Utilization: A Theory of Multi-Part Tariffs" (coauthored with Roberto Corrao and Karthik A. Sastry), we study the nonlinear pricing of goods whose usage generates revenue for the seller and of which buyers can freely dispose. The optimal price schedule is a multi-part tariff, featuring tiers within which buyers pay a marginal price of zero. We apply our model to digital goods, for which advertising, data generation, and network effects make usage valuable, but monitoring legitimate usage is infeasible. Our results rationalize common pricing schemes including free products, free trials, and unlimited subscriptions. The possibility of free disposal harms producer and consumer welfare and makes both less sensitive to changes in usage-based revenue and demand.

Thesis Supervisor: Daron Acemoglu
Title: Institute Professor

Thesis Supervisor: George-Marios Angeletos
Title: Professor of Economics

Thesis Supervisor: Stephen Morris
Title: Peter A. Diamond Professor of Economics

Acknowledgments

This thesis is dedicated to my future wife Lizzy, my mum Helen, my dad Peter, and my brother Ethan. Without their love, none of this would have been possible.

I want to thank my advisors, Daron, Marios, and Stephen, for their tireless support, extensive feedback, and unparalleled generosity in helping me develop my work and, more importantly, my thinking. It is hard to overstate the impact that Daron has had on my development as a scholar. He has constantly pushed me to develop my ideas, add rigor to my arguments, and bring forth the key economics to make my analysis transparent. Marios has been equally key in the evolution of my thought and work. His inspirational quarter in the graduate advanced macroeconomics sequence introduced me to behavioral macroeconomics and game-theoretic approaches to the analysis of general-equilibrium macroeconomic models. Since then, his advice and expertise in this area have led me to sharpen my analysis and do better work than I could have ever produced without his guidance. Stephen has also been a constant source of essential advice for both my theoretical and empirical work. His breadth of interest and economic intuition has enriched each of the chapters in this thesis and more.

The greatest deal of thanks is owed to my collaborators on this thesis, Karthik, Oğuzhan, and Roberto. Without their ideas, ingenuity, and hard work, I would be much the lesser economist. I am extremely fortunate to have worked with them, and I look forward to doing so for years to come. Beyond this thesis, I have had the privilege to work with a set of fantastic scholars, Jacob Moscona, George Nikolakoudis, Christina Patterson, John Sturm, Alexis Akira Toda, and Lawrence Schmidt, each of whom has improved my thinking. Collaboration has truly been the highlight of the last six years. I certainly could not have produced work of comparable quality and had so much fun doing so without the academic and personal brilliance of my coauthors.

My time in graduate school has also been enormously enriched by my conversations with other fantastic economists and people, including Andrea Manera, Charles Serfaty, Michele Fornino, Basil Halperin, Jackson Mejia, Charlie Rafkin, Jon Cohen, Giacomo Lanzani, Ricardo Caballero, Parag Pathak, and Rob Townsend.

Finally, thanks are due to my undergraduate mentors from King's, Aytek Erdil, Chryssi Giannitsarou, Elisa Faraglia, and Hamid Sabourian. Without their advice, support, and faith, I would not have had the ability to attend graduate school in the first place.

Contents

Part I

Essays in Behavioral Macroeconomics

Chapter 1

The Macroeconomics of Narratives

This chapter is jointly authored with Karthik A. Sastry.

1.1 Introduction

At least since Keynes (1936), economists have hypothesized that waves of "spontaneous optimism" generate business cycles. But what drives these fluctuations in beliefs and how much do they matter? The *Narrative Economics* of Shiller (2017, 2020) postulates that contagious stories and worldviews, or "narratives," induce such movements in beliefs and underlie macroeconomic events. However, the business-cycle relevance of economic narratives, and even their precise meaning, remain unclear.

In this paper, we study how movements in narratives drive aggregate fluctuations in beliefs and output. We first develop a conceptual framework in which we define narratives as subjective models of the macroeconomy that are potentially incorrect. By altering beliefs, narratives influence economic actions like hiring and investment (they are *decision-relevant*). Moreover, narratives can feed into themselves and gain or lose prevalence over time, in two distinct but complementary ways: direct feedback from their prevalence (they are *contagious*), as in as in epidemiological models, and indirect feedback from the economic activity that narratives induce (they are *associative*), as in models of learning.

We next introduce empirical strategies to test narratives' decision-relevance, contagiousness, and associativeness at the microeconomic level. We apply several natural-language-processing methods to measure textual proxies for narratives in the universe of 10-K regulatory filings, in which all US public firms discuss "perspectives on [their] business results and what is driving them" (SEC, 2011). We find that measured narratives predict hiring and investment and that their spread depends on both narrative prevalence and economic outcomes at the aggregate and industry level. Moreover, we find that measured *narrative optimism* predicts over-optimism in sales forecasts and does not correlate positively with future firm productivity growth, stock returns, or earnings growth. Thus, we interpret textual optimism as a narrative that shifts beliefs while being unrelated to future fundamentals.

To understand the importance of these results, we adopt a "micro-to-macro" approach by embedding narratives in a business-cycle model and quantifying its theoretical predictions by using our empirical estimates. We show theoretically that optimistic narratives can lead to non-fundamentally driven boom-bust cycles and hysteresis. Conditional on calibrating standard preference and technological parameters, our empirical estimates point-identify the dynamic macroeconomic effects of narratives. Quantitatively, we find that aggregate fluctuations in narrative optimism account for approximately 32% and 18% of the output reductions over the early 2000s recession and Great Recession, respectively. We therefore argue that contagious narratives may be a first-order determinant of the business cycle.

Measuring Narratives We first empirically evaluate the two premises of our framework: narratives' decision-relevance and their contagiousness and associativeness. To this end, we combine data on US public firms' adoption of textual narratives, using firms' 10-K filings and earnings call transcripts, and their decisions, using data from Compustat.

Our first method for measuring textual proxies for narratives computes the intensity of positive and negative sentiment using the 10-K-specific dictionary introduced by Loughran and McDonald (2011a). We interpret this measure as capturing overall optimism. This measure, which will be our main focus, connects our work to a large literature that studies waves of optimism and pessimism at the aggregate level, without measuring how optimism shapes economic decisions or spreads at the microeconomic level (see *e.g.,* Beaudry and Portier, 2006; Lorenzoni, 2009; Angeletos and La'O, 2013; Benhabib, Wang, and Wen, 2015).

To complement our study of narrative optimism, we apply two other techniques to measure more granular narratives. Our first such technique computes the similarity between firms' language and the language that best characterizes the *Perennial Economic Narratives* introduced by Shiller (2020) using a method that has also been applied in Hassan, Hollander, Van Lent, and Tahoun (2019) and Flynn and Sastry (2022a). These "narratively identified narratives" are motivated by the historical evidence of relevance and contagiousness provided by Shiller (2020). Our second granular technique estimates a Latent Dirichlet Allocation (LDA) model (Blei, Ng, and Jordan, 2003), which extracts an underlying set of *topics*, probability distributions over words, based on the frequency with which certain words co-occur within documents. Since we recover these "topic narratives" via an unsupervised method, they allow the data to speak flexibly about what narratives are prevalent among firms without restricting the issues to which narratives may pertain.

Empirical Results We first provide descriptive evidence about our estimated narratives. Across the three methods, almost all of our estimated narratives are persistent and cyclical. However, it is difficult to ascertain the relationship between narrative and macroeconomic dynamics from the time series alone. This is because narratives potentially serve dual roles of describing true economic fundamentals and encoding non-fundamental beliefs. This challenge for time-series analysis motivates our strategy of testing the two premises of narrative macroeconomics—that narratives are decision-relevant and that they spread contagiously and associatively—at the microeconomic level.

Using our panel data, we first test the decision-relevance of our measured narratives. We focus initially on optimism. We find that optimistic firms, defined as firms with above-median sentiment, hire 3.6 percentage points more than pessimistic firms in a given year, net of firm and sector-time fixed effects. This finding is robust to accounting for firm-level productivity and financial conditions. Moreover, we find that optimism is uncorrelated with future productivity growth and *negatively* correlated with future stock returns and profitability. These findings are inconsistent with the hypothesis that optimism predicts hiring only because it captures positive firm-level fundamentals (or news thereof). Further, we show using managerial guidance data from IBES that optimism in language predicts negative errors in sales forecasts, or over-optimistic beliefs.

We therefore interpret the association of textual optimism with hiring as arising from non-fundamental, narratively driven, and optimistic beliefs. To underscore this interpretation, we show that changes in optimism driven by plausibly exogenous changes in CEOs (*i.e.,* those caused by death, illness, personal issues, or voluntary retirement of an incumbent CEO, as coded by Gentry, Harrison, Quigley, and Boivie, 2021) lead to quantitatively similar effects on hiring. Finally, we study the relevance of the Perennial Economic Narratives and our estimated topics. Because these narratives are high-dimensional and may not be relevant for firm decisions, we use the Rigorous LASSO method of Belloni, Chernozhukov, Hansen, and Kozbur (2016) for estimating their effects on hiring. We find that two of the nine Perennial Economic Narratives and eleven of the one hundred topics are relevant for hiring.

Second, we study how our measured narratives spread across firms over time.

Focusing on optimism, we find that greater aggregate optimism and higher aggregate real GDP growth are associated with a greater probability that a firm is optimistic in the following year—that is, in our language, optimism is contagious and associative. We also find evidence of contagiousness and associativeness at the industry level when we control for aggregate conditions with time fixed effects. Moreover, both these aggregate and industry-level results are robust to controlling for future idiosyncratic and aggregate economic conditions. This finding is inconsistent with the explanation that aggregate optimism drives future optimism through its correlation with omitted positive news about measured economic conditions.

By using lagged aggregate optimism in a panel setting, our estimates are not threatened by the reflection problem of Manski (1993). Nevertheless, common shocks that are not spanned by measured aggregate and industry-level conditions may generate omitted variables bias. To address this concern, we employ strategies based on using idiosyncratic shocks to large firms (the granular IV approach of Gabaix and Koijen, 2020) and the aforementioned plausibly exogenous changes in CEOs as instruments for aggregate and industry-level optimism. We find qualitatively consistent effects. Finally, we perform similar analyses for the other decision-relevant narratively-identified and topic narratives. We find that almost all of them are contagious and that many are associative.

Model Having provided microeconomic evidence about narratives' decision-relevance, contagiousness, and associativeness, we now study their macroeconomic implications. To do this, we embed narratives in an otherwise standard Neoclassical business-cycle model with dispersed information à la Angeletos and La'O (2010, 2013). The consumption, production, and labor supply side of the model is a real variant of the standard Neoclassical model of Woodford (2003b) and Galí (2008). In particular, the model features aggregate demand externalities (Blanchard and Kiyotaki, 1987), which generate a motive among firms to co-ordinate the levels of their production. Unlike the aforementioned models, in which firms have correctly specified and rational expectations, our model features narratives that generate heterogeneous and incorrect prior beliefs about the state of aggregate productivity. In our main analysis, we specialize to a case with two narratives: *optimism* and *pessimism*. The evolution of narratives is governed by the probabilities that optimists and pessimists remain and become optimistic, respectively, as a function of aggregate output (associativeness) and the fraction of optimists in the population (contagiousness). This allows the model to accommodate the narrative dynamics that we estimated in the data.

Theoretical Results We next derive analytical results that characterize how narratives can generate non-fundamental fluctuations in aggregate output, hysteresis, and boom-bust cycles. We first establish that there is a unique equilibrium in which aggregate output is log-linear in aggregate productivity and a non-linear function of the fraction of optimists in the population. In the case of unanimous optimism, the contribution of optimism equals the partial-equilibrium effect of optimism on one firm's hiring, as we measured empirically, times a general-equilibrium multiplier. This is because optimism matters both directly for firms and indirectly through aggregate demand externalities. In this way, movements in aggregate optimism lead to non-fundamental fluctuations in aggregate output.

We next describe the dynamics of narratives and output. For a fixed level of aggregate productivity, while there always exists a steady-state level of optimism and equilibrium is unique, there may nevertheless be multiple steady-state levels of optimism. We provide a necessary and sufficient condition for a particularly extreme type of this steady-state multiplicity: if the decision-relevance, contagiousness, and associativeness of narratives are sufficiently strong, then unanimous optimism and unanimous pessimism are both stable steady states. Moreover, depending on the initial fraction of optimists, the economy is (almost everywhere) globally attracted to one of these extreme steady states. In this way, narratives can generate hysteresis: depending on how many optimists there are initially, optimism can either catch on forever ("go viral") or die out entirely.

We finally study how the economy evolves in response to shocks. Responses to unanticipated "MIT shocks" can fall into three qualitative regimes. If a shock is small, it has a fully transitory impact on aggregate output, because it fails to seed a new narrative. If a shock is medium-sized, it has potentially hump-shaped effects on aggregate output, because it seeds a new narrative that briefly persists before dying out. If a shock is large, it has a permanent effect on aggregate output, because it makes a narrative "go viral." Studying stochastic behavior, we show that the economy oscillates between extreme optimism and pessimism and provide analytical upper bounds on the expected period of these oscillations. Both the possibility of these effects and their quantitative magnitudes depend on key measurable parameters: the decision-relevance, associativeness, and contagiousness of narratives.

Quantification In the final part of the paper, we calibrate our model to quantify the extent to which fluctuations in narrative optimism explain historical business cycle fluctuations and understand the extent to which narrative dynamics generate hysteresis. In point-identifying the model, we leverage the fact that our empirical estimates identify both the partial equilibrium effects of narratives on hiring and the nature of narrative diffusion.

We first study the extent to which narratives generate non-fundamental fluctuations in output. We find that measured aggregate movements in optimism account for 32% of output loss during the early 2000s recession and 18% during the Great Recession. More systematically, we find that optimism accounts for 19% of output variance, and 34% of the short-run (one-year) and 81% of the medium-run (two-year) autocovariance in output.[1]

We next study the potential for narratively driven hysteresis. For optimism, we quantitatively reject the theoretical condition required for hysteresis in both optimism and output dynamics. But we do not reject this condition for other narratives, implying that it is possible for these narratives either to die out or "go viral" depending on initial conditions.

Finally, we study a variant model which allows multiple latent narratives to form a basis for overall optimism, to evaluate Shiller's (2020) hypothesis that many narratives may be mutually reinforcing. We find that the interaction of many simultaneously evolving, granular, and highly contagious narratives can underlie stable fluctuations in aggregate optimism.

Related Literature We relate to an empirical literature that measures narratives following Carroll (2001) and Shiller (2017).[2] Of most relevance, Andre, Haaland, Roth, and Wohlfart (2022) use surveys to understand narratives underlying inflation, Goetzmann, Kim, and Shiller (2022) analyze narratives about financial crashes in news media, Macaulay and Song (2022) study how news coverage of specific narratives affects sentiment on social media, and Bybee, Kelly, Manela, and Xiu (2021a) apply LDA to the full text of *Wall Street Journal* articles to extract narrative time series. Our approach differs in its use of text data about the cross-section of firms to extract narratives, uncover their effects on decision-making, and study their spread. Our empirical analysis therefore relates to a literature studying the relationship between firm-level outcomes and their language (Loughran and McDonald, 2011a; Hassan, Hollander, Van Lent, and Tahoun, 2019; Hassan, Schreger, Schwedeler, and Tahoun, 2021; Handley and Li, 2020). In contrast to these papers, we calibrate a model to

[1] Normatively, we show that contagious optimism can be welfare-improving even if it is unfounded. Quantitatively, we find that optimism is welfare-improving and welfare-equivalent to a 1.3% production subsidy.

[2] See Carroll and Wang (2022) for a review of the literature on "epidemiological" models of expectations formation that are centered around social interactions.

match our firm-level findings and study their general-equilibrium consequences.[3]

Our work relates to a literature that studies business cycles through time variation in agents' beliefs. First, our work relates to papers which postulate that the economy undergoes exogenous shocks to demand via news, noise, sentiment, or extrapolation, in theory (Lorenzoni, 2009; Angeletos and La'O, 2010, 2013; Benhabib, Wang, and Wen, 2015; Benhima, 2019; Caballero and Simsek, 2020) and in quantitative applications (Beaudry and Portier, 2006; Christiano, Ilut, Motto, and Rostagno, 2008; Angeletos, Collard, and Dellas, 2018; Bhandari, Borovička, and Ho, 2019; Maxted, 2020; Huo and Takayama, 2022).[4] Our work micro-founds such shocks via the endogenous evolution of narratives and corresponding degree of optimism, develops methods to measure agent-level sentiment via textual analysis, and shows how to use microeconomic evidence to quantify their macroeconomic implications.

Second, our modeling of narratives and their spread relates to the work of Carroll (2001), Burnside, Eichenbaum, and Rebelo (2016), and Shiller (2017), in which beliefs spread contagiously between agents. At the same time, our approach differs as we explicitly model narratives and provide microeconomic evidence about narratives' decision-relevance and spread.[5]

Finally, in studying the dynamics of misspecified models, we relate to a large macroeconomics and theory literature on model misspecification and learning (see *e.g.,* Marcet and Sargent, 1989a,b; Brock and Hommes, 1997; Esponda and Pouzo, 2016; Acemoglu, Chernozhukov, and Yildiz, 2016; Adam, Marcet, and Beutel, 2017; Molavi, 2019; Bohren and Hauser, 2021). This literature primarily characterizes the limit points of agents' models. Instead, we study short-run fluctuations and time variation in the models held by agents. This approach is similar to that of Kozlowski, Veldkamp, and Venkateswaran (2020), but we differ in our non-Bayesian and analytical, rather than computational, approach.

Outline The rest of the paper proceeds as follows. In Section 1.2, we develop our general framework. In Section 1.3, we describe our data and measurement. In Section 1.4, we detail our empirical strategy and results. In Section 1.5, we introduce our macroeconomic model with contagious narratives. In Section 1.6, we provide theoretical results on macroeconomic dynamics. In Section 1.7, we quantify the role

[3]Flynn and Sastry (2022a) and Song and Stern (2021a) share this approach of contextualizing the effect of language-based variables on firm-level outcomes in a macroeconomic model.

[4]Bordalo, Gennaioli, Shleifer, and Terry (2021) and Bordalo, Gennaioli, Kwon, and Shleifer (2021) respectively study how extrapolation can generate credit cycles and speculative bubbles.

[5]Thus, our model also differs from recent theoretical work in which models correspond to likelihoods (Schwartzstein and Sunderam, 2021) or directed acyclic graphs (Spiegler, 2016; Eliaz and Spiegler, 2020).

of narratives. Section 1.8 concludes.

1.2 Narratives: A Conceptual Framework

We first describe a conceptual framework that formalizes the two premises of the macroeconomics of narratives: that narratives are decision-relevant and that narratives spread contagiously and associatively. We embed these two premises in an abstract game in which narratives form the building blocks of agents' beliefs and agents care about their own actions, fundamentals, and aggregates of other agents' actions. This game nests our later macroeconomic model in Section 1.5. We then derive two regression equations that allow us to test the decision-relevance, contagiousness, and associativeness of narratives. We bring these regressions to the data in Section 1.4 and show that they obtain exactly in our macroeconomic model in Section 1.6.

1.2.1 Narratives and Beliefs

We begin by formally defining narratives and how they map into agents' beliefs. There are random aggregate fundamentals $\theta \in \Theta$. For example, these fundamentals might represent aggregate productivity or the strength of demand. An individual narrative is a model of fundamentals. We describe each narrative, indexed by $k \in \mathcal{K}$, as a probability distribution $N_k \in \Delta(\Theta)$ within the set of narratives $\mathcal{N} = \{N_k\}_{k \in \mathcal{K}}$. For example, if the fundamental θ describes the strength of productivity, then a pessimistic narrative N_P might correspond to the view that "productivity in the economy is low," while an optimistic narrative N_O may represent the view that "productivity in the economy is high."

Agents combine narratives to form priors about the fundamental by placing a vector of weights $\lambda = \{\lambda_k\}_{k \in \mathcal{K}} \in \Lambda \subseteq \Delta(\mathcal{K})$ on each narrative.[6] An agent with narrative weights λ has an induced prior distribution over fundamentals given by the following linear combination of distributions in $\mathcal{N}$:

$$\pi_\lambda(\theta) = \sum_{k \in \mathcal{K}} \lambda_k N_k(\theta) \tag{1}$$

Continuing the example, an agent who is fully pessimistic might place weight $\lambda_P = 1$

[6]The management and organizational literature also views narratives as forming a common set of stories that underpin beliefs (see, *e.g.*, Isabella, 1990; Maitlis, 2005; Loewenstein, Ocasio, and Jones, 2012; Vaara, Sonenshein, and Boje, 2016). Relatedly, Acemoglu and Robinson (2021) postulate that culture arises from the combination of latent cultural attributes. By microfounding the process by which cultural attributes combine, our analysis could be applied to this context.

on the pessimistic narrative and complementary weight $\lambda_O = 0$ on the optimistic narrative, so their subjective probabilities for each state θ are $\pi(\theta) = N_P(\theta)$. An agent who has been convinced by neither narrative might take a middle ground and consider both equally likely, which we would represent with $(\lambda_P, \lambda_O) = (\frac{1}{2}, \frac{1}{2})$ or beliefs $\pi(\theta) = \frac{1}{2}N_O(\theta) + \frac{1}{2}N_P(\theta)$.

1.2.2 Premise I: Narratives are Decision-Relevant

The first premise of the macroeconomics of narratives is that narratives are decision-relevant. To model this, we introduce a continuum of agents indexed by $i \in [0,1]$ whose beliefs are formed from narratives, as described above, and who make decisions at a sequence of time periods $t \in \mathbb{N}$.

These agents care about their own actions $x_{it} \in \mathcal{X}$, aggregate outcomes $Y_t \in \mathcal{Y}$, the fundamental state $\theta_t \in \Theta$, and an idiosyncratic preference shifter $\omega_i \in \Omega$. They have utility functions $u : \mathcal{X} \times \mathcal{Y} \times \Theta \times \Omega \to \mathbb{R}$ and information sets $\mathcal{I}_{it}$. Given their information sets, agents update each narrative belief by applying Bayes' rule and then form their posterior by placing their narrative weights on the updated narrative beliefs. Given a conjecture about the mapping from fundamental states to aggregates $\hat{Y}_t : \Theta \to \mathcal{Y}$, the agents maximize their expected utility given narrative weights λ_{it} and information $\mathcal{I}_{it}$:

$$\max_{x_{it} \in \mathcal{X}} \mathbb{E}_{\pi_{\lambda_{it}}} \left[u(x_{it}, \hat{Y}_t(\theta_t), \theta_t, \omega_i) \mid \mathcal{I}_{it} \right] \tag{2}$$

We linearize these best replies to obtain the following regression equation that allows us to test for the decision-relevance of narratives (see Proposition 25 in Appendix A.1.1 for the formal arguments):[7]

$$x_{it} = \gamma_i + \chi_t + \sum_{k \in \mathcal{K}} \delta_k \lambda_{k,it} + \varepsilon_{it} \tag{3}$$

In this equation, γ_i captures time-invariant factors (preference shifters), χ_t captures time-varying aggregate variables such as fundamentals or the overall prevalence of narratives, the δ_k correspond to the appropriately normalized expectation of both fundamentals and endogenous aggregate outcomes under narrative k, and ε_{it} corresponds to noise around these expectations caused by differences in the information sets across agents. The hypothesis that narratives are decision-relevant is that $\delta_k \neq 0$ for some $k \in \mathcal{K}$. To test this hypothesis in panel data on firms, we will use firm hiring

[7]There we also provide assumptions sufficient to guarantee a quadratic misspecification bound and show that ε_{it} is mean zero and independent from γ_i, χ_t, and λ_{it}. This latter point implies that, modulo issues of misspecification, the δ_k can be estimated consistently via OLS.

as the relevant outcome and textual measures of narrative adoption as proxies for $\lambda_{k,it}$, the belief weights in the model. An analog of this equation will hold in equilibrium without approximation in our theoretical model in Section 1.5, facilitating the use of our estimates for model calibration.

1.2.3 Premise II: Narrative Spread Is Contagious and Associative

We now formalize the second premise: that narratives spread contagiously and associatively. The extent of narrative penetration is summarized by the cross-sectional distribution of narratives in the population, $Q \in \Delta(\Lambda)$. This represents the distribution of agents' distributions of narrative weights. For example, in an economy populated by only optimists $\lambda^O = (0, 1)$, pessimists $\lambda^P = (1, 0)$ and moderates $\lambda^M = (\frac{1}{2}, \frac{1}{2})$, we would have that $Q = (Q^O, Q^P, Q^M)$ corresponds to the fraction of the population with each combination of weights over optimism and pessimism.

The evolution of the distribution of narratives over time is described by an updating rule $P : \Lambda \times \mathcal{Y} \times \Delta(\Lambda) \to \Delta(\Lambda)$, which returns the probabilities $\{P_{\lambda'}(\lambda, Y, Q)\}_{\lambda' \in \Lambda}$ that an agent with narrative weights λ changes their weights to λ' when the endogenous state is Y and the distribution of narratives in the population is Q.[8] Hence, conditional on a distribution of narratives at time t given by Q_t and realized endogenous outcomes given by Y_t, the next period's distribution of narratives is:

$$Q_{t+1,\lambda'} = \sum_{\lambda \in \Lambda} Q_{t,\lambda} P_{\lambda'}(\lambda, Y_t, Q_t) \tag{4}$$

At this level of generality, the updating function can capture Bayesian updating by agents given some latent information structure. However, we can also model behavioral phenomena such as associative learning where agents associate certain states of the economy with certain models (*e.g.,* "aggregate output is high, therefore productivity is high"), and contagiousness, wherein the distribution of narratives itself affects updating.

We linearize the narrative updating equations to obtain the following system of linear probability models (see Proposition 26 in Appendix A.1.1 for the formal arguments):

$$\mathbb{P}[\lambda_{it} = \lambda \mid \lambda_{i,t-1}, Y_{t-1}, Q_{t-1}] = \zeta_\lambda + \sum_{\lambda' \in \Lambda} u_{\lambda',\lambda} \mathbb{I}[\lambda_{i,t-1} = \lambda'] + r_\lambda' Y_{t-1} + s_\lambda' Q_{t-1} \tag{5}$$

[8] In Appendix A.2.6, we extend this setting to allow for idiosyncratic fundamentals and updating that depends on their realizations.

Here, u captures agents' stubbornness in updating (*i.e.,* their proclivity not to update), r captures associativeness in updating (*i.e.,* their association of outcomes with a direction of updating), and s captures contagiousness in updating (*i.e.,* the direct influence of peers' narrative weights). The hypotheses that narratives are contagious and associative correspond, respectively, to $s_\lambda \neq 0$ and $r_\lambda \neq 0$ for some $\lambda \in \Lambda$. Again, we will use panel data on textual narrative adoption to test these hypotheses. Equation 5 is generated without approximation in the special case of our theoretical model in Section 1.5 that we study quantitatively.

1.3 Data, Measurement, and Descriptive Statistics

We now describe how we develop a panel dataset on firms' narrative loadings and decisions. We measure textual proxies for narratives by applying several natural-language-processing techniques to two corpora of language: the universe of public firms' SEC Forms 10-K and a large sample of earnings calls. We combine these measures of narratives with data on firm fundamentals and choices. Finally, we provide descriptive facts regarding the time-series and cross-sectional properties of narratives.

1.3.1 Data

Text Our main source of firm-level textual data is SEC Form 10-K. Each publicly traded firm in the US submits an annual 10-K to the SEC. These forms provide "a detailed picture of a company's business, the risks it faces, and the operating and financial results of the fiscal year." Moreover, "company management also discusses its perspective on the business results and what is driving them" (SEC, 2011). This description is consistent with our notion that agents' narratives constitute a view of the world and its rationalization via some model.

We download the universe of SEC forms 10-K from the SEC Edgar database from 1995 to 2019. This yields a corpus of 182,259 `html` files comprising the underlying text of the 10-K, various formatting information, and tables. We describe our exact method for processing the text data in Appendix A.3.1. The three key steps are pre-processing the raw text data to isolate English-language words, associating words with their common roots via lemmatization, and fitting a bigram model that groups together co-occurring two-word phrases. We then count the occurrences of all words, including bigrams, in all documents to obtain the bag-of-words representation (*i.e.,* a vector of word counts) for each document.[9] Our final sample consists of 100,936

[9]Other machine-learning approaches use, as input, the entire document instead of its bag-of-words

firm-by-year observations from 1995 to 2018.

As an alternative source of text data, we use public firms' sales and earnings conference calls. Our initial sample consists of 158,810 documents from 2002 to 2014. We apply the same natural-language-processing techniques that we employ for the 10-Ks to this corpus. We average variables over the periods between successive 10-Ks to obtain a firm-by-fiscal-year dataset. Our final sample consists of 25,589 firm-by-year observations. We describe more details in Appendix A.3.2.

Firm Fundamentals and Choices We compile our dataset of firm fundamentals and choices using Compustat Annual Fundamentals from 1995 to 2018. This dataset includes information from firms' financial statements on employment, sales, input expenses, capital, and other financial variables. We apply standard selection criteria to screen out firms that are very small, report incomplete information, or were likely involved in an acquisition. We also ignore firms in the financial and utilities sectors due to their markedly different production and/or market structure. More details about our sample selection are in Appendix A.4.1.

We organize firms into 44 industries, which are defined at the NAICS 2-digit level, but for Manufacturing (31-33) and Information (51), which we split into the 3-digit level. To study narrative transmission at a finer level, we also define peer sets for the subset of firms traded on the New York Stock Exchange using the method of Kaustia and Rantala (2021). These authors exploit common equity analyst coverage to define peers for each firm.[10]

To measure total factor productivity, we estimate a constant-returns-to-scale, Cobb-Douglas, two-factor production function in materials and capital, for each industry. We estimate the output elasticities using the ratio of materials expenditures to total sales and an assumed revenue returns-to-scale of 0.75. More details are provided in Appendix A.4.2. We denote our estimated log-TFP variable as $\log \hat{\theta}_{it}$.

Manager and Analyst Beliefs We collect data from IBES (the International Brokers' Estimate System) on quantitative sales forecasts by companies and their equity analysts. Specifically, we use the IBES Guidance dataset which records, for specific variables, both (i) a numerical management expectation recorded from press

representation. Examples include Doc2Vec, as recently employed by Goetzmann, Kim, and Shiller (2022) to study crash narratives, and RELATIO, which was recently developed by Ash, Gauthier, and Widmer (2021). We view integrating these methods into our analysis as an interesting avenue for future study.

[10]Firm j is a peer of firm i at time t if they have more than C common analysts, where C is chosen so that the probability of having C or more common analysts by chance is less than 1% when analysts following firm i randomly choose the firms they follow among all firms with analysts in period t.

releases or transcripts of corporate events and (ii) a contemporaneous consensus value from equity analysts. We restrict to the first recorded forecast per fiscal year of that year's sales. When managers' guidance is reported as a range, we code a point-estimate forecast as the range's midpoint. We construct two variables from these data at the level of firms i and fiscal years t. The first, $\text{GuidanceOptExAnte}_{it}$, is an indicator of managers' guidance exceeding the analyst consensus. The second, $\text{GuidanceOptExPost}_{it}$, is an indicator of managers' guidance minus the realization (both in log units) exceeding the sample median.[11]

1.3.2 Measurement: Recovering Narratives from Language

We employ three techniques to measure textual narratives at different levels of granularity.

Sentiment Narratives We first measure firms' narrative sentiment. We categorize individual words as either positive or negative using the dictionaries constructed by Loughran and McDonald (2011a). These dictionaries adjust standard tools for sentiment analysis to more precisely score financial communications, in which certain words (*e.g.*, the leading example "liability") have specific definitions.[12] We first define $\mathcal{W}_P$ as the set of positive words and $\mathcal{W}_N$ as the set of negative words. For reference, we print the 20 most common words in each set in Appendix Table A.1. We calculate positive and negative sentiment as:

$$\text{pos}_{it} = \sum_{w \in \mathcal{W}_P} \text{tf}(w)_{it} \qquad \text{neg}_{it} = \sum_{w \in \mathcal{W}_N} \text{tf}(w)_{it} \tag{6}$$

where $\text{tf}(w)_{it}$ is the term frequency of all bigrams including word w in the time-t 10-K of firm i. We then construct a one-dimensional measure of net sentiment, sentiment_{it}, by computing the across-sample z-scores of both positive and negative sentiment and taking their difference. Finally, we define a firm i as being optimistic at time t if its sentiment is above the entire-sample median:

$$\text{opt}_{it} = \mathbb{I}\left[\text{sentiment}_{it} \geq \text{med}\left(\text{sentiment}_{it}\right)\right] \tag{7}$$

[11] This method corrects for the fact that, in more than half of our observations, guidance is lower than the realized value, presumably due to asymmetric incentives.

[12] Loughran and McDonald (2011a) generate the dictionaries based on human inspection of the most common words in the 10-Ks and their usage in context. We describe more details of our document scoring methodology in Appendix A.3.3.

This variable has a simple interpretation in capturing optimistic narratives, but necessarily collapses more fine-grained discussion of specific topics.

Narrative Identification of Narratives To measure more specific narratives entertained by firms, we next consider a supervised strategy based on narratively identifying a set of narratives using the text of Shiller's *Narrative Economics*. Shiller identifies a set of nine *Perennial Economic Narratives*: Panic versus Confidence; Frugality versus Conspicuous Consumption; The Gold Standard versus Bimetallism; Labor-Saving Machines Replace Many Jobs; Automation and Artificial Intelligence Replace Almost All Jobs; Real Estate Booms and Busts; Stock Market Bubbles; Boycotts, Profiteers, and Evil Businesses; and The Wage-Price Spiral and Evil Labor Unions. Each of these narratives and its history is described in its own chapter in *Narrative Economics*. We measure narrative adoption by computing the similarity between each 10-K filing and the relevant chapter of the book.

Formally, we use a method related to prior work by Hassan, Hollander, Van Lent, and Tahoun (2019) and our own implementation in Flynn and Sastry (2022a). For each narrative k, we first compute the term-frequency-inverse-document-frequency (tf-idf) score to obtain a set of words most indicative of that narrative:

$$\text{tf-idf}(w)_k = \text{tf}(w)_k \times \log\left(\frac{1}{\text{df}(w)}\right) \tag{8}$$

where $\text{tf}(w)_k$ is the number of times that word w appears in the chapter corresponding to narrative k in *Narrative Economics* and $\text{df}(w)$ is the fraction of 10-K documents containing the word. Intuitively, if a word has a higher tf-idf score, it is common in Shiller's description of a narrative but relatively uncommon in 10-K filings. We define the set of 100 words with the highest tf-idf score for narrative k as $\mathcal{W}_k$. For reference, we print the twenty most common words in each set $\mathcal{W}_k$ in Appendix Table A.2.

Finally, we score document (i, t) for narrative k by the total frequency of narrative words:

$$\widehat{\text{Shiller}}_{it}^{k} = \sum_{w \in \mathcal{W}_k} \text{tf}(w)_{it} \tag{9}$$

We compute loadings on each narrative, Shiller_{it}^{k}, by taking the z-score. These variables measure a set of more specific topics, but rely on Shiller's specific wording of narratives.

Unsupervised Recovery of Narratives Finally, to identify narratives without relying on any external references, we use Latent Dirichlet Allocation (LDA), a hierarchical Bayesian model in which documents are constructed by combining a latent set of topic narratives (Blei, Ng, and Jordan, 2003). More specifically, given our corpus of 10-Ks with M documents, we postulate that there are $K = 100$ topics. First, the number of words in each document is drawn from a Poisson distribution with parameter ξ. Second, the distribution of topics in each document is given by $\vartheta = (\vartheta_1, \ldots, \vartheta_M)$, over which we impose a Dirichlet prior with parameter $\alpha = \{\alpha_k\}_{k \in \mathcal{K}}$, where α_k represents the prior weight that topic k is in any document. Third, the distribution of words across topics is given by $\phi = (\phi_1, \ldots, \phi_K)$, over which we impose a Dirichlet prior with parameter $\beta = \{\beta_{jk}\}_{k \in \mathcal{K}}$, where β_{jk} is the prior weight that word j is in topic k. Finally, we assume that individual words in each document d are generated by first drawing a topic z from a multinomial distribution with parameter ϑ, and then selecting a word from that topic by drawing a word from a multinomial distribution with parameter ϕ_z. Intuitively, in an LDA, the set of documents is formed of a low-dimensional space of narratives of co-occurring words.

To estimate the LDA, we use the `Gensim` implementation of the variational Bayes algorithm of Hoffman, Bach, and Blei (2010), which makes estimation of LDA on our large dataset feasible, when standard Markov Chain Monte Carlo methods would be slow.[13] Given the estimated LDA, we construct the document-level narrative score as the posterior probability of that topic in the estimated document-specific topic distribution $\hat{p}$:

$$\text{topic}_{it}^{k} = \hat{p}(k|d_{it}) \tag{10}$$

For each of the eleven topics that our subsequent analysis identifies as relevant for hiring (see Section 1.4.1), we print the ten highest-weight bigrams and their weights in Appendix Table A.3. These topics are qualitatively different from the word sets used by our sentiment scoring (Appendix Table A.1) and Shiller narratives (Appendix Table A.2).

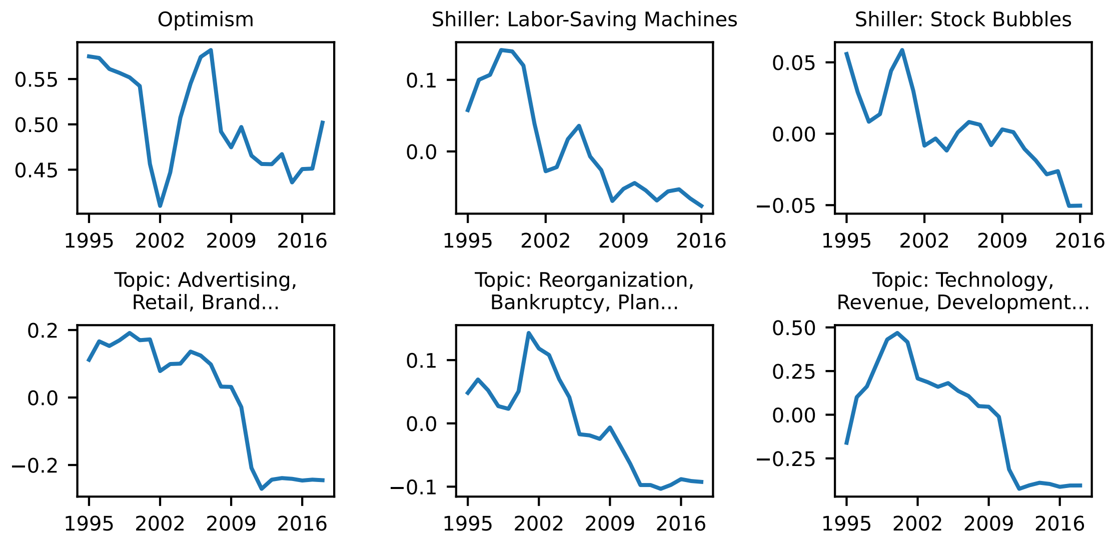
1.3.3 Descriptive Analysis of Narratives

Before our main empirical analysis, we first describe the time-series and cross-sectional structure of our measured narratives.

Time-Series Properties In Figure 1-1, we show the time path of six selected measured narratives: optimism, "Labor-Saving Machines" and "Stock Bubbles" from Shiller's perennial narratives, and three topics whose three most common terms are

[13]For computational reasons, we estimate the model using all available documents from a randomly sampled 10,000 of our 37,684 unique possible firms. We score all documents with this estimated model.

Figure 1-1: Aggregate Time Series for Six Selected Narratives



Notes: Optimism is measured as the fraction of optimistic firms. The other five time series are cross-sectional averages of z-score transformed variables (zero mean, unit standard deviation).

"Advertising, Retail, Brand"; "Reorganization, Bankruptcy, Plan"; and "Technology, Revenue, Development." Our choices among the Shiller chapters and unsupervised topics are among the set that our later analysis suggests is particularly important for explaining hiring in the cross-section. At a glance, all of these narratives are highly persistent and feature business-cycle fluctuations, and some have notable trends and breaks.[14] In Appendix Table A.4, we report summary statistics for all narratives' autocorrelation and correlation with unemployment. Almost all measured narratives are persistent, and several among the Shiller and topic sets are pro- or counter-cyclical. This observation is consistent with existing evidence in the literature on the cyclicality of aggregate text-based measures of narratives (*e.g.,* Shiller, 2020) and news coverage (*e.g.,* Baker, Bloom, and Davis, 2016; Bybee, Kelly, Manela, and Xiu, 2021a).

However, our framework implies that it is challenging to interpret these basic time-series facts for two reasons. First, it is difficult to disentangle the dual roles of narratives in driving behavior versus describing fundamentals. In the next section,

[14]In the Appendix, we report the time-series plot for Positive Sentiment, Negative Sentiment, and their difference (Figure A-1); all nine Shiller (2020) Perennial Economic Narratives (Figure A-2); and all eleven LDA topics that our later analysis identifies as relevant for hiring (Figure A-3).

we will use cross-sectional variation in narratives to isolate the impact on behavior.[15] Second, without an understanding of how narratives affect decisions and how decisions aggregate, it is difficult to understand the macroeconomic implications of even desirable time-series variation in optimism. We will later combine our macroeconomic model with our microeconomic evidence to evaluate the business-cycle impact of aggregate variation in narratives quantitatively.

Cross-Sectional Properties Our firm-level panel allows us to explore variation that is more fine-grained than the time-series variation of Figure 1-1. We perform a variance decomposition of each narrative variable by comparing the total variance of each variable with the variance after removing means at the time, industry, industry-by-time, and firm levels. In Appendix Table A.5, we present the results in units of the fraction of variance explained by each level of fixed effects, relative to the total. Time-series variation constitutes a very small percentage of the total variation in our variables—only 1.1% for optimism, 0.2% for the median Shiller narrative, and 3.5% for the median topic narrative. Adding industry-specific trends increases these percentages, respectively, to only 6.7%, 8.7%, and 9.9%. The vast majority of variation is therefore at the firm level.

1.4 Empirical Results

Moving beyond descriptive evidence, we now use our dataset of firm-level outcomes and narrative loadings to test the two premises that narratives are decision-relevant and that narrative spread is contagious and associative.

1.4.1 Testing Premise I: Narratives Are Decision-Relevant

Empirical Strategy From the conceptual framework in Section 1.2 (see Equation 3 and Proposition 25 in Appendix A.1.1), we have that firm hiring $\Delta \log L_{it}$ can be described to first order by the following regression equation:

$$\Delta \log L_{it} = \sum_{k \in \mathcal{K}} \delta_k \lambda_{k,it} + \gamma_i + \chi_t + \varepsilon_{it} \tag{11}$$

where the $\lambda_{k,it}$ are firm-specific loadings on narratives indexed by k, γ_i is a fixed effect spanning static firm characteristics, χ_t is a fixed effect spanning aggregate conditions (including both fundamentals and the distribution of narratives), and ε_{it} is a residual term arising from idiosyncratic noise in individuals' signals.

[15] In Table A.28 we show that failing to control for aggregate time-series variation leads to an upward bias in the impact of optimism on firms' hiring, as predicted by the theory.

We first operationalize this by estimating the following regression equation relating hiring to our optimism variable constructed in the 10-Ks:

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \tag{12}$$

Hiring and optimism are constructed as described in Section 1.3, at the level of firms and fiscal years. We augment our theoretically implied specification (Equation 11) with controls, including industry-by-time fixed effects and a suite of firm-level time-varying controls X_{it} (current and past TFP, lagged labor, and financial variables). We later estimate analogs of Equation 12 with other estimated narratives as independent variables.

Main Result: Optimism Drives Decisions We present our estimates of Equation 12 in Table 1.1. We first estimate the model with no additional controls beyond fixed effects and find a point estimate of $\hat{\delta}^{OP} = 0.0355$ with a standard error of 0.0030 (column 1). In column 2, we add controls for current and lagged TFP, and lagged labor $(\log \hat{\theta}_{it}, \log \hat{\theta}_{i,t-1}, \log L_{i,t-1})$. These controls proxy both for time-varying firm fundamentals and, to first order, the presence of adjustment costs in labor.[16] Our point estimate $\hat{\delta}^{OP} = 0.0305$ (SE: 0.0030) is quantitatively comparable to our uncontrolled estimate. To formalize this, in Appendix A.5.1 we report the robustness of our estimate to selection on unobservables using the method of Oster (2019). We find that our finding of a positive effect of optimism on hiring is robust by the benchmark suggested by Oster (2019) (see Table A.6).

In column 3, we add measures of firms' financial characteristics, the (log) book-to-market ratio, last fiscal year's log stock return (inclusive of dividends), and leverage (total debt over total assets). These controls proxy for both Tobin's q and firm-level financial frictions. These controls are conservative in that they may absorb variation in both omitted firm fundamentals and optimism itself. The point estimate remains positive and quantitatively similar. In column 4, we estimate a specification with the controls from column 2 but no firm fixed effects to guard against small-sample bias from strict exogeneity violations (Nickell, 1981) and find similar results.[17]

To test if optimism predicts (and does not merely describe) hiring, we finally

[16]In Appendix A.2.9, we show that the controls capture the impact of adjustment costs, to first order, for a forward-looking firm that observes current productivity. This notwithstanding, to evaluate robustness to the presence of richer adjustment dynamics, in Table A.7, we control for up to three lags of productivity and labor and our financial controls and continue to find a significant impact of optimism on hiring.

[17]In Table A.8, we report standard errors for the estimates of Table 1.1 under alternative schemes for clustering standard errors.

Table 1.1: Narrative Optimism Predicts Hiring

	(1)	(2)	(3)	(4)	(5)
	Outcome is				
	$\Delta \log L_{it}$				$\Delta \log L_{i,t+1}$
opt_{it}	0.0355	0.0305	0.0250	0.0322	0.0216
	(0.0030)	(0.0030)	(0.0032)	(0.0028)	(0.0037)
Firm FE	✓	✓	✓		✓
Industry-by-time FE	✓	✓	✓	✓	✓
Lag labor		✓	✓	✓	✓
Current and lag TFP		✓	✓	✓	✓
Log Book to Market			✓		
Stock Return			✓		
Leverage			✓		
N	71,161	39,298	33,589	40,580	38,402
R^2	0.259	0.401	0.419	0.142	0.398

Notes: For columns 1-4, the regression model is Equation 12 and the outcome is the log change in firms' employment from year $t-1$ to t. The main regressor is a binary indicator for the optimistic narrative, defined in Section 1.3.2. In all specifications, we trim the 1% and 99% tails of the outcome variable. In column 5, the regression model is Equation 13, the outcome is the log change in firms' employment from year t to $t+1$, and control variables are dated $t+1$. Standard errors are two-way clustered by firm ID and industry-year.

estimate a specification in which the outcome and control variables are time-shifted one year in advance:

$$\Delta \log L_{i,t+1} = \delta_{-1}^{OP} \text{opt}_{it} + \tau' X_{i,t+1} + \gamma_i + \chi_{j(i),t+1} + \varepsilon_{i,t+1} \tag{13}$$

where δ_{-1}^{OP} is the effect of lagged optimism on hiring and the (time-shifted) control variables $X_{i,t+1}$ are those studied in column 2. In this specification, hiring takes place in fiscal year $t+1$ after the filing of the 10-K at the end of fiscal year t. Our point estimate in column 5 is similar in magnitude to our comparable baseline estimate (column 2).[18]

Robustness As an alternative strategy to isolate plausibly exogenous variation in the narratives considered by firms, we study the effects on hiring of changes in narratives induced by plausibly exogenous CEO turnover. We provide the details in Appendix A.5.2. Specifically, we estimate a variant of Equation 12 over firm-year observations corresponding to the death, illness, or voluntary retirement of a CEO, as measured by Gentry, Harrison, Quigley, and Boivie (2021). We find quantitatively similar effects of narrative optimism on hiring as those reported in Table 1.1.

In the Appendix, we also report several additional results that probe the robustness of our main specification. We summarize them briefly here. First, Figure A-4 shows estimates of a variant of our baseline regression interacting optimism with quartiles of firm characteristics. We find that the effect of optimism is decreasing in capital intensity, essentially flat in market capitalization, and U-shaped in the book-to-market ratio (*i.e.,* high for both growth firms and value firms). Second, Table A.10 repeats the analysis of Table 1.1 with our conference-call-based optimism measure, and finds similar results. Third, Table A.11 repeats our main analysis for different measured inputs—employment (the baseline), total variable input expenditure, and investment—and demonstrates a positive and comparably sized effect of optimism on all three. Thus, optimism expands operations uniformly across inputs. Finally, we have so far studied the impact of binary optimism on hiring. To check if this construction drives our results, in Figure A-5 we re-create the regression models of the first three columns of Table 1.1 with indicators for each decile of the continuous sentiment measure. We find monotonically increasing associations of hiring with higher bins of sentiment, implying that our binary construction is not masking non-monotone

[18]In Table A.9, we report results from our baseline regression Equation 12, using $\text{opt}_{i,t-1}$ as an instrument for opt_{it}. This is robust to any identification concern arising from the simultaneous determination of opt_{it} and $\Delta \log L_{it}$, but estimates the original parameter δ^{OP} rather than δ_{-1}^{OP}. Our estimates are positive, statistically significant, and larger than our baseline estimates.

effects of the continuous measure.[19]

Inspecting the Mechanism: Narrative Optimism Does Not Predict Future Productivity Growth, Predicts Negative Stock Returns and Profitability The coefficient of interest, δ^{OP}, measures the impact of optimism on hiring if optimism is uncorrelated with any omitted fundamental factors that affect hiring. We have already demonstrated that controlling for firm-level productivity, current labor employed, and financial variables has minimal impact on the estimated value of δ^{OP}. Thus, any correlation between measured optimism and measured contemporaneous and lagged fundamentals does not generate quantitatively significant omitted variables bias. But we have not yet systematically investigated those correlations, or more formally explored whether measured optimism captures news about future fundamentals.

To investigate these issues, we estimate projection regressions of firm-fundamentals Z_{it}, either TFP growth $\Delta \log \hat{\theta}_{it}$, log stock returns R_{it}, or changes in profitability $\Delta\pi_{it}$, on optimism at leads and lags k:[20]

$$Z_{it} = \beta_k \,\mathrm{opt}_{i,t-k} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{14}$$

For negative k, β_k measures the relationship of current fundamentals with future optimism. For positive k, β_k measures the relationship of current fundamentals with past optimism.

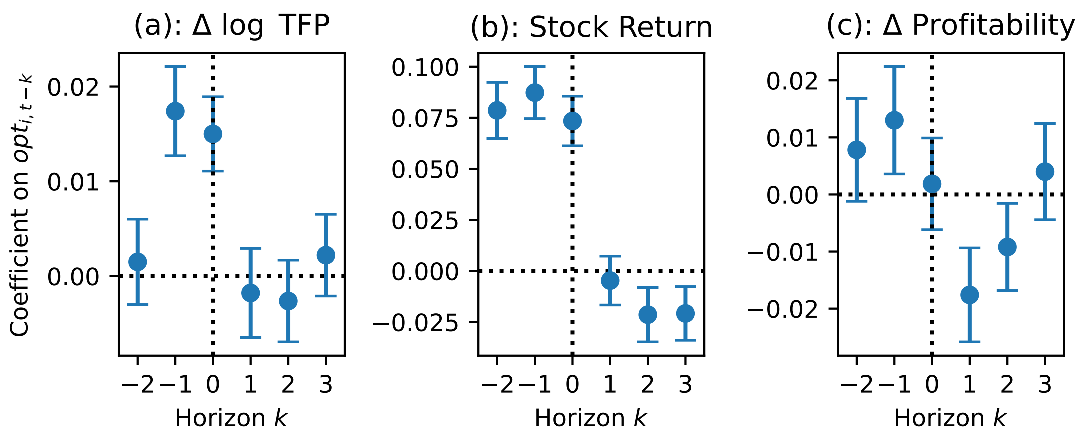
We show our findings graphically in Figure 1-2, in which each point is a coefficient from a separate estimation of Equation 14 and the error bars are 95% confidence intervals. For $k < 0$, and all three outcome variables, we find evidence of $\beta_k > 0$—that is, a firm doing well today in terms of TFP growth, stock-market returns, and/or profitability is likely to become optimistic in the future. However, for $k > 0$, and all three outcome variables, we find no positive association—that is, a firm doing well today was not on average optimistic yesterday, or a firm that is optimistic today does not on average do better tomorrow. This is consistent with our required exclusion restriction that our narrative measure of optimism is non-fundamental, and it is inconsistent with a story in which optimism is driven by news about fundamentals.[21]

[19] In Appendix A.5.4, we also check whether the effects of narrative optimism depend on the past level of narrative optimism. We find that there are larger marginal effects on average for recently pessimistic firms, but that this heterogeneity is quantitatively small.

[20] We measure profitability as earnings before interest and taxes (EBIT) divided by the previous fiscal year's total variable costs (cost of goods sold (COGS) plus selling, general, and administrative expense (SGA), minus depreciation).

[21] To further investigate the effects on stock prices, we also estimate the correlation of optimism

Figure 1-2: Dynamic Relationship of Optimism with Firm Fundamentals



Notes: The regression model is Equation 14, and each coefficient estimate is from a different regression. The outcomes are (a) the log change in TFP, calculated as described in Appendix A.4.2, (b) the log stock return inclusive of dividends, and (c) changes in profitability, defined as earnings before interest and taxes (EBIT) as a fraction of the previous fiscal year's variable costs. In all specifications, we trim the 1% and 99% tails of the outcome variables. Error bars are 95% confidence intervals, based on standard errors clustered at the firm and industry-year level.

Table 1.2: Narrative Optimism Predicts Over-Optimistic Forecasts

	(1)	(2)	(3)	(4)
	Outcome is			
	$\text{GuidanceOptExPost}_{i,t+1}$		$\text{GuidanceOptExAnte}_{i,t+1}$	
opt_{it}	0.0354	0.0561	0.0267	-0.000272
	(0.0184)	(0.0257)	(0.0231)	(0.0353)
Ind.-by-time FE	✓	✓	✓	✓
Lag labor		✓		✓
Current and lag TFP		✓		✓
N	3,817	2,159	3,044	1,718
R^2	0.173	0.193	0.161	0.192

Notes: The regression model is Equation 15. The outcomes are binary indicators for whether sales guidance was high relative to realized sales (columns 1 and 2) or high relative to contemporaneous analyst forecasts (columns 3 and 4), as defined in Section 1.3.1. Standard errors are two-way clustered by firm ID and industry-year.

We find, in sharp contrast, that optimistic firms have *negative* stock returns and decreasing profitability in the future. This is consistent with our finding that optimistic firms persistently increase input expenditure (column 5 of Table 1.1), but see no increase in productivity (panel (a) of Figure 1-2). Figures A-6 and A-7 replicate this analysis with conference-call-based optimism and the continuous measure of net sentiment, respectively, and find similar results.[22]

This analysis focuses on real, rather than financial, fundamentals. In Figure A-8, we investigate the relationship between narrative optimism, leverage, the capital structure, and payout policy. We find that narrative optimism predicts higher leverage and higher borrowing and has no effect on both equity issuance and payouts. Taken together, this provides evidence that narrative optimism is associated with tighter, instead of looser, future financial conditions. This is again inconsistent with a view that narrative optimism drives increases in hiring because it is correlated with positive news about future firm-level financial conditions.

Inspecting the Mechanism: Narrative Optimism Predicts Optimistic Beliefs In our theoretical framework, optimistic narratives increase hiring by increasing

with stock returns near the 10-K filing date (Appendix Table A.12). We find essentially no evidence of stock response on or before the filing day, and weak evidence of positive returns (about 15-25 basis points) in the four days after. The latter finding is consistent with those in Loughran and McDonald (2011a).

[22] Jiang, Lee, Martin, and Zhou (2019) relatedly find that positive textual sentiment in firm disclosures, by their own measure, predicts negative excess returns over the subsequent year.

firms' expectations about fundamentals. To test this mechanism, we investigate the relationship between narrative optimism and the extent to which firms make more optimistic forecasts. As described in Section 1.3.1, we define $\text{GuidanceOptExPost}_{i,t+1}$ and $\text{GuidanceOptExAnte}_{i,t+1}$ to indicate firms' optimism at the beginning of fiscal year $t+1$ relative to realized sales and contemporaneous sales forecasts of equity analysts, respectively. For each variable $\text{GuidanceOpt}_{i,t+1}$, we estimate the following regression model:

$$\text{GuidanceOpt}_{i,t+1} = \beta\, \text{opt}_{it} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it} \tag{15}$$

The control variables X_{it} are current and lagged TFP and lagged labor, all in log units. As we have guidance data for only a small subset of firms, we do not include firm fixed effects.

Our findings are reported in Table 1.2. For optimism relative to realizations, we find a positive correlation that increases when we add the aforementioned control variables (columns 1 and 2). This is consistent with the notion that firms producing an optimistic 10-K truly hold optimistic views about firm performance. For optimism relative to analysts, we find an imprecise positive effect in an uncontrolled model and a zero effect in the controlled model. These findings are consistent with a story in which optimism is shared between management and investors, potentially due to persuasion in communications.[23,24]

Given that we have found that guidance correlates with narrative optimism, it is natural to ask if narrative optimism affects firm decisions conditional on guidance (and *vice versa*). In Appendix A.5.3, we find that narrative optimism and measured expectations each have predictive power conditional on the other for explaining hiring and capital investment. These results suggest that textual optimism captures aspects of managers' latent beliefs that are not represented in traditional measurement of expectations (here, in guidance data).

The Narratives that Matter for Decisions We now study the decision-relevance of the measured Shiller (2020) and topic narratives. Specifically, for each of the two sets of narratives, we estimate the regression equation implied by our theoretical

[23]In Table A.13, we re-estimate this relationship with alternative measurement schemes. We find a positive relationship between *ex post* optimism and continuous sentiment, and an insignificant relationship between binary or continuous narrative sentiment with the continuous difference between guidance and realized sales.

[24]Loughran and McDonald (2011a) similarly find that, in Fama-MacBeth predictive regressions of standardized unexpected earnings, 10-K negativity predicts higher earnings surprises in the subsequent quarter.

Table 1.3: Narratives Selected as Relevant for Hiring by LASSO

Shiller (2020) Chapters	Topics
1. Labor-Saving Machines	1. Lease, Tenant, Landlord
2. Stock Bubbles	2. Business, Public, Combination
	3. Value, Fair, Loss
	4. Advertising, Retail, Brand
	5. Financial, Control, Internal
	6. Stock, Compensation, Tax
	7. Gaming, Service, Network
	8. Debt, Credit, Facility
	9. Reorganization, Bankruptcy, Plan
	10. Court, Settlement, District
	11. Technology, Revenue, Development

Notes: Each column lists the narrative variables chosen as relevant regressors in a Rigorous Square-Root LASSO (Belloni, Chernozhukov, Hansen, and Kozbur, 2016) estimation of Equation 16, with the baseline controls as unpenalized regressors. The Shiller (2020) chapters are named for the title of the corresponding book chapter. The topics are named after the three highest-weight bigrams. The corresponding post-LASSO estimates are reported in Table A.14.

framework:

$$\Delta \log L_{it} = \sum_{k \in \mathcal{K}} \delta_k \hat{\lambda}_{k,it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \tag{16}$$

We use our baseline controls, current and lagged TFP and lagged labor. Because we have many candidate narratives (9 and 100, respectively), and we expect only a few to matter for decisions, we apply the Rigorous Square-Root LASSO method of Belloni, Chen, Chernozhukov, and Hansen (2012) and Belloni, Chernozhukov, Hansen, and Kozbur (2016) to estimate the subset of hiring-relevant narratives. In Table 1.3, we list the selected Shiller (2020) and topic narratives. In Appendix Table A.14, we report the post-LASSO OLS estimates of Equation 16 with the selected variables.

Among the Shiller (2020) Perennial Economic Narratives, the LASSO methodology selects two of nine as quantitatively relevant for hiring: "Labor-Saving Machines" and "Stock Bubbles." Among the unsupervised topics, the LASSO methodology selects eleven variables out of 100. In Table A.14, we present these topics in the (essentially random) order they come out of our LDA exercise and identify them by their three highest-weight bigrams (in all cases, single words).[25] *Ex post*, based on their word combinations, we identify two as relating to demand conditions (Topics 4

[25] Appendix Table A.3 prints the top ten words per topic and their numerical weights.

and 7); two related to legal proceedings (Topics 9 and 10); one related to technology development (Topic 11); one related to real estate (Topic 1); and the remaining five related to financial conditions.

Are these selected narratives reasonable? For the Perennial Economic Narratives, we observe that "The Gold Standard" and "Boycotts and Evil Businesses" describe episodes in history that are unlikely to be relevant over our sample. Our results are therefore consistent with this "placebo test." There are no analogous tests for the topics, which pertain to the sample period by construction. Instead, we test if topics that seem to describe specific decisions are relevant for those decisions. Concretely, using analogs of Equation 12 to predict other firm-level decisions, we find that the "Advertising, Retail, Brand" narrative predicts SG&A expenditure growth ($\delta = 0.0076$, SE: 0.0022) and that the "Technology, Revenue, Development" narrative predicts growth in R&D spending ($\delta = 0.0402$, SE: 0.0044).

Going further, we examine the relationship of the selected Shiller and topic narratives with narrative optimism in two ways. First, we study if optimism has different effects when it coincides with more intense discussion of other narratives. To do this, for each of the thirteen hiring-relevant Shiller and topic narratives, we take our baseline regression (Equation 12) with controls for lagged labor and current and lagged TFP and add both the non-optimism narrative and its interaction with optimism. Out of our thirteen estimated regressions, the smallest p-value for an interaction coefficient that is different from zero is 0.039. Applying a Bonferroni correction for multiple hypothesis testing, we would not reject the null that all interactions are zero at any significance level less than 50%. Thus, we find limited evidence that optimism acts differently when it interacts with other, more specific narratives.

Given the irrelevance of the interaction between optimism and more specific narratives, we next consider the possibility that these narratives form a *basis* for narrative optimism: that is, emergent overall optimism is driven by discussion of more specific narratives. We estimate the following system of equations in which we treat optimism as an endogenous variable and the LASSO-selected Shiller and topic narratives (in sets $\mathcal{K}_S^*$ and $\mathcal{K}_T^*$) as instruments:

$$\begin{aligned} \Delta \log L_{it} &= \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \\ \text{opt}_{it} &= \sum_{k \in \mathcal{K}_S^*} \delta_{Sk} \text{Shiller}_{it}^k + \sum_{k \in \mathcal{K}_T^*} \delta_{Tk} \text{topic}_{it}^k + \tilde{\gamma}_i + \tilde{\chi}_{j(i),t} + \tilde{\tau}' \tilde{X}_{it} + \tilde{\varepsilon}_{it} \end{aligned} \tag{17}$$

where X_{it} are, again, our baseline controls. We provide coefficient estimates for Equation 17 in column 4 of Appendix Table A.14. The Shiller and topic narratives strongly predict optimism ($F = 189$), and our IV estimate of a 0.0597 log-point effect of optimism on hiring is larger than our baseline estimate of 0.0305.

1.4.2 Testing Premise II: Contagious and Associative Spread

Empirical Strategy From the conceptual framework in Section 1.2 (see Equation 5 and Proposition 26 in Appendix A.1.1), we have that narrative updating is described by a system of linear probability models that depend on agents' previous narrative weights, the previous narrative weights of the population, and economic outcomes.

To operationalize this idea in the context of our measured binary optimism, we first estimate the following model:

$$\mathrm{opt}_{it} = u\ \mathrm{opt}_{i,t-1} + s\ \overline{\mathrm{opt}}_{t-1} + r\ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it} \tag{18}$$

where $\overline{\mathrm{opt}}_{t-1}$ is average optimism in the previous period, $\Delta \log Y_{t-1}$ is US real GDP growth, and γ_i is an individual fixed effect. Following our earlier interpretation, u measures stubbornness, s measures contagiousness, and r measures associativeness.

Main Result: Optimism Spreads Contagiously and Associatively In column 1 of Table 1.4, we show our estimates. We find strong evidence of $u > 0$, $s > 0$, and $r > 0$—that is, firms are significantly more likely to be optimistic in year t if, in the previous year, they were optimistic, if other firms were optimistic, and if the economy grew.

Our estimation of Equation 18 levers only the time-series variation over our studied 23-year period. We therefore also study a model that allows for contagiousness and associativeness at the finer levels of our 44 industries and our firm-specific peer groups. Specifically, we estimate the equation:

$$\mathrm{opt}_{it} = u_{\mathrm{ind}}\ \mathrm{opt}_{i,t-1} + s_{\mathrm{ind}}\ \overline{\mathrm{opt}}_{j(i),t-1} + s_{\mathrm{peer}}\ \overline{\mathrm{opt}}_{p(i),t-1} + r_{\mathrm{ind}}\ \Delta \log Y_{j(i),t-1} + \gamma_i + \chi_t + \varepsilon_{it} \tag{19}$$

where $\overline{\mathrm{opt}}_{j(i),t-1}$ and $\overline{\mathrm{opt}}_{p(i),t-1}$ are (leave-one-out) means of optimism among a firm's industry and peer set, respectively, and $\Delta \log Y_{j(i),t-1}$ is the growth of sectoral value-added measured by linking Bureau of Economic Analysis (BEA) sector-level data to our NAICS-based classification.[26] The time fixed effect χ_t absorbs aggregate contagiousness and associativeness.

We show the results in columns 2 and 3 of Table 1.4. First, using just the industry-level data, we find strong evidence for contagiousness and weaker evidence for associa-

[26] These data are available only from 1997.

Table 1.4: Narrative Optimism is Contagious and Associative

	(1)	(2)	(3)
	Outcome is opt_{it}		
Own lag, $\text{opt}_{i,t-1}$	0.209	0.214	0.135
	(0.0071)	(0.0080)	(0.0166)
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290		
	(0.0578)		
Real GDP growth, $\Delta \log Y_{t-1}$	0.804		
	(0.2204)		
Industry lag, $\overline{\text{opt}}_{j(i),t-1}$		0.276	0.207
		(0.0396)	(0.0733)
Industry output growth, $\Delta \log Y_{j(i),t-1}$		0.0560	0.0549
		(0.0309)	(0.0632)
Peer lag, $\overline{\text{opt}}_{p(i),t-1}$			0.0356
			(0.0225)
Firm FE	✓	✓	✓
Time FE		✓	✓
N	64,948	52,258	8,514
R^2	0.481	0.501	0.501

Notes: The regression model is Equation 18 for column 1, and Equation 19 for columns 2 and 3. Aggregate, industry, and peer average optimism are averages of the narrative optimism variable over the respective sets of firms. Industry output growth is the log difference in sectoral value-added calculated from BEA data, linked to two-digit NAICS industries. Standard errors are two-way clustered by firm ID and industry-year.

tiveness within industries. Second, including the peer set optimism and restricting to the much smaller number of NYSE-listed firms, we find both a quantitatively similar industry-level effect and an independent peer-set effect. Moreover, the sum of coefficients $s_{\text{ind}} + s_{\text{peer}}$, the marginal effect of optimism in both the industry and peer set, is positive and strongly significant (estimate 0.243, standard error 0.075). In Table A.16, we report evidence of stubbornness, contagiousness, and associativeness with the continuous measure of sentiment and find consistent results, suggesting that our qualitative findings are not unduly sensitive to variable construction.[27]

Inspecting the Mechanism: Spillovers are Not Driven by Common Shocks The coefficients of interest, u, r, and s identify stubbornness, contagiousness, and associativeness, when idiosyncratic optimism, aggregate optimism, and GDP are unrelated to other factors that affect changes in optimistic sentiment at the firm level.

[27] In Table A.15, we report standard errors for Table 1.4 under alternative clustering.

By using lagged aggregate optimism, our estimates are not threatened by the reflection problem of Manski (1993). Nevertheless, our estimates may be contaminated by omitted variables bias because aggregate optimism is correlated with common shocks to the economy that are in the error term.

To test for this possibility, we augment our previous regressions to include controls for past and future fundamentals in the form of two leads and lags of real value-added growth at the aggregate and sectoral levels as well as firm-level TFP growth. Specifically, we estimate

$$\begin{aligned} \text{opt}_{it} = u\ \text{opt}_{i,t-1} + s\ \overline{\text{opt}}_{t-1} + \gamma_i + \\ + \sum_{k=-2}^{2} \left(\eta_k^{\text{agg}} \Delta \log Y_{t+k} + \eta_k^{\text{ind}} \Delta \log Y_{j(i),t+k} + \eta_k^{\text{firm}} \Delta \log \theta_{i,t+k} \right) + \varepsilon_{it} \end{aligned} \tag{20}$$

We estimate an analogous specification at the industry level, but with the aggregate leads and lags absorbed. If common positive shocks to the economy and sectors were driving some or all of the estimated spillovers, we would expect to find a severely attenuated estimate of the contagiousness coefficient s. Even under our interpretation, future output growth could be a "bad control" that is caused by optimism and absorbs some of its effect.

We report our estimates of the contagiousness coefficients in Table 1.5, adding the "bad controls" one at a time. In column 2 we find that instead of attenuating $\hat{s}$, controlling for past and future aggregate fundamentals in fact slightly increases the original point estimate reported in column 1 (within one standard error of the original value). In columns 3 and 4, when we additionally control for sectoral-level value-added growth and firm-level TFP growth, the point estimates drop slightly while standard errors increase significantly. Similarly, for our industry-level estimates, we find no statistically significant evidence of coefficient attenuation as additional controls are added (columns 5 to 7). In Table A.17, we report analogous estimates with the continuous sentiment variable and find similar results. Taken together, these estimates build confidence that our baseline contagiousness results are not driven by omitted aggregate shocks.

To further test whether our measure of contagiousness captures spillovers, and not omitted common shocks, we pursue two additional instrumental variables strategies. First, in Appendix A.5.2, we use spillovers from the same plausibly exogenous CEO changes to construct instruments for industry and peer-set optimism. We find similar point estimates as in our main analysis. Second, in Appendix A.5.5, we use size-

Table 1.5: Narrative Optimism is Contagious, Controlling for Past and Future Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Outcome is opt_{it}			
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290	0.339	0.235	0.222			
	(0.0578)	(0.0763)	(0.1278)	(0.2044)			
Ind. lag, $\overline{\text{opt}}_{j(i),t-1}$					0.276	0.241	0.262
					(0.0396)	(0.0434)	(0.0705)
Firm FE	✓	✓	✓	✓	✓	✓	✓
Time FE					✓	✓	✓
Own lag, $\text{opt}_{i,t-1}$	✓	✓	✓	✓	✓	✓	✓
$(\Delta \log Y_{t+k})_{k=-2}^{2}$		✓	✓	✓			
$(\Delta \log Y_{j(i),t+k})_{k=-2}^{2}$			✓	✓		✓	✓
$(\Delta \log \hat{\theta}_{i,t+k})_{k=-2}^{2}$				✓			✓
N	64,948	49,631	38,132	13,272	52,258	38,132	13,272
R^2	0.481	0.484	0.497	0.543	0.501	0.498	0.545

Notes: The regression model is Equation 20 for columns 1-4, and an analogous industry-level specification for columns 5-7 (*i.e.,* Equation 19 with past and future controls). Columns 1 and 5 are "baseline estimates" corresponding, respectively, with columns 1 and 3 of Table 1.4. The added control variables are two leads, two lags, and the contemporaneous value of: real GDP growth (columns 2-4), industry-level output growth (columns 3-4 and 6-7), and firm-level TFP growth (columns 4 and 7). Standard errors are two-way clustered by firm ID and industry-year.

weighted idiosyncratic shocks to firm-level optimism as an instrument for aggregate size-weighted optimism (a granular IV à la Gabaix and Koijen, 2020). While not comparable to our main estimates as the measure of spillovers is different, we recover a statistically significant contagiousness effect.

The Spread of Hiring-Relevant Narratives We repeat the estimation of our equation measuring aggregate associativeness and contagiousness, Equation 18, for the other thirteen narratives that are selected by our LASSO procedure as relevant for hiring. To allow for the greatest comparability with our estimates for optimism, we transform these narrative loadings into binary indicators for being above the sample median. We present our estimates of u, r, and s in the three panels of Appendix Figure A-9. We find significant evidence of stubbornness, or $u > 0$, in each case and significant evidence of contagiousness, or $s > 0$, in all but two cases. We find some evidence of associativeness ($r \neq 0$) for certain narratives, with "Lease, Tenant, Landlord" (relating to real estate), "Debt, Credit, Facility" (relating to financial conditions and leverage), and "Reorganization, Bankruptcy, Plan" (relating to firm restructuring) having $r < 0$, and "Court, Settlement, District" (relating to legal proceedings), "Business, Public, Combination" (relating to firm origination), and "Technology, Revenue, Development" (relating to R&D) having $r > 0$. In Appendix Table A.18, we instrument for optimism with the other 13 hiring-relevant narratives in the estimation of Equations 18 and 19. We find similar point estimates to our baseline analysis that are suggestive of increased contagiousness.

To assess whether these findings are consistent with our earlier findings about decision relevance, we can test whether the sign and magnitude of associativeness line up across narratives with the sign and magnitude of the hiring effect. That is, do narratives that increase firm hiring also spread more when the economy is growing? While our theory does not impose such a restriction, it might be a natural consequence of the unmodeled process by which narratives pick up associations with aggregate outcomes. In Figure A-10, we plot the associativeness coefficients for each narrative (including optimism) against the effect of the corresponding binary variable on hiring. Consistent with our hypothesis, the relationship is upward-sloping, with optimism itself in the top-right corner (with the second-highest hiring effect and highest associativeness).

1.5 A Narrative Business-Cycle Model

To study the implications of narratives for macroeconomic dynamics, we now specialize our abstract framework and develop a microfounded business-cycle model that embeds the decision-relevance, contagiousness, and associativeness of narratives.

1.5.1 Technology and Preferences

The consumption, production, and labor supply side of the model is intentionally standard and is a purely real variant of the models described in Woodford (2003b) and Galí (2008). Time is discrete and infinite, indexed by $t \in \mathbb{N}$. There is a continuum of monopolistically competitive intermediate goods firms of unit measure, indexed by i, and uniformly distributed on the interval $[0, 1]$. Intermediate goods firms have idiosyncratic (Hicks-neutral) productivity θ_{it}. They hire labor L_{it} monopsonistically at wage w_{it} to produce a differentiated variety in quantity x_{it} that they sell at price p_{it} according to the production function:

$$x_{it} = \theta_{it} L_{it}^{\alpha} \tag{21}$$

where $\alpha \in (0, 1]$ describes returns-to-scale in production.

A final goods firm competitively produces aggregate output Y_t by using a constant elasticity of substitution (CES) production function:

$$Y_t = \left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}} \, \mathrm{d}i \right)^{\frac{\epsilon}{\epsilon-1}} \tag{22}$$

where $\epsilon > 1$ is the elasticity of substitution between varieties.

A representative household consumes final goods C_t and supplies labor $\{L_{it}\}_{i\in[0,1]}$ to the intermediate goods firms with isoelastic, separable, expected discounted utility preferences:

$$\mathcal{U}\left(\{C_t, \{L_{it}\}_{i\in[0,1]}\}_{t\in\mathbb{N}}\right) = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\gamma}}{1-\gamma} - \int_{[0,1]} \frac{L_{it}^{1+\psi}}{1+\psi} \, \mathrm{d}i \right) \right] \tag{23}$$

where $\gamma \in \mathbb{R}_+$ indexes the size of income effects in the household's supply of labor and $\psi \in \mathbb{R}_+$ indexes their inverse Frisch labor supply elasticity to each firm.

Finally, we define the composite parameter:

$$\omega = \frac{\frac{1}{\epsilon} - \gamma}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \tag{24}$$

which indexes the strength of strategic complementarity. So that complementarity is positive but not so extreme that the model features multiple equilibria, we assume that $\omega \in [0, 1)$. This requires that income effects in labor supply do not overwhelm aggregate demand externalities and that these externalities are not too large.

1.5.2 Narratives and Beliefs

Firm productivity θ_{it} is comprised of a common, aggregate component θ_t, an idiosyncratic time-invariant component γ_i, and an idiosyncratic time-varying component $\tilde{\theta}_{it}$:

$$\theta_{it} = \tilde{\theta}_{it} \gamma_i \theta_t \tag{25}$$

Firms know that $\log \gamma_i \sim N(\mu_\gamma, \sigma_\gamma^2)$, know their own value of γ_i, and believe that $\log \tilde{\theta}_{it} \sim N(0, \sigma_{\tilde{\theta}}^2)$ and independently and identically distributed (IID) across firms and time. As in Angeletos and La'O (2010, 2013), firms receive idiosyncratic Gaussian signals about $\log \theta_t$ with noise $e_{it} \sim N(0, \sigma_e^2)$ that is IID across firms and time: $s_{it} = \log \theta_t + e_{it}$.

As in the conceptual framework from Section 1.2, and unlike in previous work, narratives form a common factor structure of agents' prior beliefs about the aggregate component of productivity θ_t. To best fit our main empirical analysis, we suppose that there are two competing narratives: an optimistic narrative and a pessimistic narrative. According to each narrative, the aggregate component of productivity follows:

$$\log \theta_t \sim N(\mu, \sigma^2) \tag{26}$$

where $\mu = \mu_P$ under the pessimistic narrative and $\mu = \mu_O > \mu_P$ under the optimistic narrative. Both of these narratives are potentially misspecified, and the true distribution for fundamentals is given by H. In Appendix A.2.4, we study a variant model in which narratives pertain to beliefs about idiosyncratic productivity and show that our analysis is unchanged.

Firms either believe the optimistic narrative or the pessimistic narrative. Hence, each firm's prior belief regarding the fundamental can be described as:

$$\pi_{it}(\lambda_{it}) = N\left(\lambda_{it}\mu_O + (1 - \lambda_{it})\mu_P, \sigma^2\right) \tag{27}$$

where $\lambda_{it} \in \{0, 1\}$, $\lambda_{it} = 1$ corresponds to a firm believing in the optimistic narrative, and $\lambda_{it} = 0$ corresponds to a firm believing in the pessimistic narrative. We let $Q_t = \int_{[0,1]} \lambda_{it} \, \mathrm{d}i$ correspond to the fraction of optimists in the population, which agents observe each period.

1.5.3 Narrative Evolution

To describe the evolution of narratives, we need to describe the probability that optimists remain optimistic, P_O, and the probability that pessimists become optimistic, P_P. We specify that both probabilities depend on aggregate output Y_t, the fraction of optimists in the population Q_t, and an aggregate *narrative* shock to how agents update ε_t, which has distribution G. Hence, the fraction of optimists evolves according to:

$$Q_{t+1} = Q_t P_O(\log Y_t, Q_t, \varepsilon_t) + (1 - Q_t) P_P(\log Y_t, Q_t, \varepsilon_t) \tag{28}$$

This aggregates the behavior of individual firms' narrative updating. As we found that narratives spread associatively and contagiously (Section 1.4.2), we assume that P_O and P_P are both increasing functions. As we found that firms are stubborn, or that optimism is persistent at the firm level, we assume that $P_O \geq P_P$. As associativeness and contagiousness do not explain all narrative updating, we add narrative shocks to the probabilities that optimists and pessimists update. Finally, for technical reasons, we assume that P_O and P_P are continuous and almost everywhere differentiable.

These conditions, motivated by the data, rule out some models for how individual firms update their beliefs. An important such model that is ruled out is one in which firms observe aggregate variables $\log Y_t$ and Q_t and use Bayes' rule to update their beliefs over models. As we formalize in Appendix A.2.1, this "Bayesian benchmark" rules out dependence of firms' updating on Q_t and ε_t conditional on $\log Y_t$ (respectively, contagiousness and shocks). Moreover, this "Bayesian benchmark" predicts that agents converge to holding the better-fitting empirical model exponentially quickly, which is at odds with our finding of cyclical dynamics for aggregate optimism (Figure 1-1). However, richer Bayesian models that are consistent with our empirical results can be nested by our reduced-form updating probabilities.

To illustrate our results, obtain closed-form expressions, and exactly match our regressions and quantitative model, we will study the following updating probabilities throughout:

Main Case (Linear-Associative-Contagious Updating Probabilities)**.** The linear-associative-contagious (LAC) specification for updating probabilities sets:

$$\begin{aligned} P_O(\log Y, Q, \varepsilon) &= \left[\frac{u}{2} + r\log Y + sQ + \varepsilon\right]_0^1 \\ P_P(\log Y, Q, \varepsilon) &= \left[-\frac{u}{2} + r\log Y + sQ + \varepsilon\right]_0^1 \end{aligned} \tag{29}$$

where $[z]_0^1 = \max\{\min\{z,1\},0\}$, $u \geq 0$ indexes stubbornness, $r \geq 0$ indexes associativeness, and $s \geq 0$ indexes contagiousness. △

1.5.4 Equilibrium

An equilibrium is a path for all variables:

$$\mathcal{E} = \left\{Y_t, C_t, Q_t, \theta_t, \varepsilon_t, \{L_{it}, x_{it}, p_{it}, w_{it}, \lambda_{it}, s_{it}, \tilde{\theta}_{it}\}_{i\in[0,1]}\right\}_{t\in\mathbb{N}} \tag{30}$$

such that (i) narrative weights λ_{it} follow a Markov process consistent with Equation 28 given Q_t and Y_t, (ii) x_{it} maximizes intermediate goods firms' expected profits given their narrative weights λ_{it}, signal s_{it}, and knowledge of $\mathcal{E}$, (iii) L_{it} is consistent with production technology (Equation 21) given x_{it} and $\tilde{\theta}_{it}$, (iv) prices p_{it} are consistent with profit maximization by the final goods firm, (v) wages w_{it} clear the labor market for each firm, (vi) Y_t aggregates intermediate good production according to Equation 22, (vii) C_t satisfies goods market clearing, $C_t = Y_t$, and (viii) Q_t evolves according to Equation 28.

1.6 Theoretical Results

We now study the equilibrium dynamics of narratives and output. We find that narratives induce non-fundamental fluctuations in the economy and have the potential to generate hysteresis. Moreover, we show that our empirical estimates identify the model. We use this mapping to the data to quantify and test the model's predictions in Section 1.7.

1.6.1 Characterizing Equilibrium Dynamics

To solve for equilibrium production, it suffices to solve for intermediates goods production. These firms maximize expected profits, as priced by the representative household:

$$\Pi_{it} = \mathbb{E}_{it}[C_t^{-\gamma}\left(p_{it}x_{it} - w_{it}L_{it}\right)] \tag{31}$$

where $C_t^{-\gamma}$ is the (unnormalized) stochastic discount factor that converts the profits of the firm into their marginal value to the household. The intermediate goods firm acts as a monopolist in the product market and a monopsonist in the labor market.

We first solve for the demand curve faced by the intermediates goods firms. The final goods firm maximizes profits taking as given the prices set by intermediates goods firms. This implies the following constant-price-elasticity demand curve:

$$p_{it} = Y_t^{\frac{1}{\epsilon}} x_{it}^{-\frac{1}{\epsilon}} \tag{32}$$

Increases in aggregate output shift out this demand curve via aggregate demand externalities. Second, we solve for the wage schedule faced by the intermediate goods firm. When facing a wage w_{it}, the intratemporal Euler equation of the representative household implies that labor supply is given by:

$$L_{it}^{\psi} = w_{it}C_t^{-\gamma} \tag{33}$$

Third, given the production technology of the firm, when it commits to producing x_{it}, its implied labor input is given by:

$$L_{it} = \theta_{it}^{-\frac{1}{\alpha}} x_{it}^{\frac{1}{\alpha}} \tag{34}$$

Finally, by imposing goods market clearing $C_t = Y_t$ and substituting Equations 32, 33, and 34 into Equation 31, we obtain that the intermediates goods firms solve the following profit maximization problems:

$$\max_{x_{it}} \mathbb{E}_{it} \left[Y_t^{-\gamma} \left(Y_t^{\frac{1}{\epsilon}} x_{it}^{1-\frac{1}{\epsilon}} - Y_t^{\gamma} \theta_{it}^{-\frac{1+\psi}{\alpha}} x_{it}^{\frac{1+\psi}{\alpha}} \right) \right] \tag{35}$$

By the first-order condition of this program, we have that optimal production solves:

$$\left(1 - \frac{1}{\epsilon}\right) \mathbb{E}_{it} \left[Y_t^{\frac{1}{\epsilon}-\gamma} \right] x_{it}^{-\frac{1}{\epsilon}} = \frac{1+\psi}{\alpha} \mathbb{E}_{it} \left[\theta_{it}^{-\frac{1+\psi}{\alpha}} \right] x_{it}^{\frac{1+\psi-\alpha}{\alpha}} \tag{36}$$

where the left-hand side is the marginal expected revenue of the firm from expanding production and the right-hand side is the marginal expected cost of this expansion. In this equation, a given firm's narrative affects their expected marginal costs of production, via the expectation of idiosyncratic productivity, and their expected marginal benefits of production, via the expectation of aggregate output (which encompasses aggregate demand externalities, asset pricing forces, and wage pressure). Moreover, in equilibrium, the distribution of narratives in the population affects the level of aggregate output and agents' expectations thereof.

We now take logarithms of all variables, and substitute this best reply into the production function of the final goods firm. From this, we obtain that the static equilibrium of the model is characterized by the solution to the following fixed-point equation:

$$\begin{aligned} \log Y_t = \frac{\epsilon}{\epsilon - 1} \log \mathbb{E}_t \Bigg[\exp \Bigg\{ \frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \Bigg(\log \left(\frac{1 - \frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}} \right) \\ - \log \mathbb{E}_{it} \left[\exp \left\{ -\frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] + \log \mathbb{E}_{it} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) \log Y_t \right\} \right] \Bigg) \Bigg\} \Bigg] \end{aligned} \tag{37}$$

where the outer expectation operator integrates over the realizations of productivity shocks $(\tilde{\theta}_{it}, \gamma_i)$, narrative loadings λ_{it}, and signals s_{it}.

By employing a functional guess-and-verify argument, we obtain that the model has a unique *quasi-linear* equilibrium in which log output depends linearly on log aggregate productivity and non-linearly, but separably, on the fraction of optimists in the population:

Proposition 1 (Equilibrium Characterization)**.** *There exists a unique equilibrium such that:*

$$\log Y(\log \theta_t, Q_t) = a_0 + a_1 \log \theta_t + f(Q_t) \tag{38}$$

for some coefficients a_0 and $a_1 > 0$, and a strictly increasing function f.

Proof. See Appendix A.1.2 □

Remark 1. This result claims uniqueness only within the quasi-linear class. As best replies and aggregation are non-linear and the spaces of actions and fundamentals are not compact, one cannot use classical arguments to ensure that the fixed point operator implicit in Equation 37 is a contraction. Nevertheless, in Appendix A.1.2, we show that there is a unique equilibrium when fundamentals are restricted to lie in a compact set (Lemma 7). Moreover, the claimed quasi-linear equilibrium is an $\varepsilon-$equilibrium for any $\varepsilon > 0$ for some sufficiently large support for fundamentals (Lemma 8). Hence, the quasi-linear equilibrium we study is the limit of the unique equilibrium with bounded fundamentals as the bound becomes large. This justifies our restriction in analyzing this class of equilibrium. △

Narratives Drive Non-Fundamental Fluctuations in Aggregate Output The coefficient a_1 and function f respectively describe how fundamentals and optimism drive aggregate output. In the proof of Proposition 1, we derive these objects as functions of the macroeconomic parameters $(\epsilon, \psi, \gamma, \alpha)$, the signal-to-noise ratio of agents' signals about productivity κ, and the extent of mean differences in the priors of optimists and pessimists $\mu_O - \mu_P$. The effect on output from going from full pessimism to full optimism is given by

$$f(1) = \frac{\alpha \delta^{OP}}{1 - \omega} \tag{39}$$

where δ^{OP} is the average partial equilibrium effect of a firm being optimistic on hiring, the returns-to-scale parameter α converts this into the effect on production, and $\frac{1}{1-\omega}$ is the general equilibrium multiplier of this effect.

The role of optimism in equilibrium has two subtle properties. First, the effect of optimism on output, $f(Q)$, is non-linear. The non-linearity arises from the fact that firms' heterogeneous priors induce heterogeneity in production conditional on productivity and hence also misallocation. Second, there is an equilibrium multiplier for optimism due to demand externalities. In particular, even a pessimistic firm will produce more if a large fraction of *other* firms is optimistic, as this optimism increases aggregate demand.

Identification of Model Parameters We now show how our empirical strategy identifies the aggregate behavior of output and optimism.

Corollary 1 (Identification of Model Parameters). *In equilibrium, firms' hiring decisions obey the following equation:*

$$\Delta \log L_{it} = c_{0,i} + c_1 \log \theta_t + c_2 f(Q_t) + c_3 \log \theta_{it} + c_4 \log L_{i,t-1} + \delta^{OP} \lambda_{it} + \zeta_{it} \quad (40)$$

where ζ_{it} is an IID normal random variable with zero mean. Thus, conditional on $(\alpha, \epsilon, \gamma, \psi)$, δ^{OP} uniquely identifies f, the equilibrium effect of optimism on aggregate output.

Proof. See Appendix A.1.4. □

This clarifies the exact interpretation of our regression model for hiring, Equation 12, in the model. The general-equilibrium effect of optimism on hiring, $c_2 f(Q_t)$, was absorbed in the regression equation as a time fixed effect. As we previously discussed, aggregate fundamentals also appear in the time fixed effect of the regression. These facts highlight formally the necessity of combining cross-sectional variation and some structural model for general-equilibrium interaction to identify the effect of optimism on economic outcomes.

The Dynamics of Optimism Finally, using Proposition 1, we express the dynamics of the economy as a first-order nonlinear stochastic difference equation for aggregate optimism:

Corollary 2 (Characterization of Dynamics). *Optimism evolves according to the following stochastic, nonlinear first-order difference equation $Q_{t+1} = T(Q_t, \nu_t)$, where $\nu_t = (\log \theta_t, \varepsilon_t)$ and*

$$T(Q_t, \nu_t) = Q_t P_O(a_0 + a_1 \log \theta_t + f(Q_t), Q_t, \varepsilon_t) + (1 - Q_t) P_P(a_0 + a_1 \log \theta_t + f(Q_t), Q_t, \varepsilon_t) \quad (41)$$

Proof. See Appendix A.1.5 □

1.6.2 Dynamics: Steady-State Multiplicity and Hysteresis

We next characterize the steady states of optimism and their stability, for fixed aggregate fundamentals. This analysis highlights how associative, contagious optimism affects dynamics even in the absence of shocks.

Steady-State Characterization Let T be the equilibrium transition map from Corollary 2 and $T_\theta(Q) = T(Q, \theta, 0)$ be the map for a fixed value of aggregate productivity when there is no narrative shock. A level of optimism Q^*_θ is a deterministic steady state for level of productivity θ if it is a fixed point of the corresponding map, $T_\theta(Q^*_\theta) = Q^*_\theta$. The following result establishes that a deterministic steady state always exists and provides necessary and sufficient conditions for extreme optimism and pessimism to be (stable) steady states.

Proposition 2 (Steady State Existence, Multiplicity, and Stability)**.** *The following statements are true:*

1. *There exists a deterministic steady-state level of optimism for every $\theta \in \Theta$*

2. *There exist thresholds θ_P and θ_O such that: $Q = 0$ is a deterministic steady state for θ if and only if $\theta \leq \theta_P$ and $Q = 1$ is a deterministic steady state for θ if and only if $\theta \geq \theta_O$. Moreover, these thresholds are given by:*

$$\theta_P = \exp\left\{\frac{P_P^{-1}(0;0) - a_0}{a_1}\right\} \quad \textit{and} \quad \theta_O = \exp\left\{\frac{P_O^{-1}(1;1) - a_0 - f(1)}{a_1}\right\} \tag{42}$$

   where $P_P^{-1}(x;Q) = \sup\{Y : P_P(Y,Q,0) = x\}$ and $P_O^{-1}(x;Q) = \inf\{Y : P_O(Y,Q,0) = x\}$.[28]

3. *Extreme pessimism is stable if $\theta < \theta_P$ and $P_O(P_P^{-1}(0;0),0,0) < 1$ and extreme optimism is stable if $\theta > \theta_O$ and $P_P(P_O^{-1}(1;1),1,0) > 0$.*

Proof. See Appendix A.1.6. □

This result establishes conditions under which extreme optimism and pessimism can be stable steady states. These conditions can be checked with only a few parameters: the responsiveness of output to productivity a_1, its baseline level a_0, the impact of all agents being optimistic on output $f(1)$, the highest level of output such that all pessimists remain pessimistic when everyone is a pessimist $P_P^{-1}(0;0)$, and the lowest level of output such that all optimists remain optimistic when all other agents are optimists $P_O^{-1}(1;1)$.

Hysteresis Proposition 2 demonstrates the possibility for hysteresis: multiple steady states of optimism that are entirely self-fulfilling. Thus, differing initial conditions of narratives in the population can lead to differing steady-state levels of narrative penetration and therefore output. The following corollary characterizes exactly when this can happen:

[28]With the convention that the infimum of an empty set is $+\infty$ and the supremum of an empty set is $-\infty$.

Corollary 3 (Characterization of Extremal Multiplicity)**.** *Extreme optimism and pessimism are simultaneously deterministic steady states for θ if and only if $\theta \in [\theta_O, \theta_P]$, which is non-empty if and only if*

$$P_O^{-1}(1;1) - P_P^{-1}(0;0) \leq f(1) \tag{43}$$

Proof. See Appendix A.1.7 □

To gain intuition for these results, and to derive a parametric condition for hysteresis that we will later empirically assess, we compute these conditions in our running LAC model:

Main Case (continuing from p. 46)**.** In the LAC special case, we can compute the sufficient statistics for narrative updating analytically. In particular, we have that extreme optimism and pessimism can coexist if and only if:

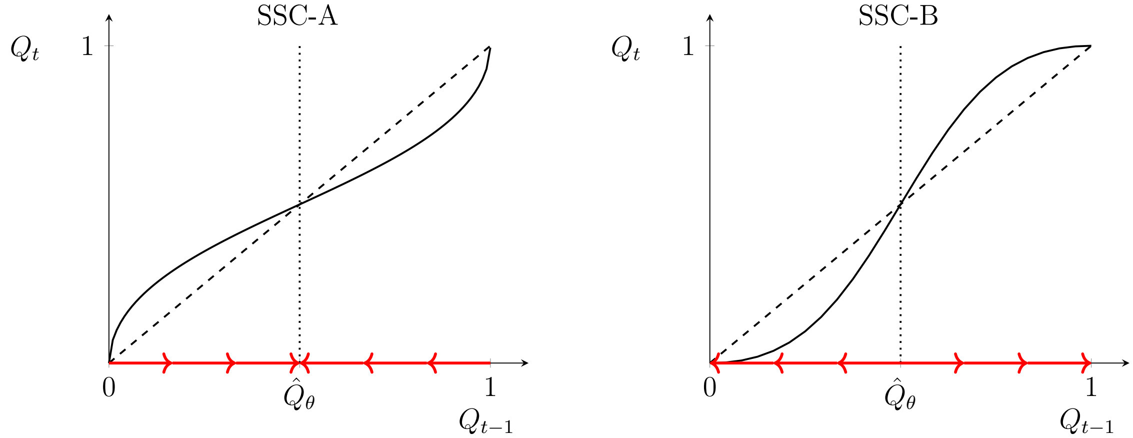
$$M = u + s + rf(1) - 1 \geq 0 \tag{44}$$

which is to say that stubbornness, associativeness, contagiousness, and the equilibrium impact of optimism on output are sufficiently large. In Section 1.7.3, we will empirically assess this condition and its quantitative implications in our calibration. △

To say more, we restrict attention to two important subclasses of updating rules that satisfy a natural single-crossing condition. We say that T is *strictly single-crossing from above* (SSC-A) if for all $\theta \in \Theta$ there exists $\hat{Q}_\theta \in [0,1]$ such that $T_\theta(Q) > Q$ for all $Q \in (0, \hat{Q}_\theta)$ and $T_\theta(Q) < Q$ for all $Q \in (\hat{Q}_\theta, 1)$. We say that T is *strictly single-crossing from below* (SSC-B) if for all $\theta \in \Theta$ there exists $\hat{Q}_\theta \in [0,1]$ such that $T_\theta(Q) > Q$ for all $Q \in (\hat{Q}_\theta, 1)$ and $T_\theta(Q) < Q$ for all $Q \in (0, \hat{Q}_\theta)$. If T is either SSC-A or SSC-B, we say that it is SSC. The left and right panels of Figure 1-3 illustrate examples of SSC-A and SSC-B transition maps as black solid lines.

Lemma 1 (Steady States under the SSC Property)**.** *If T_θ is SSC, then there exist at most three deterministic steady states. These correspond to extreme pessimism $Q = 0$, extreme optimism $Q = 1$, and intermediate optimism $Q = \hat{Q}_\theta$. Moreover, when T_θ is SSC-A: intermediate optimism is stable with a basin of attraction that includes $(0,1)$; and whenever extreme optimism or extreme pessimism are steady states that do not coincide with $\hat{Q}_\theta$, they are unstable with respective basins of attraction $\{0\}$ and $\{1\}$. When T_θ is SSC-B: whenever extreme optimism is a steady state, it is stable with basin of attraction $(\hat{Q}_\theta, 1]$; whenever extreme pessimism is a steady state it is stable*

Figure 1-3: Illustration of Steady States and Dynamics Under the SSC Property



Notes: In each subfigure, the solid line is an example transition map T_θ, the dashed line is the 45-degree line, the dotted vertical line indicates the interior steady state $\hat{Q}_\theta$, and the red arrows indicate the dynamics. The subfigures respectively correspond to SSC-A ("strict single crossing from above") and SSC-B ("strict single crossing from below"), as defined in the text.

with basin of attraction $[0, \hat{Q}_\theta)$*; and intermediate optimism is always unstable with basin of attraction* $\{\hat{Q}_\theta\}$*.*

Proof. See Appendix A.1.8 □

In the SSC-A case there is a unique, (almost) globally stable steady state (left panel of Figure 1-3). In the SSC-B class, there exists a state-dependent criticality threshold $\hat{Q}_\theta \in [0, 1]$, below which the economy converges to extreme, self-fulfilling pessimism and above which the economy converges to extreme, self-fulfilling optimism (right panel of Figure 1-3). These two classes delineate two qualitatively different regimes for narrative dynamics: one with stable narrative convergence around a long-run steady state (SSC-A) and one with a strong role for initial conditions and hysteresis (SSC-B).

To understand the determinants of this criticality threshold, we study our LAC model:

Main Case (continuing from p. 46)**.** The LAC model satisfies SSC-B if $u, r,$ and s are sufficiently large and $\theta \in (\theta_O, \theta_P)$. Moreover, in the SSC-B case, the criticality threshold is given by the unique solution to the equation:

$$\hat{Q}_\theta = \frac{u}{2}(2\hat{Q}_\theta - 1) + s\hat{Q}_\theta + r(a_0 + a_1 \log\theta + f(\hat{Q}_\theta)) \tag{45}$$

Thus, under the approximation that $f(Q) \approx kQ$, we have that:

$$\hat{Q}_\theta \approx \frac{\frac{u}{2} - r(a_0 + a_1 \log \theta)}{u + s + rk - 1} \tag{46}$$

Hence, greater contagiousness, associativeness, and decision relevance make the criticality threshold lower and therefore make it easier for an epidemic of extreme optimism to take hold. △

1.6.3 Impulse Responses and Stochastic Fluctuations

Having characterized narrative dynamics with fixed fundamentals, we now study how the economy responds to deterministic and stochastic fundamental and narrative shocks. For this analysis, we restrict attention to the SSC class, noting that this is an assumption solely on primitives.[29]

Hump-Shaped and Discontinuous Impulse Responses We consider the responses of aggregate output and optimism in the economy to a one-time positive shock to fundamentals from a steady state corresponding to $\theta = 1$:

$$\theta_t = \begin{cases} 1, & t = 0, \\ \hat{\theta}, & t = 1, \\ 1, & t \geq 2. \end{cases} \tag{47}$$

where $\hat{\theta} > 1$. We would like to understand when the impulse response to a one-time shock is *hump-shaped*, meaning that there exists a $\hat{t} \geq 2$ such that Y_t is increasing for $t \leq \hat{t}$ and decreasing thereafter. Moreover, we would like to understand how big a shock needs to be to send the economy from one steady state to another, as manifested as a discontinuity in the IRFs in the shock size $\hat{\theta}$.

In the SSC-A case, IRFs are continuous in the shock but can nevertheless display hump-shaped dynamics as a result of the endogenous evolution of optimism.

Proposition 3 (SSC-A Impulse Response Functions)**.** *In the SSC-A case, suppose*

[29] This is without a substantive loss of generality as we can always represent any non-SSC T_θ as the concatenation of a set of restricted functions that are SSC on their respective domains. Concretely, whenever T_θ is not SSC, we can represent its domain $[0,1]$ as a collection of intervals $\{I_j\}_{j\in\mathcal{J}}$ such that $\cup_{j\in\mathcal{J}} I_j = [0,1]$ and the restricted functions $T_{\theta,j} : I_j \to [0,1]$ defined by the property that $T_{\theta,j}(Q) = T_\theta(Q)$ for all $Q \in I_j$ are either SSC-A or SSC-B for all $j \in \mathcal{J}$. Thus, applying our results to these restricted functions, we have a complete description of the global dynamics.

that $Q_0 = \hat{Q}_1 \in (0,1)$*. The impulse response of the economy is given by:*

$$\log Y_t = \begin{cases} a_0 + f(\hat{Q}_1), & t = 0, \\ a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \qquad Q_t = \begin{cases} \hat{Q}_1, & t \leq 1, \\ Q_2, & t = 2, \\ T_1(Q_{t-1}), & t \geq 3. \end{cases} \tag{48}$$

Moreover, $Q_2 = \hat{Q}_1 P_O(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) + (1 - \hat{Q}_1) P_P(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) > \hat{Q}_1$*,* Q_t *is monotonically declining for all* $t \geq 2$*, and* $Q_t \to \hat{Q}_1$*. The IRF is hump-shaped if and only if* $\hat{\theta} < \exp\{(f(Q_2) - f(\hat{Q}_1))/a_1\}$*.*

Proof. See Appendix A.1.9 □

All persistence in the IRF of output derives from persistence in the IRF of optimism. There is a hump in the IRF for output if the boom induced by optimism exceeds the direct effect of the shock. This contrasts with the SSC-B case, wherein impulse responses can be discontinuous in the shock size. The following proposition characterizes the IRFs from the pessimistic steady state; those from the optimistic steady state are analogous.

Proposition 4 (SSC-B Impulse Response Functions)**.** *In the SSC-B case, suppose that* $\theta_O < 1 < \theta_P$ *and that* $Q_0 = 0$*. The impulse response of the economy is given by:*

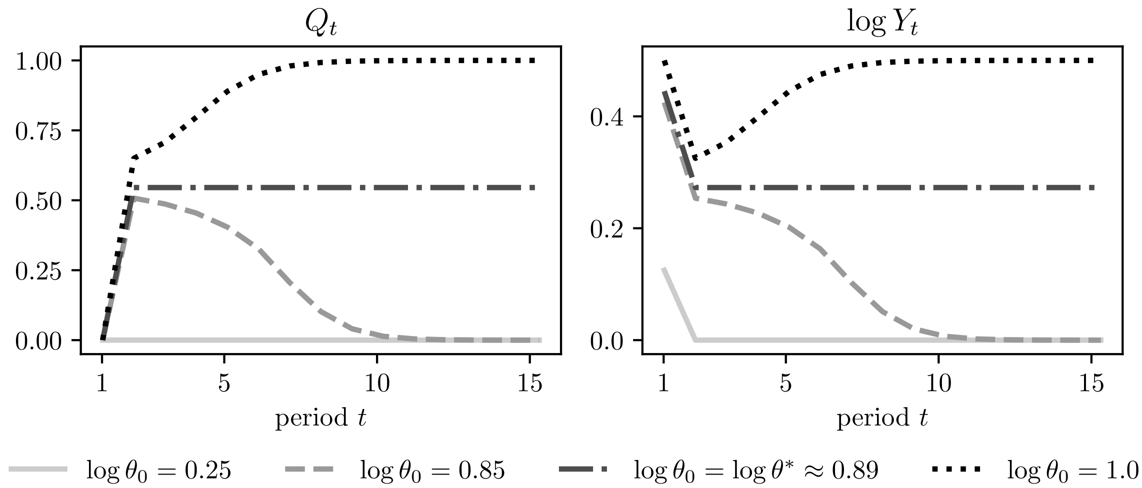
$$\log Y_t = \begin{cases} a_0, & t = 0, \\ a_0 + a_1 \log \hat{\theta}, & t = 1, \\ a_0 + f(Q_t), & t \geq 2 \end{cases} \qquad Q_t = \begin{cases} 0, & t \leq 1, \\ P_P(a_0 + a_1 \log \hat{\theta}, 0, 0), & t = 2, \\ T_1(Q_{t-1}), & t \geq 3. \end{cases} \tag{49}$$

These impulse responses fall into the following four exhaustive cases:

1. $\hat{\theta} \leq \theta_P$*, No Lift-Off:* $Q_t = 0$ *for all* $t \in \mathbb{N}$*.*
2. $\hat{\theta} \in (\theta_P, \theta^*)$*, Transitory Impact:* Q_t *is monotonically declining for all* $t \geq 2$ *and* $Q_t \to 0$*.*
3. $\hat{\theta} = \theta^*$*, Permanent (Knife-edge) Impact:* $Q_t = \hat{Q}_1$ *for all* $t \geq 1$
4. $\hat{\theta} > \theta^*$*, Permanent Impact:* Q_t *is monotonically increasing for all* $t \geq 2$ *and* $Q_t \to 1$

where the critical shock threshold is $\theta^* = \exp\{(P_P^{-1}(\hat{Q}_1; 0) - a_0)/a_1\} > \theta_P$*. In the transitory case, the output IRF is hump-shaped if and only if* $\hat{\theta} < \exp\{f(P_P(a_0 + a_1 \log \hat{\theta}, 0, 0))/a_1\}$*.*

Figure 1-4: Illustration of IRFs in an SSC-B Case



Notes: The plots show the deterministic impulse responses of Q_t and $\log Y_t$ in a model calibration with LAC updating. The four initial conditions correspond to the four cases of Proposition 4.

Proof. See Appendix A.1.10. □

To understand this result, we first inspect the IRFs. At time $t = 0$, the economy lies at a steady state of extreme pessimism with $\log \theta_0 = 0$ and so $\log Y_0 = a_0$. At time $t = 1$, the one-time productivity shock takes place and output jumps up to $\log Y_1 = a_0 + a_1 \log \hat{\theta}$ as everyone remains pessimistic. At time $t = 2$, agents observe that output rose in the previous period. As a result, a fraction $P_P(\log Y_1, 0)$ of the population becomes optimistic. For output, the one-time productivity shock has dissipated, so output is now given by its unshocked baseline a_0 plus the equilibrium output effect of optimism $f(Q_2)$. From this point, the IRF evolves deterministically and its long-run behavior depends solely on whether the fraction that became initially optimistic exceeds the criticality threshold $\hat{Q}_1$ that delineates the basins of attraction of the steady states of extreme optimism and extreme pessimism.

As a result, productivity shocks have the potential for the following four qualitatively distinct effects, described in Proposition 4 and illustrated numerically in Figure 1-4. First, if a shock is small and no agent is moved toward optimism, the shock has a one-period impact on aggregate output. Second, if some agents are moved to optimism by the transitory boost to output but this fraction lies below the criticality threshold, then output steadily declines back to its pessimistic steady-state level as optimism was not sufficiently great to be self-fulfilling. Third, in the knife-edge case, optimism moves to a new (unstable) steady state and permanently increases output. Fourth, when enough agents are moved to optimism by the initial boost to output, then the economy converges to the fully optimistic steady state and optimism is completely self-fulfilling.

The impulse responses to narrative shocks are identical to those described above. One can take the formulas in Propositions 3 and 4 from $t \geq 2$ and set Q_2 equal to the value of Q that obtains following the narrative shock ε. It follows that the qualitative nature of the impulse response to a narrative shock is identical to that of a fundamental shock.

Stochastic Boom-Bust Cycles Having characterized the deterministic impulse propagation mechanisms at work in the economy, we now turn to understand the stochastic properties of the path of the economy as it is hit by fundamental and narrative shocks.

To this end, we analytically study the period of boom and bust cycles: the expected time that it takes for the economy to move from a state of extreme pessimism to a state of extreme optimism, and *vice versa*. Formally, define these expected stopping times as:

$$T_{PO} = \mathbb{E}\left[\min\{\tau \in \mathbb{N} : Q_\tau = 1\} | Q_0 = 0\right], \; T_{OP} = \mathbb{E}\left[\min\{\tau \in \mathbb{N} : Q_\tau = 0\} | Q_0 = 1\right] \tag{50}$$

where the expectation is taken under the true data generating process for the aggregate component of productivity H, which may or may not coincide with one of the narratives under consideration, and that of the narrative shocks G.

The following result provides sharp upper bounds, in the sense that they are attained for some (H, G), on these stopping times as a function of deep structural parameters:

Proposition 5 (Period of Boom-Bust Cycles)**.** *The expected regime-switching times satisfy the following inequalities:*

$$\begin{aligned} T_{PO} &\leq \frac{1}{1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^\dagger(1;0,\varepsilon) - a_0}{a_1}\right\}\right)\right]} \\ T_{OP} &\leq \frac{1}{\mathbb{E}_G\left[H\left(\exp\left\{\frac{P_O^\dagger(0;1,\varepsilon) - a_0 - f(1)}{a_1}\right\}\right)\right]} \end{aligned} \tag{51}$$

where $P_P^\dagger(x; Q, \varepsilon) = \inf\{Y : P_P(Y, Q, \varepsilon) = x\}$ *and* $P_O^\dagger(x; Q, \varepsilon) = \sup\{Y : P_O(Y, Q, \varepsilon) =$

$x\}$. *Moreover, when* $P_O^\dagger(0;1,0) - P_P^\dagger(1;0,0) \leq f(1)$*, these bounds are tight in the sense that they are attained for some processes for fundamentals and narrative shocks* (H, G).

Proof. See Appendix A.1.11 □

This result establishes that the economy regularly oscillates between times of booms and busts. We establish this result by postulating fictitious processes for optimism and showing that they bound, path-by-path, the true optimism process. This enables us to construct stopping times that dominate the true stopping times in the sense of first-order stochastic dominance and have expectations that can be computed analytically, thus providing the claimed bounds. We establish that these bounds are tight by constructing a family of distributions (H, G) such that the fictitious processes coincide always with the true processes.[30]

We can provide insights into the determinants of the period of boom-bust cycles from these analytical bounds. Concretely, consider the bound on the expected time to reach a bust from a boom. This bound is small when the quantity $\mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^\dagger(1;0,\varepsilon)-a_0}{a_1}\right\}\right)\right]$ is large, which happens when there is a fat left tail of fundamentals, when it is relatively easier for optimists to switch to pessimism as measured by $P_O^\dagger(0;1,\varepsilon^P)$, and when co-ordination motives are weak as measured by $f(1)$.

1.6.4 Additional Results and Extensions

Before proceeding to our quantification of the model, we briefly summarize additional results and extensions contained within the Appendix.

Welfare Implications So far, we have studied the positive implications of fluctuations in optimism. In Appendix A.2.2, we study the normative implications of optimism and provide conditions under which its presence is welfare improving, despite it being misspecified. Intuitively, optimism acts as an *ad valorem* price subsidy for firms, which induces firms to hire more and can undo distortions caused by market power.

Continuous Optimism Our baseline model featured, as in our main empirical specifications, only two narratives. In Appendix A.2.3, we generalize the setting studied in this section to feature a continuum of models. We show that very similar

[30]We moreover show that elements of this family can be attained by taking the limit of normal mixtures with sufficiently dispersed means. Thus, for sufficiently dispersed μ_O and μ_P, we can therefore construct (H, G) for which the bound is attained by taking weighted averages of the optimistic and pessimistic narratives, making the uncertainty under each sufficiently small, and eliminating narrative shocks.

dynamics for both real output and narratives obtain in this specification of the model and the condition for extremal multiplicity is almost identical. Thus, the qualitative and quantitative features of the baseline model carry over to this richer setting.

Multi-dimensional Narratives and Persistent Fundamentals In the conceptual framework and our measurement, we allowed for a general set of narratives. However, in our main theoretical analysis, we restricted attention to two different narratives where agents differ only in their optimism. In Appendix A.2.5, we extend the baseline model to allow for arbitrarily many narratives regarding the mean, persistence, and volatility of fundamentals, which is essentially exhaustive within the Gaussian class. We characterize quasi-linear equilibrium in this richer setting and show how qualitatively similar dynamics obtain.

Persistent Idiosyncratic Shocks and Narrative Updating Empirically, we found that firms that experience positive idiosyncratic shocks are more likely to become optimistic. In Appendix A.2.6, we extend the multi-dimensional narrative equilibrium characterization when we allow for persistent idiosyncratic states and updating that depends on realized idiosyncratic states. When idiosyncratic shocks are fully transitory, this is of no consequence and our equilibrium characterization is identical. However, when idiosyncratic shocks are persistent, the fact that narrative updating depends on idiosyncratic shock realizations induces dependence between an agent's narrative and their idiosyncratic productivity state. This matters for equilibrium output only insofar as it induces a time-varying covariance between and optimism and productivity. We find no empirical evidence for cyclicality of this covariance. We therefore abstract from this channel in our quantitative analysis.

Contrarianism, Endogenous Cycles, and Chaos While this model generates narratively driven fluctuations, it cannot generate fully endogenous cycles and chaotic dynamics. In Appendix A.2.7, we extend this model to allow for contrarianism and the possibility that pessimists may be more likely to become optimists than optimists are to remain optimists. Allowing for these features generates the possibility of endogenous cycles of arbitrary period and topological chaos (sensitivity to arbitrarily small changes in initial conditions). This model also admits a structural test for the presence of cycles and chaos that we bring directly to the data and reject at the 95% confidence level that either cycles or chaos obtain.

Narratives in Games and the Role of Higher-Order Beliefs We have studied narratively driven fluctuations in a business-cycle model, but our insights apply to co-ordination games much more generally. In Appendix A.2.8, we study contagious narratives in beauty contests Morris and Shin (2002), in which agents' best replies are a linear function of their expectations of fundamentals and the average actions of others. Many models of aggregative games in macroeconomics and finance can be recast as such games when (log-)linearized (for a review, see Angeletos and Lian, 2016). We characterize equilibrium in this context and show how optimism percolates through the hierarchy of higher-order beliefs about fundamentals.

1.7 Quantifying the Impact of Narratives

We now combine our model and empirical results to gauge the quantitative effects of narratives on business cycles and their qualitative properties. First, combining our narrative optimism time series with our calibration of partial- and general-equilibrium effects of optimism, we find that optimism explains 32% of the reduction in GDP over the early 2000s recession, 18% over the Great Recession. Next, calibrating the dynamics of optimism to match our empirical results, we find that narrative optimism generates 19% of the variance in output. Finally, we study the macroeconomic consequences of all of the decision-relevant narratives together. We reject the condition for extremal multiplicity and hysteresis for optimism, but fail to reject it for other decision-relevant narratives. We then show in an extended model how multiple latent narratives may co-evolve and drive emergent optimism. Taken together, we therefore find that contagious narratives may explain a significant fraction of the US business cycle.

1.7.1 Calibrating the Model

To obtain numerical predictions from the model, we need to know (i) the static relationship between output and optimism; (ii) the data-generating process for fundamental shocks; and (iii) the updating process for optimists and pessimists. We provide the model calibration in Table 1.6 and additional details in Appendix A.6.

First, we have shown in Section 1.6 that, to identify the static relationship between output and optimism, we need to estimate f. In turn, f requires knowledge of: δ^{OP}, the partial-equilibrium effect of optimism on hiring; α, the returns-to-scale parameter; ϵ, the elasticity of substitution between varieties; and ω, the extent of complementarity (which itself depends on γ, indexing income effects in labor supply, and ψ, the inverse Frisch elasticity of labor supply). In our main analysis, we combine our baseline regression estimate of $\hat{\delta}^{OP} = 0.0355$ (see Table 1.1) with an external calibration of α, ϵ, γ, and ψ, which together also pin down ω. In Section 1.7.2, we study the sensitivity of our results to this external calibration, and we introduce two

Table 1.6: Model Calibration

Fixed	ϵ	Elasticity of substitution	2.6
	γ	Income effects in labor supply	0
	ψ	Inverse Frisch elasticity	0.4
	α	Returns-to-scale	1
Calibrated	$\mu_O - \mu_P$	Belief effect of optimism	0.028
	κ	Signal-to-noise ratio	0.344
	ρ	Persistence of productivity	0.086
	σ	Std. dev. of the productivity innovation	0.011
	u	Stubbornness	0.208
	r	Associativeness	0.804
	s	Contagiousness	0.290
	σ_ε	Std. dev. of the optimism shock	0.044

Notes: "Fixed" parameters are externally set. "Calibrated" parameters are chosen to hit various moments. Our specific calibration methods are described in Section 1.7.1.

other calibration strategies for complementarity: using estimates of demand multipliers from the literature and inferring a demand multiplier for optimism using our own firm-level regressions.

For the external calibration, we impose that intermediate goods firms have constant returns-to-scale or $\alpha = 1$, which has been argued by Basu and Fernald (1997), Foster, Haltiwanger, and Syverson (2008), and Flynn, Traina, and Gandhi (2019) to be a reasonable assumption for large US firms. Second, as noted by Angeletos and La'O (2010), γ indexes wealth effects in labor supply, which are empirically very small (Cesarini, Lindqvist, Notowidigdo, and Östling, 2017). Hence, we set $\gamma = 0$ for our benchmark calibration. Third, we calibrate the inverse Frisch elasticity of labor supply at $\psi = 0.4$, which is within the range of standard macroeconomic estimates (Peterman, 2016). Finally, we calibrate the elasticity of substitution to match estimated markups from De Loecker, Eeckhout, and Unger (2020) of 60%, which implies that $\epsilon = 2.6$. Hence, altogether, this calibration implies an aggregate degree of strategic complementarity of $\omega = 0.49$. Finally, we observe that the calibration of δ^{OP} puts only one restriction on the parameters κ and $\mu_O - \mu_P$, respectively the signal-to-noise ratio and the impact of optimism on agents' prior means.

Second, we calibrate the process for fundamentals. To allow for persistence in both fundamentals as well as any unmodelled factors, we calibrate a case of the model with persistent fundamentals based on the analysis in Appendix A.2.5. Concretely, we suppose that $\log \theta_t$ is a Gaussian AR(1) process with persistence ρ and IID innovations

$u_t \sim N(0, \sigma^2)$:

$$\log \theta_t = \rho \log \theta_{t-1} + u_t \tag{52}$$

To obtain the law of motion of aggregate output, we require three parameters (ρ, σ, κ). We calibrate these to match the properties of *fundamental output*, defined as

$$\log Y_t^f = \log Y_t - f(Q_t) \tag{53}$$

In Appendix A.6, we show that $\log Y_t^f$ follows an ARMA$(1, 1)$ with white noise process ζ_t. To calculate $\log Y_t^f$ in the data, we take $\log Y_t$ as band-pass filtered US real GDP (Baxter and King, 1999), Q_t as our measured time series of aggregate optimism (see Figure 1-1), and f as our calibrated function.[31] We estimate by maximum-likelihood the ARMA(1,1) process for Y_t^f and then set (ρ, σ, κ) to exactly match the three estimated ARMA parameters. Upon obtaining κ, the restriction on κ and $\mu_O - \mu_P$ imposed by δ^{OP} yields the value of $\mu_O - \mu_P$.

Third, we calibrate the process for updating probabilities. We use our regression estimates of the LAC form (see Equation 29) which corresponds to the linear probability model (see Table 1.4).[32] This yields values of $u = 0.208$ for stubbornness, $r = 0.804$ for associativeness, and $s = 0.290$ for contagiousness. We finally calibrate a shock process for ε_t, the aggregate shocks for proclivity toward optimism. In particular, we assume that $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$. Conditional on the rest of the calibration, we set σ_ε^2 to match the time-series variance of optimism.

1.7.2 How Does Optimism Shape the Business Cycle?

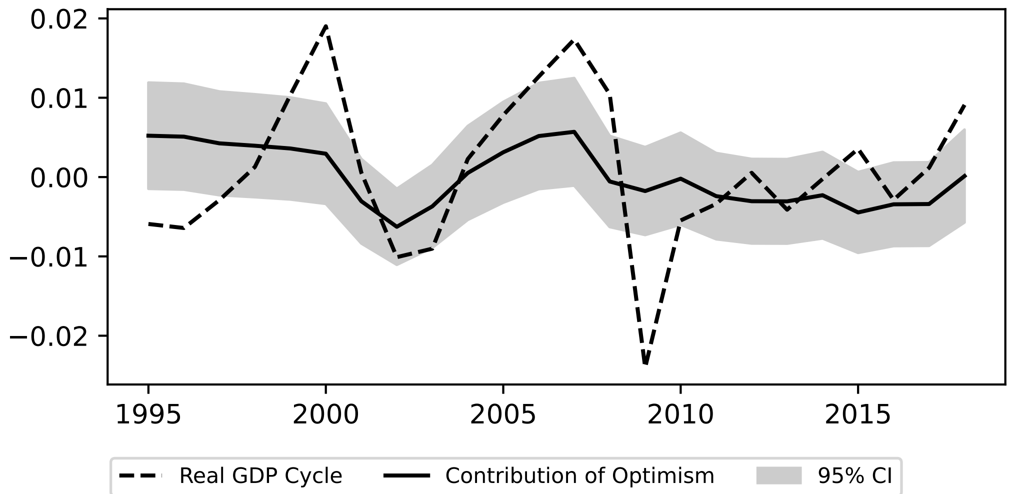
Using the calibrated model, we now study the effects of optimism on the business cycle via two complementary approaches: (i) gauging the historical effect of swings in business optimism on US GDP and (ii) exploring the full dynamic implications of contagious and associative optimism in simulations.

The Effects of Optimism on US GDP In our empirical exercise, which leveraged cross-sectional data on US firms' optimism, the general-equilibrium effect of optimism on total production was the unidentified "missing intercept." Now, equipped with the model calibration of general-equilibrium forces, we can return to the question of how

[31] We apply the Baxter and King (1999) band-pass filter to post-war quarterly US real GDP data (Q1 1947 to Q1 2022). We use a lead-lag length of 12 quarters, a low period of 6 quarters, and a high period of 32 quarters. We then average these data to the annual level.

[32] While the linear probability model does not necessarily yield probabilities between zero and one, our estimates of u, r and s imply updating probabilities that are always between zero and one so long as output does not deviate by more than 30% (holding fixed ε) or there is a five-standard-deviation optimism shock (holding fixed output at steady state).

Figure 1-5: The Effect of Optimism on Historical US GDP



Notes: The "Real GDP Cycle" is calculated from a Baxter and King (1999) band-pass filter capturing periods between 6 and 32 quarters. The "Contribution of Optimism" is the model-implied effect of optimism on log output. The 95% confidence interval incorporates uncertainty from the calibration of δ^{OP} using the delta method.

changes in optimism have historically affected the US business cycle. Concretely, we calculate the time series of $f(Q_t)$, where f is the calibrated function mapping aggregate optimism to aggregate output, which depends on the partial-equilibrium effect of optimism on hiring, returns-to-scale, and the demand multiplier, and Q_t observed annual time series for business optimism, originally reported in Figure 1-1. We take the observed time path of aggregate optimism as given, and therefore use the estimated dynamics of optimism only to determine the shocks that rationalize this observed path.

Figure 1-5 illustrates our findings by plotting the cyclical component of real GDP (dashed line) and the contribution of measured optimism toward output according to our model (solid line with grey 95% confidence interval). We observe that cyclical optimism explains a meaningful portion of fluctuations, particularly the booms of the mid-1990s and the mid-2000s and the busts of 2000-2002 and 2007-2009.

We next zoom in on the contribution of optimism toward macroeconomic crashes in 2000-02 and 2007-09. Over each of these two downturns, we calculate the percentage of output reduction explained by the dynamics of optimism as

$$\% \text{ Explained}(t_0, t_1) = 100 \cdot \frac{f(Q_{t_1}) - f(Q_{t_0})}{\log Y_{t_1} - \log Y_{t_0}} \tag{54}$$

Table 1.7: The Effect of Optimism on US Recessions

Period	Change (%) Detrended GDP	Change (%) Optimism Component $f(Q_t)$	% Explained
2000-2002	-2.91	-0.92	31.65
		(0.08)	(2.68)
2007-2009	-4.13	-0.75	18.06
		(0.06)	(1.53)

Notes: The first column gives the change in detrended, annualized real GDP over the stated periods. The second column gives the component of this change attributed to the change in aggregate optimism by the model. The third column is the fraction of the real GDP change explained by optimism, defined in Equation 54. Standard errors for columns 2 and 3 incorporate uncertainty from estimating δ^{OP} and are calculated using the delta method.

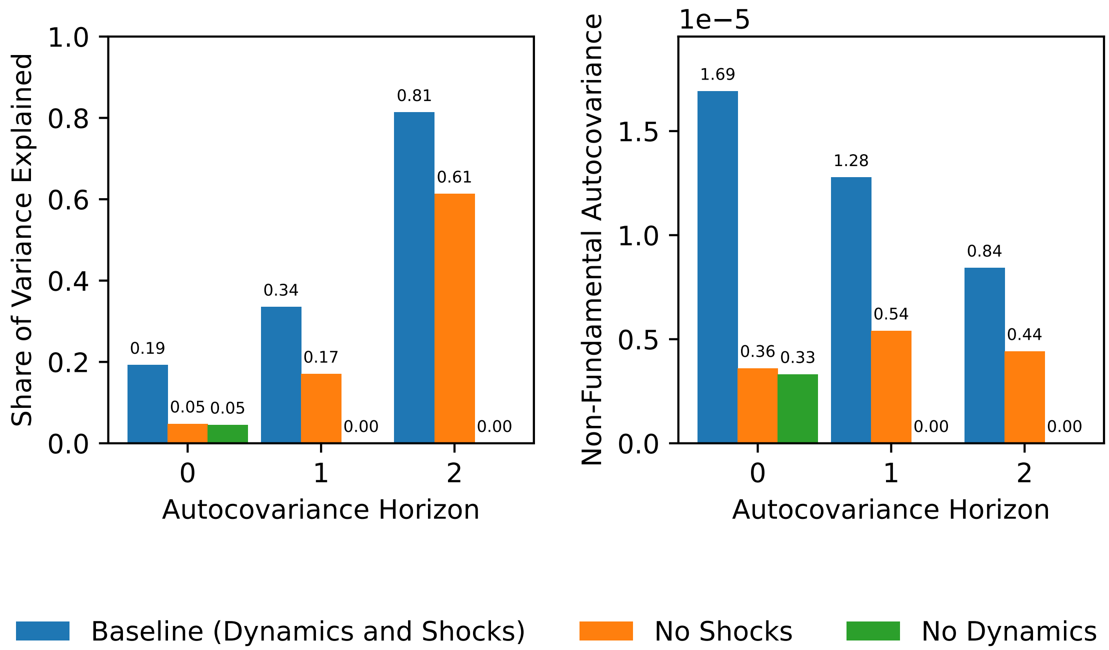
where Q is measured optimism, $\log Y$ is the measured cyclical component of log real GDP, and (t_0, t_1) are the endpoints. We report these results in Table 1.7. The decline in the optimism component of GDP explains 31.65% of the output loss between 2000 and 2002, and 18.06% of the output loss between 2007 and 2009.

To more systematically gauge the model-implied causes of the historical business cycle, we plot the sequence of fundamental output and optimism shocks that our model requires to match the realized optimism and output time series in Figure A-11. Our model accounts for the early 2000s recession with a large negative optimism shock—$\varepsilon_{2001} = -0.08$, or -1.8 standard deviations in our calibration—and a moderate-sized shock to fundamental output. For the Great Recession, our model implies a larger shock to fundamentals along with a smaller optimism shock —$\varepsilon_{2008} = -0.06$ or -1.4 standard deviations. The larger contribution of, and shock to, optimism at the outset of the early 2000s recession is consistent with a story that a break in confidence, associated with the "dot com" crash in the stock market, spurred a recession despite sound economic fundamentals. This is further consistent with independent textual evidence that "crash narratives" in financial news were especially rampant in this period (Goetzmann, Kim, and Shiller, 2022).

Contagious Narratives and Economic Fluctuations: Simulation Results To more fully describe the role of narrative dynamics in shaping the business cycle, we turn to model simulations which incorporate the fully calibrated process for how optimism spreads.

To produce a summary statistic for the contribution of optimism toward the covariance structure of output, we observe that the covariance of output at lag $\ell \geq 0$

Figure 1-6: The Contribution of Optimism to Output Variance



Notes: The left panel plots the fraction of variance, one-year autocovariance, and two-year autocovariance explained by endogenous optimism in model simulations. The right panel plots the total non-fundamental autocovariance. Both quantities are defined in Equation 56. In each figure, we plot results under three model scenarios: the baseline model with optimism shocks and optimism dynamics (blue), a variant model with no shocks, or $\sigma_\varepsilon^2 = 0$ (orange), and a variant model with shocks but no dynamics for narrative spread, or $u = r = s = 0$ (green).

can be decomposed into four terms:

$$\begin{aligned}\mathrm{Cov}[\log Y_t, \log Y_{t-\ell}] &= \mathrm{Cov}[\log Y_t^f, \log Y_{t-\ell}^f] + \mathrm{Cov}[f(Q_t), f(Q_{t-\ell})] \\ &\quad + \mathrm{Cov}[f(Q_t), Y_{t-\ell}] + \mathrm{Cov}[f(Q_{t-\ell}), Y_t]\end{aligned} \tag{55}$$

The first term captures the volatility and persistence of exogenous fundamentals (*i.e.*, the driving productivity shocks). The second term captures the volatility and persistence of the non-fundamental component of output. The last two terms capture the relationship of optimism with past and future fundamentals, which arises from the co-evolution of narratives with economic outcomes.

We therefore define non-fundamental variance as the total autocovariance arising from endogenous optimism as the sum of the last three terms, as well as its fraction of total variance, at each lag ℓ:

$$\text{Non-Fundamental Autocovariance}_\ell = \text{Cov}[\log Y_t, \log Y_{t-\ell}] - \text{Cov}[\log Y_t^f, \log Y_{t-\ell}^f]$$
$$\text{Share of Variance Explained}_\ell = \frac{\text{Non-Fundamental Autocovariance}_\ell}{\text{Cov}[\log Y_t, \log Y_{t-\ell}]} \tag{56}$$

We calculate these statistics at horizons $\ell \in \{0, 1, 2\}$ and under three model variants: the baseline model with shocks, a variant model which turns off the shocks (or sets $\sigma_\epsilon^2 = 0$), and a variant model that keeps the shocks but turns off the endogenous evolution of narratives (by setting $u = r = s = 0$).[33]

We report our findings in Figure 1-6. Optimism explains 19% of contemporary variance ($\ell = 0$), and this fraction increases with the lag. At one-year and two-year lags, optimism explains 34% and 81% of output autocovariance, respectively. Thus, most medium-frequency (two-year) dynamics are produced by contagious optimism instead of fundamentals. The model without endogenous dynameics of optimism explains only 4% of output variance and, as optimism shocks are IID, 0% of output auto-covariance. Moreover, while the model without optimism shocks matches only 5% of output variance, it accounts for 17% and 61% of one-year and two-year output autocovariance. Interestingly, the separate contributions to output variance of shocks and endogenous dynamics sum to less than one half of their joint explanatory power. This result establishes that the contagiousness and associativeness of narratives are amplifying propagation mechanisms for exogenous sentiment shocks.

Parametric Sensitivity Analysis In Table A.19, we report sensitivity analysis of the conclusions above to different calibrations for the macroeconomic parameters. We first focus on the calibration of macroeconomic complementarity and, by extension, the demand multiplier. Recall that $f(Q) \approx \frac{\alpha\delta^{OP}}{1-\omega}Q$, where $\frac{1}{1-\omega}$ is the general equilibrium demand multiplier in our economy, α indexes the returns-to-scale, and δ^{OP} is the partial equilibrium effect of optimism on hiring. Our baseline calibration implies a multiplier of $\frac{1}{1-\omega} = 1.96$. In rows 1, 2, 3, and 4 we vary the multiplier by: (i) adjusting the inverse-Frisch elasticity 2.5 to match micro estimates (Peterman, 2016), (ii) allowing for greater income effects in labor supply $\gamma = 1$, (iii) matching the empirical estimates of the demand multiplier of 1.33 from Flynn, Patterson, and Sturm (2021), and (iv) estimating the general equilibrium multiplier by optimism semi-structurally by using the extent of omitted variables bias from omitting a time fixed effect in the regression of hiring on optimism (see Appendix A.6.3 for the details). Our numer-

[33] As discussed in Appendix A.6, we always add a constant to LAC updating so 0.5 is the interior steady-state when output is at its steady state. Thus, the "no dynamics" variant sets $Q_{t+1} = 0.5 + \epsilon_t$.

ical results from adjusting the multiplier, holding fixed $(\delta^{OP}, \alpha, \epsilon)$, convey that the contribution of optimism is increasing in this number. We finally consider sensitivity to the calibrations of the elasticity of substitution ϵ (row 5 of Table A.19) and the returns-to-scale α (row 6 of Table A.19) holding fixed the multiplier (via adjustment in ψ). Changing ϵ has close to no effect on our results, due to the aforementioned near-linearity of f. Reducing α, or assuming decreasing returns to scale, dampens the effect of optimism on output because it implies a smaller production effect of our estimated effect of optimism on hiring.

1.7.3 Can Contagious Optimism Generate Hysteresis?

We have shown that the dynamics of optimism generate quantitatively significant business cycles. However, we have not yet explored the implications of narratives for hysteresis and long-run movements in output. Our theoretical analysis delimited two qualitatively different regimes for macroeconomic dynamics with contagious optimism: one with stochastic fluctuations around a stable steady state, and one with hysteresis and (almost) global convergence to extreme steady states. In which regime do the estimated dynamics lie?

For the LAC case which we have taken to the data, the necessary and sufficient condition for extremal multiplicity is given by Equation 44. We compute the empirical analog of this condition as:

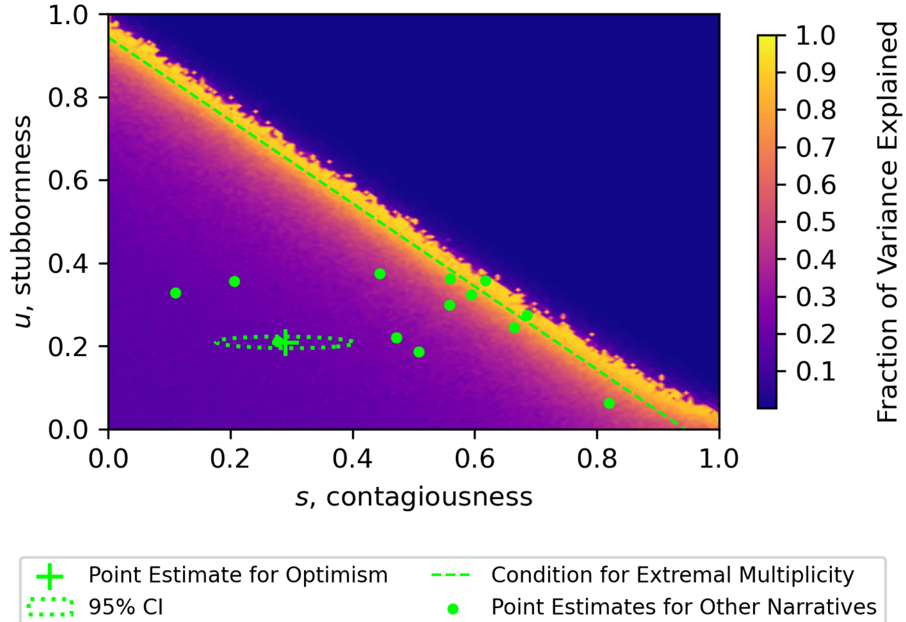
$$\hat{M} = \hat{u} + \hat{s} + \hat{r} f(1) - 1 \tag{57}$$

If $\hat{M} > 0$, the calibrated model features hysteresis in the dynamics of optimism and output; if $\hat{M} < 0$, the model features oscillations around a stable steady state. We find $\hat{M} = -0.44 < 0$ with a standard error of 0.052, implying stable oscillations and ruling out hysteresis dynamics. This reflects the fact that decision-relevance, stubbornness, contagiousness, and associativeness are sufficiently small for narrative optimism.

While we find that the model of narrative optimism is consistent with stable fluctuations, this conclusion could be overturned with higher stubbornness or higher contagiousness. Both parameters were somewhat imprecisely estimated in our empirical analysis. Moreover, we might suspect that they vary over time. For example, the rise of the internet and the corresponding increase in the speed with which ideas can spread may have increased contagiousness.

Therefore, we explore more carefully the sensitivity of our conclusions regarding optimism's role for the business cycle to the calibration of these parameters. In

Figure 1-7: Variance Decomposition for Different Values of Stubbornness and Contagiousness



Notes: Calculations vary u and s, holding fixed all other parameters at their calibrated values. The shading corresponds to the fraction of variance explained by optimism, or Share of Variance Explained$_0$ defined in Equation 56. The plus is our calibrated value from Table 1.6, and the dotted line is the boundary of a 95% confidence set. The dots are calibrated values for other narratives from Figure A-9. The dashed line is the condition of extremal multiplicity from Corollary 3 and Equation 44.

Figure 1-7, we plot our point estimate of contagiousness and stubbornness as a plus and its 95% confidence interval as a dotted ellipse. We also plot, as a dashed line, the condition for $M = 0$; to the left of this line, $M < 0$, and to the right of this line, $M > 0$. To gauge whether optimism is an "outlier" among narratives in being so far from this line, we compare its estimates to those corresponding to our other thirteen decision-relevant narratives (Table 1.3). The associated point estimates of (s, u) for the other nraratives are green dots in Figure 1-7. Several of the thirteen plotted points are close to the condition for extremal multiplicity. Two are across the threshold. Thus, if optimism were to have the stubbornness and contagiousness of either of these two narratives, the joint dynamics of optimism and output would feature hysteresis.[34]

[34]The narratives across the line are Topic 3 (Value, Fair, Loss) and Topic 6 (Stock, Compensation, Tax).

How does extremal multiplicity interact with our model's predictions for non-fundamental volatility? To isolate the role of endogenous propagation, our theoretical discussion of extremal multiplicity considered paths of the economy with shocks to neither fundamentals nor narratives. In the quantitative model, the economy is constantly buffeted with shocks that move optimism away from its steady state(s). To measure how this shapes macroeconomic dynamics, we also shade in Figure 1-7 the fraction of variance explained by non-fundamental optimism. In our baseline calibration, indicated by the plus, this is 19%. This figure is stable near the confidence interval associated with our calibration, suggesting that statistical uncertainty about narrative propagation does not overly influence our results.

Near the condition for extremal multiplicity, non-fundamental variance reaches essentially 100% of total variance. This is because even small shocks have the potential to "go viral," and the force pulling the economy toward an interior steady state (*i.e.,* balanced optimism and pessimism) is weak. If optimism were to have the propagation of some of our other observed narratives, which lie close to this line, it could induce such violent fluctuations.[35]

Finally, far to the right of the extremal multiplicity condition, contagious optimism explains little output variance. This is because the economy quickly settles into an extreme steady state, fully optimistic or fully pessimistic, and moves quickly back to this steady-state in response to shocks. Figure A-13 shows this quantitatively by plotting, against the same (s, u) grid, the fraction of time that the optimistic fraction Q_t lies outside of $[0.25, 0.75]$—this is 0% at the baseline calibration, and essentially 100% in the calibrations featuring extremal multiplicity. In this region, while sentiment does not greatly affect output dynamics, it does affect the *static* level of output—moreover, path dependence in the early history of our simulation determines whether output is permanently high (optimism goes viral) or permanently low (pessimism goes viral). Thus, even in a stochastic economy, M remains a highly predictive statistic for the nature of the dynamics of the economy.

[35] Due to the presence of shocks to optimism, this prediction is symmetric around the extremal multiplicity threshold; in the variant model which turns off optimism shocks, which is closer to what we studied in the theory, the extremal multiplicity condition sharply delineates the regime in which optimism fluctuations contribute to output variance from the regime in which there is complete hysteresis (Figure A-12).

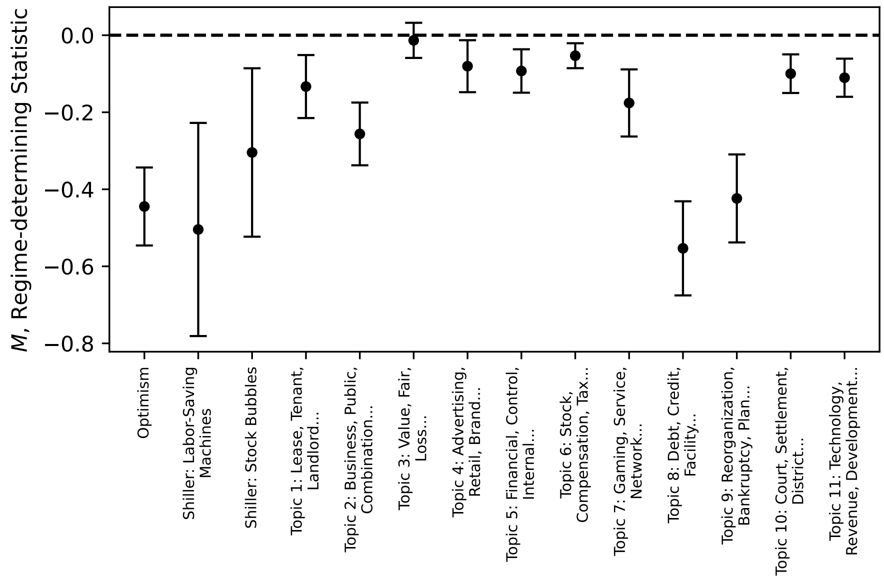
1.7.4 Multi-Dimensional Narratives: Hysteresis, Confluence, and Constellations

Our quantitative analysis thus far has focused on the overall dynamics of optimism and pessimism and their macroeconomic effects. In this final subsection, we study how multiple, more granular narratives may evolve, interact, and shape macroeconomic dynamics. To do this, we calibrate versions of the quantitative model to match the estimated properties of Shiller's Perennial Economic Narratives and our unsupervised Latent Dirichlet Allocation estimation. While this analysis is necessarily more speculative, due to the richness of possible interactions between these narratives and the number of parameters required to discipline these interactions, we highlight two overall findings: (i) the more granular narratives have a higher tendency toward hysteresis and unstable dynamics, and (ii) "constellations" of more granular narratives, with qualitatively different evolution, may underlie the overall dynamics of sentiment in the economy.

Hysteresis We first perform tests for the possibility of hysteresis for the other decision-relevant narratives. To do this, we estimate $\hat{M}^k$ (Equation 57) using the estimated stubbornness, associativeness, contagiousness, and partial equilibrium effects on hiring of the other thirteen narratives indexed by k. In models in which each of these narratives existed in isolation, the condition $M^k > 0$ would be necessary and sufficient for extremal multiplicity. We plot our estimates along with 95% confidence intervals in Figure 1-8. We find that many of our narratives have values for M that are close to zero, while one narrative's 95% confidence interval for M contains zero. Thus, while we can reject that optimism features hysteresis dynamics, we cannot do so for all of the narratives that we consider. Heuristically, this finding matches up with the corresponding time series (see Figures 1-1, A-2, A-3): while optimism appears to fluctuate in a stable fashion, other narratives have time series that appear to undergo regime shifts, with many notably happening around the Great Recession.

Confluence and Constellations in an Enriched Model Shiller (2020) argues that *constellations* of many smaller and semantically related narratives may reinforce one another to create strong economic and social effects, and that the *confluence* of seemingly unrelated narratives may explain business-cycle fluctuations. In our empirical analysis, we found two pieces of evidence suggesting that the thirteen more granular narratives discussed above behave like a constellation. First, we found that more specific narratives significantly predict movements in optimism. Second, we found that optimism about any specific narrative affects hiring no differently than

Figure 1-8: Evaluating Potential for Hysteresis for All Decision-Relevant Narratives



Notes: For the binary construction of each narrative, we estimate the parameters of the updating rule (see Figure A-9) and the partial-equilibrium effect on hiring via a variant of Equation 12 (see Figure A-10). We then calculate $\hat{M}^k = \hat{u}^k + \hat{s}^k + \hat{r}^k \hat{f}^k(1) - 1$. We report our point estimates for each narrative along with 95% confidence intervals error bars, calculated using the delta method.

optimism alone.

To understand how this constellation structure for optimism may affect our macroeconomic predictions, we calibrate a modified quantitative model in which the granular narratives drive *emergent* optimism. In this model, there is a latent space of K narratives, in which agents either believe or do not believe: $\lambda_{it} = (\lambda_{1,it}, \ldots, \lambda_{K,it}) \in \{0,1\}^K$. Based on a vector of *constellation weights* $\beta = (\beta^1, \ldots, \beta^K)$, the probability that an agent is optimistic is given by $\beta' \lambda_{it}$. Thus, the aggregate level of emergent optimism is given by $Q_t = \beta' Q_{k,t}$, where $Q_{k,t} = \int_{[0,1]} \lambda_{k,it} \, \mathrm{d}i$ is the fraction of agents that believe in narrative k. Each narrative k evolves according to LAC updating probabilities with stubbornness u^k, associativeness r^k, and contagiousness s^k, with Gaussian narrative shocks that are IID across k with common variance σ_ε^2. The rest of the model is the same as the baseline. Thus, while dynamics are the same *conditional* on the process for optimism, the process for emergent optimism through the latent evolution of narratives may differ.[36]

[36] A caveat to our approach is that we do not allow for any complementarity or substitutability in how contagious each narrative is. However, this does not mean that each narrative spreads

To calibrate the new parameters, we proceed in three steps. First, we run a firm-level regression of optimism on the loadings for the other narratives to estimate the β^k. We normalize the sign of each narrative such that its corresponding β^k is positive, for comparability. Second, we use our estimated stubbornness, associativeness, and contagiousness parameters for each decision-relevant narrative and specify an LAC updating process for each given these parameters (see Figure A-9). Finally, we calibrate the variance of the narrative shocks to match the time series variance of optimism. We describe the full details of this calibration approach in Appendix A.6.2.

We find that optimism explains 17% of total output variance, comparable to our baseline finding of 19%. In this sense, the "reduced-form" model of the entire optimism constellation delivers similar macroeconomic predictions to the more granular model. This implies that the evolution of multiple narratives can be captured through reduced-form sentiment, conditional on correctly specifying the process by which this sentiment spreads.

However, this similarity belies significant heterogeneity in how the granular narratives spread, which is in turn related to each narrative's tendency to "go viral." To illustrate this, observe that the variance in emergent optimism can be decomposed into components arising from each narrative's variance, $\text{Var}[Q_{k,t}]$, as well as a remainder arising from their covariance, which is quantitatively negligible (it equals -0.016). We find that the $\text{Var}[Q_{k,t}]$ are considerably different across narratives. In Figure A-14, we report a scatterplot of the individual $\text{Var}[Q_{k,t}]$ against the calculated $\hat{M}^k$ statistics. Narratives with high $\hat{M}^k$ have markedly higher variance contributions, reflecting their more mercurial dynamics, as predicted by the theory. Practically speaking, this means that a few narratives with a great proclivity toward dramatic swings may be significant drivers of emergent optimism. Thus, notwithstanding the fact that it is sufficient to study emergent sentiment if one wishes to understand macroeconomic dynamics, modeling more granular narratives is essential for understanding the process by which sentiment emerges.

1.8 Conclusion

This paper studies the macroeconomic implications of contagious, belief-altering narratives. We develop a conceptual framework in which narratives form building blocks

independently. As each narrative affects decisions and decisions affect output, associativeness links the dynamics of the full set of narratives. Moreover, even in the absence of associativeness, optimism that emerges from the combination of separate nonlinear processes for narrative evolution could feature different dynamics.

of agents' beliefs, affect agents' decisions, and spread contagiously and associatively between agents. We measure proxies for narratives among US firms and find evidence that narratives are decision-relevant, contagious, and associative. We find that a contagious, associative, and decision-relevant *optimism narrative* reflects a non-fundamental shifter of beliefs, corresponding to over-optimism about economic conditions. We develop a business-cycle model that embeds these findings and find that narratives can generate non-fundamentally driven boom-bust cycles, hysteresis, and impulse responses that are hump-shaped over time and discontinuous in the sizes of shocks. When we calibrate the model to match the data, we find that the business-cycle implications of narratives are quantitatively significant: we estimate that measured declines in optimism account for approximately 32% of the peak-to-trough decline in output over the early 2000s recession and 18% over the Great Recession.

An important issue that our analysis leaves unexplored is what "makes a narrative a narrative"—that is, in the language of our model, what microfounds the set of narratives and their contagiousness? A richer study of these issues would be essential to study policy issues, including both the interaction of standard macroeconomic policies with narratives and the potential effects of directly "managing narratives" via communication. Moreover, probing these deeper origins of narratives could further enrich the study of narrative constellations beyond our suggestive analysis, to account for the full economic, semantic, and psychological interactions among narratives in a complex world.

Chapter 2

Attention Cycles

This chapter is jointly authored with Karthik A. Sastry.

2.1 Introduction

Firms often make decisions which, from an observer's *ex post* perspective, are inconsistent with profit maximization. An influential explanation for this behavior, prominently articulated by Simon (1947), is that firms' managers face constraints on their attention and decisionmaking capacity. Under this view, changing microeconomic and macroeconomic conditions shape firm-level incentives in the allocation of both physical and cognitive resources. A direct implication is that the state of the economy as a whole may be central for determining the extent of firms' apparent "bounded rationality."

In this paper, we study the joint determination of *business cycles*, aggregate fluctuations in economic production, and *attention cycles*, aggregate fluctuations in cognitive effort and mistakes. Our analysis has three parts. First, we introduce two strategies to measure the attention cycle, which capture attention to the macroeconomy in "what firms say" and precision in "what firms do." We find that firms speak more about the macroeconomy and make smaller input-choice mistakes, defined relative to a model benchmark, in downturns. Second, to interpret this evidence, we develop a macroeconomic model in which firms face a cognitive cost of making precise decisions. Because firms are owned by risk-averse investors, incentives for precise profit maximization are higher when aggregate consumption is low. This *risk-pricing channel* rationalizes our finding of smaller mistakes during downturns. Third, we combine our model and evidence to quantify the effects of attention cycles on the dynamic properties of aggregate output and labor productivity. We find that state-dependent attention explains a quantitatively significant fraction of the observed asymmetry and state-dependence in the responses of these macroeconomic aggregates to shocks.

Motivating Evidence To establish the premise that firms' allocation of attention varies with the business cycle, we begin our analysis with a model-free measurement of attention to the macroeconomy in firms' language. Our dataset is the full text of all US public firms' end-of-quarter and end-of-year financial performance reports (Forms 10-Q and 10-K) from 1995-2017. We measure each filing's attention to macroeconomic developments using a natural-language-processing technique that compares the filing's word choice to that of macroeconomics references, which we take to be introductory college-level textbooks. We find that aggregate macroeconomic attention in language is counter-cyclical. This finding suggests that firms pay more attention to external conditions during downturns, but could also reflect a greater importance of the macroeconomy during downturns and/or usage of macroeconomic conditions as an excuse for poor performance. Later, to allow for a more precise interpretation of these findings, we relate our measure of attention in firms' language with a measure of mistakes in firms' input choices. To build toward that goal, we first introduce a model of firms' state-dependent attention, production and "mistake-making."

Theoretical Analysis Our model describes firms' state-dependent choices of attention and production in equilibrium. Firms face a cognitive cost in making their input choices contingent on the microeconomic and macroeconomic state, or "steadying their hand" to precisely respond to shocks in productivity, demand, or input prices. They choose state-contingent, and potentially imperfect, plans to maximize risk-adjusted profits net of this cognitive cost. Both the premise of costly planning and the emphasis on state-dependence contrasts our approach with that of the existing macroeconomic literature on decision frictions (e.g., Woodford, 2003a; Maćkowiak and Wiederholt, 2009; Angeletos and La'O, 2010; Gabaix, 2016). The environment is Neoclassical, with aggregate demand externalities (Blanchard and Kiyotaki, 1987) and an aggregate shock that shifts the productivity distribution to generate a business cycle. We show that equilibrium analysis in this setting is well-posed and tractable, proving equilibrium existence, uniqueness, and monotonicity of output in the aggregate shock. This allows us to study the causes and consequences of attention cycles in the model's unique equilibrium.

In partial equilibrium, firms pay more attention, measured by the extent of their cognitive effort, and make smaller mistakes, measured by the variance of their actions around the *ex post* ideal point, when the cost of making mistakes is high. We show that the correct metric for such costs is the curvature of agents' objective functions in their own action. When firms choose production to maximize risk-adjusted profits, the aforementioned curvature is the product of two terms. The first term is the curvature of firms' dollar profits, which is highest when aggregate output and firm-specific productivity are low in standard parameterizations. We refer to this relationship as the *profit-curvature channel*. The second term is the stochastic discount factor (SDF), which is highest when aggregate consumption is low because risk-averse households own the firms. We refer to this relationship as the *risk-pricing channel*, as it indexes the relative value of profits across states of the world.

In general equilibrium, the cyclicality of these incentives is determined by household risk aversion, the extent of aggregate demand externalities, and the elasticity of wages to real output. When household relative risk aversion exceeds an empirically modest lower bound of one plus the elasticity of real wages to output and wages are sufficiently rigid, equilibrium cognitive effort decreases in output and the average size of agents' mistakes increases in output. Put differently, market incentives push firms toward paying more attention to decisions and precisely maximizing current profits when the aggregate economy is doing poorly.

We next use the model to study the macroeconomic consequences of this mechanism. We show that output is the product of the counterfactual output under the full-attention benchmark with an *attention wedge* that is less than one. The wedge arises because inattentive, stochastic choices translate into dispersion of value marginal products ("misallocation") across firms and reduced aggregate total factor productivity (TFP). Due to counter-cyclical attention, the wedge widens when the economy is booming and firms optimize less precisely. Thus, misallocation across firms in the model is endogenously higher in booms than recessions. We show that the amplification of negative shocks via increased attention leads to asymmetric, state-dependent shock propagation and endogenous stochastic volatility.

Testing the Theory: The Misoptimization Cycle Our model's predictions for misoptimization in input choices can be taken directly to the data. In the next part of the paper, we develop and implement tests of these predictions. We use data on firm production and input choices from public firms' financial statements from 1986-2018, collected in Compustat Annual Fundamentals. We estimate firm-level TFP using conventional methods and estimate empirical policy functions for labor choice conditional on these TFP estimates, firm fixed effects, and sector-by-time fixed effects. We show in our model that the empirical policy function estimates the counterfactual "unconstrained optimal" input choice of firms and that its residuals estimate the *ex post* misoptimizations. Consistent with our interpretation, we find in the data that both positive and negative residuals have negative effects on firms' contemporaneous stock returns and profitability and that these contemporaneous negative effects do not predict higher future productivity, stock returns, or profitability. Moreover, to build confidence that our results are not driven by unmodeled features of firm behavior or by measurement error, we replicate all our empirical results for alternative constructions of misoptimizations based on richer empirical models of firm behavior (e.g., with adjustment costs or financial frictions) and alternative measurement schemes for productivity.

Our main aggregate-level finding is that the variance of firms' misoptimizations is pro-cyclical, as predicted by the main case of our model. In a linear regression of our *Misoptimization Dispersion* measure on the unemployment rate, a five percentage point increase in the latter predicts a 53% decrease in the former. The finding of pro-cyclical misoptimization holds using the aforementioned alternative constructions of misoptimizations, alternative aggregation across firms that adjusts for compositional bias, and alternative choices for the studied time period. Moreover, while the pro-cyclicality of Misoptimization Dispersion contrasts with existing evidence that firm fundamentals like TFP have counter-cyclical variance (e.g., Kehrig, 2015; Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry, 2018), we underscore that our finding is empirically and theoretically compatible with a story in which dispersion *around* fundamentals decreases in downturns while the fundamentals themselves become more volatile.

We next test the main model mechanism for counter-cyclical attention, markets' greater punishment of mistakes during downturns. Specifically, we investigate whether the negative relationship of misoptimizations with stock returns and profitability steepens in recessions. We find strong evidence in the case of returns and weak evidence in the case of profitability—that is, markets particularly reward firms for avoiding mistakes in recessions, while the profit cost of mistakes is close to constant. Interpreted via the model, this evidence points to the primacy of the risk-pricing channel (an increasing SDF) driving counter-cyclical incentives, and therefore counter-cyclical attention, relative to the profit-curvature channel (changing sensitivity of dollar profits to mistakes).

We also link back to our motivating evidence and document that macroeconomic attention in language is associated with smaller misoptimizations at the firm level. This result suggests that our initial, model-free finding of counter-cyclical "macroeconomic attention" is closely related to our subsequent, model-implied finding of pro-cyclical misoptimization.

Quantification We finally assess our findings' quantitative importance in a numerical calibration of the model. We calibrate the cognitive friction and coefficient of relative risk aversion to match the level and cyclicality of misoptimizations in the data, and use external calibrations for the substitutability between goods and elasticity of real wages to output. This approach lets the data speak directly toward our model's novel mechanism, while using standard calibrations to discipline that mechanism's interaction with other forces.

We find that negative productivity shocks have larger effects than equal-sized positive ones; all shocks have larger effects when the aggregate state is low; and macroeconomic volatility is highest in low states. In the model, negative productivity shocks have a 7% larger effect on aggregate output (5% larger on employment) than positive shocks of the same size. This amounts to 25% of the asymmetry estimated by Ilut, Kehrig, and Schneider (2018) using industry-level data on productivity and employment. Similarly, in the model, a shock that replicates a 5% peak-to-trough reduction in output, comparable to what was experienced during the Great Recession, generates also an 11% increase in the conditional volatility of output. This is 19% of the increase in statistical uncertainty about output growth between the trough of the Great Recession and the preceding peak as measured by Jurado, Ludvigson, and Ng (2015). Our model therefore explains an economically significant fraction of observed non-linearities and state-dependence of macroeconomic dynamics as consequences of state-dependent attention.

Related Literature Our study contributes to a large literature on cognitive frictions and behavioral inattention in macroeconomics.[1] Most such models (e.g., Woodford, 2003a; Maćkowiak and Wiederholt, 2009; Angeletos and La'O, 2010; Gabaix, 2016) have ignored both state-dependence of attention and its equilibrium consequences.[2] Sims (2003) and Gabaix (2014) show, in different models sharing a "rational attention" premise, that optimal costly attention should be tuned to the most payoff-relevant attributes of the decision problem. In this vein, Mäkinen and Ohl (2015), Chiang (2021), and Benhabib, Liu, and Wang (2016) share our focus on the cyclicality

[1] We model cognitive frictions via stochastic choice. The equilibrium implications of stochastic choice are discussed on a more theoretical level in Morris and Yang (2019) and Flynn and Sastry (2021). Costain and Nakov (2015, 2019) use stochastic choice models to study price-setting.

[2] Studies of costly adjustment of prices, including Gorodnichenko (2008) and Alvarez, Lippi, and Paciello (2011), or of costly adjustment of investment portfolios, including Abel, Eberly, and Panageas (2013) and Kacperczyk, Van Nieuwerburgh, and Veldkamp (2016), draw on mechanisms operating through the curvature of payoffs in different contexts.

of attention due to firms' changing incentives.[3] We, in contrast to all three studies, motivate our analysis with and quantitatively benchmark our findings against empirical evidence on cyclical misoptimization and risk-pricing incentives. Our approach of modeling stochastic choice also allows us to develop this mapping to the data and to study macroeconomic implications for state-dependent dynamics. Ilut and Valchev (2020) model cognitive constraints as costly learning of policy rules, focusing on different applications and using numerical analysis of aggregative equilibrium.

In contemporaneous work, Song and Stern (2021b) develop a similar text-based measure of firms' attention to macroeconomic news and use this measure to test a different hypothesis. Specifically, Song and Stern (2021b) show that more attentive firms, according to their text-based measure, gain more market value from expansionary monetary policy shocks and lose less value from contractionary shocks. They interpret how this heterogeneous responsiveness can drive monetary non-neutrality. Together, the two papers validate that textual measures can capture aspects of "attention" that are consistent with economic models. Our findings with textual data also relate to a literature documenting that language in regulatory filings and earnings calls contains relevant information about firm sentiment (Loughran and McDonald, 2011b) and risk exposures (Hassan, Hollander, Van Lent, and Tahoun, 2019; Hassan, Hollander, van Lent, Schwedeler, and Tahoun, 2020).

Our empirical findings relate to a literature studying the cyclicality of microeconomic dispersion. Our finding of pro-cyclical misoptimization is consistent with work by Kehrig (2015) and Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) documenting that dispersion in firm-level productivity rises in recessions in the US manufacturing sector. Concretely, we show that fundamentals are more volatile in downturns, while misoptimizations around them are less volatile.[4] Our focus on cognitive frictions and misallocation dynamics builds on prior work studying this statically (David, Hopenhayn, and Venkateswaran, 2016; David and Venkateswaran,

[3]In Mäkinen and Ohl (2015), decreasing returns to scale lead firms to demand more information when aggregate productivity is low, while we emphasize a novel risk-pricing mechanism. Contemporaneous work by Chiang (2021) has considered a complementary mechanism in the context of a model where firms acquire Gaussian signals about the state. Benhabib, Liu, and Wang (2016) predict that firms should have less demand for information during recessions, a prediction at odds with our empirical findings.

[4]Our findings are also consistent with Eisfeldt and Rampini (2006) and Bachmann and Bayer (2014), who document pro-cyclical investment-rate dispersion; Dew-Becker and Giglio (2020), who document acyclicality of implied volatility in the cross-section of firm returns; Ilut, Kehrig, and Schneider (2018), who document larger response of employment to negative versus positive productivity shocks; and Berger and Vavra (2019), who document counter-cyclical responsiveness of price-setters to nominal shocks. Macaulay (2020) similarly treats choice dispersion as evidence of misoptimization in households' cyclical attention toward savings choices.

2019; Ma, Ropele, Sraer, and Thesmar, 2020; Barrero, 2022).

Overview The rest of the paper proceeds as follows. In Section 2.2, we present our motivating evidence. In Section 2.3, we introduce our model. In Section 2.4, we present our theoretical results. In Section 2.5, we build our empirical measure of misoptimizations and test the model's predictions. In Section 2.6, we calibrate our model and analyze the impact of attention cycles on macroeconomic dynamics. Section 2.7 concludes.

2.2 Motivation: The Macroeconomic Attention Cycle

In this section, we describe our strategy to measure macroeconomic attention in firms' language and present our finding that firms discuss macroeconomic topics more in downturns.

2.2.1 Data and Measurement

Our data source is the full text of the quarterly 10-Q and annual 10-K reports submitted by all US public firms to the Securities and Exchange Commission (SEC).[5] We use data from 1995 to 2018 in our main analysis.[6] Our total sample consists of 479,403 individual documents, or about 5,000 per quarter, which we index by their date of filing.

The key challenge for identifying attention toward macroeconomic risks is to differentiate characteristic language of macroeconomics from the intrinsically economic and financial vocabulary of standard firm activities (e.g., "credit" and "costs"). To address this, we apply a simple natural language processing technique that identifies specific documents as "attentive to the macroeconomy" if their word choice is both different from the standard word choice in regulatory filings and similar to the word choice of macroeconomics-focused references. Following the method introduced by Hassan, Hollander, Van Lent, and Tahoun (2019) to study firm attention to political

[5]To give an example, the automaker General Motors in its summary report for 2009 highlighted the threats of the ongoing Great Recession to its business:"[The] deteriorating economic and market conditions that have driven the drop in vehicle sales, including declines in real estate and equity values, rising unemployment, tightened credit markets, depressed consumer confidence and weak housing markets, may not improve significantly during 2010 and may continue past 2010 and could deteriorate further."

[6]The relevant digitized documents are hosted by the Security and Exchange Commission's EDGAR (Electronic Data Gathering, Analysis, and Retrieval), which began operation in 1994. We choose 1995 as a starting point at which a nearly comprehensive sample of firms' reports are available in the system.

risks, we use introductory college-level textbooks: *Macroeconomics* and *Principles of Macroeconomics* by N. Gregory Mankiw and *Macroeconomics: Principles and Policy* by William J. Baumol and Alan S. Blinder.[7] This choice balances our considerations of keeping the relevant macroeconomic vocabulary mostly non-technical (e.g., "unemployment" instead of "tightness"), but still specific (e.g., "inflation" instead of "price").

To operationalize this method, we first define $\text{tf}(w)_{it}$ as the *term frequency* for a word w in the communication of firm i at time t, measured as the proportion of total English-language words; and $\text{df}(w)$ as the *document frequency* of a given word w among all observed (i, t) regulatory filings, measured as a proportion of total documents that use the word at least once. We define the "term frequency inverse document frequency," or tf-idf, as:

$$\text{tf-idf}(w)_{it} := \text{tf}(w)_{it} \cdot \log \left(\frac{1}{\text{df}(w)} \right) \tag{58}$$

The log functional form is a heuristic in natural language processing for scaling the relative importance of each term, and it is bounded below by 0 when a word appears in all documents.

For each word that appears in the 10-Q and 10-K corpus, we calculate the tf-idf using term frequencies in each of the three textbooks and (inverse) document frequencies among regulatory filings. We rank the top 200 words by this metric in each textbook, and take the intersection among the three books as a final set of 89 words.[8] Appendix Figure B-1 prints these words in alphabetical order, and plots their time-series frequency. Many of the words relate to common macro indicators ("unemployment","inflation"); some to the topic or profession itself ("macroeconomics","economist"); and some to policy ("Fed," "multiplier"). There are also "false positive" words that are related to pedagogy, like "question" and "equation." To allow the method to be fully devoid of direct researcher manipulation, we do not remove such words from the main analysis. We then use our set of macroeconomic words, denoted by $\mathcal{W}_M$, to calculate our firm-by-time measures of attention as the sum of

[7] We use electronic copies of the 7th, 3rd, and 12th editions of these books, respectively.

[8] Taking the intersection helps guard against the idiosyncratic language of certain books. For instance, in *Principles of Macroeconomics* by N. Gregory Mankiw, a parable about supply and demand for "ice cream" is used often enough to make "ice" and "cream" high tf-idf words in our procedure.

the (idf-weighted) macroeconomic word frequency:

$$\text{MacroAttention}_{it} = \sum_{w \in \mathcal{W}_M} \text{tf-idf}(w)_{it} \tag{59}$$

To generate an measure MacroAttention_t, we average $\text{MacroAttention}_{it}$ across firms. In our main aggregate results, we remove seasonal trends modeled as quarter-of-the-year means.

2.2.2 Attention to the Macroeconomy is Counter-Cyclical

Figure 2-1 plots the time series of our (log) macroeconomic attention metric, at the quarterly frequency and net of seasonal trends, against two macroeconomic time series: the US unemployment rate (Unemployment_t) and the linearly detrended log price of the S&P 500 ($\log \text{SPDetrend}_t$). We find that macroeconomic attention persistently rises when the macroeconomy and financial market are distressed.

To measure the cyclicality of MacroAttention, we estimate a linear regression of MacroAttention on each macroeconomic variable, or

$$\log \text{MacroAttention}_t = \alpha + \beta_Z \cdot Z_t + \epsilon_t \tag{60}$$

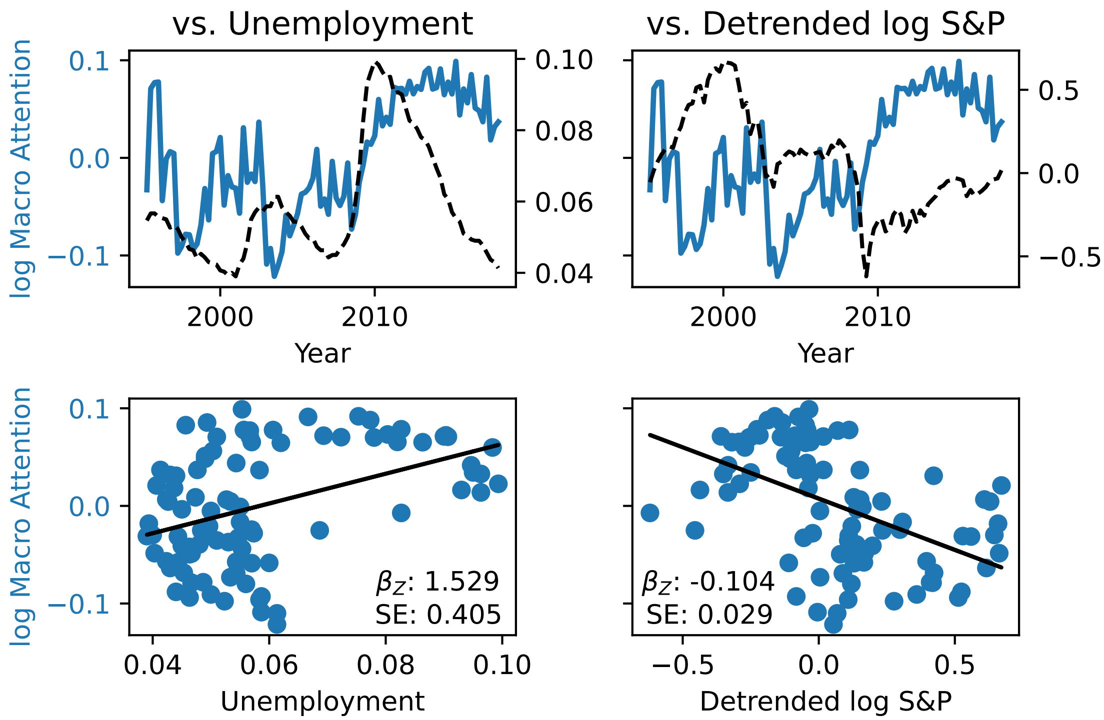
for $Z_t \in \{\text{Unemployment}_t/100, \log \text{SPDetrend}_t\}$. Our coefficient estimates are, respectively, 1.529 (SE: 0.405) and -0.104 (SE: 0.029), with R^2 values of 0.180 and 0.237. The former estimate conveys that a five percentage point swing in unemployment from peak to trough of the business cycle predicts an approximately 7.6% increase in macroeconomic attention.

To measure the persistence of Macro Attention, we estimate an AR(1) model with coefficient ρ. We find $\hat{\rho} = 0.820$ (SE: 0.050), implying a shock half-life of 3.5 quarters.

2.2.3 Discussion

We interpret the above findings as suggestive evidence that firms' attention allocation varies with the business cycle and, in particular, focuses more intensely on macroeconomic risks during downturns. But, by itself and without a model interpretation, the evidence above does not *directly* demonstrate any change in the process or outcome of firm decisionmaking over the business cycle. In Sections 2.3 and 2.4, we will write a macroeconomic model that structures the translation from "attention" to "actions," describes the macroeconomic causes and consequences of a model-consistent "attention cycle," and generates predictions that can be directly tested in firm decisions. Before proceeding, we briefly summarize robustness checks and additional exercises

Figure 2-1: Macro Attention is Counter-Cyclical



Notes: The top two panels plot log Macro Attention (blue line, left axis) along with, respectively, unemployment and the linearly detrended S&P 500 price (black dashed lines, right axis). The bottom two panels are scatterplots of log Macro Attention versus the corresponding macroeconomic aggregate. The black solid line is the linear regression fit. The standard errors are HAC robust based on a Bartlett kernel with a four-quarter bandwidth.

based on our Macro Attention measure.

Comparison with the News In Appendix Figure B-3, we compare our Macro Attention series at the quarterly frequency with the News Index of Baker, Bloom, and Davis (2016), which captures newspaper discussion of economic policy, and the "macroeconomic" sub-topics of Bybee, Kelly, Manela, and Xiu (2021b), who use machine learning to flexibly categorize the text of *Wall Street Journal* by topic. Macro Attention has, respectively, correlations of 0.21 and 0.17 with each. The news series are much more sharply peaked around turning points, while the firm attention measure is considerably more persistent. Thus, firm-level attention requires a substantially different theoretical interpretation which emphasizes more persistent incentives for attention.

Industry-Level Patterns We might expect the cyclicality of macroeconomic attention to differ across differently cyclical industries. To study this heterogeneity, we partition our sample into 44 different industries.[9] For each industry, we calculate "output cyclicality" as the correlation between sectoral GDP growth, calculated using quarterly BEA data since 2005 linked appropriately to NAICS-definition sectors, with aggregate nominal GDP growth. In Appendix Figure B-4, we plot in the cross-section of industries the relationship between this output cyclicality and the coefficient of sector-level Macro Attention on the US unemployment rate. Of the 42 industries, 36 feature counter-cyclical macroeconomic attention. The extent of counter-cyclicality increases with the industry's output cyclicality, but almost acyclical industries also have counter-cyclical attention. The model, consistent with this observation, will focus on a risk-pricing mechanism that gives firms a uniform incentive to concentrate attention in aggregate downturns regardless of the cyclicality of firms' demand.

Alternative Data Construction Appendix Figure B-1 plots the time-series behavior of each word-level component at the quarterly frequency. In our sample, 61 of the 89 words have a positive quarterly correlation with the unemployment series (Appendix Figure B-2). Appendix B.6.1 describes how we replicate our procedure using the full text of US Public Firms' sales and earnings conference calls as an alternative dataset. This produces a similar counter-cyclical pattern over a smaller time period (2004-2013). Appendix B.6.2 describes an alternative procedure for analyzing the 10-Q/Ks which uses the frequency of algorithmically determined word stems rather than full words to build our word set $\mathcal{W}_M$. This yields essentially identical results to our main procedure.

2.3 Model

We now describe our model, a Neoclassical Real Business Cycle model with a stochastic choice friction for intermediate goods firms that captures attention and mistake making.

2.3.1 Consumers and Final Goods

Time periods are indexed by $t \in \mathbb{N}$. A representative household has constant relative risk-aversion preferences over final-good consumption C_t and labor L_t. Their payoffs

[9] These classifications are based primarily on NAICS2 codes, but we separate manufacturing (NAICS 31-33) and information (NAICS 51) into three-digit categories to maintain comparable numbers of firms in each bin. The industry categorization is reviewed in more detail in Appendix B.3.1, in the context of our empirical analysis in Section 2.5.

are:

$$\mathcal{U}(\{C_{t+j}, L_{t+j}\}_{j\in\mathbb{N}}) = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j \left(\frac{C_{t+j}^{1-\gamma}}{1-\gamma} - v(L_{t+j}) \right) \tag{61}$$

where $\beta \in [0, 1)$ is the discount factor, v is an increasing and convex labor disutility, and $\gamma > 0$ is the coefficient of relative risk aversion. The aggregate final good is produced from a continuum of intermediate goods, indexed by $i \in [0, 1]$, by a representative, perfectly competitive firm using a constant-elasticity-of-substitution production function

$$X_t = X(\{x_{it}\}_{i\in[0,1]}) = \left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}} \, \mathrm{d}i \right)^{\frac{\epsilon}{\epsilon-1}} \tag{62}$$

where $\epsilon > 1$ is the elasticity of substitution. This firm buys inputs at prices $\{q_{it}\}_{i\in[0,1]}$ and sells its output at a normalized price of one. The household owns equity in firms that produce intermediate goods, receiving profits $\{\pi_{it}\}_{i\in[0,1]}$.

Wages are determined by the following wage rule:

$$w_t = \bar{w} \cdot \left(\frac{X_t}{\bar{X}} \right)^{\chi} \tag{63}$$

where $\bar{w} > 0$ and $\bar{X} > 0$ are constants, and $\chi \geq 0$ measures the extent of real wage rigidity. Households supply sufficient labor to meet firms' labor demand at the wages from Equation 63. Describing wage dynamics via a reduced-form wage rule is for technical simplicity as it allows us to study the economy via a scalar fixed-point equation that is simple to characterize.[10] Moreover, it allows our model to parsimoniously match the empirical acyclicality of real wages (Solon, Barsky, and Parker, 1994; Grigsby, Hurst, and Yildirmaz, 2019). In Appendix B.2, we microfound this wage rule when households have Greenwood, Hercowitz, and Huffman (1988) preferences and markets clear in the standard fashion. We show in Appendix B.5.1 that the quantitative results of Section 2.6 are robust to considering wages set in this manner.

2.3.2 Intermediate Goods Firms and Productivity Shocks

Each intermediate goods firm i is a monopoly producer of its own variety and faces a demand curve $d(x_{it}, X_t) = X_t^{\frac{1}{\epsilon}} x_{it}^{-\frac{1}{\epsilon}}$ from the final goods producer. They hire a labor quantity L_{it}, pay wage w_t per worker, and produce with the following linear

[10] Blanchard and Galí (2010) and Alves, Kaplan, Moll, and Violante (2020) use similar wage-rule formulations.

technology:

$$x_{it} = \theta_{it} L_{it} \tag{64}$$

where θ_{it} is a firm-level shifter of productivity, which lies in a set $\Theta \subset \mathbb{R}_+$.

We parameterize the stochastic process for firm-level productivity in the following way that captures "aggregate productivity shocks" while allowing for rich cross-firm heterogeneity. There is an aggregate productivity state $\theta_t \in \Theta$, which follows a first-order Markov process with transition density given by $h(\theta_t \mid \theta_{t-1})$. The cross-sectional productivity distribution is given in state Θ by the mapping $G : \Theta \to \Delta(\Theta)$, where we denote the productivity distribution in any state θ_t by $G_t = G(\theta_t)$ with corresponding density g_t. To give the parameter θ_t an interpretation as "aggregate productivity," we assume that the total order on θ_t ranks distributions G_t by first-order stochastic dominance, or $\theta \geq \theta'$ implies $G(\theta) \succsim_{FOSD} G(\theta')$. This can accommodate rich dynamics for all moments of the productivity distribution (e.g., state-dependent variance measured by Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry, 2018).

We proceed in the rest of our main analysis under this "Real Business Cycle" case with productivity shocks. But we will emphasize when presenting results why neither the "supply-side" view of economic dynamics nor the specific structure for productivity dynamics is essential for our main conclusions or proposed macroeconomic propagation mechanism.

2.3.3 Attention and Mistake Making

We now describe our model for attention and mistake-making by intermediate goods firms. We define the firm-level *decision state* $z_{it} = (\theta_{it}, X_t, w_t) \in \mathcal{Z}$ as the concatenation of all decision-relevant variables that the firm takes as given.[11] All firms believe that the vector z_{it} follows a first-order Markov process with transition densities described by $f : \mathcal{Z} \to \Delta(\mathcal{Z})$, with $f(z_{it}|z_{i,t-1})$ being the density of z_{it} conditional on last period's state being $z_{i,t-1}$. We denote the corresponding set of possible transition densities by $\mathcal{F}$. At time t, each firm i knows the sequence of previous $\{z_{is}\}_{s<t}$ but not the contemporaneous value z_{it}.

The Cost of Attention Intermediate goods firms choose a production level x_{it} in the feasible set $\mathcal{X}$. But, due to cognitive constraints, they struggle to match that choice to the state $z_{it} = (\theta_{it}, X_t, w_t)$ without making idiosyncratic mistakes. We model this by having them choose stochastic choice rules at a cost.

Formally, each firm chooses a stochastic choice rule $p : \mathcal{Z} \to \Delta(\mathcal{X})$ in set $\mathcal{P}$, or a

[11]As will become clear, $\mathcal{Z} = \Theta \times \mathcal{X} \times \mathcal{W}$, where $\mathcal{X}$ is feasible set of production, and $\mathcal{W}$ is the image of $\mathcal{X}$ via the wage rule (63).

mapping from states of the world to *distributions of actions* described by probability density (mass) function $p(\cdot \mid z_{it})$ when the firm-level state is z_{it}. A firm using rule p commits to delivering the realized quantity $x_{it} \in \mathcal{X}$ to the market, selling it at the (maximum) price $q_{it} = d(x_{it}, X_t)$ at which the final goods firm is willing to buy, and hiring sufficient labor in production. Helpfully, it is fully equivalent to interpret firms' choices as committing to hire $L_{it} = \frac{x_{it}}{\theta_{it}}$ workers at wage w_t, producing the maximum level $x_{it} = \theta_{it} L_{it}$, and selling at price $q_{it} = d(x_{it}, X_t)$.

We model the cost of attention via the cost functional $c : \mathcal{P} \times \Lambda \times \mathcal{Z} \times \mathcal{F} \to \mathbb{R}$ which returns how costly any given stochastic choice rule is to implement in units of utility. The cost can depend on a firm-specific type $\lambda_i \in \Lambda \subseteq \mathbb{R}_+$, by assumption independent from the decision state and distributed in the cross-section as $L \in \Delta(\Lambda)$, and the previous value of the decision state $z_{i,t-1}$, which under the Markov assumption summarizes the transition probabilities for z_{it}. The basic idea that we wish to embody, consistent with our motivation of studying costly cognition and mistake making, is that playing actions that are more *precise* in any given state is more costly.

To make this tension more clear, and also to make the analysis more tractable, we specialize to the following cost functional, which equals the negative expected entropy of the action distribution multiplied by a scaling $\lambda_i > 0$:

$$c(p, \lambda_i, z_{i,t-1}, f) = \lambda_i \int_{\mathcal{Z}} \int_{\mathcal{X}} p(x \mid z_{it}) \log(p(x \mid z_{it})) \, \mathrm{d}x \, f(z_{it} \mid z_{i,t-1}) \, \mathrm{d}z_{it} \tag{65}$$

This cost functional captures the idea that it is costly for agents to avoid "mistakes" or misoptimizations, relative to an unrestricted (costless) optimal choice.[12] In Section 2.4.5, we discuss the robustness of our results to considering alternative cost functionals that represent costly information acquisition and stochastic-choice costs denominated in units of final output.

The Firm's Problem Intermediate goods firms are owned by the representative household and maximize the product of their dollar profit, which we write as $\pi(x_{it}, z_{it})$, and the household's marginal utility, which we write as $M(X_t)$. We define "risk-adjusted profits" as the product of these terms:

$$\Pi(x_{it}, z_{it}) = \pi(x_{it}, z_{it}) \cdot M(X_t) \tag{66}$$

[12] This is an example of a *likelihood-separable cost functional* as discussed in more detail in Flynn and Sastry (2021). Also, as formalized by Fudenberg, Iijima, and Strzalecki (2015), a formulation with likelihood-separable stochastic choice is often equivalent to an additive random utility model. The formulation (65) is exactly isomorphic to a "logit demand model," or additive random utility model with Gumbel distributed perturbations, and embodies the familiar associated axioms including independence of irrelevant alternatives (IIA).

Under our assumed structure for the firms' cost and revenue structure and the household's utility function, the profit function and marginal utility are respectively

$$\pi(x_{it}, z_{it}) := x_{it}\left(x_{it}^{-\frac{1}{\epsilon}} X_t^{\frac{1}{\epsilon}} - \frac{w_t}{\theta_{it}}\right) \qquad M(X_t) = X_t^{-\gamma} \tag{67}$$

Because decisions are separable across time, and $z_{i,t-1}$ is an observed sufficient statistic for the history of state realizations, the firm can be thought to solve a series of one-shot problems of choosing a stochastic choice rule in period t, conditional on the realization $z_{i,t-1}$. The firm has a conjecture for how aggregate output and wages move over time, as embedded in their subjective prior distribution $f(z_{it} \mid z_{i,t-1})$. Given this conjecture, they play a best reply by solving the following program:

$$\max_{p\in\mathcal{P}} \left\{\int_{\mathcal{Z}}\int_{\mathcal{X}} \Pi(x, z_{it})\, p(x \mid z_{it})\, \mathrm{d}x \ f(z_{it} \mid z_{i,t-1})\, \mathrm{d}z_{it} - c\left(p, \lambda_i, z_{i,t-1}, f\right)\right\} \tag{68}$$

which is maximization of expected utility, averaged over risk in the state z_i and stochastic action x, net of the utility cost of the chosen stochastic choice rule.

2.3.4 Linear-Quadratic Approximation and Equilibrium

To tractably study equilibrium, we simplify the intermediate goods firms' objective and the final goods firm's production with quadratic approximations. Both approximations are derived in Appendix B.1.7.

Toward simplifying the intermediate goods firms' objective, we first define an intermediate firm's *ex post* optimal production level

$$x^*(z_i) := \arg\max_{x\in\mathcal{X}} \ \Pi(x, z_i) \tag{69}$$

We next define $\bar{\Pi}(z_i)$ as risk-adjusted profits evaluated at $(x^*(z_i), z_i)$ and $\Pi_{xx}(z_i)$ as the function's second derivative in x evaluated at the same point. The latter measures the *state-dependent cost* of misoptimizations relative to $x^*(z_i)$ and will be central to our analysis. The objective of the intermediate goods firm is, to the second order:

$$\tilde{\Pi}(x, z_i) := \bar{\Pi}(z_i) + \frac{1}{2}\Pi_{xx}(z_i)(x - x^*(z_i))^2 \tag{70}$$

So that this is globally defined, we will also apply the simplifying assumption that $\mathcal{X} = \mathbb{R}$.

Next, we approximate the final goods firm's production function (62) around the *ex post* optimal production levels $x^*(z_i)$ to the second order. We first define the aggregate of the *ex post* optimal production levels as

$$X^* = \left(\int_0^1 x^*(z_i)^{1-\frac{1}{\epsilon}} \, \mathrm{d}i \right)^{\frac{\epsilon}{\epsilon-1}} \tag{71}$$

We then write the approximate production function as

$$X = X^* - \frac{1}{2\epsilon} \int_0^1 \frac{(x_i - x^*(z_i))^2}{(X^*)^{-\frac{1}{\epsilon}} (x^*(z_i))^{1+\frac{1}{\epsilon}}} \, \mathrm{d}i \tag{72}$$

We now define equilibrium, up to these approximations, in terms of the stochastic choices of intermediate goods firms p, and the transition density f:[13]

Definition 1 (Equilibrium)**.** *An equilibrium is a stochastic choice rule $p \in \mathcal{P}$ and a transition density $f \in \mathcal{F}$ such that:*

1. *Intermediate goods firms' stochastic choice rules p solve program (68) given f, with $\tilde{\Pi}$, defined in (70), replacing Π.*

2. *The transition density f is consistent with p in the sense that: the marginal distribution of firm-level productivity is given by G; aggregate output is given by the aggregator (72) evaluated in the cross-sectional distribution of production implied by p and G; and the wage is derived from the wage rule (63) evaluated at aggregate output.*

2.4 Theoretical Results

We now present our main theoretical results, in four parts. First, we characterize firms' attention and misoptimization in partial equilibrium. These results show how firms' attention choice is shaped by two critical mechanisms, a *profit-curvature channel* related to the dollar cost of misoptimization and a *risk-pricing channel* related to the utility translation or risk adjustment of these costs. Second, we derive two conditions—a restriction to aggregate strategic complementarity and a lower bound on risk aversion—under which attention is counter-cyclical (and misoptimization pro-cyclical) in general equilibrium. We argue these conditions are *ex ante* reasonable based on existing macro-financial evidence. Third, we characterize equilibrium output in the economy as the product of the full-attention, frictionless benchmark and an attention wedge that is smaller than one and decreasing in the underlying productivity of the economy. This attention wedge reflects the cyclical nature of misallocation

[13] This embeds our notion of rational expectations equilibrium (REE) by requiring that firms' subjective prior about endogenous variables is "correct."

in our economy: when productivity is lower, firms optimize more precisely and productivity is closer to the frictionless benchmark. Finally, we show how endogenous variation in misallocation drives asymmetric and state-dependent shock propagation, and endogenous stochastic volatility of output growth.

2.4.1 Attention and Misoptimization in Partial Equilibrium

We begin by describing the stochastic choice behavior of firms. We observe that, under the entropic cost, firms' costly control problem is *linearly separable* across state realizations z_{it}. This implies that firms' optimal policies will be independent of the prior distribution f. In this way, consistent with our motivation, our model isolates firms' difficulties in making state-contingent plans. Leveraging this observation, the following result describes the solution to the firm's problem and allows us to characterize its comparative statics:

Proposition 6 (Firms' Optimal Stochastic Choice Rules)**.** *The production of a type-λ_i firm, conditional on realized state $z_i = (\theta_i, X, w)$, can be written as*

$$x_i = x^*(z_i) + \sqrt{\frac{\lambda_i}{|\pi_{xx}(z_i)|\, M(X)}} \cdot v_i$$

where $x^*(z_i)$ is the unconstrained optimal action, $|\pi_{xx}(z_i)|$ is the magnitude of curvature for the firms' profit function, $M(X)$ is the household's marginal utility, and v_i is an idiosyncratic, standard Normal random variable.

Proof. See Appendix B.1.1. ☐

Economically, Proposition 6 says that firms center their action around the full-attention optimum $x^*(z_i)$ but, due to costly control, make an idiosyncratic misoptimization. The variance of the misoptimization increases if the marginal cost of precision increases (higher λ_i), and decreases if either of two components of the marginal benefits of precision increases.

The first component of this marginal benefit is the state-dependent curvature of the firms' dollar profit function, $|\pi_{xx}(z_i)|$, which translates small misoptimizations into their dollar cost near the optimal production level. How this $|\pi_{xx}|$ moves as a function of the aggregate business cycle hinges on the specific assumed structure of firms' demand curves and cost structure. We will refer to this phenomenon as a profit-curvature channel affecting the extent of misoptimizations.

The second component of this marginal benefit is the household's marginal utility, $M(X)$. This translates dollar losses into utility losses, which can be directly compared to the utility cost of cognition. When households are risk-averse, this marginal utility is a decreasing function of X: the representative household is "hungrier" for a given firm's dollar profits, in utility terms, when the aggregate economy is doing poorly, and therefore less tolerant of misoptimizations. We will refer to this phenomenon as a risk-pricing channel affecting the extent of misoptimizations. We moreover observe that this channel relies on the (conventional) denomination of cognitive costs in utility, rather than final-good, units—we discuss the implications of the opposite assumption in Section 2.4.5.

The previous argument used only the structure of the cost functional and the assumption of a "Neoclassical firm" owned by a representative household. We can use the specific structure of our model to re-state the comparative statics in the discussion above in terms of firms' productivity θ_i and aggregate output X, after substituting in wages as a function of output via the wage rule. In particular, the curvature terms of interest can be written in the following way:[14]

$$|\pi_{xx}(z_i)| := v_\pi(\epsilon, \chi, \bar{X}, \bar{w}) \cdot \theta_i^{-1-\epsilon} X^{\chi(1+\epsilon)-1} \qquad M(X) = X^{-\gamma} \tag{73}$$

where $v_\pi(\epsilon, \chi, \bar{X}, \bar{w}) > 0$. We observe also that the variance of production conditional on the realized decision state, or $\mathbb{E}[(x_i - x^*(z_i))^2 \mid z_i]$, is a summary statistic for both misoptimization and "attention" measured by realized cognitive costs, which are decreasing in this variance. We summarize the comparative statics of this conditional variance, and by extension of attention and misoptimization, under the assumed payoff structure in the following result:

Corollary 4 (Comparative Statics for Mistakes)**.** *Consider a type-λ_i firm in state $z_i = (\theta_i, X, w(X))$. The extent of misoptimization, $\mathbb{E}[(x_i - x^*(z_i))^2 \mid z_i]$, increases in θ_i. Moreover, $\mathbb{E}[(x_i - x^*(z_i))^2 \mid z_i]$ strictly increases in X if and only if $\gamma > \chi(1+\epsilon) - 1$.*

Proof. Immediate from combining Proposition 6 with Equation 73. □

The (absolute) curvature of the profit function always increases in marginal costs, and therefore decreases in θ_i. The monotonicity of curvature in aggregate output depends jointly on the cyclicality of wages, which contributes a term with exponent $\chi(1+\epsilon)$; the aggregate demand externality, which contributes a term with exponent -1; and marginal utility, which contributes an exponent $-\gamma$. In particular, an economy with sufficiently cyclical wages would have misoptimization decrease in aggregate

[14] The first expression, and the associated constant, are derived in Appendix B.1.7.

output due to the profit-curvature channel, while an economy with sufficient risk aversion or sufficiently cyclical marginal utility would have misoptimization increase in aggregate output due to the risk-pricing channel. We will discuss the interpretation of this parameter condition "horse race" in the context of our general-equilibrium result in the next subsection.

2.4.2 Attention and Misoptimization Cycles in Equilibrium

We now translate the partial equilibrium behavior of the firm into general equilibrium. We first provide conditions under which equilibrium analysis is well-posed and output is a uniquely determined, monotone function of the underlying productivity state:

Proposition 7 (Existence, Uniqueness, and Monotonicity)**.** *For any $\chi > 0$, an equilibrium in the sense of Definition 1 exists. If $\chi\epsilon < 1$ and $\gamma > \chi + 1$, there is a unique equilibrium. In this equilibrium, output can be expressed via a function $X : \Theta \to \mathbb{R}$ that is strictly positive and strictly increasing.*

Proof. See Appendix B.1.2. □

To establish these properties, we derive a representation of equilibrium as a fixed point for aggregate output X as a function of the aggregate state θ. To establish uniqueness and monotonicity, we derive conditions under which the fixed-point equation is a contraction map that depends positively on productivity. The condition $\chi\epsilon < 1$ ensures that firms' production plans are on average an increasing function of aggregate output, by bounding wage pressure relative to the aggregate demand externality. The condition $\gamma > \chi + 1$ bounds the variance of actions around this optimum and ensures that, even in the presence of endogenous dispersion, there is positive but bounded complementarity.

The latter condition $\gamma > \chi + 1$ is both conservative in the model, as it ensures uniqueness and monotonicity for any possible distribution of λ_i, and highly plausible in practice. The elasticity of detrended real wages to GDP in (detrended) US data since 1987 is 0.095, and micro-level studies find similarly severe wage rigidity (Solon, Barsky, and Parker, 1994; Grigsby, Hurst, and Yildirmaz, 2019).[15] Moreover, to be consistent with the restrictions $\chi\epsilon < 1$ (upward-sloping best responses) and $\epsilon > 1$ (substitutable goods), χ cannot exceed one. The corresponding conditions $\gamma > 1.095$ or $\gamma > 2$ are likely both slack by one or two orders of magnitude, given abundant

[15] This calculation uses quarterly-frequency, seasonally-adjusted data on real GDP and median, CPI-adjusted wages of all full-time employed wage and salary workers. Both series are linearly detrended.

evidence in financial economics about the high cyclicality of the stochastic discount factor (e.g., Hansen and Jagannathan, 1991).

We now demonstrate conditions under which this economy exhibits aggregate attention and misoptimization cycles. We say that all firms "pay more attention" in a state if, averaging over idiosyncratic states (θ_i, λ_i), they pay a greater attention cost conditional on that state being realized. We say that firms "misoptimize less" in a state if, again averaging over idiosyncratic states (θ_i, λ_i), they have a lower expected mean-squared error around the *ex post* optimal action $x^*(z_i)$.[16] We now show that, under the stated assumptions for tractable equilibrium analysis, the model features counter-cyclical attention and pro-cyclical misoptimization:

Proposition 8 (Attention and Misoptimization Cycles)**.** *Assume* $\chi\epsilon < 1$ *and* $\gamma > \chi + 1$*. Intermediate goods firms pay more attention and misoptimize less in lower-productivity, lower-output states.*

Proof. See Appendix B.1.3. □

The proof of this result verifies that the stated conditions are sufficient for output X to be monotone increasing in θ, via Proposition 7; observes that the same conditions are sufficient for the idiosyncratic variance of choices to be monotone increasing in θ_i and X, via Corollary 4; and translates these into predictions for cross-firm averages.

Economically, Proposition 8 says that a calibration of the model that is consistent with existing evidence about wage rigidity and the stochastic discount factor predicts that firms should pay more attention to their decisions and make smaller misoptimizations, conditional on stochastic fundamentals, in downturns. The model's prediction about aggregate attention is qualitatively consistent with our motivating evidence on macroeconomic attention in language, with the caveat that the language-based measurement has no direct analogue in the model. The prediction about aggregate misoptimizations is a testable prediction on firm choices, which we will map to the data in Section 2.5. We will also, in the same section, develop and conduct tests that can distinguish the microeconomic mechanisms underlying this phenomenon (i.e., the risk-pricing and profit-curvature channels).

2.4.3 Misoptimization, Output, and Productivity

Having provided model conditions that generate attention and misoptimization cycles, we now study the effects of these phenomena on output and production. Despite

[16] Mathematical definitions of both notions are included in the proof of Proposition 8 in Appendix B.1.3.

unrestricted heterogeneity in the cross-sectional distributions of productivity and attention costs, we find that there are scalar sufficient statistics in equilibrium for each distribution. The cross-sectional distribution of productivity is summarized by quasi-arithmetic mean $\hat{\theta}(G) = \left(\mathbb{E}_G[\theta_i^{\epsilon-1}]\right)^{\frac{1}{\epsilon-1}}$.[17] We therefore, without loss of generality, write $\theta_t = \hat{\theta}(G_t)$ for the remainder of the analysis. The cross-sectional distribution of attention costs is summarized by the mean $\lambda := \mathbb{E}[\lambda_i]$.

We define $\log X(\log\theta)$ as a mapping from the log productivity state to log output in the economy, holding fixed all other parameters. The following result describes output in log units as the sum of an "RBC core" factor and an *attention wedge* $\log W(\log\theta)$:

Proposition 9 (Equilibrium Output Characterization)**.** *Equilibrium output follows:*

$$\log X(\log\theta) = X_0 + \chi^{-1}\log\theta + \log W(\log\theta) \tag{74}$$

where X_0 is a constant and $\log W(\log\theta) \leq 0$, with equality if and only if $\lambda = 0$. When $\chi\epsilon < 1$, $\gamma > \chi + 1$, and $\lambda > 0$, the wedge has the following properties:

1. $\frac{\partial \log W}{\partial \lambda} < 0$*, or the wedge widens with the average cost of attention.*

2. $\frac{\partial \log W}{\partial \log\theta} < 0$*, or the wedge widens as the state increases.*

Proof. See Appendix B.1.4. □

Absent inattention, output is log-linear in aggregate productivity. With inattention and under our stated conditions from Propositions 7 and 8, output is depressed by the presence of stochastic choice. The magnitude of this force increases in the extent of inattention λ and in aggregate productivity θ. Both results have a partial-equilibrium and general-equilibrium component. In partial equilibrium, both increasing λ and increasing θ make firms play more dispersed actions, as shown in Propositions 6, and this dispersion has a cost to output when the aggregate production function is concave. In general equilibrium, we iterate this logic until convergence; our comparative statics results verify that this fixed-point operation converges on a lower value of output.

To better understand this wedge, we can re-cast it in terms of labor productivity:

[17] Intuitively, heterogeneity in productivity is mediated by the elasticity of substitution: when goods are highly substitutable, this index converges to the maximum productivity; when goods are highly complementary, it converges to the minimum productivity.

Corollary 5 (The Productivity Wedge)**.** *Equilibrium labor productivity $A := \frac{X}{L}$ follows:*

$$\log A(\log\theta) = \log\theta + \chi\epsilon\log W(\log\theta) \tag{75}$$

Proof. See Appendix B.1.5. □

The productivity wedge representation allows for three useful parallels between our paper's mechanism and classic arguments in the macroeconomics literature. The first relates to a literature on the cleansing effect of recessions following Caballero and Hammour (1994). Our mechanism is like an attentional, *intensive margin* version of the same effect: conditional on a given firm operating, it is more focused on making precise and accurate choices in recessions, and this on average reduces the wedge and raises aggregate labor productivity.

The second relates to an empirical literature studying the dynamics of labor productivity in the United States, and in particular highlighting its unstable and often negative cyclicality (e.g., Galí and Gambetti, 2009; Barnichon, 2010; Galí and Van Rens, 2021). Our model accommodates a non-monotone relationship between aggregate labor productivity and aggregate output, due to the competing forces of increased productivity with increased misallocation. In fact, our quantitative analysis of Section 2.6 will generate such a non-monotone relationship.

The third relates to the literature on the aggregate effects of resource misallocation across firms, pioneered by Hsieh and Klenow (2009) and Restuccia and Rogerson (2008), and related to informational and cognitive frictions by David, Hopenhayn, and Venkateswaran (2016), David and Venkateswaran (2019), Ma, Ropele, Sraer, and Thesmar (2020), and Barrero (2022). The mechanism whereby dispersion in firm-level value marginal products depresses aggregate productivity is shared with these analyses. What is new is our prediction about the cyclicality of this force, driven by changing incentives for (in)attention, and the implications of that cyclicality for business cycles.

2.4.4 Shock Propagation and Volatility

The fact that agents are differentially attentive to shocks across states of the world makes the economy differentially sensitive to shocks. This observation is formalized in the following Corollary, which shows how the model generates endogenous stochastic volatility, state-dependent propagation of shocks and asymmetric shock propagation, features which are all absent in the benchmark with fully attentive firms or $\lambda = 0$:

Corollary 6 (Endogenous Stochastic Volatility, State-Dependent Propagation, and Asymmetric Propagation)**.** *Suppose that aggregate productivity follows the process* $\log\theta_t = \rho_\theta \log\theta_{t-1} + (1-\rho_\theta)\log\bar{\theta} + \nu_t$ *where* $Var[\nu_t] = \sigma_\theta^2$. *The model generates the following properties:*

1. *Endogenous stochastic volatility: to a first-order approximation in* ν_t*, the variance of output conditional on last period's productivity* θ_{t-1} *is given by*

$$\mathbb{V}[\log X_t|\theta_{t-1}] = \left(\chi^{-1} + \left.\frac{\partial \log W(\log\theta)}{\partial \log\theta}\right|_{\theta=\theta_{t-1}}\right)^2 \sigma_\theta^2 \tag{76}$$

2. *State-dependent shock propagation: the impact on output from a small shock* ν_t *in state* θ_{t-1} *is given by:*

$$\left.\frac{\partial \log X(\log\theta)}{\partial \log\theta}\right|_{\theta=\theta_{t-1}} = \chi^{-1} + \left.\frac{\partial \log W(\log\theta)}{\partial \log\theta}\right|_{\theta=\theta_{t-1}} \tag{77}$$

3. *Asymmetric shock propagation: the impact of a shock to second-order in* ν_t *in state* θ_{t-1} *is given by*

$$\left.\frac{\partial \log X(\log\theta)}{\partial \log\theta}\right|_{\theta=\theta_{t-1}} = \chi^{-1} + \left.\frac{\partial \log W(\log\theta)}{\partial \log\theta}\right|_{\theta=\theta_{t-1}} + \left(\left.\frac{\partial^2 \log W(\log\theta)}{\partial \log\theta^2}\right|_{\theta=\theta_{t-1}}\right)\nu_t \tag{78}$$

   where the sign of the wedge's second derivative determines the direction of asymmetry.

Proof. See Appendix B.1.6. □

To understand this result, note that the sensitivity of the economy to shocks is simply the sum of the frictionless economy's response to shocks, which is always χ^{-1}, and the response of the attention wedge to shocks. This latter response is always negative in the case studied by Proposition 9, so the economy is mechanically less responsive to shocks than the full attention benchmark. This dampened response is a familiar prediction in the literature with cognitively constrained agents (e.g., Sims, 1998, 2003; Gabaix, 2014), drawn out here in a general equilibrium context. More specific to our analysis is the fact that this dampening from the attention wedge can depend on the state, or the fact that the attention wedge need not be linear in $\log\theta$. This is a direct consequence of modeling state-dependent attention and accommodating attention cycles. This fact may generate the asymmetry, state-dependence, and stochastic volatility indicated in Corollary 6.

The ultimate macroeconomic implications of this result hinge on the concavity or convexity of the attention wedge in the state. When the attention wedge is concave (respectively, convex), the economy generates greater (smaller) volatility in low states, a larger (smaller) impact of shocks in low states, and features larger (smaller) impact of negative than positive shocks from any initial state. Owing to the intuition that when θ is very small, it is as-if the economy is operating in its "full pass-through" RBC core, the natural case appears to be a concave wedge. While we cannot establish theoretically that the wedge is globally concave, it cannot be globally convex and therefore must be concave in at least some part of the parameter space. Our quantitative analysis in Section 2.6 will feature a concave wedge. In that section we will review the implications of this finding.

2.4.5 Extensions

In the next parts of the paper, we will directly test the micro and macro predictions of the model (Section 2.5) and numerically examine the implications of a model that matches the empirical facts (Section 2.6). Before proceeding, we describe extensions of our main analysis.

Wage-Rule Shocks Our main analysis considered supply shocks operating through the shifts in the productivity distribution. Consider instead an economy with fixed productivities but stochastic $\bar{w}$, the level term in the wage rule. Shocks to $\bar{w}$ can be interpreted as a *labor wedge shock* or *demand shock* with the appropriate micro-foundation. Inspection of the formulas in Proposition 6, reveals that $\bar{w}$ is tantamount to an aggregate shifter of firms' revenue productivity. Propositions 7 and 8 provide conditions for output to be monotone *decreasing* in $\bar{w}$ and for counter-cyclical attention (pro-cyclical misoptimization) in a labor-wedge driven economy: when the labor wedge is high, and the economy is in a recession, attention is high. Proposition 9 and Corollary 6 hold as written, with $\bar{w}^{-1}$ replacing θ.

Decision Costs in Dollar Units Motivated by the nature of cognitive costs, we have denominated the cost of precise planning in utility units. An alternative choice is to specify these costs in units of final output or numeraire "dollars," which is natural if overcoming decision frictions requires investing in inputs (e.g., employees focused on planning). Since relative risk aversion enters our analysis only in translating dollar costs to utility costs, the assumption of dollar-denominated cognitive costs is nested by setting $\gamma = 0$. Proposition 6 and Corollary 4, concerning misoptimization in partial equilibrium, hold as written. The latter's condition for misoptimization to strictly increase in X is that $1 > \chi(1+\epsilon)$. Under the assumption that output is monotone in productivity, the model can feature equilibrium attention cycles under the same condition. This demonstrates that the profit-curvature channel, by itself, can induce attention cycles if wages are sufficiently rigid, a condition likely to hold based on our earlier discussion.

However, our empirical and quantitative analysis will support a strong role for the risk-pricing channel. First, we will find weak evidence of differential sensitivity of dollar profits to misoptimizations over the business cycle (Section 2.5.3). Second, we will find strong procyclicality of misoptimization variance. Thus, to rationalize these facts, our quantitative model requires a strong risk-pricing channel ($\gamma = 11.5$), rejecting the dollar-cost model nested with $\gamma = 0$ (Section 2.6.1).

Alternative Cognitive Costs We have used an entropic formulation for cognitive costs. This formulation has been axiomatized by Mattsson and Weibull (2002) as an effort disutility of controlling mistakes. Nevertheless, our main results are not sensitive to this modelling choice. In Flynn and Sastry (2021), we establish more general analogs of Propositions 7 and 8 in abstract games rather than business cycle models and allow for a more general class of *likelihood-separable* costs of stochastic choice (that includes the quadratic costs introduced by Harsanyi, 1973). The core of the argument remains the monotonicity of objective-function curvature in both the state and output.[18] In Appendix B.4.1, we further show how our results translate in a variant of our model that accommodates cognitive inertia and persistence of misoptimizations. This extension is useful in matching the patterns we uncover in the data in the next section.

We also study the robustness of our results to information-acquisition costs by examining our model with Gaussian signal extraction with costly precision (B.4.2) and mutual-information costs as studied by Sims (2003) (B.4.3). In both cases we provide conditions under which a greater cost of making mistakes, for instance due to the risk-pricing mechanism we highlight, leads to firms making smaller mistakes. We therefore argue that the conclusions of this section are robust to considering information acquisition. We do, however, note that application of unrestricted rational inattention to the model we study is not analytically possible using any known tech-

[18] Our analysis in that paper also clarifies the exact kind of supermodularity (macroeconomic complementarity) and discounting (concavity of the aggregate production function) required to achieve monotone precision in a game's unique equilibrium. A generalized version of Propositions 7 and 8 that dispenses with the approximated CES aggregator could be proven exactly using the main results of Flynn and Sastry (2021).

niques as our firms' have non-Gaussian priors and non-quadratic payoff functions, and aggregation is non-linear.

Multiple Inputs and Classical Labor Supply In Appendix B.2, we extend the model to allow for intermediate inputs, separate capital owners and laborers, and market-clearing wages rather than a wage rule. The first two features will be useful in mapping the model to the data in Section 2.5, while the third enables a more Neoclassical micro-foundation. We show under general conditions how the main results derived in this section, regarding the cyclicality of attention of misoptimization and the effects on output, continue to hold so long as the extent of the cognitive friction is not too large. Together, these extensions demonstrate the stability of our main model insights to a richer macroeconomic environment.

2.5 Testing the Model: The Misoptimization Cycle

In this section, we develop and implement tests of the model's main predictions using empirical proxies for "misoptimizations" by firms. We verify the two major predictions of the model's main case: firms on average make smaller misoptimizations in downturns, and firms face sharper market incentives to do so in the same states. We finally show that greater macroeconomic attention in language predicts lower misoptimizations at the firm level.

2.5.1 Measuring Misoptimizations

Data Our dataset for public firms' production and input choices is Compustat Annual Fundamentals. Compustat compiles information from firms' financial statements on sales, employment, variable input expenses, and capital measured via net and gross values of plants, property and equipment (PPE). It also provides a historically consistent industry classification, based on firms' main operational NAICS codes. We organize firms into 44 industries, which are defined at the NAICS 2-digit level but for Manufacturing (31-33) and Information (51), which we split into the 3-digit level.[19]

We use a sample period from 1986 to 2018, to focus on the modern macroeconomic regime after the Volcker disinflation. We view this relative stability in the macroeconomic environment as important for measuring externally valid patterns in macroeconomic attention and its relationship with firm-level decisionmaking. We apply standard filters to remove firms that are based outside the US, are insufficiently large, are likely to have been involved in a merger or acquisition, or do not report one

[19] We drop financial firms and utilities due to their markedly different production functions and profit structures.

or more important pieces of data. These filters yield a final sample of 68,198 firm-year observations, or about 2,200 per year. Appendix B.3.1 describes the procedure in detail.

From Theory to Estimation As emphasized in our discussion of Section 2.4.1, testing our main model mechanism requires data either on cognitive effort, in payoff units, or choice misoptimization. The former is at best imperfectly proxied by observable variables like the textual attention from Section 2.2. The latter, by contrast, can be constructed from standard firm-level data, and moreover maps directly to our economic predictions for cyclical misallocation and shock propagation. Thus, our main goal in this section is to derive an empirical proxy for "misoptimizations" in production or input-choice relative to an *ex post* optimal level. Our overall strategy is to measure residuals in firms' choices relative to empirically estimated production and policy functions.

We assume that each firm i at time t, within sector $j(i)$, operates a sector-specific, constant-returns-to-scale, Cobb-Douglas production technology over labor, materials, and capital, with total factor productivity θ_{it}:

$$\log x_{it} = \log \theta_{it} + \alpha_{L,j(i)} \log L_{it} + \alpha_{M,j(i)} \log M_{it} + \alpha_{K,j(i)} \log K_{it} \tag{79}$$

with the restriction, in each sector, that $\alpha_{L,j(i)} + \alpha_{M,j(i)} + \alpha_{K,j(i)} = 1$.[20] This generalizes our model from Section 2.3, with a bundle of inputs replacing the single theoretical factor of production and heterogeneity in technology across industries, as in Appendix B.2. Next, we assume that input choice follows a policy function that is log-linear in individual firm characteristics, industry-by-time trends, individual productivity, and a residual m_{it}:

$$\log L_{it} = \eta_i + \chi_{j(i),t} + \beta \log \theta_{it} + m_{it} \tag{80}$$

This is a generalization of the optimal policy in our model from Section 2.3, with input costs and demand varying at the industry-by-time and firm level. Appendix B.3.3 derives Equation 80 exactly in such an extended model, provided that firms cost-minimize over input bundles. In this mapping to the model, we can interpret the residual as the firm's percentage misoptimization from the counterfactual *ex post*

[20]The Cobb-Douglas assumption is a convenient and common step to enable production function estimation via input-cost shares (e.g., Foster, Haltiwanger, and Krizan, 2001; Foster, Haltiwanger, and Syverson, 2008; Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry, 2018). Moreover, a number of studies including Basu and Fernald (1997), Foster, Haltiwanger, and Syverson (2008), and Flynn, Traina, and Gandhi (2019) argue that constant returns to scale in physical terms is a reasonable approximation for large, US-based firms.

optimal level L^*_{it}, or $m_{it} \approx \frac{L_{it}-L^*_{it}}{L^*_{it}}$. As observed in Appendix B.3.3, and readily apparent in our main theoretical model, these are approximately equal to percentage misoptimizations in production, or $\frac{L_{it}-L^*_{it}}{L^*_{it}} \approx \frac{x_{it}-x^*_{it}}{x^*_{it}}$.

In the data, for myriad reasons unmodeled in our main analysis, these residuals are likely to be persistent. To capture this, we assume that the m_{it} are determined by an AR(1) process with persistence $\rho \in (0,1)$ and scaled innovations u_{it}, or

$$m_{it} = \rho m_{i,t-1} + \left(\sqrt{1-\rho^2}\right) u_{it} \tag{81}$$

The innovations u_{it} are mean zero and have a variance $\mathbb{E}[u^2_{it} \mid i, Z_t]$ which varies as a function of fixed individual characteristics and all aggregate state variables Z_t. We provide a formal micro-foundation for this AR(1) structure of mistakes in Appendix B.4.1.[21]

Of course, the procedure described is exactly only in a stylized model of the firm, and additional unmodeled elements like adjustment costs and financial frictions would also manifest in the residual m_{it} of Equation 80. While we focus on this model interpretation and associated data construction in our main analysis, we also estimate richer firm-level policy functions and conduct all our empirical analysis with their implied misoptimization measures. We will return to this point when discussing our results.

Estimation Procedure To empirically estimate our model, we proceed as follows. First, we estimate productivity as the residual of the production function, Equation 79, after estimating input elasticities using revenue shares in each sector. Because we model decreasing returns to scale in revenue from the product demand functions, our measure corresponds in our model to physical-quantity TFP or "TFPQ." We use the standard input definitions of Keller and Yeaple (2009) and provide more details in Appendix B.3.[22]

Given estimates of productivity, we estimate the system of Equations 80 and 81. We first estimate Equation 80 via ordinary least squares (OLS) and obtain a

[21] As long as $\mathbb{E}[u^2_{it} \mid i, Z_t]$ is sufficiently persistent, $\mathbb{E}[m^2_{it} \mid Z_t] \approx \mathbb{E}[u^2_{it} \mid Z_t]$, where both expectations average over fixed individual characteristics indexed by i. We will use this approximation in practice, as the one-step-ahead variance of Equation 81 is easier to measure with a relatively short sample than the stationary variance.

[22] We treat labor expenditures as the product of reported employees and a sector-specific wage calculated from the US County Business Patterns; materials expenditures as the sum of variable costs and administrative expenses (COGS + XSGA) net of depreciation and labor expenditures; and the capital stock as the initial gross level of plant, property, and equipment plus net investment. Instead of imputing rental rates for capital, we impose constant returns to scale and a fixed profit share of 0.75 (e.g., in the model, $\epsilon = 4$).

preliminary estimate $\hat{m}^0_{it}$ of the residual. We next estimate Equation 81 using $\hat{m}^0_{it}$ to obtain an estimate $\hat{\rho}$ of the residual persistence (in our main procedure, 0.70). We next estimate via OLS the "quasi-differenced" equation for labor choice:[23]

$$\log L_{it} - \hat{\rho} \log L_{i,t-1} = \eta_i + \chi_{j(i),t} + \beta_0 \log \hat{\theta}_{it} + \beta_1 \log \hat{\theta}_{i,t-1} + \nu_{it} \tag{82}$$

with residual $\nu_{it} = m_{it} - \hat{\rho} m_{i,t-1}$. When $\hat{\rho} = \rho$, using Equation 81, $\nu_{it} = \left(\sqrt{1-\rho^2}\right) u_{it}$. We therefore obtain an estimate of u_{it} as $\hat{u}_{it} = \frac{\hat{\nu}_{it}}{\sqrt{1-\hat{\rho}^2}}$.

Our measure of aggregate "Misoptimization Dispersion" is an estimate of the variance $\mathbb{E}[m^2_{it} \mid Z_t]$ with weights s^*_{it} proportional to firms' *predicted* sales based on fundamentals:[24]

$$\text{MisoptimizationDispersion}_t = \frac{\sum_{i \in \mathcal{I}_t} s^*_{it} \cdot \hat{u}^2_{it}}{\sum_{i \in \mathcal{I}_t} s^*_{it}} \tag{83}$$

The weights are the appropriate ones for mapping average misoptimization to misallocation in the theory, and will aid in our subsequent structural interpretation of our findings. In a nutshell, these weights give higher influence to larger, more productive firms while not "double-counting" misoptimizations in both input choice and total production. In Appendix B.1.7, we show that pro-cyclical MisoptimizationDispersion by our empirical definition is sufficient for pro-cyclical misoptimization as defined in Proposition 8, and thus constitutes a more stringent test of that model prediction.

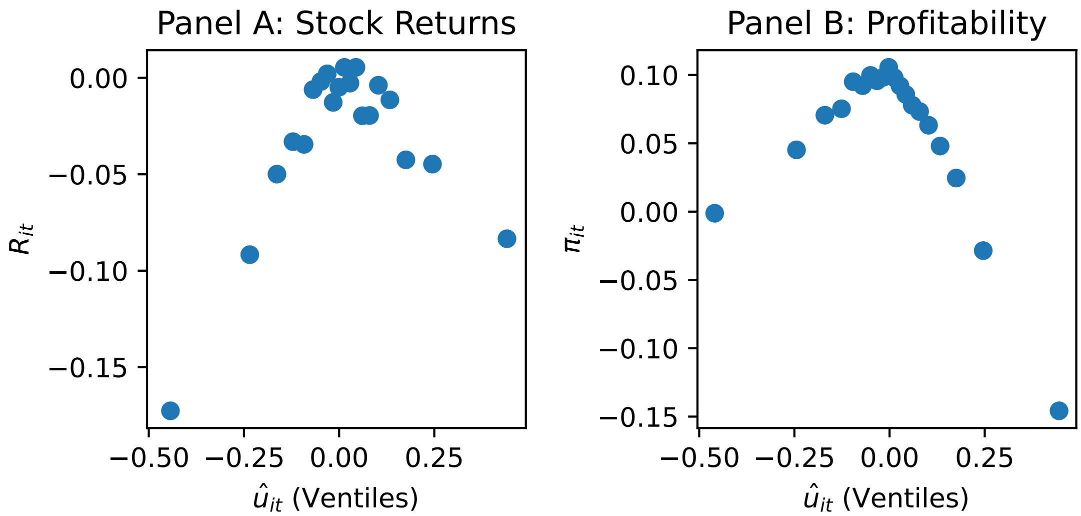
Validating the Measure: Mistakes are Bad for Returns and Profitability Before proceeding to the microeconomic and macroeconomic tests of our theory, we first check that our measured "misoptimizations" are, as required by the theory, bad for firm performance. We proxy for firm performance with two measures. The first is firms' log stock returns R_{it}. The second is firms' "profitability" π_{it}, or earnings before interest and taxes (EBIT), or income net of variable costs, divided by the last year's total variable costs.[25] We examine the non-parametric binned scatter relationship of each of these measures $X_{it} \in \{R_{it}, \pi_{it}\}$ with measured misoptimization (residuals) $\hat{u}_{it}$,

[23] Appendix Table B.9 contains our estimates of Equations 81 and 82 under the baseline procedure outlined in this section, along with several alternative choices used in robustness checks. In all estimations, we drop the top and bottom 1% tails of the TFP distribution to limit the effects of outliers. Results are quantitatively highly similar without this trimming.

[24] In particular, the weights are the exponentiated fitted values $\exp(\hat{\beta} \log \theta_{it})$ from the following regression equation: $\log \text{Sales}_{it} = \beta \log \theta_{it} + \eta_i + \chi_{j(i),t} + \epsilon_{it}$.

[25] Variable costs, by our definition, are cost of goods sold (COGS) plus administrative expenses (XSGA) net of depreciation (DP). In other words, they are the sum of what we term "materials" and labor expenses in the production function. Normalization by lagged costs, rather than current costs, limits mechanical denominator bias related to the current period's mistakes. Results are similar when normalizing by total sales or costs in the current or previous period.

Figure 2-2: Binned Scatter Plots of Misoptimization and Firm Performance



Notes: Both panels are binned scatter plots. The outcome variables are log stock returns (Panel A) and profitability (Panel B). Dots represent means of the corresponding outcome conditional on ventiles of the x-axis variable, $\hat{u}_{it}$, and industry-by-time fixed effects. The construction of $\hat{u}_{it}$ is described in Section 2.5.1.

net of industry-by-time fixed effects $\chi_{j(i),t}$:

$$X_{it} = f(\hat{u}_{it}) + \chi_{j(i),t} + \epsilon_{it} \tag{84}$$

where f is a piecewise-constant function defined over ventiles of $\hat{u}_{it}$. Figure 2-2 shows this binned scatterplot relationship for each variable. Misoptimizations in both directions (i.e., under- or over-hiring labor) hurt firm performance, measured in each way.[26]

There are two potential threats to this interpretation. First, misoptimization variance could correlate with current or future productivity growth, and only through this channel affect firm performance. Second, misoptimizations may harm profitability and returns on impact in return for future improved performance. To investigate these hypotheses, we estimate projection regressions of productivity growth, prof-

[26]In Appendix Table B.1, we show that the result of a negative relationship between $\hat{u}_{it}^2$ and $\{R_{it}, \pi_{it}\}$ is robust to different choices of control variables.

itability and returns on $\hat{u}_{it}^2$, net of industry-by-time fixed effects:[27]

$$X_{i,t+k} = \beta_{X,k} \cdot \hat{u}_{it}^2 + \chi_{j(i),t} + \epsilon_{it} \tag{85}$$

for $X_{i,t+k} \in \{\Delta \log \hat{\theta}_{i,t+k}, \pi_{i,t+k}, R_{i,t+k}\}$ and $k \in \{0, 1, 2\}$.[28] We find no evidence of a quantitatively large effect on current and future TFP growth (Appendix Table B.2). We also find strong evidence of persistent negative effects on profitability and stock returns. These results rule out the possibility that the observed negative effect of $\hat{u}_{it}^2$ operates through a channel related to current or future productivity or dynamic trade-offs between poor performance today and improved future performance.

2.5.2 Misoptimizations Rise in Booms and Fall in Downturns

We now study the model's macroeconomic prediction for the behavior of aggregate misoptimizations. Figure 2-3 plots aggregate Misoptimization Dispersion, measured at the annual frequency, against the annual average unemployment rate and (detrended) end-of-year S&P 500 price. Misoptimization Dispersion rises when the real economy and financial markets are doing well (e.g., the late 1990s), falls during recessionary or financial crisis periods (e.g., 1990, 2001, and 2008), and is approximately as persistent as the overall business cycle.

We benchmark the slope of this relationship with the business cycle by estimating a linear regression of Misoptimization Dispersion on each macroeconomic variable. Specifically, we estimate

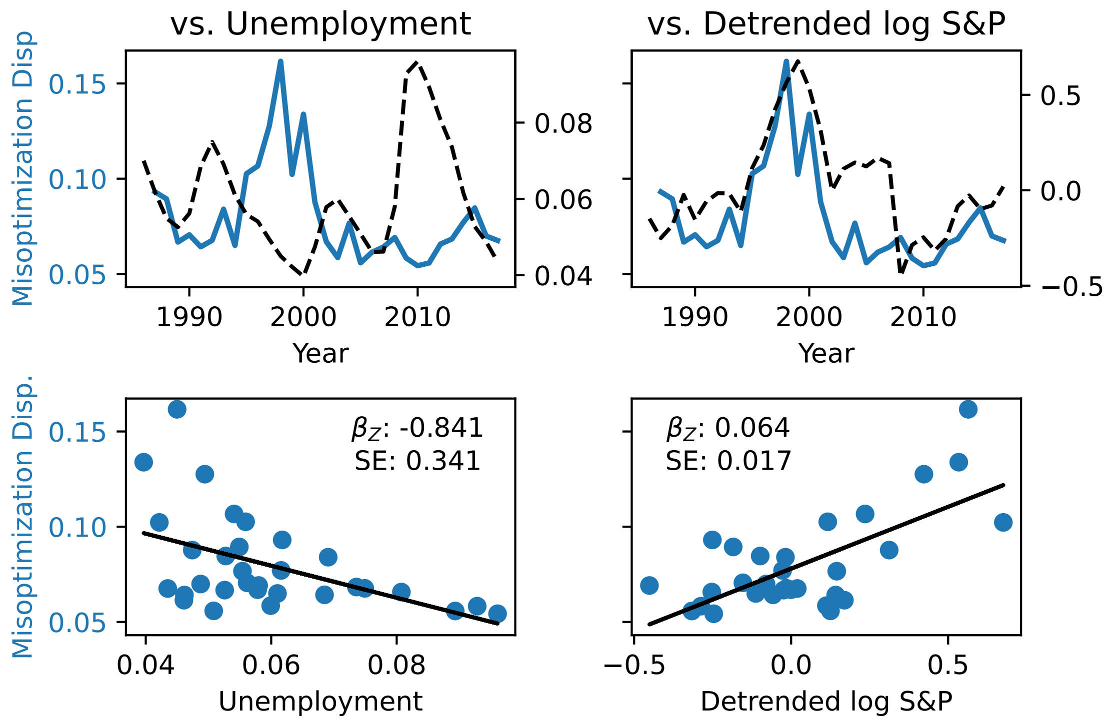
$$\text{MisoptimizationDispersion}_t = \alpha + \beta_Z \cdot Z_t + \epsilon_t \tag{86}$$

for $Z_t \in \{\text{Unemployment}_t/100, \log \text{SPDetrend}_t\}$, over our 31 annual observations. We estimate a slope of -0.841 (SE: 0.341, $p = 0.02$) with respect to unemployment and 0.064 (SE: 0.017, $p = 0.001$) with respect to the detrended S&P 500. The respective correlations are -0.493 and 0.689. The regression on unemployment implies that a five percentage point swing in unemployment is associated with an increase of Misoptimization Dispersion by 0.042 log points, or 53% of its sample mean value. Appendix Figure B-5 shows that the same pro-cyclical pattern is apparent in other

[27]For this and all subsequent panel regressions, we drop observations in the 1% tail of both the outcome and main regressor to limit the effect of outliers, which are likely driven in this context by measurement errors (both sides for symmetric variables; only the upper tail for positive variables).

[28]While misoptimizations were constructed to be orthogonal to measured total factor productivity, their squared values do not necessarily have this property; moreover, there is no mechanical relationship between misoptimizations today and productivity tomorrow.

Figure 2-3: Misoptimization Dispersion is Pro-Cyclical



Notes: The top two panels plot Misoptimization Dispersion (blue line, left axis) along with, respectively, unemployment and the linearly detrended S&P 500 price (black dashed lines, right axis). The bottom two panels are scatterplots of Misoptimization Dispersion versus the corresponding macroeconomic aggregate. The black solid line is the linear regression fit. The standard errors are HAC robust based on a Bartlett kernel with a three-year bandwidth.

measures of dispersion, the (weighted) mean of $|\hat{u}_{it}|$ and inter-quartile range of $\hat{u}_{it}$.

These findings are consistent with the model prediction for attention and misoptimization in Proposition 8. Of course, outside of our model, any time-varying misspecification of our model for production and policy functions could also generate the observed pattern, while demanding a different economic interpretation. Before proceeding to a direct test of our model's proposed microeconomic mechanism, we review evidence that supports our "attention cycles" interpretation of the finding in favor of other economic mechanisms for misoptimization or mismeasurement.

Robustness to Incorporating Other Frictions If firms pay significant physical costs to adjust ostensibly "variable" inputs like labor (e.g., as in Hopenhayn and Rogerson, 1993), then "misoptimizations" by our measure may pick up frictional adjustment and their corresponding effect on firms' optimality conditions. To capture such adjustment costs in reduced form, we add the first lag of labor choice to the policy function (80):

$$\log L_{it} = \eta_i + \chi_{j(i),t} + \beta \log \theta_{it} + \tau \log L_{i,t-1} + m_{it} \tag{87}$$

If Equation 87 were the true model of input choice, and we had instead estimated our main model Equation 80, then variation in $\tau \log L_{i,t-1}$ not spanned by current productivity (or the absorbed effects) would be part of the misoptimization m_{it}. Some of this variation would be truly spanned by previous productivity shocks and be misrepresented as a cognitive error, while the remainder of the variation would be a "bad control" spanned by persistent lagged cognitive errors. With these caveats in mind, we recalculate Misoptimization Dispersion using the same methods described above starting with the policy function Equation 87.[29]

We find that new, more adjustment-cost robust measure of Misoptimization Dispersion is strongly pro-cyclical. The regression coefficient of this version of Misoptimization Dispersion on unemployment is -0.439 (SE: 0.196), attenuated relative to our baseline estimate (-0.841) but statistically indistinguishable from it. Moreover, the new measure has an almost identical correlation with unemployment (-0.468) as our baseline does (-0.493), suggesting that the aforementioned attenuation comes from the smaller scale of the alternatively measured mistakes and not from a diminished cyclicality in normalized units.

A second prediction of a model with costly adjustment is that firms might trade-off misoptimizations today to ensure better performance in the future (e.g., by smoothing response to productivity news). This hypothesis is inconsistent with the already presented evidence that misoptimizations had *persistent* negative effects on firm stock returns and profitability, instead of mean-reverting effects (Table B.2).

We next consider financial frictions, broadly defined as state-dependent constraints or costs of financing input purchases. To capture financial frictions, we add a direct control for leverage (total debt over total assets), as constructed by Ottonello and Winberry (2020), and its interaction with TFP, in the policy function (80):

$$\log L_{it} = \eta_i + \chi_{j(i),t} + \beta \log \theta_{it} + \tau \text{Lev}_{it} + \phi \cdot (\text{Lev}_{it} \times \log \theta_{it}) + m_{it} \tag{88}$$

[29] The residuals from Equation 87 have negligable autocorrelation (AR(1) coefficient 0.016, as reported in Table B.9), and imply a lower misoptimization variance than our baseline measure (a mean of 0.036 across years, compared to 0.080 for the baseline).

This specification allows for more levered firms to face, effectively, a "TFP adjustment" (direct effect) and have different responsiveness to TFP shocks (interactive effect). If Equation 88 were the true model of input choice, and we had instead estimated the main model Equation 80, then both the direct effect and changing responsiveness to TFP (if not spanned by fixed effects) would be erroneously attributed to the cognitive friction. When Misoptimization Dispersion is recalculated with the leverage-augmented policy function, its regression coefficient on unemployment is -0.841 (SE: 0.341), almost identical to that of our main model. This suggests, consistent also with the standard narrative that financial frictions create wedges primarily in recessions, that time-varying financial frictions do not explain our results.

Robustness to Alternative Measurement Strategies To investigate the possibility that time-varying mismeasurement explains our findings, we consider the following set of alternative econometric strategies in estimating the policy and production functions: (i) estimating sector-specific policy functions, to combat against different market conditions and/or measurement error in TFP; (ii) allowing firm responsiveness to TFP to vary by year, to capture time variation in the same;[30] (iii) allowing for TFP to affect the policy function non-linearly, to capture asymmetric hiring and firing rules as found by Ilut, Kehrig, and Schneider (2018); (iv) allowing output elasticities to change over time, to capture changes in production technology like automation; (v) estimating production functions and policy functions in a pre-sample, while estimating misoptimization in a post sample, to avoid mechanical over-fitting; (vi) and calculating productivity using the proxy-variable method of Olley and Pakes (1996), to guard against over-reliance on cost shares and the necessary imputation of returns to scale. We finally consider two alternative specifications of the main regression, the first with linear and quadratic time-trend controls, which can absorb confounding low-frequency trends, and the second restricting to the manufacturing sector, which has been the focus of much of the literature on misallocation (e.g., Hsieh and Klenow, 2009; Kehrig, 2015). Appendix Table B.3 demonstrates the stability of our main finding, pro-cyclical misoptimization, under each scenario.

Robustness to Compositional Adjustment One issue affecting the interpretation of volatility time-series among US public firms is the large change in these firms' composition. In particular, previous research suggests that US public firms became more volatile in sales growth, employment growth, and stock returns over time,

[30] Decker, Haltiwanger, Jarmin, and Miranda (2020), in particular, note a secular pattern of "declining business dynamism" or declining responsiveness to TFP, which in principle could contaminate our results.

and that these trends can be statistically explained by cohort-of-entry (Davis, Haltiwanger, Jarmin, Miranda, Foote, and Nagypal, 2006) and/or firm-specific (Brown and Kapadia, 2007) effects. To guard against the possibility that these compositional trends spuriously drive our findings, we calculate a variant of macroeconomic attention based on the component of misoptimizations residualized from firm fixed effects, or $\hat{u}^2_{it,\perp}$ defined below

$$\hat{u}^2_{it} = \gamma_i + \hat{u}^2_{it,\perp} \tag{89}$$

The new measure of Misoptimization Dispersion, calculated from $u^2_{it,\perp}$, has correlation 0.91 with our original version. This measure's regression slopes against unemployment and the detrended level of the S&P 500 are -0.520 (SE: 0.206) and 0.040 (SE: 0.009), of the same sign as and of comparable magnitude to our baseline estimates.

Robustness to Studied Time Period Our main analysis uses firm data from 1986 to 2018. Our choice has two motivations. First, the compositional changes and associated drift in firm-level volatility among public firms is more pronounced earlier in the sample (Davis, Haltiwanger, Jarmin, Miranda, Foote, and Nagypal, 2006). Second, macroeconomic policy and therefore dynamics differ markedly before and after the Volcker disinflation. Nevertheless, we can extend our analysis to measure Misoptimization Dispersion for the entire Compustat sample from 1950 to 2018.[31] To allay the concerns about composition, we use the compositional adjustment strategy described in Equation 89. We plot our resulting Misoptimization Dispersion series in Figure B-6. We continue to find strong evidence of pro-cyclicality. Replicating the regressions against the unemployment rate and detrended S&P 500 price, we obtain coefficients of -0.571 (SE: 0.197) and 0.052 (SE: 0.014). Both estimates are within one standard error of our baseline estimates.

Misoptimization in Other Input Choices We focused on misoptimization in labor choice because it is the only input or output measured in quantity units in our dataset. The model implied that we could measure misoptimization in any variable input choice. As a robustness exercise, we recalculate misoptimizations and Misoptimization Dispersion using as the choice variable total variable cost expenditures (i.e., materials plus labor) and investment rates (i.e., log growth rates of the capital stock). We plot the resulting variations of Misoptimization Dispersion in Appendix Figure B-7. We find broadly similar patterns, particularly in the spike of the mid-1990s and falls in the 2002 and 2009-10 downturns. The result with investment rates

[31] One minor difference in the calculation is that, due to lack of sectoral wage data, we calculate TFP with two factors, materials (inclusive of labor) and capital.

echoes findings by Eisfeldt and Rampini (2006) and Bachmann and Bayer (2014), using comparable data on US public firms, that investment rates are more dispersed in booms than in downturns. Our model's explanation, complementary to that of the literature, is that much of the apparent "reallocation" of capital in booms is unrelated to fundamentals and therefore a cause of *misallocation*, whereas the less dispersed investments in downturns are more related to fundamentals and therefore a cause of *better allocation*.

2.5.3 Misoptimization is More Costly in Bad States

We have established that, by our measure, firms make smaller misoptimizations in downturns. We now investigate the extent to which this is explained by our model mechanism: that firms have higher financial incentives to avoid misoptimization in downturns. We will also investigate, more specifically, the extent to which such incentives are separately driven by the model's risk-pricing and profit-curvature channels.

Markets Punish Mistakes More When Aggregate Returns Are Low We first test whether misoptimizations have a state-dependent effect on stock returns—heuristically, whether the relationship between misoptimizations and stock returns (Figure 2-2) steepens in bad states. Specifically, we relate a firm's log stock return R_{it} with the firm's squared misoptimization innovation over that year interacted with the log aggregate (S&P 500) stock return R_t:[32]

$$R_{it} = \beta \cdot \hat{u}_{it}^2 + \phi \cdot \left(\hat{u}_{it}^2 \times R_t\right) + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it} \tag{90}$$

Sector-by-time fixed effects partial out industry trends. The vector of control variables X_{it} can include firm fixed effects and the growth rate of firm-level TFP, to partial out other important determinants of returns. As we have shown, the marginal effect of misoptimization on the stock price is negative. The hypothesis that the market punishes misoptimization more severely in times of distress, or low R_t, is captured by $\phi > 0$. In the expanded model of Appendix B.2, we directly derive the regression equation and the prediction $\phi \geq 0$, with equality only if investors are risk-neutral (there is no risk-pricing channel) and profit sensitivity to mistakes is state-independent (there is no profit-curvature channel).

Table 2.1 shows our estimates. In all four specifications, with and without firm fixed effects and productivity controls, we verify that $\phi > 0$: misoptimization is priced

[32] To be consistent with our earlier analysis, we run this regression using re-scaled mistake innovations $\hat{u}_{it}$. In Appendix Table B.4, we replicate the analysis using our (first-stage) estimates of the mistake "level" $\hat{m}_{it}$ and find qualitatively similar results.

Table 2.1: Markets Punish Misoptimization Harder in Low-Return States

	(1)	(2)	(3)	(4)
	Outcome: R_{it}			
$\hat{u}_{it}^2$	-0.268	-0.262	-0.097	-0.087
	(0.025)	(0.023)	(0.034)	(0.033)
$\hat{u}_{it}^2 \times R_t$	0.376	0.376	0.443	0.431
	(0.123)	(0.124)	(0.171)	(0.167)
Sector x Time FE	✓	✓	✓	✓
Firm FE			✓	✓
TFP Control		✓		✓
N	41,578	41,578	41,206	41,206
R^2	0.239	0.261	0.385	0.403

Notes: R_{it} is the firm-level log stock return. $\hat{u}_{it}^2$ is the squared firm-level misoptimization residual, constructed using the methods described in Section 2.5.1. R_t is the log return of the S&P 500. Standard errors are double-clustered at the year and firm level.

more severely in states of low aggregate returns.[33] Our estimates in column (3), in particular, suggest that mistakes have a zero price if the S&P return is 22%, close to its value in the late 1990s or the height of the dot com bubble. By contrast, in the trough of 2008 ($R_t = -0.52$), the model implies that pricing is 6.2 times more severe than in the "usual" states of $R_t \approx 0.10$.

In the Appendix, we report a number of robustness exercises. First, Appendix Table B.5 shows the stability of our main finding of $\phi > 0$ to the alternative data-construction approaches highlighted in the previous subsection. Second, Appendix Figure B-8 reports the estimates of the effect of misoptimizations on returns for every year in our sample. The effects are negative in every year and, in line with our result from the continuous-interaction regression, visibly greatest when the S&P 500 return is low. Third, in Appendix Table B.6, we show that our finding of $\phi > 0$ is robust to controlling for other plausible heterogeneities in the effects of misoptimizations on stock returns. In particular, we control for the level and $\hat{u}_{it}^2$-interaction of TFP, lagged stock returns, and financial leverage to control for the observed tendencies for negative-return firms to have higher volatility (the leverage effect) and binding financial constraints; and we control for interactions of $\hat{u}_{it}^2$ with industry and firm fixed effects to model heterogeneous exposure to aggregates.

[33]Similar results are obtained using mistake "levels" as discussed in Footnote 32 (columns 3-6 of Appendix Table B.4)

Table 2.2: Misoptimization, Profits, and Pricing

	(1) Outcome: π_{it}	(2)	(3) Outcome: R_{it}	(4)
$\hat{u}_{it}^2$	-0.114	-0.021		
	(0.020)	(0.032)		
$\hat{u}_{it}^2 \times R_t$	0.112			
	(0.089)			
π_{it}		0.400	0.421	0.690
		(0.028)	(0.034)	(0.305)
$\pi_{it} \times R_t$			-0.303	-1.642
			(0.166)	(0.632)
Firm FE	✓	✓	✓	✓
Sector x Time FE	✓	✓	✓	✓
N	50,966	40,879	40,879	40,879
R^2	0.663	0.402	0.402	
First-stage F				17.80

Notes: π_{it} is firm-level profitability and R_{it} is the firm-level log stock return. $\hat{u}_{it}^2$ is the squared firm-level misoptimization residual, constructed using the methods described in Section 2.5.1. R_t is the log return of the S&P 500. Standard errors are double-clustered at the year and firm level.

Unpacking the Mechanism: Profitability *vs.* Returns Interpreted via our model, the result in Table 2.1 implies that at least one of the risk-pricing or profit-curvature channels drive changes in the cost of mistakes.

To differentiate these explanations, we now explore the extent to which profitability is more sensitive to mistakes during downturns. To this end, we define profitability π_{it}, as before, as this year's EBIT divided by the last year's total variable costs. We study the state-dependent effects of misoptimizations on profitability using the following regression model that mirrors our previous analysis of stock returns:

$$\pi_{it} = \beta_\pi \cdot \hat{u}_{it}^2 + \phi_\pi \cdot \left(\hat{u}_{it}^2 \times R_t\right) + \chi_{j(i),t} + \gamma_i + \epsilon_{it} \tag{91}$$

Heuristically, $\phi_\pi > 0$ isolates the *dollar-profit-curvature* channel of Proposition 6 and Corollary 4. In the expanded model of Appendix B.2, we derive the prediction $\phi_\pi \geq 0$ with equality if the profits function, in dollars, has state-independent curvature.

We report the results in the first column of Table 2.2: a positive, but small and statistically insignificant, ϕ_π. Thus misoptimizations have a constant effect on firms' dollar income up to statistical precision. This suggests that the mechanism for our earlier finding of state-dependent market punishment relates specifically to the market's greater *reaction* to fixed profit effects of misoptimizations.

We next estimate a sequence of models that explore the joint effects of misoptimizations and profitability on stock returns. Our first model regresses firm stock returns on $\hat{u}_{it}^2$ and π_{it} conditional on firm and sector by time fixed effects. We find that, conditional on profitability, misoptimizations have a severely attenuated, and statistically indistinguishable from zero, effect on stock returns (column 2 of Table 2.2). This is consistent with the model interpretation that misoptimizations matter for prices by reducing current profits.

We next explore the market response to profit anomalies, relative to fixed firm profitability and industry trends, without taking a stand structurally on how those anomalies arise. Our estimating equation is the mirror of Equation 90 with profitability in place of misoptimizations:

$$R_{it} = \beta_{\pi\to P} \cdot \pi_{it} + \phi_{\pi\to P} \cdot (\pi_{it} \times R_t) + \chi_{j(i),t} + \gamma_i + \epsilon_{it} \tag{92}$$

again estimated with firm fixed effects as the control variables. We recover the pattern that the market prices more aggressively respond to profitability when aggregate returns are low, or an estimate of $\phi_{\pi\to P} < 0$ (column 3 of Table 2.2). This validates, without any intermediate structural estimation of production functions or misoptimization, the idea that the market values "firm performance" more acutely in low-return environments.

We finally present an instrumental variables (IV) estimate of cyclical market response to directly quantify the pathway from misoptimizations to state-dependently-priced profitability to stock returns. Specifically, defining Z_{it} as the vector of regressors $(\pi_{it}, \pi_{it} \times R_t)$ and W_{it} as the vector of instruments $(\hat{u}_{it}^2, \hat{u}_{it}^2 \times R_t)$, we augment the now "second-stage" Equation 92 with a "first-stage" equation

$$Z_{it} = W_{it}' A + \chi_{j(i),t} + \gamma_i + \epsilon_{it} \tag{93}$$

The first stage F-statistic is 17.8, owing to the strong relationship between misoptimizations and profitability. Our IV estimates of $(\beta_{\pi\to P}, \pi_{\pi\to P})$, while not very precisely estimated, suggest if anything a greater state-dependence of the market response to misoptimization compared to the market response to other determinants of profits (column 4 of Table 2.2). We interpret this result, along with the others in this subsection, as empirical validation of the model's microeconomic mechanism for

Table 2.3: Macro-Attentive Firms Make Smaller Misoptimizations

	(1)	(2)	(3)	(4)
		Outcome: $\hat{u}_{it}^2$		
$\log \text{MacroAttention}_{it}$	-0.0081	-0.0052	-0.0058	-0.0056
	(0.0028)	(0.0029)	(0.0044)	(0.0038)
Sector x Time FE	✓	✓	✓	✓
Firm FE			✓	✓
TFP, Return Controls		✓		✓
N	28,279	24,392	27,875	23,930
R^2	0.053	0.067	0.383	0.384

Notes: $\hat{u}_{it}^2$ is the squared firm-level misoptimization residual, constructed using the methods described in Section 2.5.1. $\log \text{MacroAttention}_{it}$ is the measure of firm-level macroeconomic attention, constructed using the methods described in Section 2.2.1. Standard errors are double-clustered at the year and firm level.

state-dependent attention and misoptimization.

2.5.4 Macro Attention Predicts Smaller Misoptimizations

We finally investigate the relationship between misoptimization, which we could link directly to the model, and Macroeconomic Attention from Section 2.2, which motivated our analysis but did not have a formal analogue in the model. To this end, we estimate the following empirical model of $\hat{u}_{it}^2$, the squared innovation of the firm's model-implied misoptimization, and the log of measured firm-level attention:

$$\hat{u}_{it}^2 = \beta_a \cdot \log \text{MacroAttention}_{it} + \chi_{j(i),t} + \Gamma' X_{it} + \epsilon_{it} \tag{94}$$

Absorbed effects at the sector-by-time level partial out all trends, including the cyclical patterns studied earlier. Additional controls X_{it} can include individual fixed effects, to isolate variation at the firm level; and log stock returns and TFP growth, to help further isolate variation in attention unrelated to firm-level fundamentals.

We find that $\beta_a < 0$: higher macroeconomic attention corresponds with smaller production misoptimization (Table 2.3). The main specification in column (1) finds a strongly statistically significant effect (p <0.01). The more controlled specifications in columns (2) to (4) estimate negative, similarly sized effects, but with considerably less precision.

Appendix Table B.7 explores the timing of this relationship. We find only weak evidence of anticipatory effects, or high attention preceding years of low misoptimizations, and strong evidence of persistent effects, or high attention following years of low misoptimizations. Appendix Table B.8 shows that results are similar using our conference-call-based measure of attention as well as all considered alternative models for the misoptimizations.

2.5.5 Discussion and Relationship with the Literature

Before proceeding, we first show our findings are consistent with those from the literature on uncertainty and productivity over the business cycle. Second, we relate our model's predictions to the cyclicality of forecast and backcast errors and the demand for information.

Dispersion in Fundamentals Versus Misoptimizations Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018), using micro-data in the manufacturing sector, estimate that the variance in total factor productivity rises in recessions or periods of negative growth.[34] Our analysis, by contrast, studies variance in input choices *conditional* on productivity, the variance of which varies over time.

To demonstrate the empirical consistency of these sets of findings, we follow Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) and estimate a first-order autoregressive model for TFP with firm and sector-by-time fixed effects:

$$\log \theta_{it} = \gamma_i + \chi_{j(i),t} + \rho_\theta \log \theta_{i,t-1} + \epsilon_{it} \tag{95}$$

We estimate "TFP Innovation Variance" as the weighted average, $\mathbb{E}[s^*_{it}\epsilon^2_{it}]/\mathbb{E}[s^*_{it}]$. Based on the model interpretation of our TFP measurement, this corresponds to the variance of TFPQ.

Appendix Figure B-9 plots our estimate of TFP Innovation Variance against the unemployment rate and detrended stock price. TFP Innovation Variance spikes markedly in the 2002 and 2007-09 recessions, as well as in the 1990s boom. Our measure is weakly and insignificantly counter-cyclical by our measure, with a regression coefficient of 0.051 (SE: 0.153) on unemployment; and it is weakly and insignificantly higher in US recessions, with a regression coefficient on an NBER recession indicator of 0.010 (SE: 0.009).

Our measure of TFP dispersion has a correlation of 0.39 with the equivalent mea-

[34] Kehrig (2015) reports that dispersion in levels of TFPR, or the marginal value product of all inputs under a Cobb-Douglas assumption, is counter-cyclical. In Appendix B.6.3, we study how this object, as well as the related calculation for the value marginal product of labor, behaves in our data.

sure from Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) over a common sample, and this correlation increases to 0.47 if we restrict in our data to the manufacturing sector.[35] This suggests that our measurement of TFP stochasticity among US public firms is consistent with the measurement of Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) using Census data in the manufacturing sector. The Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) measure is weakly and insignificantly *pro-cyclical* based on its regression coefficient on unemployment (coefficient: -0.309, SE: 1.209) and, as reported by Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018), larger in recessions (coefficient: 0.098, SE: 0.022).

Forecast Errors and Backcast Errors Our model made no specific prediction about the accuracy of firm-level forecasts over the business cycle. To illustrate this, consider a firm's expectation in period t of its production in period $t+k$, for $k > 0$, conditional on the observed history of the decision state z_{it} (firm-level TFP, aggregate output, and aggregate wages). Writing $\mathbb{E}_{it}[\cdot]$ as a shorthand for the firms' expectation conditional on this history and applying Proposition 6, we can write the firm's realized forecast error as

$$x_{i,t+k} - \mathbb{E}_{it}\left[x_{i,t+k}\right] = \left(x^*_{i,t+k}(z_{it}) - \mathbb{E}_{it}\left[x^*_{i,t+k}(z_{i,t+k})\right]\right) + \sigma_i(z_{i,t+k}) \cdot v_{i,t+k} \tag{96}$$

The first term is the forecast error about the optimal level of production, $x^*_{i,t+k}$, which itself depends on unknown firm-level productivity, aggregate output, and aggregate wages. The second term is the firms' future misoptimization, the product of state-dependent volatility $\sigma_i(z_{i,t+k})$ and the idiosyncratic noise $v_{i,t+k}$. As the noise is uncorrelated with x^*, the variance of these errors decomposes into two terms which we label below

$$\mathbb{E}\left[\left(x_{i,t+k} - \mathbb{E}_{it}\left[x_{i,t+k}\right]\right)^2\right] = \underbrace{\mathbb{E}\left[\left(x^*_{i,t+k}(z_{i,t+k} - \mathbb{E}_{it}\left[x^*_{i,t+k}\right]\right)^2\right]}_{\text{Fundamental Risk}} + \underbrace{\left(\sigma_i(z_{i,t+k})\right)^2}_{\text{Misoptimization Risk}} \tag{97}$$

Our main aggregate empirical finding was that the second term, on average across firms, is high in booms and low in downturns. The finding of both prior work in the literature (Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry, 2018), and our

[35] The common sample for comparing our measure with the one in Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) is 1987-2010. The measure from Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) that we study is the variance (square of standard deviation) of TFP innovations on the sample of establishments that are in the Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) data for 25 years.

own reconstruction of these findings above, is that the first term is high in downturns and low in booms. Because of these countervailing forces, our main hypotheses about misoptimizations cannot be tested by evaluating the cyclicality of forecast errors.[36]

By contrast, firm-level *backcasts* of specific state variables, like external macroeconomic conditions, may have a more direct interpretation as "attentiveness." In Appendix B.7, we show two pieces of evidence consistent with our analysis in the survey of firms in New Zealand conducted by Coibion, Gorodnichenko, and Kumar (2018a). First, firms that report a higher sensitivity of firm profits to their own choices (in this case, posted prices), demonstrate higher awareness of macroeconomic aggregates. This is consistent with the profit curvature case of our model. Second, firms report that they would be significantly more likely to seek out news about the macroeconomy if there were an aggregate negative shock. This is consistent with the sign of our Attention and Misoptimization Cycles.

2.6 Quantifying the Consequences of Attention Cycles

The previous section verified the model's microeconomic and macroeconomic predictions for misoptimization, which governed the model's implications for output and productivity dynamics. In this section, we quantify the consequences of attention cycles in a numerical calibration of our model. Our strategy is to match the aggregate level and cyclicality of misoptimization-induced dispersion in production, which control the size and dynamics of the attention wedge. We find that attention cycles generate: asymmetrically large amplification of negative shocks; greater amplification of shocks when output is low; and endogenously higher volatility when output is low. Each of these features arises from the endogenous adjustment of attention, and would be missing in the full-attention model benchmark.

2.6.1 Calibration

In our calibration, as in our theoretical results of Section 2.4, the aggregate state variable $\theta_t = \left(\mathbb{E}_{G_t}[\theta_{it}^{\epsilon-1}]\right)^{\frac{1}{\epsilon-1}}$ is a one-dimensional sufficient statistic for the productivity distribution. We assume that $\log\theta_t$ follows a zero-mean, Gaussian AR(1) process:

$$\log\theta_t = \rho_\theta \log\theta_{t-1} + \nu_t \tag{98}$$

[36]See Charoenwong, Kimura, Kwan, and Tan (2021) for an analysis of cyclicality in forecast error variance.

where $\nu_t \sim^{IID} N(0, \sigma_\theta^2)$. Note that, in this formulation, the shocks ν_t may reflect changes to *any* moment of the productivity distribution that induce changes in the aggregate θ_t, such as a standard shock to average productivity or an "uncertainty shock" to productivity dispersion as studied by Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018). Since we will study purely macroeconomic phenomena, we can therefore be agnostic about what shocks to the productivity distribution induce the dynamics of Equation 98.

The model has six parameters.[37] We calibrate two to standard values and match four to measured moments, as reviewed in Table 2.4 and explained below.

The first calibrated parameter is the elasticity of substitution between products. We set $\epsilon = 4$, which implies an optimal average markup of $\frac{\epsilon}{\epsilon-1} = \frac{4}{3}$. This is conservative relative to estimates by De Loecker, Eeckhout, and Unger (2020) (1.60) and Demirer (2020) (1.45) for modern markups for US public firms, and slightly larger than the estimate by Edmond, Midrigan, and Xu (2018) on the same dataset (1.25). The elasticity of substitution controls the translation of misoptimization into output and productivity, with a lower value translating a fixed level of dispersion of production levels around *ex post* optimal levels into a larger penalty for output and productivity. We next set the persistence of the productivity shock, at the quarterly frequency, to a standard value of $\rho = 0.95$.[38]

The four remaining free parameters are the elasticity of real wages to output χ, the average attention cost λ, the coefficient of relative risk aversion γ, and the volatility of the productivity process σ_θ^2. We set these four parameters in the model to match four simulated moments with exact analogues in the data, as described in Table 2.4, and fit the model exactly up to machine precision.

Two of these parameters, χ and σ_θ^2, have a relatively simple interpretation. For the first, we match directly an OLS regression of linearly detrended real wages on linearly detrended GDP, at the quarterly frequency over our studied period 1987-2018.[39] Our estimate of 0.095 lines up with recent evidence on the "flattening" of the wage Phillips curve (Galí and Gambetti, 2019). For the second, we match the variance of quarterly

[37] The constants $\bar{w}$ and $\bar{X}$ in the wage rule scale overall production in equilibrium, but are otherwise irrelevant (see Proposition 6). We set $\bar{w}$ and $\bar{X}$ to match the wage and output prevailing in a frictionless market economy with Greenwood, Hercowitz, and Huffman (1988) preferences over labor and leisure and elasticity of labor supply $\phi = 1$, evaluated at state $\log \theta = 0$.

[38] We study quarterly dynamics to compare our predictions to standard results about business-cycle asymmetries, although our measurement was annual. As our measurement is based on the long-run, cross-sectional variance of misoptimizations, the frequency of calibration is immaterial.

[39] Our real wage series is the median weekly real earnings for wage and salary workers over the age of 16, as reported by the US BLS.

Table 2.4: Parameters for Calibration

Fixed Parameters	ϵ	Elasticity of substitution	4
	ρ_θ	Persistence of productivity	0.95
Free Parameters	χ	Elasticity of real wages to output	0.095
	λ	Average weight on entropy penalty	0.406
	γ	Coefficient of relative risk aversion	11.5
	σ_θ^2	Variance of the productivity innovation	4.82×10^{-7}
Matched Moments	β_U	Slope of Misopt. Dispersion on -$\frac{\text{Unemployment}}{100}$	0.841
	$\bar{\sigma}_M^2$	Average level of Misopt. Dispersion	0.080
	χ	Regression of detrended real wages on detrended output	0.095
	σ_Y^2	Variance of quarterly output growth	0.337

Notes: "Fixed Parameters" are externally set. "Free Parameters" are chosen to fit the "Matched Moments," which are calculated from the data and matched exactly by the model.

real GDP growth over our sample period by scaling the shock variance σ_θ^2.

The key properties of inattention and misallocation are controlled by the remaining two moments. In the model, we calculate the average level of Misoptimization Dispersion, or optimal-sales-weighted dispersion in $\frac{L_{it}-L_{it}^*}{L_{it}^*}$, and the population regression coefficient of Misoptimization Dispersion on log employment.[40] We match these moments respectively to the time-series average of Misoptimization Dispersion, 0.080, and the (negative) slope of Misoptimization Dispersion in unemployment, 0.841. Intuitively, these moments identify the level of misoptimization and the extent of its cyclicality. In the model, conditional on all other parameters, they identify the average cost of attention λ, itself a sufficient statistic for the cross-sectional distribution, and the coefficient of relative risk aversion γ, which controls the risk-pricing incentive that drives cyclical misoptimization.

Our estimates of γ are thus based entirely on fitting a stochastic discount factor that justifies the observed pattern of misoptimization in our model, rather than incorporating an informed prior from external evidence in financial economics as discussed in Section 2.4.2. Our finding of $\gamma = 11.5$ (Table 2.4) is, if anything, slightly conservative relative to the modern asset-pricing literature that estimates, in variations of the consumption capital asset pricing (CCAPM) model, γ of about 15-20 with

[40] We calculate the "regression coefficient" in the model using numerical integration over the stationary distribution of states. Appendix B.1.7 shows the exact in-model formula for aggregate Misoptimization Dispersion, calculated with optimal-sales-weights, which depends only on the sufficient statistics (θ, λ) of the cross-sectional type distribution.

statistically "unfiltered" measures of consumption (Savov, 2011; Kroencke, 2017) or long-run variation in consumption growth (Parker and Julliard, 2005). As previously discussed, our finding of $\gamma > 0$ suggests that the risk-pricing channel is necessary to explain misoptimization dynamics.

2.6.2 Output, Productivity, and the Attention Wedge

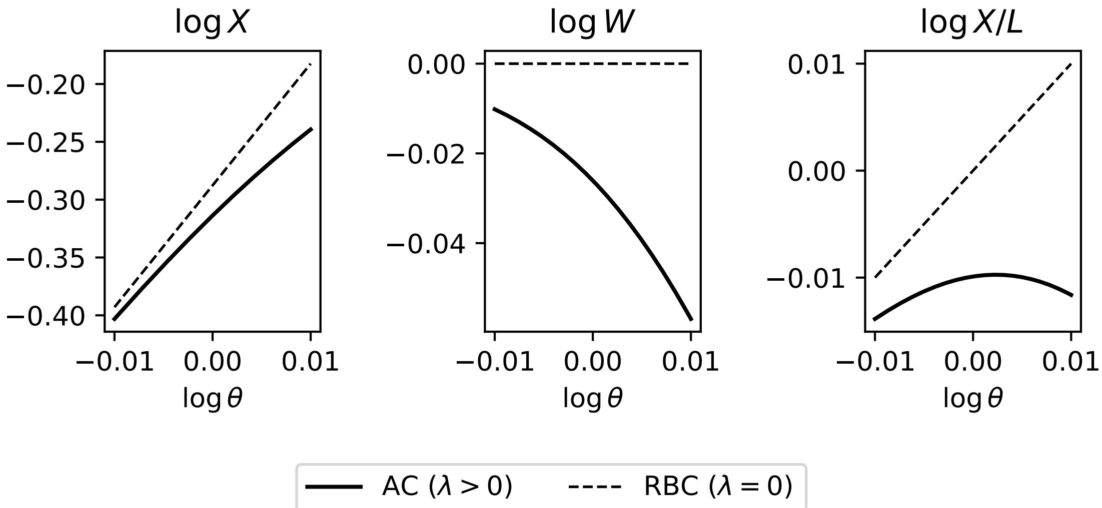
Figure 2-4 shows log output, the log attention wedge (as defined in Proposition 9), and log labor productivity in our calibration. In each case, we compare to the predictions of an otherwise identical "pure RBC model" with full attentiveness, or $\lambda = 0$, plotted as a dashed line. As implied by Proposition 9 and Corollary 5, inattention reduces output and productivity relative to the fully attentive counterfactual, and this effect grows in larger states. In the mean state, the attention wedge reduces output by 2.6% relative to the fully attentive counterfactual, labor productivity by $\chi\epsilon \times 2.6\% = 1.0\%$, and employment by $(1 - \chi\epsilon) \times 2.6\% = 1.6\%$.

To highlight the importance of studying misoptimization incentives in general equilibrium, we calculate also a "partial equilibrium" attention wedge based on firms' inattentive best-responses to the counterfactual RBC dynamics (Appendix Figure B-10). In the mean state, the "partial equilibrium" attention wedge is 1.3% in terms of output, implying that general-equilibrium interactions account for $1/2$ of the losses from inattention.

We observe two further properties of output and productivity dynamics which were not immediately clear from the theoretical results and depend on the numerical calibration. First, labor productivity is non-monotone in productivity θ (rightmost panel) and hence also in aggregate output. As we mentioned in the discussion following Corollary 5, this property results from the dueling forces of productivity shocks and induced misallocation.

Second, we find that the attention wedge is concave (Appendix Figure B-10 shows this concavity directly and provides also a partial-vs.-general-equilibrium decomposition of the finding). Corollary 6 showed how the concavity or convexity of the attention wedge leads to business cycle asymmetries. The finding of a concave attention wedge implies that, fixing shock sizes, negative shocks have a larger effect on output than positive shocks, and that overall shock responses and volatility are higher in low-output states. We explore these predictions quantitatively in the next subsection.

Figure 2-4: Output, Attention Wedge, and Labor Productivity



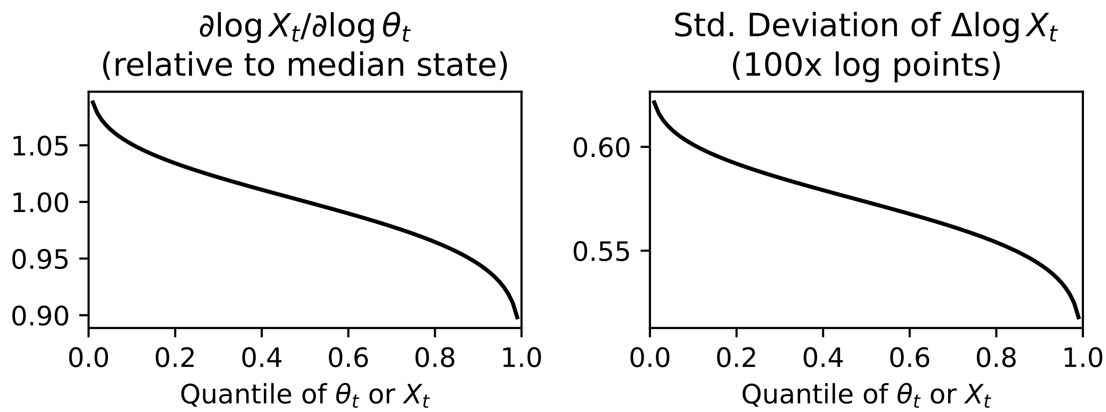
2.6.3 State-Dependence, Asymmetry, Stochastic Volatility

Figure 2-5 plots the marginal sensitivity of output to percentage productivity shocks and the standard deviation of output growth as a function of the initial quantile of the stationary distribution for productivity or output. As suggested by the discussion of the concave attention wedge above, both model objects are decreasing functions of the state.

One way to benchmark the extent of asymmetry and state-dependence in shock responses is to consider an "impulse response" thought experiment of a fixed size. Let $\log \hat{\theta}$ be the fundamental shock that induces a 3% change in output from steady state, or solves $\log X(\log \hat{\theta}) - \log X(0) = 0.03$. We compare the effect of this "Positive" shock to the effect of a "Negative" shock from $\log \theta_0 = 0$ to $\log \theta_1 = -\log \hat{\theta}$, and a "Double Dip" shock from $\log \theta_0 = -\log \hat{\theta}$ to $\log \theta_1 = -2 \cdot \log \hat{\theta}$. We find that the negative shock has a 7% larger effect on output than the positive shock, and the double dip shock has a 14% larger effect on output than the positive shock. The same results for the response of total labor are 5% and 8%, respectively. Empirically, Ilut, Kehrig, and Schneider (2018) estimate that US industries have on average a 20% larger response to negative aggregate productivity shocks than positive shocks.[41] In comparison to this benchmark, our model explains 25% of the shock response via the endogenous reallocation of attention. More generally, to the extent that linear macroeconomic models like our RBC benchmark are plagued with the descriptively

[41] This calculation is based on comparing the "data" estimates in columns 6 and 7 of Table 9 in that paper.

Figure 2-5: Asymmetric Shock Response and Stochastic Volatility



undesirable result of left-tailed shock distributions, which cannot even be explained by symmetric, exogenous stochastic volatility, the attention-cycles mechanism offers a partial solution.

To benchmark the extent of stochastic volatility, we observe that a transition from the 90th-percentile to the 10th-percentile productivity state reduces output by 4.7% and increases the conditional standard deviation of output growth by 10.6%. The peak-to-trough fall of output during the Great Recession (e.g., from early 2007 to early 2009) is comparable to this level change. Empirically, Jurado, Ludvigson, and Ng (2015) estimate that the forward-looking volatility of industrial production growth increased by 57% over three-month horizons during this episode.[42] The endogenous re-allocation of attention in our model can explain about 19% of this movement, without recourse to underlying stochastic volatility in fundamentals.

2.6.4 Parameter Robustness and Counterfactual Scenarios

In Appendix B.5, we provide additional results from our numerical exercise. We first explore the robustness of our main findings to different external calibrations of wage rigidity χ and substitutability ϵ and to introducing classical labor markets using the preferences of Greenwood, Hercowitz, and Huffman (1988). We also study the effects of calibrated attention cycles on macroeconomic dynamics under counterfactual scenarios by altering structural parameters without recalibrating the model. To sum-

[42]This calculation compares the 3-month "macro uncertainty index," as available at `https://www.sydneyludvigson.com/macro-and-financial-uncertainty-indexes`, from a low point in April 2007 to a high point in October 2008.

marize, we find that attention cycles generate more pronounced differences from the log-linear, RBC core in regimes with larger markups, greater wage rigidity, and higher attention costs.

2.7 Conclusion

This paper studies attention cycles as both a cause and consequence of macroeconomic dynamics. We develop new measures of firms' attention to the macroeconomy and misoptimization in decision making. We document that macroeconomic attention is counter-cyclical, that the extent of misoptimization is pro-cyclical, that more macroeconomically attentive firms make smaller misoptimizations, and that the market financially punishes firms for misoptimizations more in downturns. We build a macroeconomic theory to understand why cyclical attention should manifest in both partial and general equilibrium by embedding state-dependent stochastic choice in a business cycle model. We derive empirically reasonable conditions under which attention is highest in downturns owing to the higher stakes for making correct decisions, or higher curvature of firms' objectives as a function of choices, in these states. Consistent with our empirical findings, the main contributing macroeconomic force is *risk-pricing* or higher marginal utility in low-production states of the world. Calibrating the model to match our evidence on cyclical misoptimization, we uncover a quantitatively important role for attention cycles in driving macroeconomic dynamics. Attention cycles cause asymmetrically larger propagation of negative shocks, larger propagation of all shocks in low-output states and endogenously higher volatility in low-output states.

Chapter 3

Strategic Mistakes

This chapter is jointly authored with Karthik A. Sastry.

3.1 Introduction

People commonly make mistakes that affect others. Consider a monopolistically competitive firm choosing its price to maximize profits, taking into account projected demand and competitors' prices. The complexity of firms' decision-making processes makes clear that even though the problem is well-defined and an ideal solution surely exists, determining that solution is difficult. Thus, firms may fail to set the optimal price. Such a deviation from the ideal price may affect all other competitors' *benefits* from setting the right price—for instance, by altering the residual demand that they face. Moreover, the pricing of other firms may directly influence the *costs* of setting the right price—for instance, if tough competition induces managerial stress that contributes to worse decision-making. Thus, observed pricing arises from a process of *strategic mistakes*: the combination of imperfect optimization and strategic interaction that may affect both the benefits and the costs of precise decision-making.

To study such strategic mistakes, this paper introduces a model of non-parametric, state-dependent stochastic choice in continuum-player games with a continuum of actions. Agents' payoffs depend on their own action, an exogenous state and a one-dimensional aggregate of the cross-sectional distribution of others' actions. Such a setting is ubiquitous in macroeconomic models of price-setting (Woodford, 2003a; Maćkowiak and Wiederholt, 2009; Costain and Nakov, 2019), production (Angeletos and La'O, 2010, 2013; Benhabib, Wang, and Wen, 2015; Chahrour and Ulbricht, 2023), and beauty-contest games more generally (Morris and Shin, 2002; Angeletos and Pavan, 2007; Bergemann and Morris, 2013; Huo and Pedroni, 2020).

Agents face a problem of *costly control*: conditional on their conjecture for fundamentals and others' actions, they pick a stochastic choice pattern that trades off playing the best actions with a cost that penalizes playing too precisely. We introduce a new family of control cost functionals that are *state-separable*, *i.e,* total control costs are additive over states. These costs allow us to model several kinds of decision frictions that have not previously been jointly studied. The first is *ex post* misoptimization, as in the literatures on control costs (Stahl, 1990; Van Damme, 1991) and quantal response equilibrium (McKelvey and Palfrey, 1995; Goeree, Palfrey, and Holt, 2016), in which agents' imprecise play responds to strategic incentives within a given state of the world. The second is *ex ante* planning frictions, as in the literature on costly information acquisition in games (see *e.g.,* Yang, 2015; Morris and Yang, 2019; Hébert and La'O, 2022; Denti, 2023), whereby agents must weigh the benefits of precise planning for a state with the cost of that state never being realized. The third is exogenous and endogenous state-dependence in control costs, as in Hébert and La'O (2022) and Angeletos and Sastry (2023). The fourth is equilibrium determination of agents' consideration sets, *i.e,* the subset of actions that they play, as in Matějka (2015) and Stevens (2019).

We show that, despite the rich behavioral patterns that our model accommodates, equilibrium analysis remains tractable. Concretely, we provide four theoretical results that provide conditions for equilibrium existence, uniqueness, efficiency, and monotone comparative statics for actions, aggregates, and the size of agents' mistakes.

Toward establishing the existence and uniqueness of equilibrium, we first characterize equilibrium as a functional fixed-point equation for the cross-sectional distribution of actions and provide primitive conditions under which this equilibrium fixed-point operator is a contraction. This result follows from showing first that agents' state-dependent stochastic choice rules are *monotone*, *i.e.,* they are increasing in the sense of first-order stochastic dominance when aggregate actions are higher, and *discounted*, *i.e.,* increasing aggregate outcomes has a less than one-for-one effect on agents' stochastic best replies. This requires three primitive conditions: (i) that agents' actions and aggregates are jointly complementary for physical payoffs and the psychological costs of precise optimization; (ii) that this complementarity is dominated by the concavity of agents' physical payoffs relative to their psychological costs; and (iii) a technical restriction on the shape of agents' cost functionals that allows us to translate dominance in payoff units into first-order stochastic dominance in the space of stochastic choice rules. Moreover, we show that the last of these assumptions is satisfied under the two leading cost functions in the control costs literature: entropic and quadratic costs. Second, we show that, if the equilibrium aggregator is (i) increasing in agents' actions and (ii) such that level shifts of the action distribution have less than one-for-one effects on aggregates, then the equilibrium fixed-point operator is a contraction. These assumptions on aggregation are satisfied under common aggregators, such as those that take the mean or the median of the cross-sectional action distribution. Finally, since the equilibrium operator is a contraction, the existence and uniqueness of equilibrium (Theorem 1) follows.

We next study equilibrium comparative statics. First, if actions, aggregates, and the state are jointly complementary for agents' physical payoffs and psychological costs, then the unique equilibrium action distribution and aggregate are monotone in the state (Theorem 2). Under a further condition that payoffs depend only on the distance between one's own action and some optimal action, we show that the size of agents' mistakes is monotone in the state when the ratio between the stakes of misoptimization and the cost of precise optimization is monotone in endogenous and exogenous states (Theorem 3).

Turning to normative analysis, we provide a necessary condition for the efficiency of the unique equilibrium: the average marginal physical benefit of increasing the aggregate action must equal the average marginal psychological cost of so doing (Theorem 4).

We finally employ our results in two macroeconomic applications. The first application is to price-setting in a monetary economy à la Woodford (2003a) and Hellwig and Venkateswaran (2009), but where firms face *ex ante* planning frictions: firms must plan for what prices to set across contingencies for the realized level of the money supply and inflation. We derive and interpret conditions under which the aggregate price level, the distribution of prices, and the dispersion of prices are monotone increasing in an exogenous shock to the money supply in the unique equilibrium. We use these results to give a costly-planning explanation for the empirical finding that there is a positive relationship between price dispersion and aggregate inflation at rare and high, but not common and low, levels of inflation (Alvarez, Beraja, Gonzalez-Rozada, and Neumeyer, 2019; Nakamura, Steinsson, Sun, and Villar, 2018). The key mechanism is that firms set more dispersed prices in rare, highly inflationary states, because they did not invest many resources into forming precise plans for these unlikely states.

The second application is to consumption and savings in a liquidity trap, in which agents' incomes directly influence cognitive function. This is motivated by the experimental finding that individuals make worse decisions when they are poor (Mani, Mullainathan, Shafir, and Zhao, 2013) and the survey finding that individuals report being significantly distracted when near financial constraints (Sergeyev, Lian, and Gorodnichenko, 2022). We derive and interpret conditions under which the unique equilibrium features aggregate output and a consumption distribution that is monotone increasing in aggregate demand, while consumption dispersion decreases in aggregate demand as agents become less cognitively constrained. We show that this economy features a novel externality: when aggregate output is lower, agents' decision costs are higher, and they make larger consumption and savings mistakes. This mechanism provides a new explanation for the finding that consumption dispersion rises in downturns (Berger, Bocola, and Dovis, 2023), in our case because of the equilibrium effect of low income causing stress that worsens decisionmaking.

We discuss two extensions of our analysis in the Appendix. First, in Appendix C.2, we provide a detailed comparison of our model with the mutual information model of Sims (2003). Using a numerical example of a linear beauty contest (Morris and Shin, 2002), we observe that the mutual information model does *not* imply monotone and discounted stochastic choice rules and therefore opens the door to multiple equilibria defined by coordination on specific support points for the action distribution. This analysis provides a direct counter-example to the possibility that equilibrium analysis similar to ours is possible in workhorse models of unrestricted information acquisition and illustrates the tractability advantage that our model may have for specific applications. Second, in Appendix C.3, we study strategic mistakes in binary-action coordination games, which are also ubiquitous in macroeconomics and finance (Angeletos and Lian, 2016). We derive sufficient conditions on cognitive costs and payoffs to ensure unique and monotone equilibria and illustrate our results in a canonical investment game with linear payoffs (as in Yang, 2015).

Related Literature The main contribution of our paper is to provide a unified equilibrium analysis of a wide variety of decision frictions—including *ex post* misoptimization, imperfect *ex ante* planning, endogenous cognitive constraints, and endogenous consideration sets—in aggregative games of the kind that are common in macroeconomics and finance (see Angeletos and Lian, 2016, for a review). To our knowledge, comparable results on uniqueness, efficiency, and monotone comparative statics for these games do not exist in the literatures on the two most comparable decision frictions, random utility and costly information acquisition. We detail our connection to these literatures below.

An influential model of equilibrium with non-vanishing "mistakes" induced by random utility is the Quantal Response Equilibrium (QRE) of McKelvey and Palfrey (1995). These authors add type-I extreme value noise to agents' utility functions to smooth best responses into "better responses" (see the review by Goeree, Palfrey, and Holt, 2016). Subsequent work generalizes this analysis by allowing for different noise distributions that imply different shapes of best-replies (see *e.g.,* Melo, 2022; Fosgerau, Melo, De Palma, and Shum, 2020; Allen and Rehbeck, 2021). Most related to us, Melo (2022) studies games with a finite number of players and actions and general noise distributions and, using convex analysis techniques, shows that QRE are unique if agents' payoffs are sufficiently concave relative to the extent of strategic complementarity. Our analysis differs from this literature's in four important ways. First, we consider games with a continuum of agents and actions. This is important because such games are common in macroeconomics and finance and, outside of Melo's (2022) analysis with discrete actions and players, little remains understood about the uniqueness of QRE in games with a large number of players and/or actions (Goeree, Palfrey, and Holt, 2016). Second, we provide monotone comparative statics results for action distributions, in terms of both first-order stochastic dominance and dispersion. We are not aware of any analogous results in the random-utility literature for the class of games that we study. Third, we can accommodate additional decision frictions which are not well captured by fixed payoff noise—for instance, costly *ex ante* planning and endogenous cognitive constraints. Finally, our analysis has meaningfully different normative properties because we model control costs.

With unrestricted costly information acquisition, we are aware of few equilibrium results that apply to our setting. Hébert and La'O (2022) provide sufficient conditions for equilibrium existence and efficiency in a setting with costly information acquisition under restrictions, relative to our set-up, to consider only mean-critical payoff functions and only the mean aggregator. Hébert and La'O (2022) provide an equilibrium uniqueness result *only* when equilibria are efficient, while our result applies to both efficient and inefficient equilibria under appropriate restrictions on complementarity arising from both payoffs and endogenous cognitive costs. Yang (2015) and Morris and Yang (2019) study equilibrium existence and uniqueness in *binary-action* settings, to which we extend our analysis in Appendix C.3. To our knowledge, no references study monotone comparative statics at our level of generality.

Our paper contributes to the theoretical literature on aggregative games (see Jensen, 2018, for a review) by studying these games under general decision frictions. Our analysis also relates to a large literature on uniqueness in games with strategic complementarity (*e.g.,* Morris and Shin, 1998, 2002; Weinstein and Yildiz, 2007; Yang, 2015). Our proof strategy is most closely related to Frankel, Morris, and Pauzner (2003) and Mathevet (2010), in that we use contraction-mapping techniques, but differs in our use of variational techniques to derive necessary conditions for best responses that imply monotonicity and discounting. Our results on comparative statics are similar in spirit to those of Van Zandt and Vives (2007), but differ in that we study different games, with decision frictions, and provide comparative statics for action distributions.

Finally, our paper contributes to the literature on control costs and stochastic choice by proposing a new class of state-separable cost functionals and applying them in games. This builds upon the analysis of Harsanyi (1973), Stahl (1990), and Van Damme (1991) who introduce specific control cost functionals that penalize the playing of sharply peaked stochastic choice rules, and Mattsson and Weibull (2002), who axiomatize entropic costs. Most relatedly, in decision problems, Fudenberg, Iijima, and Strzalecki (2015) axiomatize the class of additive perturbed utility cost (APU) functionals which penalize the expected utility of a mixed action with *any* convex function of the distribution of the mixed action. Our cost function is a weighted sum of APU cost functionals over states, with a weighting function that can depend arbitrarily on both exogenous and exogenous states. Concretely, with weights given by the agents' priors, our cost functional reduces to a state-by-state APU control cost functional that models *ex post* misoptimization. With uniform weights across states, our cost functional captures *ex ante* planning, as control costs must be incurred *ex ante*, while the benefits of plans only realize with probabilities given by the agents' priors. With state-dependent weights, our cost functional allows for exogenous and endogenous state-dependence in the difficulty of choosing precise stochastic choice rules. As we later argue (see Section 3.2.3), capturing this broad range of behavior enables our class of cost functions to be consistent with the empirical regularities from the psychometrics literature (see Woodford, 2020, for a review), the literature on stress and decision-making (Mani, Mullainathan, Shafir, and Zhao, 2013; Sergeyev, Lian, and Gorodnichenko, 2022), and the perceptual tests performed by Dean and Neligh (2022).

Outline Section 3.2 introduces the model. Section 3.3 presents our main results on equilibrium properties. Section 3.4 discusses applications of our main results. Section 3.5 briefly discusses two extensions, a detailed comparison of state-separable and mutual information costs and an analysis of binary-action games. Section 3.6 concludes.

3.2 Model

3.2.1 Basic Set-up: Aggregative Games with Stochastic Choice

A continuum of identical agents is indexed by $i \in [0,1]$. They take actions $x_i \in \mathcal{X} = [\underline{x}, \overline{x}] \subset \mathbb{R}$. Cross-sectional distributions of actions are aggregated by an aggregator functional $X : \Delta(\mathcal{X}) \to \mathbb{R}$. There is an underlying and payoff-relevant state of the world $\theta \in \Theta \subset \mathbb{R}$. The state space Θ is a finite set, over which the agent has full-support prior $\pi \in \Delta(\Theta)$. Agents have identical utility functions $u : \mathcal{X} \times \mathbb{R} \times \Theta \to \mathbb{R}$, where $u(x, X, \theta)$ is an agent's utility from playing x when the aggregate is X and the state is θ. We assume that u and X are continuous and bounded.[1]

Given a conjecture that the aggregate follows a law of motion $\tilde{X} : \Theta \to \mathbb{R}$, which lies in the space of bounded functions $\mathcal{B} = \{\tilde{X} | \tilde{X} : \Theta \to \mathbb{R}\}$, each agent chooses a stochastic choice rule $P : \Theta \to \Delta(\mathcal{X})$ with $P(x|\theta)$ describing the cumulative distribution of actions x taken by the agent in state θ. When this admits a density function, we denote a stochastic choice rule by $p(x|\theta)$. We call the set of measurable stochastic choice rules $\mathcal{P}$. We model the cost of "controlling mistakes" via a cost functional $c : \mathcal{P} \times \mathcal{B} \to \overline{\mathbb{R}}$. This cost may depend on both the conjectured mapping from states to aggregates (as indicated) and on the prior for the state of nature (suppressed, as this prior is fixed in our analysis). In the next subsection, we specialize these costs to a specific class for our analysis.

The agent maximizes expected utility net of the control cost given their conjecture for how aggregate outcomes depend on the state. This is summarized in the following program:

$$\max_{P \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \tilde{X}(\theta), \theta) \, \mathrm{d}P(x|\theta) \, \pi(\theta) - c(P, \tilde{X}) \tag{99}$$

An equilibrium in this context is a Nash equilibrium: agents' play is optimal given aggregate outcomes, and aggregate outcomes are those that are implied by agents' play.[2]

Definition 2 (Equilibrium)**.** *An equilibrium is a collection of stochastic choice rules $\{P_i^*\}_{i \in [0,1]}$ and an equilibrium law of motion for aggregates $\hat{X} : \Theta \to \mathbb{R}$ such that:*

1. *All agents solve Program 99 under the conjecture that $\tilde{X}(\theta) = \hat{X}(\theta)$ for all*

[1] Throughout, our notion of continuity for functionals is the sup norm.

[2] Methodologically, our setup recasts the game with incomplete information in the interim as an *ex ante* game with complete information and a strategy space sufficiently rich to embed all profiles of state-dependent mixed strategies. Morris and Yang (2019) use this approach to study *binary-action* games, as we also do in Appendix C.3.

$\theta \in \Theta$

2. *The equilibrium law of motion is consistent with agents' play, or* $\hat{X} = X \circ \int_{[0,1]} P_i^* \,\mathrm{d}i$

An equilibrium is symmetric if $P_i^* = P^*$ *for all* $i \in [0,1]$.

3.2.2 State-Separable Cost Functionals

We now specialize to a new class of cost functionals that we introduce:

Definition 3 (State-Separable Cost Functional)**.** *A cost functional* c *has a state-separable representation if there exists a strictly convex function* $\phi : \mathbb{R}_+ \to \mathbb{R}$ *and a weighting function* $\lambda : \mathbb{R} \times \Theta \to \mathbb{R}_{++}$ *such that for any stochastic choice rule* P *with density* p*:*

$$c(P, \tilde{X}) = \sum_{\Theta} \lambda(\tilde{X}(\theta), \theta)\pi(\theta) \int_{\mathcal{X}} \phi(p(x|\theta)) \,\mathrm{d}x \tag{100}$$

with the convention that the cost is ∞ *if* P *does not have a density.*

Formally, state-separable cost functionals are a weighted sum across states of the APU cost functionals of Fudenberg, Iijima, and Strzalecki (2015). Informally, state-separable cost functionals capture the idea that it is costly for agents to control "mistakes." The costs of controlling mistakes in different states potentially depend on both the identity of the states and the endogenous outcomes predicted for those states via the weighting function $\lambda(X, \theta)$.

In the remainder of this subsection, we give four specific examples of state-separable costs that capture *ex post* misoptimization, *ex ante* planning frictions, endogenous cognitive constraints, and endogenous consideration sets. In the next subsection, we discuss how these costs are consistent with empirical evidence on decision frictions.

***Ex Post* Misoptimization with Entropy Costs** As a first example, we consider a case in which $\phi(p) = p \log p$ and $\lambda(X, \theta) \equiv \bar{\lambda} > 0$. These costs equal the expectation of the negative entropy of the conditional action distributions. Expected entropy costs encode that precise choice is costly and, therefore, that agents will *ex post* misoptimize. The expected entropy cost model is often applied in macroeconomics to study *ex post* misoptimization (*e.g.,* Costain and Nakov, 2019; Macaulay, 2020; Flynn and Sastry, 2022b). Expected entropy costs imply optimal action distributions of the following "logit" form:

$$p(x|\theta) = \frac{\exp\Big(\bar{\lambda}^{-1} u(x, \tilde{X}(\theta), \theta)\Big)}{\int_{\mathcal{X}} \exp\Big(\bar{\lambda}^{-1} u(z, \tilde{X}(\theta), \theta)\Big) \, \mathrm{d}z} \tag{101}$$

When the set of actions is discrete, these choice patterns are identical to those generated in the model of McFadden (1973) in which agents perceive the perturbed utility function $\tilde{u}(x, X, \theta) = u(x, X, \theta) + \varepsilon_x$, where ε_x is distributed type-I extreme value and IID across agents and actions. This model is ubiquitous for modeling consumer demand in industrial organization (see, *e.g.*, Berry and Haile, 2021). The same model for choice is applied in game theory by McKelvey and Palfrey (1995) to define Quantal Response Equilibrium. However, our entropy-cost case differs from what is studied in these references in two key ways. First, actions in our model are continuous. Second, our model's normative analysis is much different. Control costs model the fact that avoiding mistakes has real costs, while random utility treats random choices as *ex post* optimal.

Finally, Matějka and McKay (2015) show that logit choice can be obtained as a limit case of a model of information acquisition with mutual information costs plus a restriction that all actions are *ex ante* exchangeable. We revisit this last connection in Appendix C.2, which studies the difference between our model and the mutual information model.

Prior-Dependence and Imperfect *Ex Ante* Planning As a second example, consider an arbitrary kernel, but now set $\lambda(\theta) = \pi(\theta)^{-1}\bar{\lambda}$. In this case, the agent's cost functional is given by:

$$c(P) = \bar{\lambda} \sum_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta)) \, \mathrm{d}x \tag{102}$$

This captures costly planning, where the agent plans for each state in advance, and then implements these plans when states realize. Thus, costs of planning actions are incurred *ex ante*, and are therefore proportional to the *number of required plans* (*i.e.*, states contemplated) and not their likelihood of occurring. Under the entropy kernel, this process generates the following choice probabilities:

$$p(x|\theta) = \frac{\exp\Big(\bar{\lambda}^{-1} \pi(\theta) u(x, \tilde{X}(\theta), \theta)\Big)}{\int_{\mathcal{X}} \exp\Big(\bar{\lambda}^{-1} \pi(\theta) u(z, \tilde{X}(\theta), \theta)\Big) \, \mathrm{d}z} \tag{103}$$

The agent optimally chooses to form better plans in states that they believe to be more likely. Concretely, the agent trades off the benefits of precise planning in a state against the cost that the state will not be realized and the plan will be useless. This allows us to capture the idea that agents rationally may prepare for very rare events, even if actions during those events is very important (an idea also proposed by Maćkowiak and Wiederholt, 2018). In Section 3.4.1, we apply this model to study equilibrium price-setting by monopolistically competitive firms in a monetary macroeconomic model.

Endogenous Cognitive Constraints We next consider an example that sets $\lambda(X,\theta) = \tilde{\lambda}(X)$, for some decreasing function of X. Combined with the normalization that u is monotone in X, this embodies the possibility that more favorable aggregate outcomes decrease decision costs while less favorable aggregate outcomes increase decision costs. A leading example studied by Mani, Mullainathan, Shafir, and Zhao (2013) and Mullainathan and Shafir (2013) is that poverty impedes cognitive ability and induces mistakes in decisions. Our framework can model the possibility that this force is endogenous to others' actions and/or mistakes, insofar as income is determined in equilibrium. Under the entropy kernel, choice probabilities follow:

$$p(x|\theta) = \frac{\exp\Big(\tilde{\lambda}(\tilde{X}(\theta))^{-1} u(x, \tilde{X}(\theta), \theta)\Big)}{\int_{\mathcal{X}} \exp\Big(\tilde{\lambda}(\tilde{X}(\theta))^{-1} u(z, \tilde{X}(\theta), \theta)\Big)\, \mathrm{d}z} \tag{104}$$

In states with low weights $\tilde{\lambda}(X)$, when aggregate outcomes are good and stress is low, choices are more precisely concentrated on high-payoff choices; in states with high weights $\tilde{\lambda}(X)$, when aggregate outcomes are bad and stress is high, the opposite is true. Thus, in both cases, the characteristics of aggregate states and their psychological effects shape choice in ways that are not summarized by physical payoffs. In Section 3.4.2, we study a macroeconomic model in which endogenous stress shapes the determination of aggregate demand and income.

Consideration Sets with Quadratic Costs We now consider the quadratic kernel $\phi(p) = \bar{\lambda}\frac{p^2}{2}$ studied by Rosenthal (1989). Like the entropy kernel, the quadratic kernel penalizes action distributions that are more sharply peaked and rewards those that are more thinly spread. Unlike the entropy kernel, the quadratic kernel allows for agents to put exactly zero probability on certain actions. In the marketing literature, this phenomenon of agents playing only a strict subset of possible actions is sometimes referred to as a "consideration set" (*e.g.,* Hauser and Wernerfelt, 1990).

In the context of rational inattention models, Jung, Kim, Matějka, and Sims (2019), Caplin, Dean, and Leahy (2019), and Fosgerau, Melo, De Palma, and Shum (2020) study this phenomenon.

We now illustrate how consideration sets emerge. Choice probabilities follow:

$$p(x|\theta) = \frac{1}{\lambda}(u(x, \tilde{X}(\theta), \theta) - \bar{u}(\tilde{X}(\theta), \theta)) \cdot \mathbb{I}[u(x, \tilde{X}(\theta), \theta) \geq \bar{u}(\tilde{X}(\theta), \theta)] \tag{105}$$

where $\mathbb{I}[\cdot]$ is the indicator function and $\bar{u}(\tilde{X}(\theta), \theta)$ is defined such that $\int_{\mathcal{X}} p(x|\theta)\,\mathrm{d}x = 1$. The consideration set of actions in state θ is therefore given by:

$$\mathcal{X}(\theta, \tilde{X}) = \{x \in \mathcal{X} : u(x, \tilde{X}(\theta), \theta) \geq \bar{u}(\tilde{X}(\theta), \theta)\} \tag{106}$$

If $\bar{u}(\tilde{X}(\theta), \theta) > \min_{\mathcal{X}} u(x, \hat{X}(\theta), \theta)$, then a strictly positive (Lebesgue) measure of actions is chosen with zero probability in state θ. In general, without further assumptions, this set can contain many disjoint intervals. However, if u is quasiconcave in x, then $\mathcal{X}(\theta; \hat{X})$ is a closed interval. Finally, observe that these consideration sets are endogenous to equilibrium outcomes as they depend on the equilibrium aggregate.

More generally, away from the quadratic kernel, consideration sets can obtain when ϕ does not satisfy an Inada condition, *i.e*, when $\lim_{p\to 0} \phi'(p) > -\infty$.

3.2.3 Experimental Evidence and Comparisons to the Literature

Having illustrated the model's capacity to generate a rich set of decision frictions, we now assess the model's ability to match experimental evidence. We compare and contrast this with the ability of random utility and costly information acquisition models to do the same. We organize this discussion around five key stylized facts that emerge from the classic literature in experimental economics and experimental psychology surveyed by Woodford (2020), the state-of-the-art perceptual study by Dean and Neligh (2022), and the cognitive experiments of Mani, Mullainathan, Shafir, and Zhao (2013).

Fact 1: Choice is Random People make inaccurate and random judgments in decision problems. These imperfect random choices are often measured in experiments that ask participants to pick which of two stimuli is larger (*i.e.,* which noise is louder) and summarized as *psychometric functions* that plot the probability of choosing the correct option against objective differences in the stimuli that are varied across experiments. These typically reveal a smooth, monotone relationship that is interior to $(0, 1)$ (see *e.g.,* Figure 1 of Woodford, 2020, and each of the experiments in Dean and Neligh, 2022). The state-separable, random utility, and information acquisition models all rationalize random choice. The state-separable model does so by making precise optimization costly.

Fact 2: Choice Responds to Incentives People make more accurate and precise choices when the payoffs from doing so are higher. In perceptual tasks, error rates decrease in rewards (see, *e.g.,* Figure 2b of Woodford, 2020, and Experiment 2 in Dean and Neligh, 2022). The state-separable and rational inattention model this as a rational response to higher returns to cognitive effort; the random utility model generates a similar prediction because larger payoff differences drown out fixed payoff noise.

Fact 3: Choice Depends on Prior Beliefs People's random choice responds to the probabilities of states, in repeated experiments where it may be reasonable to interpret these as prior beliefs. In repeated perceptual tasks, average error rates are lower in states that recur more often (see, *e.g.,* Figure 2a of Woodford, 2020, and Experiment 3 of Dean and Neligh, 2022). This is consistent with state-separable costs that capture *ex ante* planning, as agents have incentives to exert more effort to prepare for more likely states. This result is also natural in many models of costly information acquisition. However, this result cannot be understood through the lens of random utility models (or, in games, QRE), as they embody no notion of *ex ante* planning and agents' priors are irrelevant.

Fact 4: Choice Depends on Decision Context The accuracy and precision of choice vary with the "context" of decision problems, such as the action space and the state space.

First, Dean and Neligh (2022) show the importance of the action space. In Experiment 1, the authors first ask participants to pick between two options. The authors then introduce a third choice (*i.e.,* expand the action space). They find that this *increases* the probability of one of the initial actions. This is consistent with state-separable costs exactly when different action spaces affect the difficulty of making choices (modeled through a change in the value of the weighting function). As observed by Dean and Neligh (2022), this is also consistent with models of costly information acquisition, but *inconsistent* with models of random utility, which predict that larger action spaces decrease the probabilities that all actions are played.

Second, three examples demonstrate the impact of changes in the *state space*. Experiment 4 in Dean and Neligh (2022) shows that choice probabilities are more inaccurate when participants are asked to distinguish states that look more similar. Woodford (2020) surveys two related results in the psychometric literature. First, when laboratory participants are asked to reproduce a set of unknown distances from memory, they overestimate the shorter distances and underestimate the longer distances on average (Figure 4 of Woodford, 2020). Second, the extent of bias can depend systematically on the scale of stimuli (Figure 5 of Woodford, 2020). All of these results are consistent with the state-separable model where the weighting function depends on the state space, capturing the idea that some problems are easier to solve than others. These results are also consistent with information acquisition models that emphasize that the topology of the state space matters (e.g., Hébert and Woodford, 2020). However, they are inconsistent with information acquisition models that satisfy the Invariance Under Compression Axiom (Caplin, Dean, and Leahy, 2022), such as the canonical mutual information cost proposed by Sims (2003).

Fact 5: Choice Depends on Decision-Irrelevant Context The accuracy and precision of decisions can also depend on context that is *not* decision relevant. For example, Mani, Mullainathan, Shafir, and Zhao (2013) show that performance on abstract cognitive tasks declines when individuals are reminded of the difficulty of making financial decisions under poverty or, for predictable reasons, have higher or lower income from a seasonal cycle. In each case, except for the interaction with the (small) financial incentives, income could be viewed as irrelevant for the decision problem solved.

As mentioned earlier, our state-separable model can embody this property directly via appropriate specification of how the weighting function depends on endogenous states in a game (see Equation 104). This directly embodies the idea expressed in the title of the Mani, Mullainathan, Shafir, and Zhao (2013) that "Poverty Impedes Cognitive Function," no matter what decision problem agents are solving (*i.e.,* what are their payoffs, action space, or state space).

A model of costly information acquisition has the flexibility to explain this sort of finding, mathematically speaking. But this has an important caveat. The ability of this model to generate more "mistakes" in a poverty state relies on the premise of imperfect *observation* of income, or a heightened inability to determine income when it is low overall.[3] But this would be hard to square with the findings of Mani, Mullainathan, Shafir, and Zhao (2013) in tasks for which income is (almost) decision-

[3] Concretely, one could apply a variant of the Hébert and Woodford (2020) neighborhood-based cost or the Pomatto, Strack, and Tamuz (2023) log-likelihood-ratio cost in which states corresponding to poverty are harder to distinguish from others.

irrelevant. More broadly, the notion that imprecise choice *must* arise through imperfect learning places significant restrictions on how decision frictions vary across contexts.

Summary State-separable costs are consistent with Facts 1 to 5. Random utility models can explain only Facts 1 and 2. All models of costly information acquisition can explain Facts 1, 2, and 3; some models can be consistent with Fact 4 (but not mutual information); but none could easily explain Fact 5. On the basis of this, we argue that state-separable costs provide a flexible way of modeling a variety of decision frictions in a way that is consistent with our best experimental evidence.

However, there are potentially testable implications of information acquisition models which with state-separable costs would not be consistent. In particular, information acquisition models make predictions for the joint properties of beliefs and actions. This notwithstanding, it has been customary in the decision-theoretic literature to ignore these predictions, and instead to focus entirely on predictions for choice, under the premise that internal mental states are unobservable (*e.g.,* Caplin and Dean, 2015; Caplin, Dean, and Leahy, 2022). Moreover, existing tests of information acquisition models derived from the analysis of Caplin and Dean (2015) and Caplin and Martin (2015) and performed by Dean and Neligh (2022) are one-sided: they reveal that information acquisition is consistent with the data, but not that non-informational models are inconsistent with the data.

3.3 Main Results

We now prove existence, uniqueness, efficiency and equilibrium monotone comparative statics for both the aggregate and the cross-sectional action distribution. Our approach will be to establish that the correct notion of a "best response function" for the aggregate action X is a contraction map that satisfies certain properties.

3.3.1 Assumptions: Payoffs and Aggregator

We first identify conditions on payoffs, aggregators and stochastic choice functionals sufficient to guarantee uniqueness. For payoffs, we first require complementarities in the underlying game in the form of supermodularity in *cost-normalized payoffs* between an agent's own action and the aggregate. Second, we require that these complementarities are not too strong in the sense that payoffs are sufficiently concave to outweigh them:

Assumption 1 (Supermodularity and Sufficient Concavity)**.** *The payoff function u and weighting function λ are such that the following holds for all $x' \geq x, X' \geq X$, and* *θ:*

$$\frac{u(x', X', \theta) - u(x, X', \theta)}{\lambda(X', \theta)} \geq \frac{u(x', X, \theta) - u(x, X, \theta)}{\lambda(X, \theta)} \tag{107}$$

Moreover, for all $\alpha \in \mathbb{R}_+$, $x' \geq x$, X, and θ, the following holds:[4]

$$\frac{u(x' - \alpha, X, \theta) - u(x - \alpha, X, \theta)}{\lambda(X, \theta)} \geq \frac{u(x', X + \alpha, \theta) - u(x, X + \alpha, \theta)}{\lambda(X + \alpha, \theta)} \tag{108}$$

Informally, the former part of the assumption ensures that when aggregate actions go up, agents have an incentive to increase their own action. The latter part of the assumption ensures that agents' actions are less than one-for-one sensitive to the aggregate.

To gain a stronger intuition for the role of this assumption, and to provide easily verifiable conditions under which it holds, we characterize it with twice continuously differentiable payoffs u and weighting functions λ:

Lemma 2. *When $u(\cdot, \theta)$ is twice continuously differentiable in (x, X) and $\lambda(\cdot, \theta)$ is twice continuously differentiable in X for all θ, Assumption 1 is equivalent to the following:*

$$0 \leq u_{xX}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_X(X, \theta)}{\lambda(X, \theta)} \leq -u_{xx}(x, X, \theta) \tag{109}$$

for all x, X and θ.

Proof. See Appendix C.1.1. □

When cognitive constraints are exogenous, this condition reduces to the requirement that $0 \leq u_{xX} \leq -u_{xx}$, which is a standard condition for unique equilibrium in supermodular games (see *e.g.,* Weinstein and Yildiz, 2007). Intuitively, this condition requires that the slope of agents' optimal actions to changes in aggregate actions are bounded between zero and one.

When cognitive costs are endogenous, strategic complementarity now has both a physical payoff complementarities component u_{xX} and a *cognitive complementarities* component $-u_x \frac{\lambda_X}{\lambda}$. To understand why cognitive complementarities take this form, suppose that aggregate actions increase and this raises cognitive costs by $\frac{\lambda_X}{\lambda}$ percent. This gives the agent an incentive to spread out their actions around any locally optimal action. If the agent is playing an action greater than any locally optimal action, their marginal utility from increasing their own action is negative ($u_x < 0$). However, as cognitive costs have gone up, the agent is now more willing to accept such a

[4] In stating this assumption, we are implicitly extending the domain of u so that it is well-defined under such translations.

negative marginal payoff, and so has incentives to further increase the likelihood that their action lies further from the locally optimal point. Thus, when u_x is negative, when cognitive costs increase, the agent has an incentive to play higher actions and there is strategic complementarity, *i.e.*, $-u_x \frac{\lambda_X}{\lambda} > 0$. When u_x is positive, the reverse logic is true, and increased cognitive costs make actions strategic substitutes. Thus, with cognitive strategic externalities, we require that (i) any strategic substitutability through cognitive costs never outweighs strategic complementarities in physical payoffs, and (ii) agents' payoff functions are sufficiently concave to outweigh both physical payoff complementarities and cognitive complementarities.

Having identified conditions on payoffs, we now turn to the aggregator. To retain the ordering between actions and aggregates, we assume that the aggregator is monotone in the sense of first-order stochastic dominance. We further assume that the aggregator satisfies discounting, which is to say that it is sub-linear in level shifts of the cross-sectional action distribution (see Cerreia-Vioglio, Corrao, and Lanzani, 2020, for a discussion of monotone and (sub-)linear aggregators):

Assumption 2 (Monotone and Discounted Aggregator)**.** *For all $g, g' \in \Delta(\mathcal{X})$:*

$$g' \succeq_{FOSD} g \implies X(g') \geq X(g) \tag{110}$$

Moreover, there exists $\beta \in (0,1)$ such that for any distribution $g \in \Delta(\mathcal{X})$ and any $\alpha \in \mathbb{R}_+$:

$$X(\{g(x-\alpha)\}_{x\in\mathcal{X}}) \leq X(\{g(x)\}_{x\in\mathcal{X}}) + \beta\alpha \tag{111}$$

We moreover, note that the assumption that $\beta < 1$ can be relaxed to allow $\beta = 1$ if the second inequality in Assumption 1 (Equation 108) is made strict. In the interests of concreteness, the following Lemma (the proof of which is immediate, and therefore omitted) provides several important and natural aggregator functions that satisfy Assumption 2.

Lemma 3. *The following aggregators satisfy Assumption 2:*

1. *Linear aggregators:*
$$X(g) = \beta \int_{\mathcal{X}} f(x) g(x) \, \mathrm{d}x \tag{112}$$
where $\beta \in [0,1)$ is a parameter controlling discounting and $f : X \to \mathbb{R}$ is a differentiable function such that $f' \in [0,1]$.

2. *Quantile aggregators:*
$$X(g) = \beta G^{-1}(l) \tag{113}$$

where $\beta \in [0, 1)$ is a parameter controlling discounting, $G(x) = \int_{\underline{x}}^{x} g(\tilde{x})\,\mathrm{d}\tilde{x}$ is the CDF of the cross-sectional action distribution, G^{-1} is its left-inverse, and $l \in (0, 1)$.

Linear aggregators with polynomial kernels $f(x) = a_0 + a_1 x + \ldots + a_l x^l$ (subject to the monotonicity and discounting constraints that $f'(x) \in [0, 1]$ on $[\underline{x}, \overline{x}]$) allow the aggregator to depend on *all* moments of the cross-sectional distribution of actions. The mean aggregator, $X(g) = \beta \int_{\mathcal{X}} x g(x)\,\mathrm{d}x$, is a special case of this class when $l = 1$. Thus, our analysis nests the common assumption in macroeconomics that interactions take place through the mean action (see Angeletos and Lian, 2016, for a review). Moreover, the polynomial sub-class allows for higher moments of the action distribution to enter agents' payoffs. This allows the dispersion $l = 2$, skewness $l = 3$, and kurtosis $l = 4$ of other agents' actions to matter for agents' strategic incentives. Such aggregators also have natural macroeconomic applications. For example, in Flynn and Sastry (2022b), the fact that dispersion reduces aggregate outcomes generates important general equilibrium forces. Quantile aggregators include the median when $l = \frac{1}{2}$. Such aggregators are relevant when agents care about what an average agent does, rather than what other agents do on average.

Assumption 2 rules out aggregators that do not preserve the monotonicity of actions, *e.g.,* linear aggregators with a negative slope, or those that are more than one-for-one sensitive to translations of actions, *e.g,* linear aggregators with a slope greater than one. Intuitively, such aggregators break either strategic complementarity or sufficient concavity.

3.3.2 Intermediate Result: Properties of Stochastic Choice

Assumption 2 suggests a path toward ensuring that equilibrium is described by a contraction map if, in response to level shifts in the aggregate, the optimal stochastic choice pattern increases in the sense of first-order stochastic dominance (monotonicity) but remains dominated by the level shift itself (discounting). These are intuitive properties given the supermodularity and concavity of payoffs, which encode that level shifts in the (conjectured) aggregate globally increase the attractiveness of playing higher x, but in a way that is less than one-for-one. We now show an interpretable sufficient condition within the state-separable class which guarantees that monotonicity and discounting translate appropriately to stochastic choice.

We first define a new property of a function that we label quasi-monotone-likelihood-ratio-property (quasi-MLRP). This condition allows us to relate the underlying cost functional to the distribution of actions induced by optimality.

Definition 4 (Quasi-MLRP)**.** *A function $f : \mathbb{R}_+ \to \mathbb{R}$ satisfies quasi-MLRP if for any two distributions $g', g \in \Delta(\mathcal{X})$:*

$$\Big(\, f(g'(x')) - f(g'(x)) \geq f(g(x')) - f(g(x)) \ \forall x' \geq x \, \Big) \implies g' \succeq_{FOSD} g \tag{114}$$

With this definition in hand, we can now state our final technical assumption on stochastic choice functionals, which ensures that we can always translate dominance in payoff units to dominance in terms of distributions:

Assumption 3 (Quasi-MLRP Kernel)**.** *Costs have a differentiable kernel ϕ such that ϕ' satisfies quasi-MLRP.*

It is important to note that the two workhouse kernels in the literature on control costs satisfy this assumption:

Lemma 4. *The entropy kernel $\phi(p) = p \log p$ and the quadratic kernel $\phi(p) = \frac{1}{2}p^2$ satisfy Assumption 3.*

Proof. See Appendix C.1.2. □

We can now state a Lemma using this assumption and our earlier assumptions on payoffs to establish monotonicity and discounting of the solution of the stochastic choice problem:

Proposition 10 (Monotone and Discounted Stochastic Choice)**.** *Consider the stochastic choice program with payoffs satisfying Assumption 1 and cost functional satisfying Assumption 3. Then,*

1. *The optimal stochastic choice rule p^* is weakly increasing in the sense that if $\hat{X}' \geq \hat{X}$ then $p^*(\theta; \hat{X}') \succeq_{FOSD} p^*(\theta; \hat{X})$ for all $\theta \in \Theta$.*

2. *The optimal choice profile is discounted in the sense that when $\hat{X}$ and $\hat{X}' = \hat{X} + \alpha$ for $\alpha \in \mathbb{R}_+$, we have that $p^*_{-\alpha}(\theta; \hat{X}) \succeq_{FOSD} p^*(\theta; \hat{X}')$ for all $\theta \in \Theta$, where $p^*_{-\alpha}$ denotes the translation of p^* to the right by α.*

Proof. See Appendix C.1.3. □

The key to both parts is that quasi-MLRP allows us to "invert" dominance relationships in payoffs to obtain dominance relationships between distributions. For the first part, we show that the dominance of payoffs for playing higher x from supermodularity implies dominance of distributions under quasi-MLRP. For the second part, we use the property off payoffs from (108) that concavity of utility exceeds strategic complementarity, to show the optimal stochastic choice rule is dominated by the claimed level shift in the rule.

Proposition 10 is the core of our environment's tractability. It is in principle the ingredient that might be replaced in an alternative model of stochastic choice, like a form of unrestricted information acquisition. But, to our knowledge, such monotonicity and discounting results do not exist for any form of information acquisition in general environments. Moreover, this is not merely a technical glitch. A very relevant mechanism, anchoring toward frequently played actions, fights such monotonicity and discounting in information acquisition models. In a numerical example with the mutual-information cost (Sims, 2003) in Appendix C.2, we show that violations of monotonicity and discounting obtain in the single-agent problem and how this leads to non-uniqueness and non-monotone comparative statics in the equilibrium of an example game.

3.3.3 Existence and Uniqueness

We can now state our main existence and uniqueness result:

Theorem 1 (Existence, Uniqueness, and Symmetry)**.** *Under Assumptions 1, 2 and 3, there exists a unique equilibrium. The unique equilibrium is symmetric.*

Proof. See Appendix C.1.4. □

As alluded to above, we show this result by defining an equilibrium operator that maps the law of motion of the aggregate in the state to the resulting optimal stochastic choice rule and then maps this back to a law of motion of the aggregate, and then determining that said operator is a contraction map. More formally, let $\mathcal{B} = \{\hat{X} | \hat{X} : \Theta \to \mathbb{R}\}$ be a space of (bounded) functions endowed with the sup norm. We define the operator $T : \mathcal{B} \to \mathcal{B}$:

$$T\hat{X} = X \circ p^*(\hat{X}) \tag{115}$$

To show uniqueness of the equilibrium law of motion of aggregates, it then suffices to prove that T is a contraction map. We prove this by showing that, under the given assumptions, T satisfies both of Blackwell's conditions of monotonicity and discounting. Given the unique equilibrium-consistent law of motion which satisfies $T\hat{X} = \hat{X}$, the equilibrium stochastic choice rule is then the unique solution of the stochastic choice problem given that law of motion, or $p^*(\hat{X})$. This extends classic uniqueness results to the realm of stochastic choice.[5]

3.3.4 Monotone Comparative Statics

Once we lie in the realm of unique equilibria, it is well-posed to consider comparative statics in equilibrium. We provide two such results, showing when the action distribution and aggregate action are monotone in the state and when the precision of agents actions is monotone in the state.

Monotonicity of Action Distributions

To show monotonicity of distributions and aggregates, we require a stronger supermodularity assumption that not only are individual actions and aggregate actions complements, but so too is the underlying state itself a complement to both individual actions and aggregates in cost-adjusted payoffs:

Assumption 4. *The payoff function u and weighting function λ are such that the following holds for all $\theta' \geq \theta, X' \geq X, x' \geq x$:*

$$\frac{u(x', X', \theta') - u(x, X', \theta')}{\lambda(X', \theta')} \geq \frac{u(x', X, \theta) - u(x, X, \theta)}{\lambda(X, \theta)} \tag{116}$$

As before, to gain a stronger intuition and provide an easily verifiable condition, we characterize this assumption when u and λ are twice continuous differentiable:

Lemma 5. *When u is twice continuously differentiable in (x, X, θ) and λ is twice continuously differentiable in (X, θ), Assumption 4 is equivalent to*

$$u_{xX}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_X(X, \theta)}{\lambda(X, \theta)} \geq 0 \quad \text{and} \quad u_{x\theta}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_\theta(X, \theta)}{\lambda(X, \theta)} \geq 0 \tag{117}$$

for all x, X and θ.

The proof follows from the same steps as in the proof of Lemma 2, simply relabelling X as θ, and is therefore omitted. The first inequality ("complementarity with X") is identical to that in Lemma 2, and the second is its mirror image for "complementarity with θ". When cognitive costs do not depend on exogenous states $\lambda_\theta = 0$, this second condition reduces to $u_{x\theta} \geq 0$. When cognitive costs depend on exogenous states, the intuition for the additional term echoes the discussion of complementarity

[5] One could dispense with Assumptions 1, 2, and 3 and prove existence in our setting only by applying the Schauder fixed-point theorem. We omit this result as it is simple, and because our analysis will proceed afterward under Assumptions 1, 2, and 3.

with X. The presence of this additional term underscores the fact that state-varying control costs affect agents' incentives to shift their entire distribution of stochastic choice upward in higher states.

Under this assumption, we show the following result:

Theorem 2 (Monotone Actions and Aggregates)**.** *Under Assumptions 1, 2, 3, and 4, the unique equilibrium action distribution is monotone increasing in the sense of FOSD and the law of motion of the aggregate is increasing in the underlying state.*

Proof. See Appendix C.1.5. □

The intuition for this result is simple: higher θ makes higher actions more desirable so that the distribution of actions in higher states dominates the distribution in lower states. This is complicated by the fact that agents may face higher cognitive costs in higher states. Hence, the relevant notion of complementarity is complementarity in cost-adjusted payoffs. The proof strategy makes use of the contraction mapping property used in the uniqueness proof. In particular, it shows that monotonicity is preserved by the fixed point operator and therefore that the fixed point must itself be monotone.

This result has the following immediate implication for the supports of action distributions:

Corollary 7 (Monotone Consideration Sets)**.** *Under the conditions of Theorem 2, in the unique equilibrim agents' consideration sets $\mathcal{X}(\theta) = cl_{\mathcal{X}}\{x \in \mathcal{X} : p^*(x|\theta, \hat{X}(\theta)) > 0\}$ are increasing in the strong set order.*[6]

As optimal distributions increase in the sense of first-order stochastic dominance, the supports must move in the sense of the strong set order. The result is vacuous if the cost kernel satisfies an Inada condition and $\mathcal{X}(\theta) = \mathcal{X}$ for all θ. The result has bite if agents, for example, have costs with the quadratic kernel, which does not satisfy an Inada condition and may result in agents' optimally playing only a subset of available actions. In this case, the result puts structure on the endogeneity of consideration sets—agents consider larger actions in higher states in equilibrium.

Monotonicity of Action Precision

We now turn to establishing when the precision of, or extent of mistakes in, agents' actions is monotone in the state in equilibrium. To this end, in our context with flexible stochastic choice, we first need a non-parametric notion of precision:

[6] Where $\mathrm{cl}_{\mathcal{X}} A$ denotes the closure of set A within $\mathcal{X}$.

Definition 5 (Precision)**.** *Fix an $h : \mathbb{R} \to \mathbb{R}$. A symmetric distribution g is more precise about a point x^* than g' about $x^{*'}$ under h if $h \circ g(|x - x^*|)$ is faster decreasing in $|x - x^*|$ than is $h \circ g'(|x' - x^{*'}|)$ in $|x' - x^{*'}|$.*[7]

Informally, this definition requires that a distribution is more precise than another if its density is more rapidly decreasing away from the point about which precision is being considered. This definition generalizes the property that Gaussian distributions are more precise about their mean when they have a lower standard deviation to cases with non-Gaussian densities by exactly capturing the idea that a distribution is more precise if its tails decay faster from the point about which a distribution is centered.[8]

Having defined precision, we now state sufficient assumptions on payoffs for precision to be monotone. To show this result, we specialize to a distance-based payoff environment, which we refer to throughout as *generalized beauty contest* payoffs:

Assumption 5 (Generalized Beauty Contests)**.** *The utility function is given by:*

$$u(x, X, \theta) = \alpha(X, \theta) - \beta(X, \theta)\Gamma(|x - \gamma(X, \theta)|) \tag{119}$$

where Γ is monotone increasing and such that $\Gamma(0) = 0$, $\gamma(X, \theta)$ is monotonically increasing in (X, θ), and $\beta(X, \theta)$ is positive, for every (X, θ).

Under distance-based payoffs with distance function Γ, an agent cares only about how far their action is from an optimal action conditional on others' play X and the state θ, $\gamma(X, \theta)$. The extent to which they care is governed by $\beta(X, \theta)$, with larger values inducing greater losses from failing to match the optimal action.

We note that this formulation nests the quadratic payoff functions, which can be justified via a second-order approximation of any smooth, concave utility function around its maximum value $\gamma(X, \theta)$:

[7]On an asymmetric support, we call a distribution g symmetric if $g(x) = g(-x)$ whenever both $g(x)$ and $g(-x)$ are defined. For any symmetric functions $\xi, \hat{\xi} : A \to \mathbb{R}$, we say that ξ is faster decreasing than $\hat{\xi}$ in their arguments if $\xi(0) - \xi(|x|) \geq \hat{\xi}(0) - \hat{\xi}(|x|)$ for all $x \in A$.

[8]To see this, recall that a Gaussian random variable with mean μ and standard deviation σ has pdf:

$$g(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x - \mu}{\sigma}\right)^2\right\} \tag{118}$$

Thus, for two Gaussian distributions with means μ, μ' and standard deviations σ, σ' such that $\sigma < \sigma'$, we have that $h \circ g(|x - \mu|)$ is faster decreasing than $h \circ g'(|x - \mu'|)$ whenever h is monotone. Thus, under monotone h, we have that Gaussian distributions with lower standard deviations are more precise about their mean under our definition of precision.

Lemma 6. *Consider a payoff function $u : \mathcal{X} \times \mathbb{R} \times \Theta$ that is twice differentiable, strictly concave in its first argument, and maximized for every $(X, \theta) \in \mathbb{R} \times \Theta$ at some $x^*(X, \theta) \in \mathcal{X}$. Then, up to a term that is on the order of $|x - x^*(X, \theta)|^3$, payoffs conditional on each (X, θ) take the form of Equation 119 with $\alpha(X, \theta) = u(x^*(X, \theta), X, \theta)$, $\beta(X, \theta) = \frac{1}{2}|u_{xx}(x^*(X, \theta), X, \theta)|$, $\gamma(X, \theta) = x^*(X, \theta)$, and $\Gamma(x) = x^2$.*

This result follows immediately from taking a Taylor expansion of u around its optimal value in each state, observing that the first-order term is zero because of the first-order condition for optimality, and using Taylor's Theorem to describe the residual error. In this interpretation, $\gamma(X, \theta)$ is the optimal action conditional on (X, θ) and $\beta(X, \theta)$ measures the curvature of payoffs, or second-order loss of mis-optimization, around that point.

We now state the result, which encapsulates the idea that precision is higher when the losses from mis-optimization are higher for endogenous or exogenous reasons:

Theorem 3 (Monotone Precision)**.** *Under Assumptions 1, 2, 3, 4, and 5, $p^*(\theta) \in \Delta(\mathcal{X})$ is more precise about $\gamma(\hat{X}(\theta), \theta)$ than $p^*(\theta')$ about $\gamma(\hat{X}(\theta'), \theta')$ under ϕ'*

1. *For any $\theta \leq \theta'$ if $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is monotone decreasing in both arguments.*
2. *For any $\theta \geq \theta'$ if $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is monotone increasing in both arguments.*

Proof. See Appendix C.1.6. □

The proof of this result shows that the agents' incentives to transfer probability mass from the ideal point $\gamma(\hat{X}(\theta), \theta)$ to any other $x \in \mathcal{X}$ are strictly lower when $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is larger, which translates directly to our notion of precision. Note that this combines the incentives for precision from the curvature in the utility function, β, and from the scaling of the cost function, λ. This calculation relies on the symmetry of distance-based payoffs around $\gamma(\hat{X}(\theta), \theta)$. It then leverages the fact that $\hat{X}$ is monotone in θ in equilibrium, because of Theorem 2, which in turn implies monotonicity of the mapping $\theta \mapsto \frac{\beta(\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)}$, decreasing in case 1 and increasing in case 2. Put differently, the "endogenous" and "exogenous" stakes of making good choices both move in the same direction in equilibrium. Thus, precision is monotone in the state.[9]

This result has the following immediate implication for the size of agents' equilibrium consideration sets:

[9]Unsurprisingly, we cannot state a general result when $\frac{\beta(X,\theta)}{\lambda(X,\theta)}$ is not strictly monotone in its two arguments; but we could of course still use part of the previous argument to compare precision in any two states $(\theta, \hat{X}(\theta))$,$(\theta', \hat{X}(\theta'))$ after solving for equilibrium.

Corollary 8 (Monotone Size of Consideration Sets)**.** *Under the conditions of Theorem 3, if 1. (resp. 2.) holds, then the Lebesgue measure of $\mathcal{X}(\theta)$ is increasing (resp. decreasing) in θ.*

Thus, as is intuitive, in states where agents' cost-adjusted states are higher, agents choose from smaller consideration sets.

3.3.5 Efficiency

A further question of interest is when equilibria of our model are efficient. As our agents are symmetric, *ex-ante* Pareto efficiency and utilitarian efficiency are equivalent. We therefore say that a stochastic choice rule is efficient if it maximizes utilitarian welfare:

Definition 6. *A stochastic choice rule $P^E \in \mathcal{P}$ is efficient if it solves the following program:*

$$P^E \in \arg\max_{P\in\mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u\left(x, X(P(\theta)), \theta\right) \,\mathrm{d}P(x|\theta)\, \pi(\theta) - c(P, X(P)) \tag{120}$$

An efficient stochastic choice rule both fully internalizes the effect choices have on aggregates and the costs of stochastic choice. Moreover, this notion of efficiency takes seriously that agents do incur the cognitive cost of constraining their mistakes. We now ask, when will equilibrium be efficient? The following result relates the answer to this question to the *balancing of aggregate externalities* in physical and payoffs. To derive a variational necessary condition, we make technical assumptions sufficient to guarantee differentiability:

Assumption 6 (Regularity Conditions for Efficient Program)**.** *Suppose that the planner's objective in Equation 120 is strictly concave in P, u is differentiable in its second argument X, λ is differentiable in its first argument X, and the aggregator is linear:*

$$X(g) = \int_{\mathcal{X}} f(x) \,\mathrm{d}G(x) \tag{121}$$

for some nowhere-constant function f.

Theorem 4. *Under Assumption 6, a necessary condition for efficiency of an equilibrium stochastic choice rule p^* is that:*

$$\int_{\mathcal{X}} u_X(\tilde{x}, X(p^*(\theta)), \theta) p^*(\tilde{x}|\theta) \,\mathrm{d}\tilde{x} = \lambda_X(X(p^*(\theta)), \theta) \int_{\mathcal{X}} \phi(p^*(\tilde{x}|\theta)) \,\mathrm{d}\tilde{x} \tag{122}$$

for all $\theta \in \Theta$.

Proof. See Appendix C.1.7. □

To understand this result, we first consider the case in which there are no payoff externalities in costs or $\lambda_X = 0$. In this case, the condition requires that the *average externality of increasing the aggregate is zero*. This condition is *evaluated* at the equilibrium stochastic choice pattern, but does not depend directly on the structure of cognitive costs. Thus, to evaluate such a condition (under the assumption that $\lambda_X = 0$), an observer needs only to know about payoff externalities and the observed distribution of choices.

We next consider the case in which $\lambda_X \neq 0$. In this case, efficiency obtains when the aforementioned payoff externality balances with a *cognitive externality*, to use the language of Angeletos and Sastry (2023), operating directly through costs. Consider our recurring example of cognitive costs that decrease with the value of X because of poverty-induced stress ($\lambda_X < 0$) and assume that the utility costs of cognition are positive in all states.[10] The *cognitive externality* is that increasing X directly decreases every agent's cognitive cost. A non-paternalistic planner, who takes cognitive costs into account, considers also this externality. Thus, an optimal allocation (if it exists) tolerates a *negative* marginal payoff externality to achieve a *positive* marginal cognitive externality. We return to this specific point in a concrete example in Section 3.4.2 and Corollary 10.

Relative to the literature, our analysis therefore identifies a new channel through which rational decision frictions can create equilibrium externalities and induce inefficiency. This supplements the findings of Hébert and La'O (2022) for aggregative games with information acquisition and Angeletos and Sastry (2023) for Arrow-Debreu economies with information acquisition. Relative to the related results in those papers, our Theorem 4 has three substantial differences. First, it clarifies how cognitive externalities can operate outside of information acquisition models. Second, it sheds light on the *nature* of inefficiency—in particular, the direction in which a social planner would want to perturb aggregates—in inefficient equilibrium. By contrast, due to the intractable structure of general cognitive externalities in information-acquisition models, Hébert and La'O (2022) and Angeletos and Sastry (2023) can say

[10] Note, of course, that in our model these costs need not be positive. The following more perverse model would also be consistent with empirical evidence that scarcity reduces *decision quality* (*e.g.*, Mani, Mullainathan, Shafir, and Zhao, 2013): low X increases the *scale* of cognitive costs, reducing *relative* incentives for precise optimization, but has a positive *level* effect on welfare. In this case, our intuition for why there is a role of cognitive externalities would be the same, and Theorem 4 would still hold; but the intuition for the sign of effects would flip.

relatively little about the same in their settings. Third, leveraging our state-separable structure, it provides a testable condition to compare the extent of these externalities with "standard" payoff externalities to gauge efficiency.

3.4 Applications

We now apply our model to make equilibrium predictions in two macroeconomic settings. We first study price-setting by monopolistic firms, a cornerstone of the "supply block" of modern macroeconomic models. In our model, firms imperfectly price their goods because of *ex ante* planning frictions. We show how to make equilibrium predictions for the aggregate price level and price dispersion that take into account the aggregate consequences of "pricing mistakes" and firms' differential incentives to rein in these mistakes in different aggregate states. We next study consumption and savings decisions in a liquidity trap, a cornerstone of the "demand block" of modern macroeconomic models. In our model, consumption plans are imperfect because of costly control. Moreover, these costs increase when households have low income, capturing the possibility that psychological stress impairs decisionmaking in these states. We show how to make predictions for aggregate income and consumption inequality and characterize a novel equilibrium externality that arises because one agent's lack of consumption increases others' costly stress.

3.4.1 Price-Setting with Planning Frictions

Set-up Each agent $i \in [0,1]$ is a firm that produces a differentiated variety in quantity q_i at price $p_i \in [\underline{p}, \overline{p}]$ with $\underline{p} > 0$. These firms use intermediate goods z_i, with marginal cost k, to produce according to the production technology $q_i = z_i$. The outputs of these firms are consumed by a representative household, with constant elasticity of substitution (CES) consumption bundle:

$$C = \left(\int_0^1 q_i^{\frac{\eta-1}{\eta}} \, \mathrm{d}i \right)^{\frac{\eta}{\eta-1}} \tag{123}$$

where $\eta > 1$. As is standard (see *e.g.,* Hellwig and Venkateswaran, 2009), the household's preferences over consumption and real money balances $\frac{M}{P}$ are given by:

$$V\left(C, \frac{M}{P}\right) = \frac{C^{1-\sigma}}{1-\sigma} + \ln \frac{M}{P} \tag{124}$$

where $\sigma \geq 0$. The money supply is an exogenous shock in the discrete set $\mathcal{M}$ with minimal and maximal elements $\underline{M}$ and $\overline{M}$, such that $\underline{M} > \underline{p}$ and $\overline{M} < \overline{p}$. Moreover, we make the standard simplifying assumption (Alves, Kaplan, Moll, and Violante, 2020; Flynn and Sastry, 2022b) that real marginal costs are a log-linear function of aggregate output:

$$\frac{k}{P} = C^{\chi} \tag{125}$$

where $\chi > 0$ represents "factor price pressure", *i.e.,* the extent to which real marginal costs are increasing in the level of output in the economy.

To study how planning frictions matter, we subject the firm to a state-separable cost function with any kernel satisfying Assumption 3 (*e.g.,* the entropy kernel $\phi(p) = p \log p$) and a weighting function inversely proportional to how likely the firm thinks each realization of the money supply $\pi(M)$ is, *i.e.,* $\lambda(M) = \frac{1}{\pi(M)}$. This captures a situation in which firms must plan for contingencies (realizations of the money supply) in advance and then implement these plans when the state is realized. Moreover, we assume that $\pi(M) \propto M^{\delta}$, where $\delta > 0$ corresponds to high money supply states being more likely and $\delta < 0$ means that low money supply states are more likely.

Recasting the Economy as a Game Given the CES aggregator, the firm faces an isoelastic demand curve:

$$q_i = \left(\frac{p_i}{P}\right)^{-\eta} C \tag{126}$$

where P is the ideal price index under CES production:

$$P = \left(\int_0^1 p_i^{1-\eta} \, \mathrm{d}i\right)^{\frac{1}{1-\eta}} \tag{127}$$

The firms' profits are moreover priced according to the real stochastic discount factor (the household's marginal utility from consumption) $C^{-\sigma}$. Thus, the firm's objective function is:

$$\pi(p_i, P, C, \theta) = C^{-\sigma} \frac{p_i - k}{P} \left(\frac{p_i}{P}\right)^{-\eta} C = C^{1-\sigma} P^{\eta-1} (p_i - k) p_i^{-\eta} \tag{128}$$

Substituting in the equilibrium conditions that $k = PC^{\chi}$ (factor supply) and $C = \left(\frac{M}{P}\right)^{\frac{1}{\sigma}}$ (money demand), we obtain that the firm's payoff function is:

$$u(p_i, P, M) = M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left(p_i - M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right) p_i^{-\eta} \tag{129}$$

To apply all of our results, we perform the standard approximation (as per Lemma 6) of the firm's objective function to second-order around the optimal price in each state. This yields the payoff function:

$$u(p_i, P, M) = \alpha(P, M) - \beta(P, M)(p_i - \gamma(P, M))^2 \tag{130}$$

where:

$$\begin{aligned}
\alpha(P, M) &= \frac{1}{\eta - 1}\left(\frac{\eta}{\eta - 1}\right)^{-\eta} M^{\frac{1-\sigma+\chi(1-\eta)}{\sigma}} P^{\eta - \frac{1}{\sigma} + (1-\eta)(1-\frac{\chi}{\sigma})} \\
\beta(P, M) &= \frac{\eta}{2}\left(\frac{\eta}{\eta - 1}\right)^{-(\eta+2)} M^{\frac{1-\sigma-\chi(\eta+1)}{\sigma}} P^{(\eta+1)\frac{\chi}{\sigma} - \frac{1}{\sigma} - 1} \\
\gamma(P, M) &= \frac{\eta}{\eta - 1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}
\end{aligned} \tag{131}$$

and we impose that this game has complementarity in optimal actions by assuming factor price pressure is weaker than income effects in money demand, or $\chi < \sigma$. Finally, as is also standard, we approximate the aggregator to first order as:

$$P = \int_0^1 p_i \,\mathrm{d}i \tag{132}$$

which simply says that the aggregate price-level is the average price set by firms.

Results and Interpretation To build intuition, we first characterize equilibrium in this model in the absence of *ex ante* planning frictions. In this case, the optimal price that a firm sets is given by:

$$p = \frac{\eta}{\eta - 1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \tag{133}$$

which is a constant markup over marginal cost. Thus, we observe that there is a unique equilibrium in which all firms set the same price (there is no price dispersion) and the aggregate price level is given by:

$$P = \frac{\eta}{\eta - 1} M \tag{134}$$

In this equilibrium, the elasticity of prices to the money supply is one, *i.e.,* a 1% increase in the money supply leads to a 1% increase in the price level.

We now apply our general results to this economy when firms face *ex ante* planning frictions. Specifically, we ask when this price-setting economy has a unique equilibrium, when the aggregate price level is increasing in money, when the distribution of prices is increasing (in the sense of FOSD) in money, and when the dispersion in prices is highest when inflation is high.

Corollary 9. *There is a unique equilibrium if, for all $p, P \in [\underline{p}, \overline{p}]$ and $M \in \mathcal{M}$:*

$$-\left(\frac{P}{\gamma(P,M)\left(1-\frac{\chi}{\sigma}\right)}-1\right) < \frac{-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}}{\left(1-\frac{\chi}{\sigma}\right)}\left(\frac{p}{\gamma(P,M)}-1\right) < 1 \tag{135}$$

The unique aggregate price level and distribution of prices are both increasing in the money supply if, in addition:

$$\frac{1-\sigma-\chi(\eta+1)+\delta\sigma}{\chi}\left(\frac{p}{\gamma(P,M)}-1\right) < 1 \tag{136}$$

Moreover, price precision is decreasing in the money supply and the price level if, in addition:

$$\chi(\eta+1) \in (1+\sigma(\delta-1), 1+\sigma) \tag{137}$$

Proof. See Appendix C.1.9. □

To understand this result, we go through each condition in turn. The uniqueness condition (Equation 135) comprises two inequalities. The inequality on the right ensures that the game is one of complementarities. The inequality on the left ensures that utility has sufficient concavity relative to complementarity. In the special case where the losses from mispricing do not depend on the aggregate price level ($(1+\eta)\frac{\chi}{\sigma} - \frac{1}{\sigma} - 1 = 0$), the middle term is equal to zero and the complementarity condition always holds. This is because we have assumed that factor price pressure is weaker than income effects ($\chi < \sigma$), which makes the optimal price increase in the aggregate price. When the losses from mispricing are instead endogenous ($(1+\eta)\frac{\chi}{\sigma} - \frac{1}{\sigma} - 1 \neq 0$), there is a new effect that must be accounted for. Intuitively, suppose without loss of generality that an agent is setting a price that is less than the optimal price and an increase in the aggregate price-level increases (decreases) the losses from misspricing, then the agent now has a greater (lesser) incentive to reduce the magnitude of this mistake and increase (decrease) their price. In the former case, this endogeneity of the costs of misspricing induces greater strategic complementarity. In the latter case, it induces strategic substitutability. The exact inequality we derive disciplines the magnitudes of these effects in a verifiable way that ensures that strategic complementarity always obtains.

The sufficient concavity condition similarly has a "simple" and "complex" interpretation. When the losses from mispricing are exogenous, the condition requires that the optimal price has a slope less than one in the aggregate price level. This condition that "best responses have a slope less than one" is familiar from games without decision frictions. More generally, when the aggregate price matters for the losses from mispricing, the "slope" that needs to be bounded depends directly on the deviation of the price from the optimal price and the considerations described above.

The monotonicity condition (Equation 136) requires that higher levels of the money supply are complementary with higher prices for firms. When the losses from misspricing do not depend on the money supply $\frac{1-\sigma-\chi(\eta+1)}{\sigma} + \delta = 0$, this condition always holds as factor price pressure from higher money supply (which increases demand, which increases production, which increases marginal costs) makes optimal prices higher. More generally, as above for the endogenous price-level, the monotonicity condition ensures complementarity between pricing and the exogenous money supply when the losses from mispricing depend on the money supply.

Finally, the condition for monotone precision (Equation 137) conveys, in terms of deep parameters, when price-setters optimally respond to lower money and prices by making more precise decisions. To understand this, it is useful to first turn off factor price pressure or set $\chi = 0$. In this case, the condition corresponds to $\delta + \left(\frac{1}{\sigma} - 1\right) < 0$. The first term isolates the role of costly planning—when high-money states are less likely ($\delta < 0$), firms optimally put in less effort to plan for them, and their pricing decisions in these states are less precise. The second term conveys the roles of aggregate demand externalities, which have elasticity $1/\sigma$ with respect to the money supply, and the stochastic discount factor, which has elasticity -1 with respect to the money supply. The demand externality pushes toward high precision in high-demand states, because any price mistake leads to more lost sales. The stochastic discount factor pushes toward high precision in low-demand states, since profits are more valuable in these states (Flynn and Sastry, 2022b). Finally, when factor price pressures are re-introduced, they loosen the constraint corresponding to incentives from the money supply and tighten the constraint corresponding to incentives from aggregate prices.

Economic Lessons Our finding can be used to rationalize empirical findings on the cyclicality of price dispersion. Empirically, Alvarez, Beraja, Gonzalez-Rozada, and Neumeyer (2019) find that price dispersion among firms in Argentina has an elasticity of about 1/3 to the inflation rate in high-inflation periods (*e.g.,* an annual rate above 50%) and an elasticity that is positive, but close to zero, in low-inflation periods. Nakamura, Steinsson, Sun, and Villar (2018) find limited evidence that the dispersion of US prices increased in the "Great Inflation" of the 1970s and 1980s, during which annual US inflation was regularly between 5% and 10%. This evidence is consistent with the costly planning mechanism, if firms believe that hyperinflation states (in Argentina and in the US) are relatively unlikely ($\delta < 0$) and are far in the tail of the distribution for M. Intuitively, this allows the possibility that price dispersion is especially high in hyperinflations precisely because firms have not precisely formulated plans for these unlikely states.

3.4.2 Consumption and Savings with a Stress Externality

Set-up Each agent is a consumer that lives for infinite periods, indexed by $t \in \mathbb{N}$. They choose consumption levels $c_{it} \in \mathbb{R}$ and labor levels $n_{it} \in \mathbb{R}$ and have quadratic payoffs. They maximize expected discounted utility:

$$U(\{c_{it}, n_{it}\}_{t=0}^{\infty}, \theta_d) = (1+\theta_d)c_{it} - \frac{c_{it}^2}{2} - \chi\frac{n_{it}^2}{2} + \sum_{t=1}^{\infty} \delta^t \left(c_{it} - \frac{c_{it}^2}{2} - \chi\frac{n_{it}^2}{2} \right) \tag{138}$$

where $\delta \in (0,1)$ is a discount factor, χ is a parameter controlling the labor-leisure trade-off, and θ_d is a *demand shock* in a discrete set Θ_d, with maximum element $\overline{\theta}_d \le 0$ and minimal element $\underline{\theta}_d > -1$, that reduces the household's relative preference to consume in period 0. Each agent can save in a risk-free bond with interest rate $R = 1/\delta$, fixed and unresponsive to demand as in a small open economy. Each agent receives income $w_t n_{it}$ in each period, where $w_t \in \mathbb{R}_+$ is a wage and $n_{it} \in \mathbb{R}$ is the amount that agent i works. Therefore, for each period t, the agent faces a budget constraint $c_{it} + b_{it} \le w_t n_{it} + R b_{i,t-1}$, where b_{it} is savings and $b_{i,-1} = 0$ for all agents.

Goods are produced by a representative firm, at which all of the agents work. The firm produces output via a linear production technology, $y_t = \int_{[0,1]} n_{it}\,\mathrm{d}i$. In all periods, the output market clears as $y_t = \int_{[0,1]} c_{it}\,\mathrm{d}i = y_t$ and the bond market clears as $0 = \int_{[0,1]} b_{it}\,\mathrm{d}i$. At $t = 0$, given the fixed interest rate, these conditions would be incompatible with equilibrium in the labor market. We therefore make the conventional assumption that, in this period, the firm commits to satisfying demand at the (fixed) price, households lie off their labor supply curve, and households all work an equal amount. We refer to this period as a *liquidity trap*, since the market failure is caused by the inability of interest rates to adjust downward to accommodate the negative demand shock.

We are interested in how equilibrium at $t = 0$ is affected by demand shocks, under the assumptions that households imperfectly optimize and that their cost is affected by financial stress. To simplify our analysis, we assume that all choices for $t \ge 1$, after the economy exits the liquidity trap, are made frictionlessly. At $t = 0$, households choose $c_{i0} \in [\underline{c}, \overline{c}]$, where $\overline{c} < \frac{\delta}{1-\delta}$, to maximize expected utility net of cognitive costs, given rational expectations about future aggregates and their future behavior.[11]

We introduce the idea that stress may lead to lower-quality decisions in low-income states via the cost functional. This idea is motivated by the experimental findings of Mani, Mullainathan, Shafir, and Zhao (2013) suggesting that poverty, transitory or persistent, reduces performance in cognitive tasks. Mullainathan and Shafir (2013) hypothesize that involuntary capture of attention toward contemplating negative outcomes in these states reduces the available *bandwidth* to make decisions, and therefore makes people more prone to "forgetfulness" and "cognitive slips" (p. 14). We model this by letting $\lambda(y, \theta_d) = y^{-\tau}$, where y is consumers' period-0 income and $\tau \geq 0$ is a parameter controlling how quickly decision costs increase when income is low. This cost may capture in reduced-form the distraction of cognitive resources away from the decision of interest for consumption and savings, and hence the *scarcity* of attention available for the decision problem of interest.[12] We let the cost-functional kernel be any ϕ that satisfies Assumption 3.

Recasting the Economy as a Game We now analyze the consumer's problem to reduce the equilibrium determination of first-period consumption to a game to which our results can be applied. It is simple to show that, for $t \geq 1$, aggregate output is fixed at a level $\bar{y} > 0$ and each agents' consumption is fixed at a specific level which depends on their period 0 savings. This exact consumption result smoothing follows from the intertemporal Euler equation and the simplifying assumption that $\delta R = 1$ (see, *e.g.,* Hall, 1978). Next, because the payoff for $t \geq 1$ is always increasing in period 0 savings, the agent saves all unspent income at $t = 0$: $b_{i0} = y_0 - c_{i0}$. Using these observations, and defining $c_i = c_{i0}$ and $y = \int_{[0,1]} c_i \, \mathrm{d}i$, we can re-write the objective as

$$u(c, y, \theta_d) = \alpha(y, \theta_d) - \beta(y, \theta_d)(c - \gamma(y, \theta_d))^2 \tag{139}$$

where[13]

$$\begin{aligned} \gamma(y, \theta_d) &= (1-m)(\theta_d + \bar{y}) + my \\ \beta(y, \theta_d) &= \frac{1}{2(1-m)} \end{aligned} \tag{140}$$

[11] The condition $\overline{c} < \frac{\delta}{1-\delta}$ ensures that consumption in periods $t \geq 1$ does not exceed the bliss point.

[12] As a different, and complementary formalization that is consistent with their novel survey evidence, Sergeyev, Lian, and Gorodnichenko (2022) formalize the Mullainathan and Shafir (2013) hypothesis as an involuntary use of time that could otherwise be allocated to labor or leisure.

[13] A more cumbersome expression for $\alpha(y, \theta_d)$ is given in Appendix C.1.10.

and $m = \frac{\chi(1-\delta)}{\chi+\delta} \in (0,1)$ is the agent's marginal propensity to consume (MPC), which itself depends positively on labor disutility χ and negatively on the discount factor δ. In the limit where labor supply is inelastic, or $\chi \to \infty$, then $m \to 1-\delta$ as is familiar from the permanent income hypothesis. Finally, note that the payoff representation is exact, not approximate, since the original payoffs were quadratic.

Results and Interpretation Applying our general results, we can provide conditions under which a generalized beauty contest has a unique equilibrium with a number of economically relevant properties:

Corollary 10. *If the following condition holds for all $c, y \in [\underline{c}, \overline{c}]$ and $\theta_d \in \Theta_d$:*

$$0 < m - \frac{\tau}{y}\left(c - (1-m)\left(\bar{y} + \theta_d\right) - my\right) < 1 \tag{141}$$

then there exists a unique equilibrium in which (i) the distribution of consumption and aggregate output are monotone increasing in the demand shock θ_d and (ii) the precision of consumption is monotone decreasing in the demand shock θ_d and in aggregate output y. Moreover, if the planner's problem is strictly concave, a necessary condition for the efficiency of the unique equilibrium is that:

$$y = \bar{y} + \frac{\tau}{\chi} y^{-\tau-1} \int_{\underline{c}}^{\overline{c}} \phi(p^*(c \mid \theta_d))\, \mathrm{d}c \tag{142}$$

Thus, whenever cognitive costs are positive, an efficient allocation in an economy with $\tau > 0$ has higher output than an efficient allocation in an economy with $\tau = 0$.

Proof. See Appendix C.1.10. □

The conditions in Equation 141 follow from the calculation in Lemmas 2 and 5. These conditions are trivially satisfied if $\tau = 0$ as higher demand increases income which increases consumption (as $m > 0$), but less than one-for-one since the household discounts the future and therefore has an MPC strictly less than one (as $m < 1$). If $\tau > 0$, then there are potentially countervailing forces that affect strategic complementarity. Concretely, when aggregate output increases, stress decreases, and agents' costs of precise optimization fall. If an agent is consuming more than the optimal level, this makes them prone to consume closer to the optimal level and lower their consumption, inducing strategic substitutability. Conversely, if an agent is consuming less than the optimal level, this makes them prone to consume more and induces additional strategic complementarity. The condition provides the precise conditions under which these concerns do not upset total strategic complementarity (consumption increases when income increases) and sufficient concavity (consumption increases less than one-for-one). Under these conditions, we know that higher demand increases aggregate output (point (i)); that higher demand shifts the entire distribution of consumption upward of first-order stochastic dominance (point (ii)); and that agents' actions are more precise in high states, due to their experiencing lower stress and, therefore, (endogenously) lower costs of attention in these states.

The second result, the necessary condition for efficiency, conveys that the introduction of the stress mechanism increases the optimal level of output. The reason is that the stress mechanism creates an externality operating through cognitive costs: if one agent consumes more, increasing aggregate demand and output, they reduce stress (cognitive costs) for all other agents. The extent of this externality in state θ_d is proportional to the cognitive cost paid *ex post* in that state. Thus, the externality would disappear were there no cost of cognition. And the extent of cognitive costs would not affect the optimal allocation were there no stress and, by implication, no externality operating purely through cognition.

Economic Lessons Our prediction for endogenous precision, or higher consumption "mistakes" in low output states, is consistent with the evidence from Berger, Bocola, and Dovis (2023) that the cross-sectional distribution of US consumption becomes more dispersed in recessions. In our case, this result arises because of the equilibrium effect of low income causing stress that worsens decisionmaking. Our psychological explanation complements mechanisms studied in the literature related to cyclicality of income risk and the role of financial constraints.[14]

In the emerging literature on how household stress affects decisionmaking, our results complement those in the study of Sergeyev, Lian, and Gorodnichenko (2022), who use original survey evidence to measure the extent of financial stress among US households and to calibrate a macroeconomic model in which financial stress distracts from productive labor supply. Our mechanism is different (stress reducing decision *quality*) and makes a different prediction, potentially in line with the data, about consumption dispersion.

Our normative results clarify how the stress channel may translate into inefficiency at the macro level. In particular, our results rationalize a "paradox of scarcity" logic: by not spending, households contribute toward lower overall output, which induces further financial stress for others and has psychological costs. This mechanism relies

[14]Moreover, Sergeyev, Lian, and Gorodnichenko (2022) find in their original survey that liquidity constraints exacerbate reported psychological stress related to making economic decisions. Therefore, in practice, the psychological and liquidity-constraint channels may reinforce one another in a richer model that accommodates both.

crucially on the *endogeneity* of income, and hence stress, to others' decisions.

Finally, we note that our analysis contrasts from abstract results in examples studied by Hébert and La'O (2022) and Angeletos and Sastry (2023) in two ways. First, we isolate a cognitive externality that may be difficult to formalize in a model of costly information acquisition (see the discussion of Fact 5 in Section 3.2.3). Second, we can precisely characterize equilibrium, its comparative statics properties, the equilibrium externality, and the optimal direction of policy response in an *inefficient* setting.

3.5 Extensions

3.5.1 State-Separable vs. Mutual Information Costs

Although its foundations are in information theory, the mutual information model of Sims (2003) also makes predictions for stochastic choice or "imperfect optimization." Decision-theoretic work by Caplin, Dean, and Leahy (2022) characterizes these behavioral predictions, and Woodford (2012) and Dean and Neligh (2022) discuss how they match some, but not all, features of imperfect perception and choice in the lab. Moreover, in many applications in macroeconomics in finance, information choice is unobserved and/or not the focus of predictions *per se*. Instead, the focus is on the aforementioned predictions for imperfect optimization and how they play out in equilibrium.

In Appendix C.2, we contrast the predictions of state-separable and mutual information costs as alternative models of stochastic choice in large games. First, extending a result in Matějka and McKay (2015), we give abstract conditions under which the predictions of a version of the strategic mistakes model with logit costs gives identical predictions to a twin model with mutual information costs and a restriction of agents' (subjective) priors. Relaxing this condition isolates the key difference between the models—the mutual information model naturally allows agents to anchor toward commonly played actions as if they were "default points."

Next, we numerically explore a linear beauty contest game (Morris and Shin, 2002; Angeletos and Pavan, 2007) under both state-separable and mutual information costs. The model with state-separable costs predicts a unique equilibrium in which aggregate quantities are monotone in a driving shock, consistent with our abstract results. The mutual information model opens the door to multiple equilibria, via coordination on specific support points of action distributions. We show how the equilibrium operator in the mutual information model is not a contraction map, thus providing an explicit counterexample to the possibility of using this paper's analytical tools to show similar results in a mutual-information setting.

We conclude that, while the information-acquisition underpinning and "anchoring" observation may be realistic for individual behavior in some applications, these components of the mutual information model open up the door to somewhat pathological equilibrium predictions and preclude sensible comparative statics analysis. Thus, in situations where researchers are concerned primarily with stochastic choice, the strategic mistakes model may be a tractable alternative that is still behaviorally rich enough to capture important, experimentally verified features of behavior (see Section 3.2.3).

3.5.2 State-Separable Costs in Binary-Action Games

In Appendix C.3, we study strategic mistakes in binary-action games, which are used in many applications to capture an extensive margin of adjustment and/or to simplify analysis.[15] We derive sufficient conditions on cognitive costs and payoffs to ensure unique and monotone equilibria and illustrate our results in the context of a simple investment game with linear payoffs (as in Yang, 2015). Unlike the continuous-action games studied in our main analysis, binary-action supermodular games may have multiple equilibria with small stochastic choice frictions. This result hinges on agents' ability to waver between options that have similar payoffs, but are far apart in the action space and induce very different equilibrium externalities. This result offers the following insight for researchers interested in well-posed comparative statics and not multiplicity *per se*: a "more complex" continuous-action model, by smoothing out aggregate best-response functions, may admit simpler analysis than a comparable binary-action model.

3.6 Conclusion

This paper introduces a new class of *state-separable* control costs in large games. We show how these costs accommodate a rich class of decision frictions. We provide results on equilibrium existence, uniqueness, efficiency, and monotonicity of equilibrium distributions, aggregates and mistakes. We apply these results to make robust equilibrium predictions in two macroeconomic applications, respectively to price-setting in a monetary economy and consumption and savings in a liquidity trap.

This paper's analysis of decision frictions in large games may be applicable to many additional settings in macroeconomics and finance. In Section 3.4, we show

[15]See Angeletos and Lian (2016) (in particular, Section 5) for a review of this literature.

how to recast price-setting in a monetary economy and consumption-savings choice in a liquidity trap as games with common payoff-relevant states (the money supply or aggregate demand shock) and strategic complementarity summarized in payoffs by an aggregator (the price level or real GDP). Angeletos and Lian (2016) surveys other settings in macroeconomics and finance with similar characteristics, including asset pricing and strategic firm investment.

The following "practical guide" generalizes the steps of Section 3.4 and may be useful to researchers in macroeconomics and finance who want to make general equilibrium statements about the properties of economies that feature decision frictions. First, micro-found payoffs and aggregation in the setting of interest. Second, based on an understanding of how imperfect optimization varies across states, specify an appropriate weighting function λ (or a class of plausible candidates, whose predictions one wants to contrast). Third, algebraically verify the conditions underlying our main results for equilibrium existence, equilibrium uniqueness, monotone comparative statics, and equilibrium efficiency. Fourth, use these conditions to generate *theoretically robust* and, potentially, empirically testable predictions.

Part II

Essays in Mechanism Design

Chapter 4

Priority Design in Centralized Matching Markets

This chapter is jointly authored with Oğuzhan Çelebi and has been published in *The Review of Economic Studies* as Çelebi and Flynn (2022).

4.1 Introduction

In recent years, across countries such as the US, England and Chile, ever more school districts, national university admissions boards and public authorities have adopted centralized matching mechanisms to allocate objects to agents. In many such markets, the property rights that agents receive over such objects are given by *priorities* derived from various criteria such as academic attainment, income or distance. Informed by the extensive academic literature on matching, these authorities have often introduced *stable* matching mechanisms.[1] Stable mechanisms respect these priorities in a natural sense by guaranteeing that no given agent strictly prefers the assigned object of another agent with lower priority.[2] This makes priorities (and the property rights they encode) critical to the realized distribution of outcomes in these markets across race, socioeconomic group, and space.

In this context, a large matching literature takes priorities as primitive and studies the design of the allocation mechanism (Roth, 2002). However, priorities often appear to be designed by the relevant authorities as a function of some other, underlying score.

[1] See, for example, Balinski and Sönmez (1999), Abdulkadiroğlu and Sönmez (2003), Roth, Sönmez, and Ünver (2004), and Abdulkadiroğlu, Pathak, and Roth (2005).

[2] In the context of student assignment, stable mechanisms are both individually rational and eliminate justified envy.

Prominently, Boston Public Schools (BPS) wished to ensure that students were able to attend schools close to their own homes in order to both reduce transportation costs and improve community cohesion (Dur, Kominers, Pathak, and Sönmez, 2018). However, they did not introduce a priority that ranked students strictly according to their distance from each school. Instead, in the walk-zone assignment system employed by BPS until 2013, at each school students were partitioned into two groups: walk-zone students who lived within a certain radius of the school and the others who did not. Moreover, the New York City Housing Authority (NYCHA) has the goal of providing public housing to those who cannot afford adequate housing in the absence of assistance (Collinson, Ellen, and Ludwig, 2015; Arnosti and Shi, 2017). However, agents' priorities are not strictly decreasing in their income. Instead, a household is only eligible for public housing if its income is less than a certain fraction of New York's Area Median Income (AMI).[3] Similar design concerns that trade off diversity and admitting the most academically qualified students are present in the Chicago Public Schools (CPS) system which uses scores derived from the academic merit of students.[4]

From these examples, it is clear that these authorities not only design priorities, but also choose coarse priority structures that do not reverse an underlying score. However, there exists little theoretical work that approaches the problem of optimal priority design.[5] Therefore, we study the problem of a mechanism designer that is faced with an underlying score (such as distance, income or academic achievement in our running examples) and has the power to design priorities. Following the priority designs we see in our applications, we restrict attention to priority designs that do not reverse the given underlying score.[6] We call such designs *priority coarsenings*, as

[3] In both BPS and NYCHA, the mechanism used during the matching process requires a strict priority order. In both markets, this strict order is obtained by assigning a random tie-breaker number for the agents and prioritizing according to the random tie-breaker number within each priority class.

[4] To balance these competing objectives, CPS chose a policy that divided students into four socioeconomic tiers, reserved 70% of the capacity at exam schools for these tiers in equal proportion and left the remaining 30% open to students from all groups (Ellison and Pathak, 2016).

[5] Exceptions to this include the work by Echenique and Yenmez (2015) and Erdil and Kumano (2019a) who discuss issues related to substitutable priorities. Relatedly, Hafalir, Yenmez, and Yildirim (2013) and Ehlers, Hafalir, Yenmez, and Yildirim (2014) study the design of reserves as implemented through the use of differing priorities in different slots and affirmative action policies with upper and lower bounds, while Dur, Pathak, and Sönmez (2020) study the design of precedence orders over these slots in the CPS context.

[6] Beyond the relevance in our applications of restricting attention to priority coarsenings, in certain contexts such as CPS where the underlying score is derived from academic performance, reversing scores may give rise to incentive compatibility issues whereby students intentionally perform worse on exams. Sönmez (2013) documents a case in the US Military Academy and Reserve Officer Training

they result in a coarser ordering of students relative to the initial scores. Our main results show that, under any stable matching mechanism, the set of implementable allocations can be attained by designs that split agents into at most three object-specific indifference classes (Theorem 5). Moreover, when the mechanism designer has a continuous objective function, an optimal design exists (Theorem 6) and therefore requires only three indifference classes.

Concretely, we establish a general continuum matching market framework in the spirit of Abdulkadiroğlu, Che, and Yasuda (2015) and Azevedo and Leshno (2016) for assessing the question of optimal priority design. In the *ex-ante stage* of the model, the mechanism designer knows the joint distribution of agents' preferences and rankings according to an underlying score and chooses a rule that coarsens the underlying score to maximize some arbitrary, continuous objective function. Types are then realized and in the *interim stage* agents are matched to objects according to a stable matching mechanism. Our framework features a unique stable matching, which makes it possible to express the type-contingent probability of assignment without reference to the specific stable mechanism under consideration.[7] This object functions as an *allocation* in the ex-ante stage, specifying the probability that each type of agent is assigned to each object as a function of the coarsening chosen by the designer.

Using this framework, we establish our main theoretical contribution that the optimal design can be attained by partitioning agents into at most three indifference classes at each object. The intuition for this result depends crucially on specific types of equivalence classes that emerge following the selection of a coarsening, *lottery classes*. Lottery classes have the property that some agents within them have positive probability of receiving that object and an object they rank lower in their preferences. Therefore, some agents in a lottery class have probability of being assigned to that object strictly between zero and one. We show that under a stable mechanism, there can be at most one lottery class of agents per object. This result holds for the following simple reason: if there is more than one lottery class, there will be at least two equivalence classes with interior probabilities of assignment and it must be that some agents in a lower equivalence class will be assigned the object while some agents in the higher class will not, violating stability (see the example below for a simple demonstration of this in the discrete context). Furthermore, all agents in higher

Corps where having a lower score might be more advantageous for the graduates. As a result, cadets tried to intentionally lower their scores to get assigned to a more preferred army branch.

[7] We obtain a unique stable matching through the assumption that there is full support of all student types in our economy. We show that our results are robust to relaxation of this assumption in Appendix D.6.

priority classes are guaranteed to receive the object if they prefer it, and all agents in lower priority classes never receive the object. Thus, all classes above and below the lottery class can be merged without affecting the allocation, so the outcome of any coarsening can be replicated by a coarsening that splits agents into at most three indifference classes.[8]

We leverage the simplifying power of our theoretical results in our applications. First, we study the design of distance-based priorities in the BPS system. The following discrete example provides a simple illustration of the trinary optimality result and explores how the trade-off between diversity and distance travelled shapes the optimal policy in this context. Assume there are four students, i_1, i_2, i_3, i_4, and one school with two slots. Each student i_n lives n miles away from the school and i_3 belongs to an underrepresented group of society. The mechanism designer gains $\alpha \in [0, 1]$ utility from admitting the student from the underrepresented group, i_3, and loses utility equal to $1 - \alpha$ times the total distance that admitted students travel. Mirroring the coarse walk-zone priorities in BPS, the mechanism designer decides how to coarsen students' distance to maximize this objective.

Owing to Theorem 5, it is without loss of optimality to consider priority coarsenings that divide students into at most three groups. In this particular example, it can be shown that an optimal policy is one of three such coarsenings: a two-zone priority with a cut-off at 2.5 miles, a two-zone priority with a cut-off at 3.5 miles and a three-zone priority with cut-offs at 1.5 and 3.5 miles. The first option guarantees that i_1 and i_2 are admitted for sure, the second option assigns i_1, i_2, and i_3 to the school with probability 2/3 each and the third option assigns i_1 with unit probability and i_2 and i_3 with probability half.

When are each of these optimal? A quick calculation shows that for $\alpha \leq 1/2$, the first option is optimal. This is intuitive as the first option guarantees i_1 and i_2 are admitted and for low values of α, minimizing distance travelled is much more important than diversity. For $\alpha \geq 3/4$, the second option is optimal as it maximizes the probability that the underrepresented student i_3 is admitted. This is also intuitive as in this case the mechanism designer values diversity much more than distance. For $\alpha \in (1/2, 3/4)$, the third option is optimal. Here, the mechanism designer wants to have a diverse student body but also cares about distance travelled, and hence entails a strong preference for admitting i_1.

[8]It is clear that this construction requires knowledge of the structure of the economy on the part of the planner. We discuss the validity of this assumption in our applications and robustness of our results to aggregate uncertainty in Section 4.3.4 and Appendix D.3.

While we here provided a discrete example for simplicity and to exemplify our theoretical results, in our BPS application we employ the continuum framework to provide analytical results on the optimality of various walk-zone policies. In particular, we argue that the pursued policy of two zones is compatible with a planner who places a large weight on neighborhood assignment. However, were diversity concerns to dominate, there remains additional policy latitude for a planner to adopt a three-zone policy – which our trinary optimality result ensures is the only potential welfare-improving deviation from a simple two-zone policy.

Second, we study how priority design could be used in the design of exams in the CPS system. We show how coarser grading can increase the admissions of minority groups who score less well on exams and show how the trade-off between diversity and admitting the best scoring students shapes the structure of the optimal exam design. In particular, we show that when diversity concerns are sufficiently strong, pooling students' exam scores in up to three groups constitutes optimal policy.

Finally, we study the design of income-based priorities in public housing allocation by NYCHA. Relative to the other applications, this features complications as we must consider the dynamic nature of public housing allocation. Nevertheless, in the steady state of the dynamic matching model we develop, our Theorems 5 and 6 apply directly: the planner need only introduce two income cutoffs. We show how the trade-off between widening eligibility for public housing and targeting the allocation to the most needy shapes the optimal policy. In particular, when there is a sufficiently strong relationship between income and outside options, the optimal policy excludes the richer agents from eligibility for public housing – rationalizing the policy pursued by NYCHA. However, we also show a three-tiered system may improve welfare in the case of sufficient heterogeneity in outside options.

Related Literature The theory and design of matching markets was pioneered by Gale and Shapley (1962). Balinski and Sönmez (1999) and Abdulkadiroğlu and Sönmez (2003) introduced the problems of student assignment and school choice, respectively. Much of the literature following these seminal papers focused on allocation mechanisms that take property rights (encoded in priorities) as given. While there is a small literature that studies the effects of certain classes of priorities on certain mechanisms (Erdil and Kumano, 2019a; Echenique and Yenmez, 2015),[9] this paper

[9] In particular, Erdil and Kumano (2019a) study substitutable priorities with ties and propose an algorithm to improve the widely used Deferred Acceptance algorithm from an efficiency point of view. Echenique and Yenmez (2015) emphasize that substitutability is in conflict with schools' preferences for diversity and study different rules for incorporating such preferences into the assignment mechanism.

develops the novel idea that priorities can be viewed as choice objects.

Our analysis is closely related to the rapidly growing literature on matching in the presence of distributional constraints and affirmative action. Kojima (2012) studies the widely used majority quotas and shows that such policies can harm all minority students. Hafalir, Yenmez, and Yildirim (2013) approach this question from a mechanism design perspective and introduce minority reserves to overcome the shortcomings of quota policies. Ehlers, Hafalir, Yenmez, and Yildirim (2014) generalize these reserves to incorporate multiple priority levels and accommodate further policies used in practice, such as floors and ceilings, and develop new mechanisms for hard or soft floors and ceilings. Reserve policies have also been generalized by Kominers and Sönmez (2016), who introduce and study matching with slot-specific priorities.[10] In an alternative approach to the literature with quotas and reserves, Kamada and Kojima (2017, 2018) and Goto, Kojima, Kurata, Tamura, and Yokoo (2017) study stability and efficiency in matching-with-constraints models. Finally, Çelebi and Flynn (2021) study the trade-offs between using minority quotas (and reserves) and score subsidies in affirmative action.

The common approach of papers studying priority structures, distributional constraints and affirmative action in their respective contexts is to take the priority structure as given and analyze the properties of different mechanisms. The key difference between our paper and these literatures is our introduction of, and focus on, priority coarsenings that introduce indifferences into underlying scores as a tool to design priorities.[11] In this context, in contrast to the existing literature, we compare different priority structures for a fixed stable mechanism and investigate the relationship between the allocation, welfare and the priority structure.

We apply our general results regarding the design of priorities to three different settings, which have themselves been the subject of previous research that fixes priorities and compares different mechanisms in specific contexts. First, we study distance-based priorities and the design of walk-zones with reference to BPS, a set-

[10] Further papers that analyze reserve-like policies include Doğan (2016), who proposes an assignment rule that never harms all minority students. Fragiadakis and Troyan (2017) propose a dynamic quota mechanism to improve allocations under hard bounds.

[11] As the main policy tool available to the designer in our paper is the "coarseness" of the priorities to be used in the mechanism, our paper is implicitly related to the literature that studies matching markets under indifferences. Following our applications, we require stability with respect to the tie-broken priorities and abstract away from the issues studied in that literature such as alternative stability criteria (Kesten and Ünver, 2015), computation of stable and efficient matchings (Erdil and Ergin, 2008, 2017), correlated lotteries (Ashlagi and Shi, 2014), random assignments under constraints (Budish, Che, Kojima, and Milgrom, 2013), and efficiency improving lottery mechanisms (Kesten, Kurino, and Nesterov, 2017).

ting that has been studied prominently by Dur, Kominers, Pathak, and Sönmez (2018). Second, motivated by the CPS assignment system outlined in Dur, Pathak, and Sönmez (2020), we study how diversity considerations in an environment with priorities based upon student achievement affect optimal priority design. Finally, we study income-based priorities and the allocation of public housing in NYCHA, a context studied by Arnosti and Shi (2017) in their analysis of the design of lotteries and waitlists under fixed priorities.[12] Furthermore, while not explicitly featured as applications in this paper, the design concerns we highlight are not unique to the US context. For example, Sönmez and Yenmez (2021, 2020a,b) and Aygun and Bó (2021) study affirmative action policies in India and Brazil, respectively, and propose mechanisms for each context. In all of these contexts, priority coarsenings offer a new policy lever that could be useful in cases where current policies, such as quotas in affirmative action, are controversial.

Outline The rest of the paper proceeds as follows. Section 4.2 introduces the matching model and priority coarsening. Section 4.3 studies optimal priority design and provides our main results (Theorems 5 and 6). Motivated by the BPS context, Section 4.4 applies our results in the case of distance-based scores and considers optimal walk-zone design. Section 4.5 uses our results to study priority design with test-based scores, as in CPS. In Section 4.6, we augment our framework to analyze the design of income-based scores and the allocation of public housing in a dynamic matching model. Section 4.7 concludes.

4.2 Model

In this section, we develop a model of a matching market with a continuum of students as in Abdulkadiroğlu, Che, and Yasuda (2015) and Azevedo and Leshno (2016) to study how and when priority design can affect the allocation of objects and welfare. We proceed with the standard matching literature language of matching students and schools. However, as we later show, our analysis is of relevance beyond this context. In the ex-ante stage, the mechanism designer has a prior over the distribution of student types (that comprise preferences over schools, underlying scores at each school, and other identifying information) in the population and chooses a rule that coarsens the underlying scores of students into the priorities they will hold in the interim stage. In

[12] Relatedly, Leshno (2017) and Bloch and Cantala (2017) study models of dynamic waitlists and argue that randomized assignments will improve welfare. Geyer and Sieg (2013), Waldinger (2018) and Sieg and Yoon (2020) estimate empirical models of public housing allocations and compare different mechanisms.

the interim stage, types are realized, and students submit their preferences to a stable matching mechanism that uses these priorities. Finally, the students are matched to the schools and payoffs are realized. The model timeline is shown in Figure 4-1.

4.2.1 Ex-ante Stage and Priority Coarsening

There are a finite set of schools, denoted by $\mathcal{C} = \{c_0, c_1, \ldots, c_n\}$ where c_0 is a dummy school that corresponds to a student going unmatched, and a unit measure of students. Let $\theta = (\{u_c^\theta, s_c^\theta\}_{c\in\mathcal{C}}, \kappa)$ denote the type of a student whose utility from going to school c is u_c^θ, who has score $s_c^\theta \in S_c = [0,1]$ in school c, where S_c denotes the set of possible scores at school c. For example, S_c could contain possible distances from a school, or students' scores in an exam. Finally, κ denotes any other information about the socio-economic situation or minority status of the student.[13] We use u^θ, s^θ and κ^θ to, respectively, denote the utility profile, score profile and additional information of a student with type θ. The set of student types is denoted by Θ, over which there is a probability measure F. $Q = (Q_0, \ldots, Q_n)$ denotes the capacities of schools.[14] The economy can therefore be summarized by the triple $\Omega = (F, Q, \Theta)$.

In the ex-ante stage, the designer transforms the students' scores into the priorities that will be used in the matching mechanism. Formally, a priority design at school $c \in \mathcal{C}$ is a function $\Xi_c : S_c \to P_c$ that maps students' scores $s_c \in S_c$ into their priority $\Xi_c(s_c) \in P_c$, where $P_c \subseteq S_c = [0,1]$. A *priority design* is then a function that collects each school's design $\Xi(s) \equiv (\Xi_1(s_1), \ldots, \Xi_n(s_n))$, with corresponding domain $S = \prod_{c=1}^n S_c$ and range $P = \prod_{c=1}^n P_c$. As we have motivated, we will restrict attention to coarsening rules: priority designs that coarsen, but do not reverse students' scores.

Definition 7. *A coarsening rule is a priority design $\Xi : S \to P$ such that for all $c \in \mathcal{C}$ and all $s, s' \in S$ such that $s_c \geq s'_c$, we have that $\Xi_c(s_c) \geq \Xi_c(s'_c)$. Moreover, for each $c \in \mathcal{C}$, P_c is either finite or Ξ_c is the identity function.*

The final condition that either the set of priorities at each school P_c is finite or the priority design leaves scores unchanged is a technical one that ensures ties resulting from coarse priorities can be broken while maintaining a well-defined economy. A natural example of coarsening that many will be familiar with is the conversion of fine numerical exam scores (ranging from 1 to 100) into letter grades (ranging from F to A). Example 1 demonstrates a coarsening from strict scores over $[0,1]$ to a priority

[13] We will suppress κ until the applications sections, as it is irrelevant for allocations.

[14] As school c_0 is a dummy school representing outside options, it is without loss of generality to set $Q_0 = 1$.

structure with three indifference classes, while Example 2 shows the planner may choose to use the scores as priorities without transforming them.

Example 1. *There is one school,* $|\mathcal{C}| = 1$*, scores lie in* $S = [0, 1]$*, and we coarsen these strict scores into three indifference classes* $P = \{\frac{1}{3}, \frac{2}{3}, 1\}$ *according to the rule:*

$$\Xi(s) = \begin{cases} \frac{1}{3}, & s \in [0, 1/3), \\ \frac{2}{3}, & s \in [1/3, 2/3), \\ 1, & s \in [2/3, 1]. \end{cases} \tag{143}$$

Specifically, Ξ *takes any student who had an initial score lower than* $1/3$ *and gives them a priority of* $\frac{1}{3}$*, students with initial score between* $1/3$ *and* $2/3$ *are given priority* $\frac{2}{3}$*, and students with initial score greater than* $2/3$ *are given priority* 1*.*

Example 2. *Let* $S = P$ *and* Ξ_c *be the identity function for all* $c \in \mathcal{C}$*. Then for any* s*,* $\Xi_c(s) = s$*, i.e. the priorities are identical to the scores.*

We argue that non-reversal is a relevant and natural property to demand from a priority design in our setting. From a practical perspective, such interventions seem feasible from a political economy point-of-view and have occurred in markets with an established score structure, such as distance-based priorities (as in BPS) and priorities that depend on a measurable statistic such as income (as in NYCHA).[15] Beyond the main applications of this paper, coarse priorities have also been advocated for in the 2018 US Centers for Disease Control Vaccine Allocation Guideline, which divides the general population into four tiers based on their age (CDC, 2018). Furthermore, many states in the US have recently adopted priority systems in the allocation of ventilators based on the Sequential Organ Failure Assessments (SOFA) score (Piscitello, Kapania, Miller, Rojas, Siegler, and Parker, 2020; Pathak, Sönmez, Unver, and Yenmez, 2021), which maps continuous measures of patient health to a discrete set of values for six organ systems. Moreover, when scores are based on achievement (as in CPS) or can otherwise be gamed, a transformation that does not satisfy non-reversal may incentivize students to obtain lower scores, which is clearly undesirable.[16]

We argue that requiring stability with respect to the designed priorities is reasonable for three reasons. First, if students only know the coarsened scores, a student within a given priority class would not be able to block a match between a student in the same priority class and a school. This lack of knowledge seems natural,

[15] See Sections 4.4 and 4.6 and the references therein for more detail on these contexts.

[16] As discussed in Footnote 6, Sönmez (2013) provides a concrete example of such incentive compatibility issues in the context of the US Military Academy.

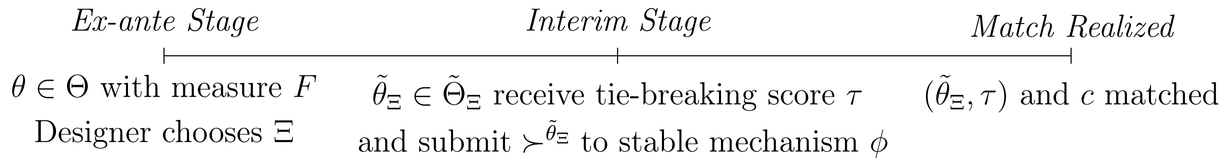


Figure 4-1: Model Timeline

for instance, in the exam schools context where two students who receive the same grade do not have access the underlying raw score from which their grades were constructed. Moreover, school boards in centralized matching markets have the ability to enforce matches and prevent schools and students from matching outside of the system. Second, if we interpret stability as encoding procedural fairness and preventing legal challenge, then stability with respect to coarsened scores retains these properties. Finally, as our examples in the introduction from BPS, CPS and NYCHA attest, authorities do in fact engage in the design of priorities and employ matching mechanisms that yield stable outcomes.

4.2.2 The Interim Stage: Matching Model

We now study how to map the choice of coarsening by the designer Ξ to a matching in the interim stage. To do this, we first construct from our economy in the *ex-ante* stage an ordinal economy that transforms utility values of the agents to ordinal preferences for the matching stage. To this end, for any type $\theta = (u^\theta, s^\theta)$, we define the corresponding ordinal type $\tilde{\theta} = (\succ^{\tilde{\theta}}, s^{\tilde{\theta}})$ by computing their induced ordinal preferences by ranking the schools in decreasing order according to u_c^θ and imposing that $s^\theta = s^{\tilde{\theta}}$.[17] Defining R as the set of ordinal preference relations over $\mathcal{C}$, we see that the set of induced types has product structure $\tilde{\Theta} = R \times S$. The distinction between a type and an induced type is subtle. As we consider an ex-ante stage for the purposes of performing welfare calculations in our later analysis, types refer to both a student's v.N-M preferences and their scores, while a student's induced type refers to the student's induced ordinal preferences and their scores.[18] An economy $\Omega = (F, Q, \Theta)$ thereby results in an ordinal economy $\tilde{\Omega} = (\tilde{F}, Q, \tilde{\Theta})$, where $\tilde{F}$ is the probability measure over $\tilde{\Theta}$ induced by F.[19] We further make the following technical

[17] For example, if $u_{c_1}^\theta = 1$, $u_{c_2}^\theta = 2$ and $u_{c_3}^\theta = 3$, the ordinal preferences are $\succ^\theta = c_3, c_2, c_1$. If $u_c^\theta = u_{c'}^\theta$ for $c \neq c'$, then students break the ties randomly.

[18] Despite being irrelevant for allocations, the cardinal utility will later matter for the welfare analysis we perform in our applications.

[19] To obtain $\tilde{F}$ from F, for each induced type $\tilde{\theta} = (\succ^{\tilde{\theta}}, s^{\tilde{\theta}}) \in \tilde{\Theta}$, simply compute the measure of types $\theta = (u^\theta, s^\theta) \in \Theta$ such that u^θ induces the ordinal preferences $\succ^{\tilde{\theta}}$.

assumption that $\tilde{F}$ admits a density $\tilde{f}$ that has full support and has no mass points:[20]

Assumption 7. *The density of all ordinal types $\tilde{f}$ is well-defined and $\tilde{f}(\tilde{\theta}) > 0$ for all $\tilde{\theta} \in \tilde{\Theta}$.*

Second, we now show how the planner's choice of coarsening Ξ affects the ordinal economy in the interim stage. This transforms each ordinal type $\tilde{\theta} = (\succ^{\tilde{\theta}}, s^{\tilde{\theta}}) \in \tilde{\Theta}$ into a new ordinal type $\tilde{\theta}_\Xi = (\succ^{\tilde{\theta}}, \Xi(s^{\tilde{\theta}})) \in \tilde{\Theta}_\Xi$ by replacing the score vector $s^{\tilde{\theta}}$ with the priority $s^{\tilde{\theta}_\Xi} \equiv \Xi(s^{\tilde{\theta}})$, and changes the set of students from $\tilde{\Theta}$ to $\tilde{\Theta}_\Xi$ and the probability measure from $\tilde{F}$ to $\tilde{F}_\Xi$.[21] Priority coarsening introduces indifferences, the existence of which necessitates tie-breaking to compute matchings. To this end, we augment the model with tie-breakers. Each student $\tilde{\theta}_\Xi$, in addition to her ordinal preferences and priority, receives a tie-breaker number $\tau \in [0,1]$, where $\tau \sim U[0,1]$. Thus, the distribution over types in the economy with tie-breakers $\tilde{F}^\tau_\Xi$ on $\tilde{\Theta}^\tau_\Xi = \tilde{\Theta}_\Xi \times [0,1]$ is almost surely such that $\tilde{f}^\tau_\Xi(\tilde{\theta}_\Xi, \tau) = \tilde{f}_\Xi(\tilde{\theta}_\Xi)$ for all $\tau \in [0,1]$.[22] This results in the coarsened ordinal economy with tie-breakers $\tilde{\Omega}^\tau_\Xi = (\tilde{F}^\tau_\Xi, Q, \tilde{\Theta}^\tau_\Xi)$, which lies in the set of all strict ordinal economies $\mathcal{O}$.

We are now ready to define the matching mechanism that applies in the coarsened ordinal economy with tie-breakers. A matching in this environment is a function $\mu : \mathcal{C} \cup \tilde{\Theta}^\tau_\Xi \to 2^{\tilde{\Theta}^\tau_\Xi} \cup \mathcal{C}$, where $\mu(\tilde{\theta}_\Xi, \tau) \in \mathcal{C}$ is the school that any ordinal type $\tilde{\theta}_\Xi$ with tie-breaker τ is assigned and $\mu(c) \subseteq \tilde{\Theta}^\tau_\Xi$ is the set of students assigned to school c.[23] Let $\mathcal{M}$ be the set of all matchings. A matching mechanism ϕ is a function $\phi : \mathcal{O} \to \mathcal{M}$ that assigns a matching to each ordinal economy. Blocking and stability are defined as follows. A student-school pair $(\tilde{\theta}_\Xi, \tau, c)$ *blocks* a matching μ if the student prefers c to her match under μ and either the school c does not fill its capacity or the school c is matched to another student who has strictly lower score than $(\tilde{\theta}_\Xi, \tau)$. Formally, $(\tilde{\theta}_\Xi, \tau, c)$ blocks μ if $c \succ^{\tilde{\theta}_\Xi} \mu(\tilde{\theta}_\Xi, \tau)$ and either (i) $\tilde{F}^\tau_\Xi(\mu(c)) < Q_c$, or (ii) there exists $(\tilde{\theta}'_\Xi, \tau') \in \mu(c)$ with (a) $s_c^{\tilde{\theta}_\Xi} > s_c^{\tilde{\theta}'_\Xi}$, or (b) $s_c^{\tilde{\theta}_\Xi} = s_c^{\tilde{\theta}'_\Xi}$ and $\tau > \tau'$. A matching μ is *stable*

[20] See Footnote 19 to see how this condition can be easily translated to a primitive condition on F.

[21] See that $\tilde{\Theta}_\Xi = R \times P$, where P is the range of Ξ. To construct $\tilde{F}_\Xi$ from $\tilde{F}$, for all types $\tilde{\theta}_\Xi = (\succ^{\tilde{\theta}_\Xi}, s^{\tilde{\theta}_\Xi}) \in \tilde{\Theta}_\Xi$, we compute the measure under $\tilde{F}$ of all types $\tilde{\theta} \in \tilde{\Theta}$ such that $\succ^{\tilde{\theta}} = \succ^{\tilde{\theta}_\Xi}$ and $\Xi(s^{\tilde{\theta}}) = s^{\tilde{\theta}_\Xi}$.

[22] See Lemma 17 in Appendix D.2 for a formal statement and proof. Also note that whenever a coarsening is not the identity, $\tilde{f}_\Xi$ is a probability mass function.

[23] The mathematical definition of a matching for the strict continuum economy we study (with ordinal types $\tilde{\Theta}^\tau_\Xi$) follows Azevedo and Leshno (2016) and requires that μ satisfies the following four properties: (i) $\mu(\tilde{\theta}_\Xi, \tau) \in \mathcal{C}, \forall (\tilde{\theta}_\Xi, \tau) \in \tilde{\Theta}^\tau_\Xi$; (ii) $\mu(c) \subseteq \tilde{\Theta}^\tau_\Xi$ is measurable, $\tilde{F}^\tau_\Xi(\mu(c)) \leq Q_c, \forall c \in \mathcal{C}$; (iii) $c = \mu(\tilde{\theta}_\Xi, \tau) \iff (\tilde{\theta}_\Xi, \tau) \in \mu(c)$; (iv) $\{(\tilde{\theta}_\Xi, \tau) \in \tilde{\Theta}^\tau_\Xi : c \succ^{\tilde{\theta}_\Xi} \mu(\tilde{\theta}_\Xi, \tau)\}$ is open $\forall c \in \mathcal{C} \setminus \{c_0\}$. This last requirement imposes that the set of students that prefer their match to any given school (excluding the outside option) is open.

if there are no blocking pairs. A mechanism ϕ is stable if it returns a stable matching for all economies.

In this environment, any coarsening Ξ and the matching mechanism ϕ together induce a probability distribution for each student type over the school that they are ultimately assigned. We call this distribution an *allocation* $g_{(\Xi,\phi)} : \Theta \times \mathcal{C} \to [0,1]$, with the probability that type θ is assigned to school c given by $g_{(\Xi,\phi)}(c|\theta)$.[24] We denote the set of potential allocations by $\mathcal{G}$.[25] We construct this allocation from the matching and distribution of tie-breakers by taking, for each tie-breaker realization of each student, the match the student receives and then integrating this over the uniform distribution of tie-breakers (see Lemma 18 in Appendix D.2.1). These steps ensure that $g_{(\Xi,\phi)}$ is a well-defined distribution and respects the constraints imposed by the mechanism, including that no school is over capacity and no student has probability exceeding one of attending all potential schools.

Throughout the paper, as we have motivated, we will assume the matching mechanism is stable.

Assumption 8. *The matching mechanism ϕ is stable.*

The importance of Assumptions 7 and 8 for our analysis is that Assumption 7 implies that there is a unique stable matching. Thus, after the planner fixes the coarsening in the *ex-ante* stage, Assumption 8 pins down the matching in the interim economy uniquely (see Lemma 17 in Appendix D.2.1). Therefore, it is not important for us to specify which stable matching mechanism ϕ is. We will correspondingly suppress dependence on ϕ for the remainder of the analysis and write allocations as g_Ξ. A strong justification for these assumptions (which rule out multiple stable matchings) is that, empirically, the set of stable matchings has been found to be very small in large markets, including BPS.[26] Nevertheless, in Appendix D.6, we relax this full-support assumption and allow multiple stable matchings to exist and we summarize the robustness of our main results to relaxing Assumption 7 in Section 4.3.5.

[24] Abdulkadiroğlu, Angrist, Narita, and Pathak (2017) use a similar representation to obtain the propensity score that any student is matched to a school to estimate treatment effects, albeit not as a function of any design tool of the policymaker.

[25] Allocations will lie in (a subset of) the space of measurable functions with finite integral, so that $\mathcal{G} \subset L^1(\Theta \times \mathcal{C})$. See Appendix D.2.2 for details on the measure space with respect to which we demand that g is measurable.

[26] See Roth and Peranson (1999) for evidence from National Resident Matching Program and Pathak and Sönmez (2008) for evidence from BPS. In particular, in BPS for school years 2005-2006 and 2006-2007, there is only one stable matching for either year.

4.3 Priority Design

Having established a framework for the analysis and justified our restriction of the policy space of the designer to be that of priority coarsenings, we can now prove our main results (Theorems 5 and 6) and establish the existence of an optimal trinary coarsening. These results are stark as they reduce the complexity of finding the optimal coarsening from an infinite-dimensional problem to a $2|\mathcal{C}|$-dimensional problem and place a simple structure on optimal policies. We later leverage these results directly in applications to provide concrete insight into a number of important problems in market design.

4.3.1 Trinary Replication

We now turn to proving the main result of the paper: any allocation achievable via a coarsening can be replicated with a trinary coarsening. Formally, we define a trinary coarsening as a coarsening such that the priority structure at each school features at most three equivalence classes of students.

Definition 8. *A coarsening $\Xi : S \to P$ is trinary if $|P_c| \leq 3$ for all schools $c \in \mathcal{C}$.*

Our main implementation result is stated formally as Theorem 1.

Theorem 5. *Suppose that a coarsening Ξ induces the allocation g_Ξ. There exists a trinary coarsening Ξ' that induces g_Ξ.*

Proof. See Appendix D.1.1. □

The basic intuition for this result and why it follows from stability is easily seen in the following example. Consider a model with an outside option and a school that all students prefer to the outside option. There is positive density of all scores between zero and one (in view of Assumption 7) at the school and the scores are coarsened into finitely many indifference classes. The matching mechanism is stable (by Assumption 8). Thus, if there is a positive probability that a student with lower priority is admitted to the school, then the higher priority student must not be allocated to the outside option, as that will violate stability. As a result, under any coarsening, there is at most one class of students (the lottery class) who have probability strictly between zero and one of being admitted. All students in higher priority classes are admitted with probability one while all students in lower priority classes are admitted with probability zero.[27] Thus, there can exist at most one lottery class, and the combination of stability and full support of ordinal types (Assumptions 8 and 7, respectively) pins down this uniqueness. The outcome of the coarsening can then be generated by an alternative coarsening that preserves the lottery class from the first coarsening, maps all students above this class into one class, and maps all students below the lottery class into another class. Hence, the outcome of any coarsening can be replicated by another coarsening with at most three indifferences classes at each school.

4.3.2 The Planner's Objective

To discuss optimal priority design, it is necessary to have an objective function for the planner. We assume that the planner has a complete and transitive preference over allocations g in the set of all potential allocations $\mathcal{G}$ represented by a utility function $Z : \mathcal{G} \to \mathbb{R}$.[28] Moreover, we make the technical assumption that the utility function of the planner is continuous:

Assumption 9. *The social planner's objective function $Z : \mathcal{G} \to \mathbb{R}$ is continuous in g.*[29]

In the interests of clarity, we now discuss three natural specifications of planner utility that satisfy this assumption and will be used later in our applications: a utilitarian planner; a planner who cares about student utility with some penalty for deviating from the underlying score; and a planner who has affirmative action concerns.

1. A utilitarian social planner has utility function given by:

$$Z(g) = \int_{\Theta} \sum_{c \in \mathcal{C}} \lambda(\theta) u_c^{\theta} g(c|\theta) dF(\theta) \tag{144}$$

[27] For concreteness, suppose instead that there were two equivalence classes A and B, where students in both have interior probabilities of assignment – so that there are two lottery classes – and students in A have higher priority than those in B. As these probabilities are interior, necessarily some students in A will not be allocated the school, while some students in B are. As there is full support of ordinal types, some students in A who are not admitted most prefer the given school. Moreover, they have higher priority than all assigned students from B. Thus, the outcome is unstable, as these rejected students in A and the school form a blocking pair. As the mechanism is assumed to be stable, this is a contradiction.

[28] Note that this rules out preferences that depend non-instrumentally on the coarsening itself. Namely, this rules out a preference for 'simple' policies. We argue that this restriction is unimportant given the fact that optimal policies will be simple insofar as they are trinary in any case.

[29] Recall from Footnote 25 that $\mathcal{G} \subset L^1(\Theta \times \mathcal{C})$, so continuity is here meant with respect to the associated L^1-norm. See Appendix D.2.2 for more details.

for some function yielding welfare weights $\lambda : \Theta \to \mathbb{R}_+$.

2. A priority-augmented social planner has utility function given by:

$$Z(g) = \int_\Theta \sum_{c \in \mathcal{C}} \left[u_c^\theta + \lambda(\theta) h(s_c^\theta) \right] g(c|\theta) dF(\theta) \tag{145}$$

   where $h : [0,1] \to \mathbb{R}$ is a monotonically increasing function that determines the base cost of the score not being met and $\lambda : \Theta \to \mathbb{R}$ is the weight of that loss for each underlying type θ.

3. An affirmative-action-concerned social planner has utility function given by:

$$Z(g) = \int_\Theta \sum_{c \in \mathcal{C}} u_c^\theta g(c|\theta) dF(\theta) + h\Big(\int_\Theta 1\{\kappa^\theta \in D\} g(\hat{c}|\theta) dF(\theta) \Big) \tag{146}$$

   where $\kappa^\theta \in D$ means that type θ is a student in the group which the planner wishes to ensure is more represented (recall that κ^θ corresponds to any non-preference or score information corresponding to a student), $\hat{c}$ is some given school, and h is a continuous and monotonically increasing function. This specification therefore rewards the planner for admitting more students in group D to the particular school $\hat{c}$. In practice, one might imagine that $\hat{c}$ is a high-quality school and D is an underrepresented minority group.

Given this structure, the planner's problem is to choose a coarsening such that the induced allocation maximizes the planner's utility function over the set of potential coarsening rules. By Theorem 5, it is without loss of optimality for the planner to restrict attention to trinary coarsenings. That is, they can simply select two cutoff values for each school $v = \{\underline{P^c}, \overline{P^c}\}_{c \in \mathcal{C}}$ where $v \in \mathcal{V} = \{v \in [0,1]^{2|\mathcal{C}|} | \overline{P^c} \geq \underline{P^c}, \forall\, c \in \mathcal{C}\}$. In this representation, the $\overline{P^c}$ represent the score cutoffs for membership of the highest priority class and the $\underline{P^c}$ represent the score cutoffs for membership of the middle indifference class.[30] Hence, coarsening rules reduce simply to points in a closed subset of the unit hypercube, $\mathcal{V}$. Thus, the planner's problem can be stated as:

$$\mathcal{V}^* = \arg\max_{v \in \mathcal{V}} Z(g_v) \tag{147}$$

[30] Following Azevedo and Leshno (2016), one can gain an interpretation of these thresholds in terms of the budget sets of students. If a student's score at a school exceeds $\overline{P}_c$ then that school is in their budget set with certainty. If a student's score lies between $\underline{P}_c$ and $\overline{P}_c$, then the school is in their budget set with some probability between zero and one. If a student's score lies below $\underline{P}_c$, then the school never lies in their budget set.

where g_v is the allocation induced by cutoff vector $v \in \mathcal{V}$.

4.3.3 Optimal Priority Design

Having established that any coarsening requires only three equivalence classes per school and set up the problem of the planner, we now show that there exists an optimal coarsening. First, we prove that g_v is continuous in v (Lemma 19 in Appendix D.2). In view of the fact that the domain $\mathcal{V}$ is compact, and the objective function is continuous in g (Assumption 9), it follows that an optimal coarsening exists. This is formalized as Theorem 6 below.

Theorem 6. *$\mathcal{V}^*$ is non-empty. That is, there exists a trinary coarsening that is optimal.*

Proof. See Appendix D.1.2. □

The implications of Theorem 2 are significant. In particular, it reduces the dimensionality involved in finding the optimal coarsening from an infinite-dimensional problem to a $2|\mathcal{C}|$-dimensional problem, as now we only need to choose two numbers per school to attain any optimum. This is interesting as it not only implies that problems of priority design for school districts are substantially simpler than one might expect but also facilitates simple computation of the value of a given policy even in cases with a large number of schools. We later leverage this result to provide insights into the structure of the optimal priority design in each of our three leading applications: design of walk-zone policies under distance-based priorities in BPS; design of diversity policies under achievement-based priorities in CPS; and the allocation of affordable housing under income-based priorities by NYCHA.

4.3.4 The Impact of Aggregate Uncertainty

In our analysis, we have assumed that the planner both knows the distribution of student types and that there is a continuum of students. In view of these assumptions, there is no aggregate uncertainty in the market and the planner knows that their choice of coarsening will lead to a particular, deterministic allocation. As most of the markets we study (BPS, CPS, and NYCHA) are large, have used centralized assignment mechanisms for a number of years, and are arguably likely to have similar distributions of types from year to year (they are stationary), we argue that this is a reasonable assumption.[31]

[31] In particular, from 2010-15 in CPS, we compute that the admissions cutoff for any school in the merit slots is within 3% of that school's average merit slot cutoff over this time period 96% of

Nevertheless, to investigate the robustness of our results to aggregate uncertainty, in Appendix D.3 we study the same problem considered in the main text augmented with uncertainty on the part of the planner regarding the distribution of student types in the population. Formally, we suppose that there is a finite set of probability measures $\mathcal{F}$ that the planner entertains as possible. In this context, Theorems 5 and 6 can be extended to show that an optimal coarsening still exists but that it may involve up to $2|\mathcal{F}|$ cutoffs at each school (Proposition 45 in Appendix D.3). As a result, the presence of uncertainty can substantially complicate the problem of priority design and give rise to a less coarse priority structure (see Example 7 in Appendix D.3 for an explicit example of this). We can further characterize when uncertainty causes a welfare loss to the planner relative to the benchmark without aggregate uncertainty. In particular, uncertainty induces no welfare loss if and only if the *ex post* optimal lottery classes either coincide or never overlap across all states of the world (Proposition 46 in Appendix D.3). Intuitively, it is exactly when aggregate uncertainty makes it impossible for the planner to target the same optimal lottery class across states of the world that this uncertainty has bite.

While we maintain that the assumption of no aggregate uncertainty is appropriate for our applications, these results suggest that our main results may have less bite in settings with appreciable aggregate uncertainty, as might be the case when markets change a great deal from year to year, or one is designing priorities in an unfamiliar market.

4.3.5 Extensions: Homogeneous Coarsenings and Multiple Stable Matchings

In Appendix D.4, we study an extension of the general analysis in this section where a planner is constrained to use the same coarsening at every school. We prove that suitably revised versions of Theorems 5 and 6 continue to hold in this setting, but now the designer needs to potentially specify up to $2|\mathcal{C}|$ cutoffs that are the same for each school (Proposition 47 in Appendix D.4). We characterize when the imposition of homogeneity leads to a loss in welfare: there is no resulting loss in welfare when the cutoffs for the lottery classes of each school either coincide exactly or do not overlap at all (Proposition 48 in Appendix D.4). Intuitively, as students who always or never gain admission to a school can receive the same allocation under homogeneity, the imposition of homogeneity leads to losses insofar as it makes it impossible to have

the time, providing strong evidence of approximate stationarity in this market. Cutoff score data is publicly available from CPS.

the same regions of students (the lottery classes) who have fractional assignment probability to each school.

On the technical side, we proved our theoretical results under the condition that there is full support of ordinal types. When this assumption fails, there can be multiple stable matchings and one must address a number of technical details. To this end, in Appendix D.6, we relax this condition and show that suitably modified versions of Theorems 5 and 6 continue to hold when the mechanism-designer optimal selection from the set of stable matchings is used (Theorems 13 and 14 in Appendix D.6) and that Theorem 5 continues to hold under the student-optimal selection (Theorem 15 in Appendix D.6). However, the student-optimal selection can cause the mechanism designer's objective to jump down, and so optima can fail to exist. Thus, Theorem 6 fails to hold under the student-optimal selection (see Example 9 in Appendix D.6).

4.4 Application: Distance-Based Priorities and Walk-Zone Design

In many school districts, such as Boston, San Francisco, Denver and much of the UK, the distance between students and schools plays an important role in the assignment process. One widely studied example, which provides the concrete motivation for the theoretical exercise in this section, is the walk-zone assignment system Boston Public Schools (BPS) utilized until 2013. Under this policy, students were partitioned into two sets at all schools: the walk-zone students who live sufficiently close to the school and the others who do not.[32] In the language of our model, this corresponds closely to a situation where the underlying score is distance and the pursued policy is a coarsening that splits students into two groups.[33]

The main issue for BPS in designing its school admissions policy was the trade-off between two competing desires. On the one hand, it is desirable to have students attend schools that are closer to their homes on grounds of decreasing transportation costs for the school district and improving community cohesion. Indeed, Landsmark (2009) notes the costs of school transportation are very large for the district, at around \$70 million.[34] On the other hand, it is also desirable to ensure that schools

[32]See Dur, Kominers, Pathak, and Sönmez (2018) for a more detailed account of this setting.

[33]To account for additional dimensions such as sibling-based priorities, one need only construct a composite underlying score comprising distance and sibling status. All of our analysis then applies to the model with this composite score.

[34]Moreover, Mayor Menino stated that (Goldstein, 2012): "Pick any street. A dozen children probably attend a dozen different schools... Parents might not know each other; children might not play together. They can't carpool, or study for the same tests."

have a diverse student body and for families to have greater choice over the schools they are able to attend. This concern is particularly relevant in communities that are socioeconomically segregated, such as those in Boston.[35] That this problem of conflicting objectives is at the heart of the design problem is attested to by Daley (1999), who notes that the walk-zone policy was created with the aim of "striking an uneasy compromise between neighborhood school advocates and those who want choice."

Thus motivated, we study the optimal distance-based priority design from the perspective of a mechanism designer who cares both about assigning students to schools they prefer and the distance students have to travel to their school. Our main results, Theorems 5 and 6, apply directly in this environment and imply that the optimal design can be attained via the use of at most three zones per school. We also show how the trade-off between distance and diversity shapes the structure of the desirable walk-zone policy and show when a simple walk-zone policy, corresponding to that pursued by BPS, is optimal.

4.4.1 Model

There is one school G with capacity $Q \in (0,1)$ and an outside option B. There is a unit measure of students who have bounded and positive Bernoulli utility $u \in \mathcal{U}$ from attending school G. The utility from attending B is normalized to zero. Students have underlying score $s \in [0,1]$ at school G. Students are indexed by their type $\theta = (u,s)$ and there is a joint distribution over the set of types $\Theta = \mathcal{U} \times [0,1]$ given by $f(\theta)$ such that there is a uniform distribution of underlying scores.[36] There is a continuous cost of students of score s attending G given by $C(s)$. This function can be interpreted as capturing transport costs, community cohesion or fairness costs associated with a student of score s attending school G. Finally, for this section, we assume that the school board is utilitarian and has no distributional preferences:

$$Z(g) = \int_0^1 \int_{\mathcal{U}} \big(u - C(s)\big) g(s) dF(u,s) \tag{148}$$

where g is the probability that a student with score s attends school G.

We denote a priority coarsening as a vector of cutoffs $v = (v_1, \ldots, v_n)$ where

[35] Indeed, Levinson, Noonan, Fay, Mantil, Buttimer, and Mehta (2012) note that increasing the priority of students who live closer to a school, as was the case under the walk-zone system, reduced the quality of schools certain socioeconomic groups could attend, making the assignment less equitable.

[36] This is without loss in this environment as it simply redefines the scores over students.

$0 \equiv v_0 \leq v_1 \leq \cdots \leq v_n \leq v_{n+1} \equiv 1$. Students with $s \in [v_i, v_{i+1})$ for $i < n$ or $s \in [v_n, v_{n+1}]$ have the same priority at school G as all other students with scores within the same interval prior to tie-breaking. We label the set of such vectors for a given natural number n as $\mathcal{V}^n$. We further denote the probability that a student with $s \in [v_i, v_{i+1})$ goes to G under uniform tie-breaking by g_i^v. It is further useful to define the expected contribution to social welfare of a student with score s being assigned to school G:

$$W(s) = \mathbb{E}[u|s] - C(s) \equiv B(s) - C(s) \tag{149}$$

where $B(s)$ captures the benefit to social welfare of student with score s attending G which we assume to be continuous and $C(s)$ is the cost. Using this notation, the school district's value from a policy (n, v) is:

$$Z(n, v) = \sum_{i=1}^{n} g_i^v \int_{v_i}^{v_{i+1}} W(s)ds \tag{150}$$

Thus, the school district faces the problem:

$$Z^* = \max_{n, v \in \mathcal{V}^n} Z(n, v) \tag{151}$$

It is important to note that this problem remains non-trivial as the choice object is of arbitrarily high-dimension. Our Theorems 5 and 6 makes this problem tractable:

Corollary 11. *Under any stable mechanism, Z^* exists and there exists an optimal policy (n^*, v^*) such that $n^* = 2$.*

Proof. See Appendix D.1.3. □

With Corollary 11 in hand, we can restrict attention to considering coarsening rules $v = (v_1, v_2) \in \mathcal{V}^2$. We now apply this result to solve the problem of the school district.

4.4.2 Solving the School District's Problem

In view of Corollary 11, one notes that there are three types of regions that can arise depending on the priorities used. The first is an *acceptance region*, where students are assigned to the school with probability one ($s \geq v_2$). The second is a *lottery region*, where students are assigned to the school if they have a high enough lottery number ($s \in [v_1, v_2)$). The third is a *rejection region*, where students are never assigned to the school ($s < v_1$). Depending on the existence of such regions or not, there are five possible types of priorities that can be optimal:

i. A ‘double walk-zone’: an interior case where all three regions exist

ii. A ‘small walk-zone’: a semi-interior case with an acceptance region and lottery region

iii. A ‘large walk-zone’: a semi-interior case with a lottery region and rejection region

iv. ‘full coarsening’: a corner case with just a lottery region

v. ‘no coarsening’: a corner case with an acceptance region and a rejection region

Moreover, in any of the (semi-)interior cases (i.e. any case excluding no coarsening or full coarsening), the optimality condition for the cutoffs is simple:

$$W(v_1) = W(v_2) = \frac{1}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds \tag{152}$$

That is to say, whenever a cutoff is on the interior, the cutoff is simply set to equalize the marginal contribution to social welfare of the student at the cutoff to the average contribution of all students in the lottery zone. As a result, under differentiability of W each optimal cutoff either satisfies Equation 152 or it is on the boundary.[37]

Proposition 11. *Any solution to the school district's problem (Equation 151)* $v^* = (v_1^*, v_2^*)$ *must satisfy one of the following conditions for each* v_i^**:*

1. *The cutoff is an interior optimum:*

$$W(v_i^*) = \frac{1}{v_2^* - v_1^*} \int_{v_1^*}^{v_2^*} W(s)ds \tag{153}$$

   Moreover, the above equation is sufficient for v^* *to be a local optimum whenever* $W'(v_i^*) > 0$ *for each cutoff to which this condition pertains.*

2. *The cutoff is on a boundary of the relevant constraint set:*

$$v_1^* \in \{0, 1 - Q\}, \quad v_2^* \in \{1 - Q, 1\} \tag{154}$$

Proof. See Appendix D.1.4. □

[37] Differentiability of W can be ensured by mild assumptions on primitives.

This proposition shows which policies can be optimal and how to compute optima given a parametric environment, but it is otherwise silent on the forces that govern the structure of the optimum. In the next section, we study these questions in a more specialized environment.

4.4.3 Optimal Walk-Zones in a Parametric Environment

To study the key trade-off faced by BPS policymakers between assigning students to schools close to where they live and ensuring both choice and diversity, we now examine how the structure of a simple parametric environment with these features determines the optimal walk-zone structure. We further assume that there are two utility types of students $u \in \{u^R, u^P\}$. Moreover, we assume that conditional on these utility types, there is a higher density of u^R types with higher scores. Specifically, we assume that score and utility have following joint distribution:

$$f(u^R, s) = s \quad f(u^P, s) = 1 - s \tag{155}$$

Typically, we will take $u^R < u^P$ and interpret this environment as one where poor students derive greater benefit from attending the good school relative to their outside option and the school is located in a neighborhood primarily featuring rich students. The average marginal benefit to social welfare of students with score s is therefore given by:

$$B(s) = \mathbb{E}[u|s] = u^R s + u^P(1 - s) \tag{156}$$

On the cost side, we specify a parametric cost function of admitting a student of score s given by:

$$C(s) = \frac{\alpha}{1 + \delta \exp\{\varepsilon(s - \bar{s})\}} + \mathbb{I}[s \leq \bar{s}]\beta(\bar{s} - s)^\gamma \tag{157}$$

where $\alpha, \beta, \delta, \varepsilon \geq 0$, $\bar{s} \in (1 - Q, 1)$ and $\gamma \geq 1$. See that this function accommodates the following two features. First, there is a score $\bar{s}$ below which there is a 'sharp' increase in the cost of a student attending the school, whenever ε is large. This can be thought of as there being a distance $1 - \bar{s}$ within which students can walk to school, beyond which walking becomes infeasible for students and transportation is required. Second, there are steadily increasing and perhaps convex costs of students having a low score, with β controlling the slope and γ the normalized convexity. This captures increasing costs of transportation and potentially fairness and community cohesion costs associated with admitting students who live far from the school.

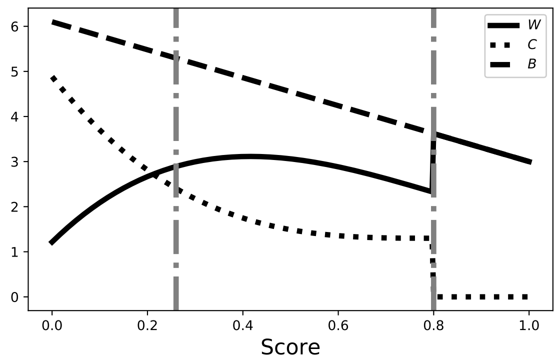


Figure 4-2: The W, B and C Functions With an Optimal 'Double' Walk-zone

As in the general analysis, the key object of interest for computing the optimal score cutoffs is the function:

$$W(s) = B(s) - C(s) \tag{158}$$

which captures the overall contribution to social welfare of a student who has underlying score of s. Moreover, given the parametric structure of B and C, W features the following trade-off: admitting poorer students increases the quality of the assignment but is more costly as those poorer students live further away.

To understand how these objects vary with the score in this parametric environment, Figure 4-2 plots the W, B and C functions for a case where:

1. Students who live further away derive higher utility from attending the school $u^P > u^R$. This can be seen in the figure as the downward sloping benefit line.
2. Students who live beyond $\bar{s} = 0.8$ have rapidly increasing transport costs (large ε).
3. There is a convex cost of students beyond $\bar{s}$ attending the school, $\gamma > 1$.

Despite the parametric structure, this model is still sufficiently rich to demonstrate each of the five classes of walk-zone policy This is stated formally in Proposition 12:

Proposition 12. *For each of the five classes of policy, there exist open sets of* $(\alpha, \beta, u_P, u_R)$ *such that each class of policy is uniquely optimal for these parameters.*

Proof. See Appendix D.1.5. $\square$

Guided by this result, we now provide an intuitive discussion of when each of these classes of policy is optimal. First, we consider the double walk-zone case and refer

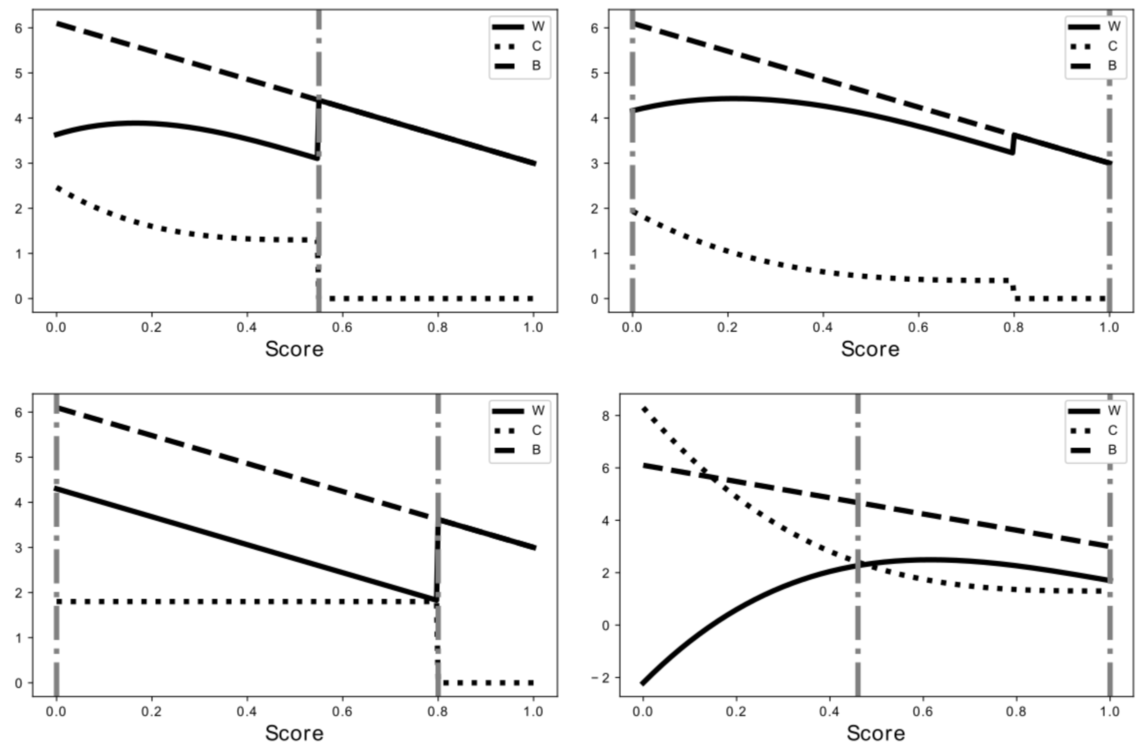


Figure 4-3: No coarsening (top-left pane), full coarsening (top-right pane), a 'small' walk-zone (bottom-left Pane) and a 'large' walk-zone (bottom-right pane)

the reader back to Figure 4-2, wherein the dashed grey lines in represent the optimal score cutoffs. See that in this example, the upper cutoff is close to $\bar{s} = 0.8$ and the lower cutoff is below the school's capacity. As a result, there is an acceptance region, a lottery region and a rejection region. Intuitively, the fixed cost being sufficiently large incentivizes the creation of the acceptance zone. Moreover, having a variable cost that is sufficiently large but not too large creates a region after $\bar{s}$ where the benefit increases more rapidly than cost, inducing coarsening. Finally, the convexity of cost eventually implies that cost exceeds benefit and it is optimal to create a rejection zone. Intuitively, the optimum simply balances the benefit of there being a higher average utility from students who are further from the school with the increasing cost of these students.

Second, we demonstrate the two corner cases that feature full coarsening and no coarsening. A simple case in which the designer will pursue full coarsening is a case where $u^P > u^R$ and there is no cost of students who live further away attending the school $\alpha = \beta = 0$. On the other hand, no coarsening will obtain in any case where $u^P < u^R$ or α and β are sufficiently large as it is optimal to simply admit all students who live as close to the school as is possible. The corner policies in these cases are shown in Figure 4-3.

Finally, the two simple walk-zone cases are also intuitive. When there is a very large fixed cost of students below $\bar{s}$, the mechanism designer wants to admit students who are sufficiently close with certainty. Moreover, when $u_P > u_R$ and β is sufficiently small, the mechanism designer wants to coarsen the remaining priority below $\bar{s}$ as much as possible. This is because the students who are furthest away contribute the most to social welfare as u^P types are relatively most dense. The case for a large walk-zone is similar. When the fixed cost is small but the cost of being further away is large, it is optimal to coarsen up to the point where the cost of being further away becomes too large. The simple walk-zone policies in these cases are shown in Figure 4-3.

Having shown how each policy can obtain depending on the strength of the trade-off between match quality and neighborhood assignment, we supplement this analysis by deriving comparative statics in the optimal cutoffs as the relative strength of these motives changes. Concretely, we see how, in the case of an optimal double walk-zone, changes in transportation costs and relative utilities of students move the optimal cutoffs:

Proposition 13. *Suppose that the solution to the planner's problem is interior and unique. The following comparative statics hold:*

1. *An increase in transportation costs enlarges the acceptance region and shrinks the lottery region:*

$$\frac{\partial v_1^*}{\partial \beta} > 0, \quad \frac{\partial v_2^*}{\partial \beta} < 0 \tag{159}$$

2. *An increase in relative utility of students who live farther away shrinks the acceptance region and enlarges the lottery region:*

$$\frac{\partial v_1^*}{\partial u_P} < 0, \quad \frac{\partial v_2^*}{\partial u_P} > 0 \tag{160}$$

Proof. See Appendix D.1.6. □

The first of these results shows how increasing transportation costs for students who live furthest from the school will disincentivize the admission of poorer students. On the contrary, when the benefits of admitting poorer students increase, the reverse is true. Even though we do not explicitly model a preference for diversity here, this increase can be thought as a preference for such students from the perspective of mechanism designer. If one interprets the mechanism designer as embodying the aggregate preferences of society, an increase in u_P may represent an improvement in the political organization of such parents or an increase in public pressure they exert. As one would expect, our comparative statics suggest that such an improvement will increase their representation in higher quality schools.

4.5 Application: Designing Exams for Diversity

In our second application, we study how to optimally design exams in an environment where a planner cares about both admitting higher quality students and the diversity of the student body in a competitive exam school. Concretely, our analysis is motivated by exam schools in the Chicago Public Schools (CPS) system, where the district uses student achievement in entrance exams as the basis for its priorities but also has a long history of controversial diversity-based affirmative action policies.[38]

The key trade-off faced by CPS is admitting the students with the highest academic qualifications while having racially and socioeconomically diverse student bodies.[39] Presently, CPS uses a strict ranking based on a composite academic score, which we take as the underlying score. In our model, the mechanism designer can also construct a coarser ranking that partitions students into different achievement levels, without releasing the strict ranking they obtained in the exam. A coarser ranking may help students with lower scores, who are potentially from less-advantaged socioeconomic backgrounds. However, this comes at the cost of admitting students with lower levels of achievement in the entrance exam. In this section, we employ Theorem 5 to characterize when such a coarse grading policy can be used to improve the allocation. We therefore explore the possibility of using priority coarsening to increase diversity without any other explicit affirmative action policies, thereby adding an additional policy lever for bodies such as CPS to consider.

4.5.1 Model

There is one school with capacity Q, to be interpreted as a desirable exam school. The exam school gives all students common utility u, which exceeds the utility they receive at their outside option. Students have exam scores $s \in [0, 1]$. There are two different socioeconomic groups of equal size, κ_1 (the underrepresented group) and κ_2, where the vector (s, κ) summarizes the type of any given student. The score densities

[38]For more detail on the institutional setting of exam schools in the CPS system, we refer the reader to Dur, Pathak, and Sönmez (2020) and Ellison and Pathak (2016).

[39]This is attested to by the Blue Ribbon Commission (BRC), which was appointed to review CPS's policy remarks (Dur, Pathak, and Sönmez, 2020): "The BRC wants these programs, [exam schools], to maintain their academic strength and excellent record of achievement, but also believes that diversity is an important part of the historical success of these programs."

of each type of students are given by $f_{\kappa_1}(s)$ and $f_{\kappa_2}(s)$, where $f_{\kappa_1}(s) + f_{\kappa_2}(s) = 1$. Motivated by affirmative action concerns of CPS, we assume the mechanism designer has the following utility function:

$$Z(g) = \sum_{\kappa_i} \int_0^1 sg(c|s,\kappa_i)dF_{\kappa_i}(s) + h\Big(\int_0^1 g(c|s,\kappa_1)dF_{\kappa_1}(s)\Big) \tag{161}$$

where h is a strictly increasing function. The first term represents the benefit to the mechanism designer from assigning students with higher scores to the school. The second term represents the benefit of assigning students from the underrepresented socioeconomic group.

4.5.2 Optimal Exam Design

We now study when designing the exam leads to welfare improvements. In this context, an exam design takes the form of score cutoffs $0 \equiv v_0 \leq v_1 \leq \cdots \leq v_n \leq v_{n+1} \equiv 1$ such that students with $s \in [v_i, v_{i+1})$ for $i < n$ or $s \in [v_n, v_{n+1}]$ have the same priority as all other students with scores within the same interval prior to tie-breaking. It is natural to interpret such a design as providing a coarse grading of an entrance exam rather than simply ranking the students. Owing to Theorem 1, we can restrict attention to trinary coarsenings. Hence it is without loss of generality to assume the mechanism designer picks two numbers $v_1 < r$ and $v_2 \geq r$ as score cut-offs where $r \equiv 1 - Q$. Under a grading policy (v_1, v_2), we have that the allocation is given by:

$$g(c|s,\kappa_i) = \begin{cases} 1 & \text{if } s \geq v_2 \\ p_L(v_1,v_2) & \text{if } s \in [v_1, v_2) \\ 0 & \text{if } s < v_1 \end{cases} \tag{162}$$

where $p_L(v_1, v_2) = \frac{v_2 - r}{v_2 - v_1}$. The utility of the mechanism designer under any coarsening (v_1, v_2) is then given by:

$$\begin{aligned} Z(v_1,v_2) = &\sum_{\kappa_i} \left(\int_{v_2}^1 sdF_{\kappa_i}(s) + p_L(v_1,v_2)\int_{v_1}^{v_2} sdF_{\kappa_i}(s)\right) \\ &+ h\left(\int_{v_2}^1 dF_{\kappa_1}(s) + p_L(v_1,v_2)\int_{v_1}^{v_2} dF_{\kappa_1}(s)\right) \end{aligned} \tag{163}$$

We can now ask when exam design leads to welfare improvements. Formally, we say that exam design leads to welfare improvements if there exists a pair (v_1, v_2) such that $Z(v_1, v_2) > Z^{NC}$, where Z^{NC} is the utility the planner receives from not coarsening exam scores.[40] Using the structure of payoffs, one can achieve the following characterization of when exam design leads to welfare improvements:

Proposition 14. *Exam design leads to welfare improvements if and only if there exists a pair* (v_1, v_2) *where* $v_2 > r$ *and* $v_1 < r$ *such that:*

$$\frac{1}{2}[(v_2 - r)(r - v_1)] < h\Big(\frac{1}{2} - F_{\kappa_1}(v_2) + (F_{\kappa_1}(v_2) - F_{\kappa_1}(v_1))p_L(v_1, v_2)\Big) - h\Big(\frac{1}{2} - F_{\kappa_1}(r)\Big) \tag{164}$$

Moreover, if h *is linear with slope* $\alpha > 0$*, this inequality reduces to:*

$$(r - v_1) < 2\alpha\Big(\frac{F_{\kappa_1}(v_2) - F_{\kappa_1}(v_1)}{v_2 - v_1} - \frac{F_{\kappa_1}(v_2) - F_{\kappa_1}(r)}{v_2 - r}\Big) \tag{165}$$

Proof. See Appendix D.1.7. □

This condition simply compares the loss from having students with lower scores to the benefit of (potentially) increasing the diversity of the student body through admitting a different composition of students. A simple sufficient condition for a priority coarsening to increase diversity is that the majority score distribution dominates the minority score distribution in the sense of first-order stochastic dominance. The trade-off between student exam scores and affirmative action is clear in the case with linear h. As v_1 decreases, students with lower scores gain admission, reducing the overall student quality. However, there is a benefit from increasing diversity if the ratio of underrepresented students with scores in $[v_1, v_2)$ is larger than the ratio in $[r, v_2)$. The total benefit from diversity then depends the difference of these ratios and the preference for diversity, which is measured by α. In particular, the mechanism designer is more likely to improve the allocation via exam design if there are more underrepresented students that are close to the no-affirmative-action cut-off r or she has a stronger preference for diversity.

4.6 Application: Income-Based Priorities and the Allocation of Public Housing

With more than 1 million housing units, and 1.6 million households on waiting lists, public housing programs in United States are very important for the welfare of low income households (Collinson, Ellen, and Ludwig, 2015). Because of this, many housing authorities employ various restrictions on the eligibility of applicants and

[40] The allocation in this case is such that students are allocated the school if their score exceeds r, and not otherwise. Plugging this allocation into Equation 161 yields Z^{NC}.

then allocate the units via lotteries to those deemed eligible. Concretely, to be eligible for public housing provided by the New York City Housing Authority (NYCHA), a household must have income less than a certain fraction of New York's Area Median Income (AMI).[41]

As in our previous applications, determining eligibility requirements for housing assistance is a contentious topic. The main trade-off for policymakers appears to be between making a larger part of society eligible and targeting households that will gain most from the assistance – those with the worst outside housing options. That it is desirable to ensure wide eligibility in this context is reflected in the words of NYC Mayor Bill de Blasio:[42]

> "Affordable housing initiatives cannot just be for the lowest income folks, ... There has to be a place for work force housing and middle-class housing as well."

However, the targeting of those with low outside options that would cause them to have inadequate housing in the absence of intervention is also extremely important. To this end, NYC officials have recently been developing laws that would reserve 15% of affordable-housing projects for the homeless (Stewart, Mays, and Haag, 2019).

Thus motivated, we study the problem of a designer who determines the priority of households as a function of their income. To model this market requires a significant departure from the previous analysis: we must consider the dynamic nature of public housing allocation. To this end, we develop a dynamic matching model to study how these trade-offs affect priority design in the allocation of public housing. We leverage our general results (Theorems 5 and 6) to place a simple three-income-tier structure on the optimal policy and deliver insights regarding the optimal design.

4.6.1 Model

There is a continuum of agents $i \in [0,1]$ that differ in their income $s \in [0,1]$, where $s = 1$ is the lowest level of income and $s = 0$ is the highest level of income, and their outside options $\kappa \in \mathcal{K}$. Income and outside options have joint distribution $F(\kappa, s)$ such that the marginal distribution of incomes is uniform. Time is discrete and infinite $t \in \mathbb{N}$. Agents have discount factors β and die at rate $(1-\delta)$. Each period $(1-\delta)$

[41] This fraction differs from development to development. Moreover, as Waldinger (2018) and Arnosti and Shi (2017) note, in some markets, there is also a minimum income level that applicants must clear. The reason for this is to make sure the applicant can pay the rent, which is a feasibility constraint rather than a policy with distributional motives. We abstract away from such issues in the present analysis.

[42] Stewart, Mays, and Haag (2019)

new agents are born from distribution F so that the population of agents always has unit measure.

A priority design is a vector $v = (v_1, v_2, \ldots, v_n)$ such that $0 \equiv v_0 \leq v_1 \leq \cdots \leq v_n \leq v_{n+1} \equiv 1$. Agents with $s \in [v_i, v_{i+1})$ for $i < n$ or $s \in [v_n, v_{n+1}]$ have the same priority as all other agents with incomes within the same interval prior to tie-breaking. A stock of measure Q of *ex-ante* identical houses are available. Agents are allocated to houses via a stable mechanism with tie-breaking within any tier to determine the order. When an agent is allocated a house, they receive $\tilde{v} \sim \Lambda$ per period they inhabit that house, where $\tilde{v} \in [v_{\min}, v_{\max}]$. Once an agent accepts a house, they inhabit that house until their death. Each period an agent goes unmatched, they receive their outside option.

Upon birth, agents are not allocated to public housing. Hence, their expected lifetime utility at birth is given by their value function when unmatched $V(\kappa, s, v)$. We suppose that the social planner has inequality-averse preferences given by the function:

$$Z(v) = \int_{\mathcal{K}} \int_0^1 \left(\frac{V(\kappa, s, v)}{1-\gamma} \right)^{1-\gamma} dF(\kappa, s) \tag{166}$$

where γ indexes the degree of inequality aversion. Of particular interest is the Rawlsian limit $\gamma \to \infty$ where the planner seeks to maximize the welfare of worst-off agent.

In Appendix D.5, we characterize the steady state of this dynamic matching model and derive a simple expression for the welfare of any priority design in terms of the steady state reservation values of each type (Proposition 49). Importantly, we can apply Theorems 5 and 6 directly in this steady state to show the following Corollary:

Corollary 12. *In the steady state of the dynamic matching model, there exists an optimal priority design with two cutoffs* $v^* = (v_1^*, v_2^*)$.

Proof. See Appendix D.1.8. □

This result greatly simplifies the analysis as we know that we need to specify at most two income cutoffs to find the optimal design.

4.6.2 Optimal Priority Design

Having understood the structure of the model, we now explore the problem of priority design from the perspective of the social planner. Given the nature of the fixed point equations for the equilibrium density of unmatched agents, finding an analytical characterization of the optimal cutoffs is challenging. Nevertheless, one can still establish interesting properties of the optimum and the trade-offs involved in the construction of optimal policy.

Given the system employed by NYCHA and elsewhere, it is of particular interest to study when we can rationalize a policy with the following feature: an income threshold below which agents have some probability of being allocated a house and above which they are ineligible. To this end, we show that when outside options are sufficiently increasing in income and the planner is sufficiently inequality averse that the richer agents are optimally excluded entirely from public housing (as in NYCHA). Conversely, if outside options are sufficiently similar across income groups, a sufficiently inequality averse social planner would like to give all agents positive probability of receiving public housing. Proposition 15 formalizes these statements:

Proposition 15. *Suppose that agents have outside options given by a decreasing and differentiable function of their underlying score, i.e. a strictly increasing function of their income, $\kappa = h(s)$. In the limit of $\gamma \to \infty$:*[43]

1. *If h is sufficiently steep, then an optimal policy features a threshold of income above which agents have zero probability of receiving public housing*

2. *If h is sufficiently flat, then an optimal policy gives all agents a positive probability of receiving public housing*

Proof. See Appendix D.1.9. □

This result highlights the key trade-off facing the planner between effective targeting and eligibility and shows how the strength of the relationship between outside options and incomes governs this trade-off. In particular, when h is very steep, those agents with the highest incomes also have relatively relatively high outside options. In this case, the poorest agents, even if they were to receive public housing in each period with certainty would have lower welfare than the richer agents. As a result, the targeting motive dominates the eligibility motive and the richer agents are excluded entirely from public housing. On the other other hand, when h is very flat, all agents have very similar outside options and excluding any agent from receiving public housing will give rise to them being worse off than all agents who have some chance at

[43] We say that a statement is true if h is sufficiently steep if there exists $\alpha < 0$ such that the statement is true whenever:

$$h'(s) < \alpha \quad \forall s \in (0, 1) \tag{167}$$

Likewise, we say that a statement is true if h is sufficiently flat if there exists $\alpha \leq 0$ such that the statement is true whenever:

$$h'(s) \geq \alpha \quad \forall s \in (0, 1) \tag{168}$$

public housing. To a sufficiently inequality-averse planner, this is unacceptable, and so the eligibility motive dominates the targeting motive.

The model moreover suggests that such an eligibility cutoff in income is optimal for similar reasons to those given by advocates in the New York public housing debate: the richest are sufficiently well off that we should reserve housing only for those who are needy. However, as Theorems 5 and 6 showed, the planner has additional latitude to introduce a tiered system with three tiers: one with unit probability of assignment, one with interior probability of assignment and one with zero probability of assignment. In Example 8 in Appendix D.5.1, we construct an explicit example of when this is desirable with three groups of agents: rich, middle class and poor agents. If poor agents have sufficiently low outside options relative to middle class agents who have sufficiently low outside options relative to rich agents, three priority tiers are strictly optimal. Intuitively, it is optimal to assign poor agents as soon as they are unmatched as their outside option is so bad (they may be homeless), while we wish to exclude rich agents from assignment altogether as in Proposition 15. Thus, when there are enough homes relative to poor agents, it is optimal to allocate the remaining homes via lottery to middle class agents.

4.7 Conclusion

Motivated by the clear design of priorities in centralized matching markets, we introduce and study the problem of optimal priority design subject to a constraint that an underlying score cannot be reversed. In our main results, we show that it is without loss of optimality for a mechanism designer to split agents into at most three indifference classes for each object (Theorem 5) and that an optimal policy exists (Theorem 6).

We apply these results and our framework to provide concrete insight into a number of important and widely studied centralized matching markets: BPS, CPS and NYCHA. In each case, we study the trade-offs highlighted by the relevant policymakers and provide normative insights as to the nature of optimal priority structures and positive rationalizations of the policies pursued in practice.

Chapter 5

Adaptive Priority Mechanisms

This chapter is jointly authored with Oğuzhan Çelebi.

5.1 Introduction

Authorities that allocate resources such as school seats, university places, and medical supplies often face conflicting objectives. On the one hand, they want to maximize match quality or appear fair by allocating resources to the highest-scoring agents according to various criteria such as academic attainment, mortality risk, or distance. On the other hand, they want to achieve diversity across a range of socioeconomic attributes including race, religion, and gender. Resolving this conflict is complicated, especially in new markets, due to uncertainty regarding the distribution of individuals' scores, characteristics, and preferences.

To balance these trade-offs, when the use of prices is seen as infeasible or unethical, authorities have broadly used two classes of policies: *quotas*,[1] where a certain portion of the resource is set aside for given groups; and *priorities*, where individuals in given groups receive higher scores. These policies have been applied across many different markets in many different countries, for example: the Indian government reserves some government jobs for disadvantaged groups; Chicago Public Schools employs quotas for students from different socioeconomic groups at its competitive exam schools; the University of California, Davis instituted a quota system for minority students; many countries gave differential priority to healthcare workers in the receipt of Covid-19 vaccines; church-run schools in the UK give explicit priority points to students from various religious groups; and the University of Michigan and the University of Texas have used different priority scales for minority students.

[1]We use *quota* as a general term that includes the widely used reserve policies (see Definition 12).

But what mechanism *should* such an authority use? Despite its revealed practical importance, we currently possess no formal understanding of this question. Thus, we do not know if (and under what circumstances) an authority should use a priority mechanism, a quota mechanism, or something else entirely.

In this paper, we formulate and solve the optimal mechanism design problem of an authority that allocates a resource to agents who are heterogeneous in their individual scores and belong to different groups. The authority cares about individuals' *scores*, through some aggregate index, and *diversity*, through the numbers of agents from different groups who are allocated the resource.[2] Moreover, they are uncertain about the market they face and have some beliefs about the joint distribution of scores and groups in the population.

We propose a new class of *adaptive priority mechanisms* (APM) that adjust agents' scores as a function of the number of assigned agents with the same characteristics and that allocate the resource to the set of agents with the highest adjusted scores. With a single authority, we derive an APM that is optimal, implements a unique outcome, and can be specified solely in terms of the *preferences* of the authority (*i.e.,* it is optimal regardless of their beliefs). By contrast, we show that priorities and quotas are optimal if and only if risk aversion over diversity is extremely low or high, respectively. Moreover, optimally set priority and quota policies depend on both the preferences and beliefs of the authority. Thus, the optimal APM both improves outcomes and requires less information. When there are many authorities, it is dominant for each of them to implement this APM and this leads to the unique stable outcome but generates inefficiency. To remedy this, we propose a centralized allocation mechanism, an *adaptive priority mechanism with quotas* (APM-Q), that restores efficiency. Finally, we benchmark the quantitative gains from APM using data from Chicago Public Schools and find that they are substantial.

Single-Authority Model We begin our analysis by studying a setting with a single authority that has some amount of a homogeneous resource (*e.g.,* seats at a school, medical resources) that it can allocate to a continuum of agents.[3] Agents

[2] This diversity preference can be interpreted more generally as encoding a preference of the authority over the composition of assigned agents across a range of attributes, *e.g.,* when allocating medical resources, the authority may care about ensuring that frontline medical workers are treated. Moreover, when scores represent individuals' property rights over objects (*e.g.,* higher-scoring students *deserve* better schools), we can interpret the preference for higher scores as a preference for procedural fairness. Whenever a lower-scoring agent obtains the resource while a higher-scoring agent does not, the latter agent has *justified envy* towards the former. In the two-sided matching literature, justified envy is often seen as inimical to fairness (see *e.g.,* Balinski and Sönmez, 1999).

[3] In Appendix E.3, we generalize our analysis and results to a setting with discrete agents.

differ in their scores (*e.g.,* exam score, clinical need) and discrete attributes (*e.g.,* socioeconomic status, if they are a frontline health worker). The authority cares separably about some index of the score distribution (*e.g.,* the average score) of those to whom it allocates the resource and the numbers of agents from different groups. As a result, the preferences of the authority over agents depend on the joint distribution of agents' scores and groups. We assume that this distribution is potentially unknown and varies arbitrarily across states of the world. The authority's problem is to design an allocation mechanism that is optimal regardless of their beliefs, a property that we call *first-best optimality.*

Adaptive Priority Mechanisms To this end, we introduce the class of adaptive priority mechanisms (APM), which proceed in two steps. First, each agent is given an *adaptive priority* that is a function of their own score and the number of agents from the same group to whom the resource is assigned. Second, APM allocate the resource to agents in order of adaptive priorities, subject to fully allocating the available amount. This class of mechanisms allows the implicit preference for agents from different groups to depend upon the ultimate allocation. The allocation under an APM is defined as the fixed point of the above operation: an allocation is implemented by APM if the adaptive priority of all agents who are allocated the resource (evaluated at the allocation) is higher than those who are not allocated the resource. When an agent's adaptive priority is increasing in their own score and decreasing in the number of agents with the same attributes that are assigned the resource – a property we call *monotonicity* – the APM implements a unique allocation. Moreover, this allocation can be computed greedily by prioritizing agents according to their adaptive priority, evaluated at the number of higher-scoring agents in their group.

Most importantly, we derive a particular, monotone APM that is first-best optimal. Under this optimal APM, an agent's priority is equal to the contribution of their own score plus their marginal contribution to diversity utility. Intuitively, this mechanism equates the benefits and costs of allocating to the marginal agent, regardless of the ultimate joint distribution of agents' scores and groups. Moreover, this APM can be described solely as a function of the authority's preferences, without any reference to its beliefs.

(Sub)Optimality of Priorities and Quotas We next establish that priority and quota mechanisms are generally dominated by APM. We do so by characterizing the conditions on the preferences of the authority such that priorities and quotas attain first-best optimality. Concretely, we find that priorities and quotas are first-best optimal if and only if (i) the authority is risk-neutral over diversity, in which case priorities are optimal, or (ii) the authority is extremely risk-averse over diversity, in which case quotas are optimal. Hence, outside of extreme cases, APM deliver strict improvements relative to the *status quo*.

A Price-Theoretic Intuition To both illustrate and develop the intuition behind these results, we study a detailed example that allows for a closed-form comparison of priorities, quotas, and the optimal adaptive priority mechanism. We do this in the spirit of the seminal analysis of Weitzman (1974), who compares price and quantity regulation in product markets. In the example, the resource corresponds to seats at a school and there are two groups of students (minority and majority students). The authority is uncertain over the relative scores of minority and majority students, and has linear-quadratic preferences over the scores of admitted students and the number of minority students admitted to the school.

The preference of the authority between priority and quota mechanisms is governed by its risk aversion over the number of admitted minority students: there is a cutoff value such that quotas are preferred when risk aversion exceeds this threshold and priorities are otherwise preferred. On the one hand, by mandating a minimal level of minority admissions, quotas *guarantee* a level of diversity. On the other hand, as relatively more minority students receive the resource in the states in which they have relatively higher scores, priorities *positively select* minority students. Adaptive priority mechanisms optimally exploit the guarantee effects of quotas and the positive selection effects of priorities, and are always optimal.

Dominance and Stability with Multiple Authorities While the single-authority model is relevant for studying settings with a single resource, in many markets there are multiple authorities who control heterogeneous resources (*e.g.,* school seats) over which agents have heterogeneous preferences. We generalize our analysis to this setting and show that APM arise under both cooperative (stability) and non-cooperative (dominant-strategy equilibrium) solution concepts. Concretely, we show that there is a unique stable allocation and a mechanism is consistent with stability if and only if it coincides with the single-authority-optimal APM. Moreover, when authorities sequentially admit agents, each authority using its single-authority-optimal APM is a dominant strategy and implements the unique stable matching. Thus, one could advise authorities to use APMs with confidence that outcomes will be stable and that they could do no better under any alternative mechanism.

Inefficiency of APM and an Efficient Multi-Authority Mechanism However, decentralized outcomes under APMs are generally inefficient for the authorities. This is because authorities do not internalize the "pecuniary externalities" they generate by over-admitting agents that have a preference for them but that the other authorities value more. To remedy this, we propose a centralized allocation mechanism, an *Adaptive Priority Mechanism with Quotas* (APM-Q). An APM-Q first constructs a fictitious aggregate authority, decides the aggregate levels of admissions of each group according to an optimal APM, and then allocates these groups across authorities according to optimally set quotas. This mechanism fixes the pecuniary externalities by creating a pseudo-market in which each group has a "price" and authorities are assigned agents only if they would be willing to "pay" for them.

Benchmarking the Gains from APM in Chicago Exam Schools Finally, we benchmark the improvements from APM using application and admission data from 2013-2017 on the selective exam schools of Chicago Public Schools (CPS), a setting also empirically studied by Angrist, Pathak, and Zárate (2019) and Ellison and Pathak (2021). CPS uses a reserve system to increase the admissions of underrepresented groups. In this system, as we later detail, academic scores and the socioeconomic tiers of the census tracts in which students live determine the schools that students can attend. Estimating preference parameters to best rationalize the pursued reserve policy, we find that the gains from using the optimal APM are equivalent to eliminating 37.5% of the loss to CPS' payoffs from failing to admit a diverse class of students. This gain is 2.3 times larger than the estimated gain from a 2012 policy change that increased the size of all reserves. This exercise shows both that APM could be practically implemented and that the gains from so doing may be considerable.

Related Literature The market design literature has largely studied the comparative statics and axiomatic foundations of mechanisms. In this context, our paper relates to the literature on matching with affirmative action concerns initiated by Abdulkadiroğlu and Sönmez (2003) and Abdulkadiroğlu (2005). For example, in the study of quotas, Kojima (2012) shows how affirmative action policies that place an upper bound on the enrollment of non-minority students may hurt all students, Hafalir, Yenmez, and Yildirim (2013) introduce the alternative and more efficient minority reserve policies, Ehlers, Hafalir, Yenmez, and Yildirim (2014) generalize reserves to accommodate policies that have floors and ceilings for minority admissions, and Doğan (2016) shows that stronger affirmative action can (weakly) harm all minority students under reserve policies and proposes a new rule that fixes this issue. The quota policies studied in this paper are a special case of the slot-specific priorities introduced in Kominers and Sönmez (2016). Further related papers study quota policies in university admissions in India (Aygün and Turhan, 2020; Sönmez and Yenmez, 2022a,b), in Germany (Westkamp, 2013) and in Brazil (Aygun and Bó, 2021). Kamada and Kojima (2017, 2018) and Goto, Kojima, Kurata, Tamura, and Yokoo (2017) study stability and efficiency in more general matching-with-constraints models. Echenique and Yenmez (2015) characterize a class of substitutable choice rules under diversity preferences and Erdil and Kumano (2019b) study tie-breaking rules under substitutable priorities under stable matching mechanisms and distributional constraints. Çelebi (2022) studies when affirmative action policies, including quota policies, can be rationalized by diversity preferences.

In this paper, we instead pursue the methodological approach of mechanism design and welfare economics by analyzing optimal mechanisms from the perspective of an authority with some given preferences over allocations. Chan and Eyster (2003) share this perspective in their analysis of the costs and benefits of banning affirmative action.[4] In this vein, we have previously analyzed the narrower problem of how to optimally coarsen agents' scores into priorities (Çelebi and Flynn, 2022) in a continuum matching market framework in the style of Abdulkadiroğlu, Che, and Yasuda (2015) and Azevedo and Leshno (2016). This analysis nevertheless restricted authorities to use a priority mechanism that does not consider agents' characteristics and implement only allocations that are stable with respect to these priorities. Thus, our focus on comparing priorities, quotas, and optimal mechanisms distinguishes our analysis from our prior work and the previous literature, which study the properties of each policy in isolation and without an explicit treatment of uncertainty.

Outline Section 5.2 exemplifies our main results. Section 5.3 studies optimal mechanisms with a single authority. Section 5.4 studies equilibrium mechanisms with many authorities. Section 5.5 analyzes efficient mechanisms with multiple authorities. Section 5.6 quantifies the gains from APM using data from Chicago Public Schools. Section 5.7 concludes.

5.2 Comparing Mechanisms: An Example

The Setting A single school has capacity q. Students are of unit total measure, have scores in $[0, 1]$, and are either minority or majority students. The authority has

[4] Other analyses of this issue include Epple, Romano, and Sieg (2008) and Temnyalov (2021).

linear-quadratic preferences $\xi : \mathbb{R}^2 \to \mathbb{R}$ over students' total scores $\bar{s}$ and the measure of admitted minority students x:

$$\xi(\bar{s}, x) = \bar{s} + \gamma\left(x - \frac{\beta}{2}x^2\right) \tag{169}$$

where $\gamma \geq 0$ indexes their general concern for admitting minority students relative to ensuring high scores and $\beta \geq 0$ indexes the degree of risk aversion regarding the measure of admitted minority students.

The minority students are of measure κ and have scores that are uniform over $[0, 1]$. The majority students are of measure $1 - \kappa$ and all have common underlying score $\omega \in [\underline{\omega}, \overline{\omega}] \subseteq [0, 1]$ with distribution Λ. The score of the majority students, ω, parameterizes how well the majority students score relative to the minority students. Finally, we assume that the affirmative action preference is neither too small nor too large with the following: $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$, $\kappa(1 - \underline{\omega}) < \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta}$. These conditions ensure that optimal affirmative action policies will neither be so large as to award all slots to minority students in some states nor so small that there is no affirmative action in some states.

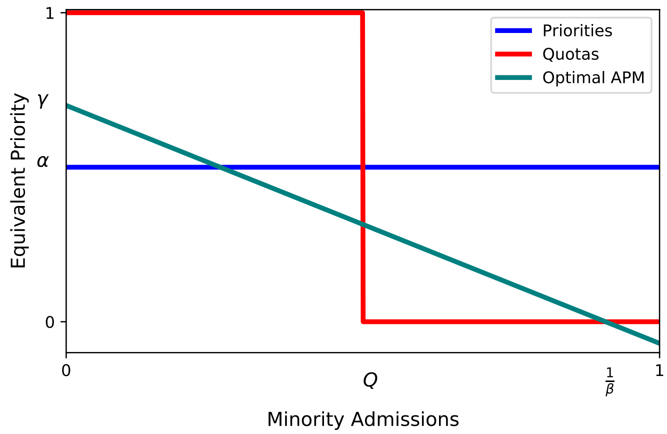
The authority can implement an APM, a priority mechanism, or a quota mechanism. An APM awards a score boost of $A(y)$ to a minority student when measure y other minority students are admitted, and allocates the seats to the highest-scoring minority students who have transformed scores higher than ω.[5] An (additive) priority mechanism $\alpha \in \mathbb{R}_+$ increases uniformly the scores of minority students: the score used in admissions becomes uniform over $[\alpha, 1 + \alpha]$. The authority then admits the highest-scoring measure q students. A quota policy $Q \in [0, \min\{\kappa, q\}]$ sets aside measure Q of the capacity for the minority students. The measure Q highest-scoring minority students are first allocated to quota slots, and all other agents are then admitted to the residual $q - Q$ places according to the underlying score.[6]

We illustrate how these three policies prioritize minority students in Figure 5-1. Priority mechanisms award a constant score boost of α. Quota mechanisms give enough points to always ensure admission until measure Q is reached and then give no advantage. APM allow any pattern of prioritization as a function of minority admissions (we plot only the optimal APM, which turns out to be linear in this context).

[5] Formally, this happens when $s(x(\omega)) + A(x(\omega)) = \omega$, where $s(x(\omega))$ denotes the score of the marginal minority student when the highest-scoring $x(\omega)$ minority students are admitted.

[6] This corresponds to a precedence order that processes quota slots first. We discuss the importance of precedence orders in Section 5.2.1 and in Appendix E.2.3.

Figure 5-1: How Priorities, Quotas, and APM Prioritize Minority Students



Notes: Illustration of the equivalent priority given to a minority student as a function of the measure of admitted minority students under: the optimal APM (see Proposition 16), a priority mechanism α, and a quota mechanism Q.

Comparing Mechanisms Let the authority's expected utility be V^* under any optimal (expected utility maximizing) mechanism, V_A under an optimal adaptive priority mechanism, V_P under an optimal priority mechanism, and V_Q under an optimal quota mechanism. The following proposition characterizes the relationships between these mechanisms:

Proposition 16. *The following statements are true:*

1. *The APM $A(y) = \gamma(1 - \beta y)$ is optimal, $V^* = V_A$*

2. *The comparative advantage of priorities over quotas is given by:*

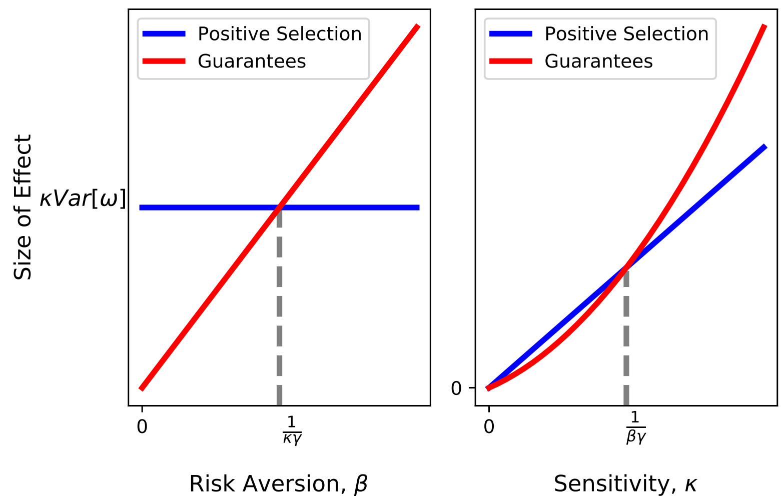
$$\Delta \equiv V_P - V_Q = \frac{\kappa}{2}\left(1 - \kappa\gamma\beta\right)\mathrm{Var}[\omega] \tag{170}$$

3. *The comparative advantage of APM over priorities and quotas is given by:*

$$\Delta^* \equiv \min\{V^* - V_P, V^* - V_Q\} = \begin{cases} \frac{1}{2}\left(\kappa\gamma\beta\right)^2 \frac{\kappa \mathrm{Var}[\omega]}{1+\kappa\gamma\beta}, & \kappa\gamma\beta \leq 1, \\ \frac{1}{2}\frac{\kappa \mathrm{Var}[\omega]}{1+\kappa\gamma\beta}, & \kappa\gamma\beta > 1. \end{cases} \tag{171}$$

The proofs of all results are in Appendix E.1. We now develop intuition for the comparative advantage of priorities over quotas. First, observe that a quota of Q admits measure Q minority students in all states of the world under our assumptions.

Figure 5-2: Comparative Statics for the Positive Selection and Guarantee Effects



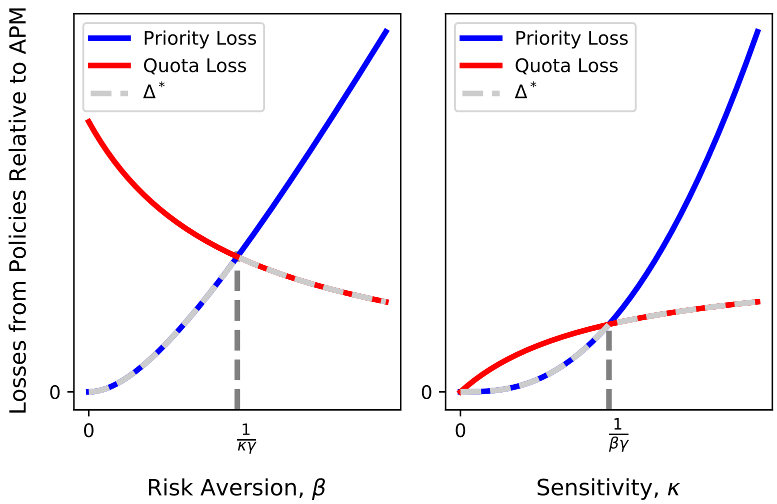
Notes: Illustration of the comparative statics for the trade-offs between priority and quota mechanisms. Positive Selection plots the positive selection effect, $\kappa\mathrm{Var}[\omega]$, and Guarantee plots the guarantee effect, $\frac{\kappa}{2}\left(1+\kappa\gamma\beta\right)\mathrm{Var}[\omega]$. As per Equation 170 in Proposition 16, priorities dominate quotas if and only if $1 \geq \kappa\gamma\beta$, where the point of indifference is denoted by the dashed grey line.

However, a priority policy induces variability in the measure of admitted minority students across states of the world. We call the gain to quota policies in eliminating this variation the *guarantee effect* and find mathematically that it is equal to $\frac{\kappa}{2}\left(1+\kappa\gamma\beta\right)\mathrm{Var}[\omega]$ in payoff terms.

Second, the optimal priority policy admits more minority students when minority students score relatively well and fewer when minority students score relatively poorly. To demonstrate this, we show that minority admissions in state ω under the optimal priority policy are $x(\alpha,\omega) = \bar{x}(\alpha) + \varepsilon(\omega)$ where $\bar{x}(\alpha) = \kappa(1+\alpha-\mathbb{E}[\omega])$ and $\varepsilon(\omega) = \kappa\left(\mathbb{E}[\omega]-\omega\right)$. Thus, in the states where minority students score relatively better ($\omega < \mathbb{E}[\omega]$), we have that $\varepsilon(\omega) > 0$ and $x(\alpha,\omega) > \bar{x}(\alpha)$. We call this effect the *positive selection* effect and find that this benefits a priority policy by $-\mathrm{Cov}[\omega,\varepsilon(\omega)] = \kappa\mathrm{Var}[\omega]$ in payoff terms.

The ultimate preference between priority and quota mechanisms is determined by which of the guarantee and positive selection effects dominates. This is itself determined by the extent to which the authority values diversity γ, the risk preferences of the authority β, and the measure of minority students κ. We illustrate how risk aversion and the measure of minority students affect the sizes of the positive selection

Figure 5-3: Comparative Statics for the Losses from Priorities and Quotas



Notes: Illustration of the comparative statics for the losses from optimal priority and quota policies relative to the optimal APM (as presented in Equation 171 in Proposition 16). The lower envelope of the losses, Δ^*, corresponds to the comparative advantage of the optimal APM over priorities and quotas. The point of indifference between priorities and quotas is denoted by the dashed grey line.

and guarantee effects in Figure 5-2. If the authority is close enough to risk-neutral (*i.e.*, $\frac{1}{\kappa\gamma} > \beta$), then priorities are strictly preferred as positive selection dominates guarantees. If the authority is sufficiently risk-averse (*i.e.*, $\frac{1}{\kappa\gamma} < \beta$), then quotas are strictly preferred as the guarantee effects dominate positive selection. The threshold for risk aversion scales inversely with the measure of minority students κ. Because minority students' scores are uniform, κ corresponds to the density of minority students' scores. Hence, the change in minority admissions from a small change in their priority equals κ. Thus, κ indexes the *sensitivity* of minority admissions to the state under priority policies. As a result, higher κ favors quota policies by increasing the magnitude of the guarantee effect relative to the positive selection effect. Finally, the extent of uncertainty $\mathrm{Var}[\omega]$ may intensify an underlying preference but never determines which regime is preferred.

An APM is optimal and overcomes the limitations posed by both priorities and quotas. In this case, the optimal APM is linear in the measure of admitted minority students, with slope given by the authority's risk aversion over minority admissions, awarding each minority student a subsidy equivalent to their marginal contribution to the diversity preferences of the authority. This allows the adaptive priorities to optimally balance the positive selection and guarantee effects, and implement the first-best allocation in every state. In Figure 5-3, we illustrate how the losses from priority mechanisms and quota mechanisms vary with risk aversion and sensitivity. As risk aversion moves, the loss from priority and quota policies relative to the optimum is greatest when the authority is indifferent between the two regimes. The loss from restricting to priority or quota policies is zero when the authority is risk-neutral or there is no uncertainty regarding relative scores, and decreases as the authority becomes extremely risk-averse. As sensitivity increases, the scope for affirmative action increases and so the gains from APM also increase. Thus, we should expect there to be large gains from switching to APM precisely when authorities have intermediate levels of risk aversion and/or the scope for implementing affirmative action is significant.

Finally, optimal APM have an advantage with respect to priority and quota mechanisms that we have not yet highlighted: they depend only on the authority's preferences, γ and β, and not their beliefs about ω, Λ. This contrasts with the optimal priority and quota policies, which depend on Λ.[7] As a result, APM improve outcomes while using *less* information.

5.2.1 Discussion

Before moving to the general analysis, we discuss three additional findings that emphasize the broader economics and scope of these results.

A Price-Theoretic Intuition This comparison of *priorities vs. quotas* echoes the comparison of *prices vs. quantitities* by Weitzman (1974). We show in Appendix E.2.1 that there is a formal mapping between the two.[8] Intuitively, the positive selection effect is equivalent to the effect that price regulation gives rise to the greatest production in states where the firm's marginal cost is lowest. Moreover, the guarantee effect is equivalent to the ability of quantity regulation to stabilize the level of production. An APM corresponds in the Weitzman (1974) setting to a regulator setting neither a price nor a quantity, but completely specifying the optimal demand curve. As in Weitzman's analysis, this allows the authority to implement the optimal allocation, regardless of the firm's realized marginal costs. Thus, in this context, the comparison of mechanisms for allocating goods without prices boils down to similar

[7] The optimal quota policy is given by $Q^* = \frac{1+\gamma-\mathbb{E}[\omega]}{\frac{1}{\kappa}+\gamma\beta}$, while the optimal priority policy sets the expected measure of minorities to Q^*. The policies depend on Λ through $\mathbb{E}[\omega]$.

[8] Mapping Weitzman's curvature of production costs $C''^{-1} \mapsto \kappa$, curvature of benefits to consumers $B'' \mapsto -\gamma\beta_m$, and uncertainty over marginal costs $\mathrm{Var}[\alpha(\theta)] \mapsto \mathrm{Var}[\omega]$, we have that Weitzman's Δ coincides with our own. See Appendix E.2.1 for more details.

trade-offs between well-understood price-based mechanisms for goods allocation.

Medical Resource Allocation In Appendix E.2.2, we apply this model to understand the trade-offs between priority and quotas in the context of medical resource allocation. This topic received enormous attention during the Covid-19 pandemic (see *e.g.,* Pathak, Sönmez, Unver, and Yenmez, 2021). Our analysis provides a formal justification for the idea that priorities may lose out relative to quotas from ignoring some groups or ethical values in the allocation of scarce resources (the guarantee effect). However, we also uncover a benefit of priorities that was not previously understood: they induce positive selection. Thus, if we care mostly about treating the neediest (β is low), priorities may yet be optimal.[9]

Optimal Precedence Orders Thus far we have modelled quotas by first allocating minority students to quota slots and then allocating all remaining students according to the underlying score. However, we could have done the opposite. The orders in which quotas are processed are called *precedence orders* in the matching literature and their importance has been the subject of a growing literature (see *e.g.,* Dur, Kominers, Pathak, and Sönmez, 2018; Dur, Pathak, and Sönmez, 2020; Pathak, Rees-Jones, and Sönmez, 2020). In Corollary 17 in Appendix E.2.3, we show that processing quotas second is equivalent to using a priority policy in this setting. Thus, processing quotas first is better than processing them second if and only if $1 \leq \kappa\gamma\beta$. We emphasize that this equivalence merely illustrates the similarity between priority policies and processing quotas second and does not hold in the more general model we study in the remainder of the paper. This notwithstanding, the main aspect of this conclusion is robust: in Theorem 8, we show that for any quota policy to be optimal in the presence of uncertainty, it must process quotas first. In the absence of uncertainty, in Appendix E.5, we show that every quota mechanism is equivalent to a priority mechanism (and *vice versa*) and use this fact to quantify the impact of changes in precedence orders for US H1-B visa allocation.

5.3 Optimal Mechanisms with a Single Authority

We begin our general analysis by studying the resource allocation problem of a single authority. In this context, we define APM and derive an optimal APM that attains the

[9]Inspired by the uncertainty over the number of medical professionals and the general population who were expected to need scarce medical resources at the onset of the Covid-19 pandemic, we consider an extension where there is also uncertainty over the number of medical professionals who get sick, κ. We show that an increase in the uncertainty regarding the need of frontline workers $\mathrm{Var}[\kappa]$ leads to a greater preference for quotas. This can be seen diagrammatically in Figure 5-2 as the guarantee effect is convex in κ.

first-best. We moreover provide necessary and sufficient conditions for the optimality of the ubiquitous priority and quota mechanisms and find that they are extremely restrictive, implying that there are likely gains from switching to APM.

5.3.1 Model

An authority allocates a single resource of measure $q \in (0, 1)$ to a unit measure of agents. Agents differ in their type $\theta \in \Theta = [0, 1] \times \mathcal{M}$ comprising their scores $s \in [0, 1]$ and the group to which they belong, $m \in \mathcal{M}$, where their score denotes their suitability for the resource and $\mathcal{M}$ is a finite set comprising potential attributes such as race, gender, or socioeconomic status. We denote the score and group of any type θ by $s(\theta)$ and $m(\theta)$, respectively. The true distribution of types is unknown to the authority. The authority's uncertainty is parameterized by $\omega \in \Omega$, where Ω is the set of all distributions over Θ that admit a density. The authority believes that ω has distribution $\Lambda \in \Delta(\Omega)$. In state of the world ω, we denote the type distribution by F_ω with density f_ω.[10] In Appendix E.3, we generalize our analysis and results to the discrete context.

An allocation $\mu : \Theta \to \{0, 1\}$ specifies for any type $\theta \in \Theta$ whether they are assigned to the resource.[11] Two allocations μ and μ' are *essentially the same* if they coincide up to a measure zero set. The set of possible allocations is $\mathcal{U}$. An allocation is feasible if it allocates no more than measure q of the resource. A mechanism is a function $\phi : \Omega \to \mathcal{U}$ that returns a feasible allocation for any possible distribution of types. A mechanism ϕ implements an *essentially unique* allocation if all allocations implemented by ϕ are essentially the same.

As motivated, authorities often have preferences over scores and diversity. To model this, we define the aggregate score index of any allocation as:

$$\bar{s}_h(\mu, \omega) = \int_\Theta \mu(s, m) h(s) \mathrm{d}F_\omega(s, m) \tag{172}$$

for some continuous, strictly increasing function $h : [0, 1] \to \mathbb{R}_+$, which determines the extent to which the authority values agents with higher scores. To capture diversity, we compute the measure of agents of each group allocated the resource $x(\mu, \omega) = \{x_m(\mu, \omega)\}_{m \in \mathcal{M}}$ as:

$$x_m(\mu, \omega) = \int_{[0,1]} \mu(s, m) f_\omega(s, m) \mathrm{d}s \tag{173}$$

[10] Formally, we mean that $f_\omega(s, m) = \frac{\partial}{\partial s} F_\omega(s, m)$ exists for all $s \in [0, 1]$ and $m \in \mathcal{M}$.

[11] Formally, μ is a measurable function with respect to the Borel $\sigma-$algebra of the product topology in Θ.

To separate the roles of scores and diversity, we impose that their utility function over these dimensions $\xi : \mathbb{R}^{|\mathcal{M}|+1} \to \mathbb{R}$ satisfies the following separability assumption:[12]

Assumption 10. *The authority's utility function can be represented as:*

$$\xi\left(\bar{s}_h, x\right) \equiv g\left(\bar{s}_h + \sum_{m\in\mathcal{M}} u_m(x_m)\right) \tag{174}$$

for some continuous, strictly increasing function $g : \mathbb{R} \to \mathbb{R}$ and differentiable and concave functions $u_m : \mathbb{R} \to \mathbb{R}$ for all $m \in \mathcal{M}$.

We also assume that the authority always prefers to allocate the entire resource.[13] The preference of the authority is a monotone transformation of a quasi-linear utility index comprised of scores and a diversity preference. Intuitively, u_m determines the preference for assigned agents of group m, with its concavity following from a preference for diversity.[14] The function g determines their risk preferences over their utility over scores and diversity across states of the world.

As we later show, this assumption allows for particularly simple functional forms for optimal mechanisms. We explore the robustness of our results to relaxing this assumption in Appendix E.4. We show that our results are essentially unchanged when preferences are non-separable over diversity, *i.e.*, when $\sum_{m\in\mathcal{M}} u_m(x_m)$ is replaced with $u(x)$. The essential assumption for our results is the separability of diversity and score preferences. When this fails, it is no longer possible to specify optimal mechanisms without explicitly conditioning the allocation on the realized distribution of agents. As a result, our analysis will not apply to situations in which there is complementarity or substitutability in preferences over match quality and diversity.

We define the value of a mechanism ϕ under distribution Λ as the authority's

[12] The assumption of separable preferences over scores and diversity is common in the literature on affirmative action concerns (see e.g., Athey, Avery, and Zemsky, 2000; Chan and Eyster, 2003; Ellison and Pathak, 2021).

[13] A necessary and sufficient condition for this is: $h(0) + u'_m(q) \geq 0$ for all $m \in \mathcal{M}$. This condition is clearly sufficient, the lowest utility the authority can get from allocating the resource is always positive. It is also necessary: if $h(0) + u'_m(q) < 0$ for some m, in the state of the world where there are only measure q of group m agents with uniform score distribution, the authority would prefer not to allocate a portion of the resource to the lowest-scoring agents.

[14] Note that u_m depends on m, so our specification allows the designer to have different preferences for allocating the resource to agents from different groups. For example, this allows for a designer with affirmative action motives who prefers to assign the resource to some particular group m: $u'_m(x) > u'_{m'}(x)$ for all x or a designer who prefers a balanced composition of allocated agents: $u'_m(x) = u'_{m'}(x)$ for all $m \in \mathcal{M}$.

expected utility of the allocations induced by that mechanism:

$$\Xi(\phi, \Lambda) = \int_{\Omega} \xi(\bar{s}_h(\phi(\omega), \omega), x(\phi(\omega), \omega)) \mathrm{d}\Lambda(\omega) \tag{175}$$

We say that a mechanism is first-best optimal if it maximizes the authority's expected utility for all possible *distributions of* distributions of agents' characteristics.

Definition 9 (First-Best Optimality)**.** *A mechanism ϕ^* is first-best optimal if:*

$$\Xi(\phi^*, \Lambda) = \sup_{\phi} \Xi(\phi, \Lambda) \tag{176}$$

for all $\Lambda \in \Delta(\Omega)$.

This is a demanding property for a mechanism to possess. Indeed, as the example from Section 5.2 shows, priority and quota mechanisms can fail to be first-best optimal while APM can attain first-best optimality. In the remainder of this section, we formally define APM, show that (when suitably designed) they are first-best optimal, and characterize the conditions under which priorities and quotas are first-best optimal.

5.3.2 Adaptive Priority Mechanisms

Toward deriving a first-best optimal mechanism, we introduce APMs. To this end, we first introduce an *adaptive priority policy* $A = \{A_m\}_{m \in \mathcal{M}}$, where $A_m : \mathbb{R} \times [0, 1] \to \mathbb{R}$. The adaptive priority policy assigns priority $A_m(y_m, s)$ to an agent with score s in group m when measure y_m of agents of the same group is allocated the object. Given an adaptive priority policy, an APM implements allocations in the following way:

Definition 10 (Adaptive Priority Mechanism)**.** *An adaptive priority mechanism, induced by an adaptive priority A, implements an allocation μ in state ω if the following are satisfied:*

1. *Allocations are in order of priorities: $\mu(\theta) = 1$ if and only if for all θ' with $\mu(\theta') = 0$, we have that:*

$$A_{m(\theta)}(x_{m(\theta)}(\mu, \omega), s(\theta)) > A_{m(\theta')}(x_{m(\theta')}(\mu, \omega), s(\theta')) \tag{177}$$

2. *The resource is fully allocated:*

$$\sum_{m \in \mathcal{M}} x_m(\mu, \omega) = q \tag{178}$$

With some abuse of terminology, we will often refer to an APM as the adaptive priority A that induces it. By way of illustration, we provide a simple example of the flexibility of APM to act like a hybrid of priority and quota policies.

Example 3. *Let $\mathcal{M} = \{m, n\}$. Both groups have measure 0.5 and capacity is $q = 0.5$. We consider the adaptive priority policy $A = \{A_m, A_n\}$ given by:*

$$A_m(x,s) = s, \qquad A_n(x,s) = \begin{cases} s+1 & \text{if } x \leq 0.1 \\ s+0.1 & \text{if } x \in (0.1, 0.25) \\ s & \text{if } x \geq 0.25 \end{cases} \tag{179}$$

This leaves the score of group m agents unchanged and gives agents of group n a score boost of: 1 if less than measure 0.1 group n agents is assigned, 0.1 if between measure 0.1 and 0.25 group n agents is assigned, and no score boost at all if measure greater than 0.25 group n agents is assigned.

To understand the properties of this adaptive priority policy, observe that the highest-scoring measure 0.1 group n agents is guaranteed the resource, even in states where they score badly. Therefore, A_n practically embeds a quota of 0.1. For admissions levels between 0.1 and 0.25, the APM acts like a priority policy and boosts the scores of group n agents by 0.1, increasing the admissions of group n when group n agents score moderately well. For admissions levels beyond 0.25, group n agents are given no further advantage. Thus, when diversity is attained, this APM "phases out" and no longer advantages any group.

At this point, we have not established that a given APM implements any allocation at all, or that it implements a unique allocation. However, there is a natural subclass of APM that do implement a unique allocation: those that are monotone. An APM A is *monotone* when (i) $A_m(\cdot, s)$ is a decreasing function for all $m \in \mathcal{M}, s \in [0,1]$ and (ii) $A_m(y_m, \cdot)$ is a strictly increasing function for all $m \in \mathcal{M}, y_m \in \mathbb{R}$.[15]

Proposition 17. *Any Monotone APM A implements an essentially unique allocation.*

Moreover, the unique outcome of a monotone APM can be implemented "greedily:"

Algorithm 1 (Greedy Algorithm for Implementation of APM)**.** *The greedy APM algorithm proceeds in the following four steps:*

[15] Observe that monotone adaptive priority mechanisms are fair in the sense that they preserve the ranking of agents within any group and assign higher priority to an agent whenever there are fewer agents from her group who are allocated the resource.

1. *For each θ, define*

$$\overline{x}(\theta) = \int_{s(\theta)}^{1} f_\omega(s, m(\theta))\mathrm{d}s \tag{180}$$

as the measure of agents who have higher scores than θ and belong to same group.

2. *Construct a ranking of the agents as*

$$R(\theta) = A_{m(\theta)}(\overline{x}(\theta), s(\theta)) \tag{181}$$

3. *Define the cutoff ranking for the agents as $\overline{R}$ by*

$$\int_{\Theta} \mathbb{I}\{R(\theta) \geq \overline{R}\}\mathrm{d}F_\omega(\theta) = q \tag{182}$$

4. *Allocate the resource to all θ with $R(\theta) \geq \overline{R}$.*

Intuitively, this algorithm works by ranking all agents by their score within each group m and assigning agents in order of their transformed scores evaluated at the measure of *already assigned* agents of the same group, conditional on their admission. Informally, the algorithm greedily moves down the ranking of agents until the resource is exhausted.

5.3.3 Adaptive Priority Mechanisms Achieve the First-Best

Having shown that monotone APM implement a unique allocation and provided an algorithm to compute this allocation, we now show that a certain, monotone APM is first-best optimal:

Theorem 7. *The following APM is monotone and first-best-optimal:*

$$A^*_m(y_m, s) \equiv h^{-1}(h(s) + u'_m(y_m)) \tag{183}$$

Moreover, if a mechanism is first-best-optimal, then it implements essentially the same allocations as A^*.

Observe that A^* is not only uniquely first-best optimal, it also requires only that the authority knows its preferences over scores h and diversity u_m. Importantly, it need have no knowledge of the underlying distribution of agents and can be fully specified even without any knowledge of the nature or extent of uncertainty, Λ. Moreover, this mechanism does not depend at all on the authority's across-state risk preferences,

g. This is because it achieves the *ex post* optimal allocation in all states and so there is no need to trade-off gains and losses across states (the preferences over which are exactly determined by g).

To gain intuition for the form of this mechanism, suppose that the authority has linear utility over scores $h(s) \equiv s$. In this case, $A^*_m(y_m, s) = s + u'_m(y_m)$, so an agent in group m is awarded a boost of $u'_m(y_m)$ when there are y_m higher-scoring agents of the same group, their direct marginal contribution to the diversity preferences of the authority. This is optimal because this boost precisely trades off the marginal benefit of additional diversity with the marginal costs of reduced scores. Moreover, failing to award this precise level of boost would result in a suboptimal allocation. Thus, any optimal mechanisms must be essentially identical to the optimal APM we have characterized. To generalize this beyond linear utility of scores, consider the following observation: we can map agents' scores from s to $h(s)$, and consider the optimal boost in this space. As h is monotone, this preserves the ordinal structure of the optimal allocation, and the authority has linear preferences over $h(s)$. Thus, in this transformed space, the optimal boost remains additive and given by $u'_m(y_m)$. To find the optimal transformed score in the original space, we simply invert the transformation h and apply it to the optimal score in the transformed space, yielding the formula for the optimal mechanism in Theorem 7.

5.3.4 (Sub)Optimality of Priorities and Quotas

We have shown that APM are optimal. However, the primary classes of mechanisms that have been used in practice are priority and quota mechanisms. Therefore, it is important to understand whether (and when) these mechanisms are also optimal. We now establish that APM generally provide a strict improvement over priority and quota mechanisms.

We first formally define priority and quota mechanisms. A *priority policy* $P : \Theta \to [0,1]$ awards an agent of type $\theta \in \Theta$ a priority $P(\theta)$.

Definition 11 (Priority Mechanisms)**.** *A priority mechanism, induced by a priority policy P, allocates the resource in order of priorities until measure q has been allocated, with ties broken uniformly and at random.*

We define a *quota policy* as (Q, D), where $Q = \{Q_m\}_{m \in \mathcal{M}}$ and $D : \mathcal{M} \cup \{R\} \to \{1, 2, \ldots, |\mathcal{M}| + 1\}$ is a bijection. The vector Q reserves measure of the capacity Q_m for agents in group m, with residual capacity $Q_R = q - \sum_{m \in \mathcal{M}} Q_m$ open to agents of all types. The bijection D (often called the precedence order) determines the order in which the groups are processed.

Definition 12 (Quota Mechanisms). *A quota mechanism, induced by a quota policy (Q, D), proceeds by allocating the measure $Q_{D^{-1}(k)}$ agents from group $D^{-1}(k)$ (if there are sufficient agents from this group) to the resource in ascending order of k, and in descending order of score within each k. If there are insufficiently many agents of any group to fill the quota, the residual capacity is allocated to a final round in which all agents are eligible.*

We now characterize when priority and quota mechanisms are (sub)optimal by providing simple necessary and sufficient conditions on the preferences of the authority that allow us to characterize the optimality of priority and quota mechanisms. To do this, we first provide some definitions. Authority preferences are *non-trivial* if for all $m, n \in \mathcal{M}$, we have that:

$$h(1) + u_n'(0) > h(0) + u_m'(q) \tag{184}$$

Intuitively, the authority's preferences are non-trivial when their concerns for representation of certain groups do not always outweigh the consideration of scores.[16] The authority is *risk-neutral* over diversity if for all $m \in \mathcal{M}$, $u_m' : [0, q] \to \mathbb{R}$ is constant, *i.e.,* there are constant marginal returns to admitting more agents from all groups. If there are decreasing marginal returns, then the authority's preferences feature risk aversion. To define extremely risk-averse preferences, let $\tilde{u}$, $\tilde{h}$ and $\{x_m^{\text{tar}}\}_{m \in \mathcal{M}}$ be as follows: (i) $\tilde{u}_m'(x_m) = 0$ for all $x_m > x_m^{\text{tar}}$ (ii) $\tilde{u}_m'(x_m) \geq \tilde{h}(1) - \tilde{h}(0)$ for $x_m \leq x_m^{\text{tar}}$ and (iii) $\sum_{m \in \mathcal{M}} x_m^{\text{tar}} \leq q$. Intuitively, an authority whose preferences are represented by $\tilde{u}$ and $\tilde{h}$ is very risk-averse as the condition that $\tilde{u}_m'(x_m) \geq \tilde{h}(1) - \tilde{h}(0)$ implies that the loss from being below the target level for a group x_m^{tar} dominates any benefit from increased scores. Thus, they are infinitely risk-averse to failing to meet this target. We say that the authority is *extremely risk-averse* if the authority's preferences over the optimal allocations can be represented by $(\tilde{u}, \tilde{h})$.[17]

Theorem 8. *Suppose that the authority has non-trivial preferences. The following statements are true:*

1. *There exists a first-best optimal priority mechanism if and only if the authority is risk-neutral. Moreover, this mechanism is given by $P(s, m) = h^{-1}(h(s) + u_m')$.*

[16] Note that failure of non-triviality means there exists m and n such that $h(1) + u_n'(0) \leq +h(0) + u_m'(q)$, *i.e.,* a group n agent with the maximum score is less preferred than a group m agent with the minimum score even when all of the entire capacity is allocated to group m agents.

[17] More formally, this means that (u, h) are such that the optimal allocation under (u, h) is also optimal under $(\tilde{u}, \tilde{h})$ for all $\omega \in \Omega$.

2. *There exists a first-best optimal quota mechanism if and only if the authority is extremely risk-averse. Moreover, this mechanism is given by* $Q_m = x_m^{tar}$ *and* $D(R) = |\mathcal{M}| + 1$.

Theorem 8 provides precise conditions on preferences such that the inability of priorities and quotas to adapt to the state is not problematic. That risk-neutrality and high risk aversion are sufficient for the optimality of priority and quota mechanisms is intuitive. On the one hand, if the authority is risk-neutral over the measure of agents from different groups, then they can perfectly balance their score and diversity goals without regard for the state of the world. This is because, under risk-neutrality, there is a constant "exchange rate" between the two: how the authority compares any two agents does not depend on the final allocation and thus can be specified *ex ante* by a priority policy. On the other hand, if the authority is extremely risk-averse as to the prospect of failing to assign x_m^{tar} agents from group m, then a quota allows them to always achieve this target level of allocation in all states of the world while minimally sacrificing scores. It is less obvious that risk-neutrality and high risk aversion are necessary. We prove this result by constructing certain adversarial distributions of agents that render any priority or quota mechanism suboptimal unless the authority is risk-neutral or extremely risk-averse, respectively. Importantly, this result also shows that the only optimal quota mechanisms are those that process open slots last.

This result highlights the fragility of priority mechanisms to uncertainty absent the strong assumption of risk-neutrality over diversity. Intuitively, this is because they feature no guarantees as to how many agents of different groups will be assigned. Indeed, the unfortunate interaction between priority mechanisms and unforeseen market realizations has led to public backlash against priority mechanisms. For example, in the Vietnamese university admissions system, which combines exam scores with priority boosts for students from disadvantaged groups, a year of unexpectedly easy exams led to "top-scoring students missing out on the opportunity to attend their university of choice" and generated backlash against the system (Tuoi Tre News, 2017). Moreover, in the Boston Public Schools system, a priority policy that is set to award students bonus points for high school admissions based on the level of disadvantage of their middle school has made it impossible for students from certain middle schools to get into certain high schools, no matter their grades. This led the Boston Herald (2022) to write that "in Boston, hard work and good grades will only set your child back." As APM dynamically adjust priorities, they have the potential to remedy these practical deficiencies of priority mechanisms.

Moreover, our result highlights that quota mechanisms similarly fail to achieve the first-best away from high levels of risk aversion as they do not take advantage of the potential for positive selection. Our quantitative analysis in Section 5.6 in the context of quota mechanisms in Chicago Public Schools suggests that the substantial variation in the distribution of characteristics across years generates economically meaningful welfare gains from switching to APM.

To formalize the connection between uncertainty and the importance of the adaptability of APM, we consider a setting with *no uncertainty*, where Λ is a Dirac measure. In this context, we say that a mechanism is optimal without uncertainty if it is a utility maximizer.

Proposition 18. *If there is no uncertainty, then there exist optimal priority and quota mechanisms.*

This result shows that if an authority is certain about the market, then appropriately constructed priority and quota mechanisms would be optimal. This formalizes the idea that the suboptimality of priority and quota mechanisms stems from their inability to adapt to the state. Of course, in practice, an authority is always somewhat uncertain of the market they face. Thus, absent the strong conditions on authority preferences that we have characterized in Theorem 8, APM dominate priority and quota mechanisms.

5.4 Equilibrium Mechanisms with Multiple Authorities

The single-authority model is relevant for many resource allocation contexts, such as the medical resource allocation problem of a hospital. However, in other settings such as school or university admissions, multiple authorities must decide upon their admissions policies and rules. In this section, we generalize our single authority model to a setting with multiple authorities. We define *stability* in this setting and show that there is a unique stable allocation. Moreover, we show that a mechanism is consistent with stability if and only if it coincides with the single-authority-optimal APM. We then consider a model where agents sequentially apply to the authorities, who decide which agents to admit. We show that the optimal APM is a dominant strategy. Moreover, we show that in any equilibrium in which authorities use the optimal APM, the resulting allocation corresponds to the unique stable matching of the economy. Taken together, our results provide cooperative (stability) and non-cooperative (dominance) foundations for recommending the use of APM in multi-authority settings.

5.4.1 The Multi-Authority Model

There are authorities $c \in \mathcal{C} = c_0 \cup \bar{\mathcal{C}} = \{c_0, c_1, \dots, c_{|\mathcal{C}|-1}\}$ with capacities q_c, where c_0 is a dummy authority that corresponds to an agent going unmatched. The agents differ in their authority-specific scores, the group to which they belong, and their preferences over the authorities, $\succ$. We index agents by their type $\theta = (s, m, \succ) \in [0,1]^{|\mathcal{C}|} \times \mathcal{M} \times \mathcal{R} = \Theta$, where $\mathcal{R}$ is set of all complete, transitive, and strict preference relations over $\mathcal{C}$ such that c_0 is less preferred than all $c \in \bar{\mathcal{C}}$. For each type θ, $s_c(\theta)$ denotes the score of θ at authority c and $m(\theta)$ denotes the group of θ. From now, to economize on notation, we suppress indexing states by $\omega \in \Omega$ and let the measure of types be F, with density f.[18] We assume that f has full support over Θ (*i.e., $f > 0$*) and that the capacity of c_0 is greater $F(\Theta)$.

Each authority has preferences over the agents they are assigned of the form introduced in the previous section:

$$\xi_c(\bar{s}_{h_c}, x_c) = g_c\left(\bar{s}_{h_c} + \sum_{m \in \mathcal{M}} u_{m,c}(x_{m,c})\right) \tag{185}$$

where the extent to which they care about risk g_c, scores h_c, and diversity $\{u_{m,c}\}_{m \in \mathcal{M}}$ are potentially specific to each authority.

A matching is a function $\mu : \mathcal{C} \cup \Theta \to 2^\Theta \cup \mathcal{C}$ where $\mu(\theta) \in \mathcal{C}$ is the authority any type θ is assigned and $\mu(c) \subseteq \Theta$ is the set of agents assigned to authority c.[19] Given a matching μ, $\bar{s}_{h_c,c}(\mu)$ and $x_c(\mu) = \{x_{m,c}(\mu)\}_{m \in \mathcal{M}}$ denote the score indices and measures of agents from different groups matched to c at μ. We say that c *prefers* μ to μ', which we denote by $\mu \succ_c \mu'$, if $\xi_c(\bar{s}_{h_c,c}(\mu), x_c(\mu)) > \xi_c(\bar{s}_{h_c,c}(\mu'), x_c(\mu'))$.

Definition 13. *A matching μ is a cutoff matching if there exist cutoffs $S = \{S_{m,c}\}_{m \in \mathcal{M}, c \in \mathcal{C}}$ such that $\mu(\theta) = c$ if (i) $s_c(\theta) \geq S_{m(\theta),c}$ and (ii) for all c' with $c' \succ_\theta c$, $s_{c'}(\theta) < S_{m(\theta),c'}$.*

Given S, the *demand* of an agent θ is their favorite authority among those for which they clear the cutoff:

$$D^\theta(S) = \{c : s_c(\theta) \geq S_{m(\theta),c} \text{ and } c \succeq_\theta c' \text{ for all } c' \text{ with } s_{c'}(\theta) \geq S_{m(\theta),c'}\} \tag{186}$$

The *aggregate demand* for authority c is the set of agents who demand it $D_c(S) =$

[18] Formally, this density is given by $f(s, m, \succ) = \frac{\partial}{\partial s} F(s, m, \succ)$.

[19] The mathematical definition of a matching for the continuum economy we study follows Azevedo and Leshno (2016) and requires that μ satisfies the following four properties: (i) for all $\theta \in \Theta$, $\mu(\theta) \in \mathcal{C}$; (ii) for all $c \in \mathcal{C}$, $\mu(c) \subseteq \Theta$ is measurable and $F(\mu(c)) \leq q_c$; (iii) $c = \mu(\theta)$ iff $\theta \in \mu(c)$; (iv) (open on the right) for any $c \in \mathcal{C}$, the set $\{\theta \in \Theta : c \succ_\theta \mu(\theta)\}$ is open.

$\{\theta : D^\theta(S) = c\}$, while $\tilde{D}_c(S_{-c}) = D_c((0,\ldots,0), S_{-c})$ returns the set of all agents who would demand c if offered admission admission when other authorities' cutoffs are S_{-c}.

5.4.2 Characterization of Stable Allocations

We first characterize the set of stable allocations. Our context presents two challenges in this regard. First, the priorities which are typically used to define stability are not primitives of our model. Therefore, to define stability, we will use the preferences of the authorities induced by Equation 185. Second, unlike discrete models, a single agent does affect the preferences of an authority. Therefore, we need to consider a positive mass of agents to define blocking.

For each matching μ, authority $c \neq c_0$, and two sets of agents $\tilde{\Theta}$ and $\hat{\Theta}$, we let $\hat{\mu}_{(\hat{\Theta},\tilde{\Theta},c,\mu)}$ denote the matching that maps $\hat{\Theta}$ to c and $\tilde{\Theta}$ to c_0 and otherwise coincides with μ.[20] A set of agents $\hat{\Theta}$ *blocks* matching μ at authority c by $\tilde{\Theta}$ if (i) for all $\theta \in \hat{\Theta}$, $c \succ_\theta \mu(\theta)$, (ii) $\tilde{\Theta} \subseteq \mu(c)$, (iii) $F(\tilde{\Theta}) = F(\hat{\Theta})$, and (iv) $\hat{\mu}_{(\hat{\Theta},\tilde{\Theta},c,\mu)} \succ_c \mu$. A matching μ is *not blocked* if there does not exist such a $(\hat{\Theta}, \tilde{\Theta}, c)$. A matching μ satisfies *within-group fairness* if for all $\theta, \theta' \in \Theta$ such that $m(\theta') = m(\theta)$ and $s_{\mu(\theta)}(\theta') > s_{\mu(\theta)}(\theta)$, it holds that $\mu(\theta') \succeq_{\theta'} \mu(\theta)$.[21] A matching μ is *stable* if it satisfies within-group fairness, is not blocked, and all non-dummy authorities fill their capacity. The following result establishes that there exists a unique stable matching and that this is a cutoff matching.

Theorem 9. *There is a unique stable matching. This matching is a cutoff matching.*

This result extends Theorem 1.1 of Azevedo and Leshno (2016) to our setting in which the preferences of authorities are not exogenously fixed and rather depend on the composition of the admitted agents. The stable allocation is characterized by cutoffs S, where each agent θ is matched to $D^\theta(S)$.

To gain intuition for this result, first imagine that there is only one group of agents $|\mathcal{M}| = 1$, so that authorities' preferences are determined by the scores of the agents. Given a set of cutoffs S_{-c}, a cutoff t_c clears the market for c if $F(D_c(t_c, S_{-c})) = q_c$.

[20] Formally,

$$\hat{\mu}_{(\hat{\Theta},\tilde{\Theta},c,\mu)}(\theta) = \begin{cases} c_0 & \text{if } \theta \in \tilde{\Theta} \\ c & \text{if } \theta \in \hat{\Theta} \\ \mu(\theta) & \text{otherwise} \end{cases} \tag{187}$$

[21] Within-group fairness simply requires an authority not to reject a agent if it is admitting a agent from the same group with lower score. Under our assumption that authorities prefer higher scores (h_c is strictly increasing), within-group fairness is satisfied if there is no blocking in discrete models.

When $|\mathcal{M}| = 1$, for a given S_{-c}, there is a unique t_c that clears the market since a smaller cutoff will exceed the capacity while a larger one will leave a positive measure of the capacity empty. Define $T = \{T_c\}_{c \in \mathcal{C}}$, where $T_c(S)$ is the function that maps each S to the market-clearing cutoff t_c under S_{-c}. The result then follows from (i) showing the fixed points of T correspond to market-clearing cutoffs of stable matchings, (ii) establishing that T is monotone, (iii) applying Tarski's fixed point theorem to show that the set of market-clearing cutoffs is a lattice, and (iv) showing that there can only be one market-clearing cutoff as, if there were two, one would strictly exceed the capacities of at least one authority.

When $|\mathcal{M}| > 1$, there is a potential continuum of cutoffs that would clear the market for authority c. A selection from this set is provided by the cutoffs induced by the optimal APM, $A^*_{m,c}(y_m, s) \equiv h_c^{-1}(h_c(s) + u'_{m,c}(y_m))$. We show that the APM cutoffs are unique among the market-clearing cutoffs in being compatible with stability. This is because, for any other t'_c, there is a set $\hat{\Theta}$ of agents (with positive measure) who have scores lower than the cutoff for their group and a set $\tilde{\Theta}$ of agents (with positive measure) who have scores higher than the cutoff for their group, but the authority is strictly better of by admitting $\hat{\Theta}$ and rejecting $\tilde{\Theta}$. We define $T_c(S)$ as the market-clearing cutoffs induced by the optimal APM, show that the fixed points of T_c correspond to market-clearing cutoffs of stable matchings, and follow the same steps as above to demonstrate uniqueness.

This hints at a connection between the stable allocation and the allocation induced by all authorities pursuing the optimal APM, which we now make explicit. The demand set of c at μ, $D_c(\mu)$, is the set of agents who prefer c to their allocation under μ. A mechanism is *consistent with stability* if for all F with stable matching μ_F, it chooses $\mu_F(c)$ from $D_c(\mu_F)$. In other words, evaluated at the set of agents who demand an authority, this mechanism chooses the set of agents with which the authority is already matched. Moreover, we say that a mechanism ϕ is *equivalent* to ϕ' if it chooses the same agents under all full support distributions. We now establish that single-authority-optimal APMs (and equivalent mechanisms) comprise the full set of mechanisms that are consistent with stability.

Theorem 10. *A mechanism ϕ is consistent with stability if and only if it is equivalent to A^*_c.*

Thus, not only is the optimal APM A^* inherent to the structure of stable allocations, but it also characterizes stability in this setting in the sense that any deviation from A^* would result in a violation of stability.

5.4.3 APM Are Dominant Under Decentralized Admissions

We now consider a model where the agents apply to the authorities sequentially, who decide which agents to admit. We index the stage of the game by $t \in \mathcal{T} = \{1, \ldots, |\mathcal{C}|-1\}$. Each stage corresponds to a (non-dummy) authority $I(t)$, where $I : \mathcal{T} \to \mathcal{T}$. At each stage t, any unmatched agents choose whether to apply to authority $I(t)$. Given the set of applicants, authority $I(t)$ chooses to admit a subset of these agents, who are then matched to the authority. Given this, histories are indexed by the path of the measure of agents who have not yet matched, $h^{t-1} = (F, F_1, \ldots, F_{t-1}) \in \mathcal{H}^{t-1}$. Given each history h^{t-1} and set of applicants $\Theta_c^A \subseteq \Theta$, a strategy for an authority returns a set of agents $\Theta_c^G \subseteq \Theta$ whom they will admit such that $\Theta_c^G \subseteq \Theta_c^A$ and $F_t(\Theta_c^G) \leq q_c$ for each time at which they could move $t \in \mathcal{T}$, $a_{c,t} : \mathcal{H}^{t-1} \times \mathcal{P}(\Theta) \to \mathcal{P}(\Theta)$, where $\mathcal{P}(\Theta)$ is the power set over Θ.[22] A strategy for an agent returns a choice of whether to apply to authorities at each history and time for all agent types $\theta \in \Theta$, $\sigma_{\theta,t} : \mathcal{H}^{t-1} \to [0, 1]$.

Within this context, our notion of equilibrium is that of subgame perfect equilibrium:

Definition 14 (Equilibrium)**.** *A strategy profile* $\Sigma = \{\{a_{c,t}\}_{c \in \bar{\mathcal{C}}}, \{\sigma_{\theta,t}\}_{\theta \in \Theta}\}_{t \in \mathcal{T}}$ *is a subgame perfect equilibrium if* $a_{c,t}$ *maximizes authority utility given* Σ *for all* $c \in \bar{\mathcal{C}}$ *and* $t \in \mathcal{T}$ *and* $\sigma_{\theta,t}$ *is maximal according to agent preferences for all* $\theta \in \Theta$ *and* $t \in \mathcal{T}$*.*

We moreover say that a strategy $a_{\tilde{c},t}$ for an authority $\tilde{c}$ at time t is *dominant* if it maximizes authority utility regardless of the strategies of all other authorities and agents, $\{\{a_{c,t}\}_{c \in \bar{\mathcal{C}} \setminus \{\tilde{c}\}}, \{\sigma_{\theta,t}\}_{\theta \in \Theta}\}_{t \in \mathcal{T}}$, and the order in which authorities admit agents, I. Moreover, an equilibrium Σ is in *dominant strategies* if $a_{c,t}$ is dominant for all $c \in \bar{\mathcal{C}}$ and $t \in \mathcal{T}$. We denote the unique probabilistic allocation of agents to authorities induced by Σ as $\mu_\Sigma : \Theta \to \Delta(\mathcal{C})$. A probabilistic allocation μ_Σ is deterministic if $\mu_\Sigma(\theta)$ is a Dirac measure on some authority $c \in \mathcal{C}$ for all $\theta \in \Theta$. A deterministic allocation μ_Σ corresponds to a matching μ if $\mu_\Sigma(\theta)$ is a Dirac measure on $\mu(\theta)$ for all $\theta \in \Theta$.

We now establish that the single-authority optimal APM characterizes dominance.

Theorem 11. *A mechanism implements a dominant strategy for an authority if and only if it implements essentially the same allocations as* A_c^**.*

The intuition behind this result is that each authority takes as given the set of agents that will accept it. Thus, given this measure of agents, they can do no better

[22] Formally, so that $F_t(\Theta_c^G)$ is well defined, we require that authorities' strategies be measurable in the Borel sigma algebra over Θ.

than to employ the same APM that a single authority would, which is A_c^* by Theorem 7.

Theorem 11 provides a powerful rationale for focusing on APMs in decentralized markets at both positive and normative levels. Normatively, this result allows an analyst to advise an authority regarding how it should conduct its admissions. This is important because any policy that does not coincide with the APM we derive — such as the popular priority and quota mechanisms outside of the cases delimited by Theorem 8 — will disadvantage an authority. Positively, this result allows a sharp prediction that the equilibrium matching between agents and authorities will be the unique stable matching (as per Theorem 9):

Proposition 19. *For all equilibria Σ^* where authorities use A^*, the allocation μ_{Σ^*} is deterministic and corresponds to the unique stable matching of this economy.*

The intuition for this result is that if an equilibrium matching under A^* was not the unique stable matching, then it must be that some agents are applying suboptimally and failing to select the best authority they can attend (according to their preferences), which contradicts that the outcome is an equilibrium.

5.5 Efficient Mechanisms with Multiple Authorities

We have so far characterized the decentralized outcome, but two natural questions remain. First, is the decentralized outcome, which corresponds to the stable allocation, efficient? Second, if not, what kind of centralized solution can remedy any inefficiency? We show that the decentralized outcome is generally inefficient and that a modified, centralized APM mechanism restores efficiency.

5.5.1 Inefficiency of the Decentralized Outcome

The notion of efficiency that we will consider is utilitarian efficiency over authorities. A mechanism in the multi-authority setting is a function $\phi : \Omega \to \mathcal{U}$, where $\mathcal{U}$ is the set of matchings (which, by definition, encodes feasibility requirement imposed in the single authority setting). We define the total authority value Ξ_T of a mechanism ϕ under distribution $\Lambda \in \Delta(\Omega)$ as the total expected utility of the allocations induced by that mechanism:

$$\Xi_T(\phi, \Lambda) = \sum_{c \in \vec{\mathcal{C}}} \Xi_c(\phi, \Lambda) \tag{188}$$

A mechanism is efficient if it maximizes total authority value for all possible distributions:

Definition 15 (Efficiency)**.** *A mechanism ϕ^* is efficient if:*

$$\Xi_T(\phi^*, \Lambda) = \sup_{\phi} \Xi_T(\phi, \Lambda) \tag{189}$$

for all $\Lambda \in \Delta(\Omega)$.

For the remainder of the paper, so that scores are directly comparable across authorities and allocations are interior, we impose the following assumption:

Assumption 11. *Scores and preferences are such that $s_c(\theta) = s_{c'}(\theta)$, $h_c = h$ and $g_c = Id$, where Id is the identity function, for all $c, c' \in \bar{\mathcal{C}}$ and $\theta \in \Theta$. Moreover, $\lim_{x \to +0} u'_{m,c}(x) = \infty$ and $u_{m,c}$ is strictly concave for all $m \in \mathcal{M}$ and $c \in \bar{\mathcal{C}}$.*

Assumption 11 makes scores a common numeraire across authorities and is akin to the standard quasi-linearity assumption in mechanism design. For example, it may be suitable in settings where the score is derived from a common index of academic attainment, such as in Chicago Public Schools. This assumption does not impose that all authorities have common marginal rates of substitution between scores and diversity, as they are allowed unrestricted heterogeneity in diversity preferences. We add the Inada condition for analytical tractability. We argue that it is also reasonable to assume that failing to admit any individuals from a given group is intolerable for authorities.

With the efficiency benchmark defined, we can now demonstrate that the decentralized equilibrium outcome can fail to be efficient.

Proposition 20 (Equilibrium Inefficiency)**.** *All authorities using the privately optimal APMs $\{A_c^*\}_{c \in \mathcal{C}}$ is not necessarily efficient.*

We prove this result via an explicit example with two authorities, c and c' of capacity $\frac{1}{4}$, and two groups of agents, m and m' of measure $\frac{1}{2}$. All agents in group m prefer c' to c and all agents in group m' prefer c to c'. Authority c values admitting group m agents more on the margin than group m' agents, and authority c' values admitting group m' agents more on the margin than group m agents. Using the optimal APMs, both authorities admit more agents of the group whose admissions they value relatively less than the efficient benchmark. The intuition for this is that both authorities "steal" the high-scoring agents of the group whom they relatively less value from the other authority, an externality that they do not internalize.

5.5.2 An Efficient Centralized Mechanism

The inefficiency of each authority using a decentralized APM stems from the implicit incompleteness of markets: if we added the ability for authorities to pay each other for agents in the equilibrium allocation, then they would have willingness-to-pay to do so. A centralized mechanism can remedy this issue by ensuring the cross-sectional allocation of agents to authorities is optimal.

We propose the following augmentation of an APM to solve this problem, an *adaptive priority mechanism with quotas* (APM-Q). The idea behind this hybrid mechanism is to use aggregate, market-level priorities with authority-specific quotas. To this end, an APM-Q comprises an aggregate non-separable APM $\tilde{A} = \{\tilde{A}_m\}_{m\in\mathcal{M}}$ with $\tilde{A}_m : \mathbb{R}^{|\mathcal{M}|} \times [0,1] \to \mathbb{R}$ and a profile of quota functions $Q = \{Q_{m,c}\}_{m,c\in\mathcal{M}}$ with with $Q_{m,c} : \mathbb{R}^{|\mathcal{M}|} \to \mathbb{R}_+$. Intuitively, the aggregate APM pins down the aggregate measures of allocations of each group to *any* authority $\{x_m\}_{m\in\mathcal{M}}$, where $x_m = \sum_{c\in\bar{\mathcal{C}}} x_{m,c}$. The non-separability of this APM simply means that the measures of all groups matter for the adaptive priority of any agent. Given the aggregate measure of allocation for group m, the quota function for authority c assigns $Q_{m,c}(\{x_m\}_{m\in\mathcal{M}})$ agents of type m to authority c.

Definition 16 (Adaptive Priority Mechanism with Quotas)**.** *An adaptive priority mechanism with quotas $(\tilde{A}, Q)$ comprises a non-separable APM $\tilde{A}$ and a quota function Q. An APM-Q implements allocation μ if the following are satisfied:*

1. *Aggregate allocations are in order or priorities: $\mu(\theta) \in \bar{\mathcal{C}}$ if and only if for all θ' with $\mu(\theta') = c_0$, we have that:*

$$\tilde{A}_{m(\theta)}(\{x_m(\mu)\}_{m\in\mathcal{M}}, s(\theta)) > \tilde{A}_{m(\theta')}(\{x_m(\mu)\}_{m\in\mathcal{M}}, s(\theta')) \tag{190}$$

2. *Authority-level allocations are given by the corresponding quota functions:*

$$x_{m,c}(\mu) = Q_{m,c}(\{x_m(\mu)\}_{m\in\mathcal{M}}) \tag{191}$$

3. *The resources are fully allocated:*

$$\sum_{m\in\mathcal{M}} x_{m,c}(\mu) = q_c \tag{192}$$

By appropriate choice of the APM and quota functions, we can derive an APM-Q that is efficient. To this end, define the optimally-allocated aggregate utility from diversity:

$$\begin{aligned}\tilde{u}(\{x_m\}_{m\in\mathcal{M}}) &= \max_{\{x_{m,c}\}_{c\in\mathcal{C}}} \sum_{c\in\mathcal{C}} \sum_{m\in\mathcal{M}} u_{m,c}(x_{m,c}) \\ \text{s.t.} \sum_{c\in\mathcal{C}} x_{m,c} &\leq x_m, \ \sum_{m\in\mathcal{M}} x_{m,c} \leq q_c, \ \forall m \in \mathcal{M}, c \in \mathcal{C}\end{aligned} \tag{193}$$

Moreover, define the marginal value of aggregate group m admissions $\tilde{u}^{(m)}(y) = \frac{\partial}{\partial y_m}\tilde{u}(y)$ and the marginal value of authority capacity $\tilde{u}_{q_c}(y) = \frac{\partial}{\partial q_c}\tilde{u}(y)$. Using these marginal values, we can design an efficient APM-Q that combines market-level APMs with authority-level quotas:

Theorem 12 (Efficient APM-Q)**.** *Every allocation induced by the following APM-Q* $(\tilde{A}, Q)$ *is efficient:*

1. *The non-separable APM is given by* $\tilde{A}_m(y,s) = h^{-1}\left(h(s) + \tilde{u}^{(m)}(y)\right)$
2. *The quota functions are given by* $Q_{m,c}(y) = \left(u'_{m,c}\right)^{-1}\left(\tilde{u}^{(m)}(y) + \tilde{u}_{q_c}(y)\right)$

The proof of this result constructs a fictitious aggregate authority in our single object setting. The claimed APM is optimal for this aggregate authority by a non-separable adaptation of Theorem 7. The substantial step in this proof establishes that $\tilde{u}$ is increasing, concave, and differentiable by employing the restrictions provided by Assumption 11. Then, given the allocation induced by this APM, we construct the quota function to optimally allocate the level of aggregate admissions induced by the APM.

Intuitively, this mechanism remedies inefficiency by "completing markets." There is a common "market price" for each group given by $\mathcal{P}_m = \tilde{u}^{(m)}(x)$ and an authority-level "shadow price of admissions" $\mathcal{P}_c = \tilde{u}_{q_c}(x)$. Authorities are allocated agents so that the marginal benefit of additional agents equals the sum of the market price and shadow price of admissions $u'_{m,c}(x_{m,c}) = \mathcal{P}_m + \mathcal{P}_c$. Hence, through the completion of markets, a centralized planner can allocate agents efficiently and internalize the externalities that prevented efficiency under the decentralized outcome. Notice that this market involves relatively few prices as it involves only $|\mathcal{M}| + |\mathcal{C}|$ shadow prices rather than the full set of $|\mathcal{M}| \times |\mathcal{C}|$ marginal values.

5.6 Benchmarking the Quantitative Gains from APM

We have so far shown theoretically that APM outperform conventional priority and quota mechanisms. In this section, we attempt to benchmark the magnitude of the gains from implementing APM relative to the reserve system employed by Chicago Public Schools (CPS). To do this, we use application and admission data from CPS for the 2013-2017 academic years. Estimating preference parameters to best rationalize the pursued reserve policy, we find that the gains from using the optimal APM are equivalent to removing 37.5% of the loss to CPS' payoffs from failing to admit a diverse class of students. Our analysis therefore suggests that the gains from APM are considerable.

5.6.1 Institutional Detail on Chicago Public Schools

We first describe the institutional context of CPS. Under current policy, CPS admits students to its selective exam schools based on two criteria. First, CPS ranks students according to a composite score which combines the results of a specialized entrance exam, prior standardized test scores, and grades in prior coursework. This composite score ranges from 0 to 900 and higher-scoring students are admitted before lower-scoring ones, reflecting CPS's desire to allocate seats in exam schools to the students with the best academic standing. In our model, these are the students' scores. Second, CPS divides the census tracts in the city into four *tiers* based on socioeconomic characteristics.[23] Tier 1 tracts are the most disadvantaged, while Tier 4 tracts are the most advantaged. This is reflected in the composite scores of students from Tier 1, who represent 25% of the city's population but comprise relatively few of the high-scoring students (Ellison and Pathak, 2021). As a result, Tier 1 students would have a very small share in the city's top exam schools without affirmative action. To ensure more equal representation across socioeconomic status in these schools, between 2013-2017 CPS implemented a quota policy that reserves 17.5% of the seats for each tier, yielding a total of 70% reserve seats and 30% merit slots that are open to students from all tiers. CPS allocates the seats by first assigning the highest-scoring students (regardless of their tier) to the merit slots and then the highest-scoring students from each tier to the 17.5% reserve seats.

We focus on the most selective CPS school, Walter Payton College Preparatory High School (Payton), which has the highest cutoffs for each tier in each year in our

[23] Concretely, 800 census tracts are divided into four tiers based on six characteristics of each census tract: (i) median family income, (ii) percentage of single-parent households, (iii) percentage of households where English is not the first language, (iv) percentage of homes occupied by the homeowner, (v) adult educational attainment, and (vi) average Illinois Standards Achievement Test scores for attendance-area schools. These characteristics are then combined to construct the socioeconomic score for the tract. Finally, the tracts are ranked according to socioeconomic scores and partitioned into 4 tiers with approximately the same number of school-age children. See Ellison and Pathak (2021) for a more detailed account of the CPS system.

Table 5.1: Admissions Cutoff Scores for Payton

Cutoff Score	2013	2014	2015	2016	2017
Tier 1	801	838	784	769	771
Tier 2	845	840	831	826	846
Tier 3	871	883	877	853	875
Tier 4	892	896	891	890	894

Notes: The table reports the score of the lowest-scoring student that was admitted to Payton in each of the four tiers and five years.

data and would have very few tier 1 students without affirmative action.[24] Table 5.1 presents the cutoff scores of each tier (*i.e.,* the composite score of the last admitted student from each tier).

We make two observations. First, the cutoff students from less advantageous tiers face is lower than the cutoff for students from more advantageous tiers. Therefore, CPS has a *revealed* preference (and not merely a stated preference) for a diverse student body. Second, cutoff scores vary across years. This implies that the distribution of applicant characteristics varies from year to year. Given this uncertainty and the fact that CPS uses a policy that processes quotas after open slots, we know by Theorem 8 that CPS' baseline policy cannot be rationalized as optimal (even if they are extremely risk-averse). Nevertheless, it is always possible that the gains from APM could be small.

5.6.2 Preferences and Estimation Methodology

We perform our analysis in two steps: establishing a parametric framework for evaluating welfare gains and losses and then estimating its parameters.

Preferences We assume a parametric form for CPS's preferences to evaluate the gains from APM. In particular, we impose that the preferences of CPS over the scores and diversity of the student body are represented by the following parametric utility

[24]This approach follows the analysis in Ellison and Pathak (2021), who focus on the two most competitive schools, Northside College Preparatory High School (Northside) alongside Payton. In the years we study, the cutoff scores for Northside are below some other schools frequently, which is why we restrict attention to Payton.

function[25]

$$\xi(\bar{s}, x; \beta, \gamma) = \bar{s} + \sum_{t=1}^{4} \beta |x_t - 0.25|^\gamma \tag{194}$$

where $\bar{s}$ is the average score of admitted students, x_t is the percentage of tier t students. Motivated by CPS' desire to allocate the highest-scoring students, ξ is increasing in $\bar{s}$. To model the diversity preferences of CPS, we assume that CPS loses as the gap from equal representation in each tier increases. We do this through the functional form $\beta|x_t - 0.25|^\gamma$. The parameters β and γ index the slope and curvature of utility in losses from unequal representation and are the two free parameters of our framework.

Estimation We estimate β and γ to best rationalize the choice of 17.5% reserves for each tier as optimal. We believe this to be a reasonable approach, as the size of the reserves is an important issue that is decided only after much deliberation.[26] Moreover, CPS has used the size of the reserves as a policy tool, increasing them from 15% to 17.5% in 2012 and is currently deliberating another change that would further boost the representation of tier 1 and tier 2 students (Chicago Public Schools, 2022).

Given our functional form, the optimality of the chosen reserve sizes yields moment conditions that we use to estimate the parameters β and γ. Formally, we index reserve mechanisms by the reserve sizes of the four socioeconomic tiers $r = (r_1, r_2, r_3, r_4)$. We let $\bar{s}(r, y)$ and $x(r, y)$ denote the average scores and tier percentages that would be obtained in year y, with distribution F_y, under reserve policy r. The payoff of the policymaker under reserve policy r is given by $\Xi(r, \Lambda; \beta, \gamma)$, as per Equation 175:

$$\Xi(r, \Lambda; \beta, \gamma) = \mathbb{E}_\Lambda[\xi(\bar{s}(r, y), x(r, y); \beta, \gamma)] \tag{195}$$

where the expectation is taken over distributions of agents' characteristics F_y under the subjective probability measure Λ. Define the expected marginal benefit of

[25]In Appendix E.6.2, we consider two other parametric utility functions that estimate separate coefficients for underrepresented and overrepresented tiers and only considers loss from underrepresented tiers.

[26]These points are emphasized in Dur, Pathak, and Sönmez (2020): "This change was made at the urging of a Blue Ribbon Commission (BRC, 2011), which examined the racial makeup of schools under the 60% reservation compared to the old Chicago's old system of racial quotas. They advocated for the increase in tier reservations on the basis it would be "improving the chances for students in neighborhoods with low performing schools, increasing diversity, and complementing the other variables."

increasing reserve i and decreasing reserve j as:

$$G_{ij}(r, \Lambda; \beta, \gamma) = \frac{\partial}{\partial r_i}\Xi(r, \Lambda; \beta, \gamma) - \frac{\partial}{\partial r_j}\Xi(r, \Lambda; \beta, \gamma) \tag{196}$$

Any (interior) optimal reserve policy r^* must equate the expected marginal benefit of increasing reserve i and decreasing reserve j at r^* to zero for all (i, j) pairs, *i.e.,* $G_{ij}(r^*, \Lambda; \beta, \gamma) = 0$ for all $\{i, j\} \subset \{1, 2, 3, 4\}$ such that $j > i$. These six first-order conditions yield six moments.

We take empirical analogs of the theoretical moments and estimate preference parameters by minimizing the sum of squared deviations of these moments from zero. We take CPS' pursued reserve policy as optimal, $\hat{r}^* = (0.175, 0.175, 0.175, 0.175)$. We estimate the empirical joint distribution of students' scores and tiers in CPS in each year $\hat{F}_y$ for $y \in \{2013, 2014, 2015, 2016, 2017\}$ and estimate $\hat{\Lambda}$ as a distribution that places equal probability on each of these five measured distributions. We plug these sample estimates into the theoretical moment functions. This yields six empirical moment functions that depend only on the preference parameters, $G_{ij}(\hat{r}^*, \hat{\Lambda}; \beta, \gamma)$. Motivated by the theoretical necessity of $G_{ij}(r^*, \Lambda; \beta, \gamma) = 0$, we estimate the preference parameters by minimizing the sum of squared deviations of the empirical moments from zero:

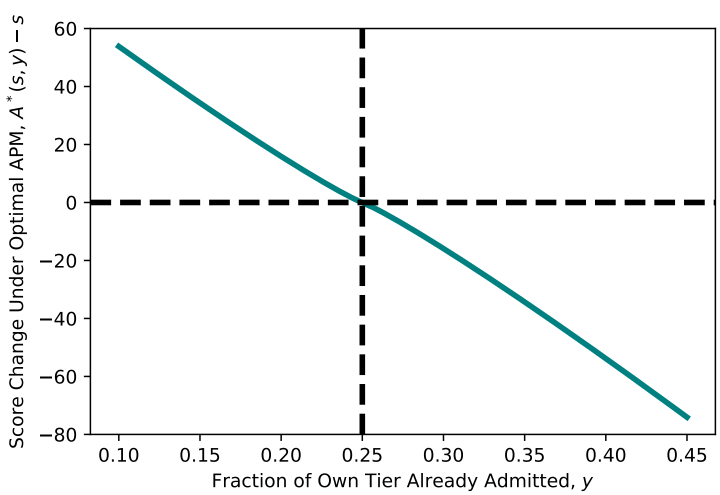
$$(\beta^*, \gamma^*) \in \arg\min_{\beta,\gamma} \sum_{i=1}^{4} \sum_{j>i} G_{ij}(\hat{r}^*, \hat{\Lambda}; \beta, \gamma)^2 \tag{197}$$

Performing this estimation yields estimated parameter values of $\beta^* = -209.5$ and $\gamma^* = 2.11$.

5.6.3 The Estimated Gains from APM

We now use our estimated model to quantify the welfare gains from using APM. To do this, we compare the empirical payoff $\Xi(\phi, \hat{\Lambda}, \beta^*, \gamma^*)$ under two mechanisms: the pursued quota policy, r^*, and the optimal APM from Theorem 7, A^*. In Figure 5-4, we illustrate how the estimated optimal APM changes students' scores to arrive at their ultimate priorities. In accordance with the preferences we have assumed, students receive a score boost when their tier is underrepresented and a score penalty when their tier is overrepresented. As we found $\gamma^* = 2.11$, the estimated diversity preference is very close to quadratic. Thus, the optimal APM is very close to linear. From Theorems 7 and 8, we know that this APM achieves the first-best allocation in each year while the implemented quota policy does not. However, our theorems do not guarantee that the gains from APM are economically meaningful.

Figure 5-4: The Estimated Optimal APM



Notes: This figure plots the change in a student's score when fraction y of students in their own tier has already been admitted under the estimated optimal APM, A^*. At $y = 0.25$ (the vertical dashed black line), the score in unchanged. For $y < 0.25$, students receive a score boost. For $y > 0.25$, students receive a score penalty. The range of the x-axis, $[0.1, 0.45]$, is chosen to cover the full range of fractions of admitted students under both the optimal and the CPS reserve policy from all tiers in all years of our sample (see Figure 5-5).

The empirical payoff under APM is 876.9, while it is 874.8 under the CPS reserve policy. Thus, the gains from APM are equivalent to increasing average scores by 2.1, holding diversity fixed. To benchmark the size of the gains, we require units in which they can be meaningfully expressed. We define the *loss from underrepresentation* as the payoff lost by CPS under its baseline policy from not admitting a fully balanced class, while holding fixed the average score of the class. This is equal to 5.6 points under our estimated parameters. Thus, the gains from APM are equal to 37.5% of the loss from underrepresentation incurred under the CPS policy.[27] We define *score cost of diversity* as the difference between the average scores of admitted students without any affirmative action (893.8) and the average score under the CPS policy (880.5). Thus, the gains from APM are equal to 15.7% of the score cost of diversity. Finally, we can compare the gains from switching to the optimal APM to the gains from the 2012 (the year before our sample) CPS reform, which increased the size of all reserves from 15% to 17.5%. Under the estimated preferences, the empirical payoff under the 15% reserve rule is 873.9, and so the gains from the reform are equivalent to increasing average scores by 0.9. Thus, the gains from switching to the optimal APM are 2.3 times larger than the gains from this recent reform.

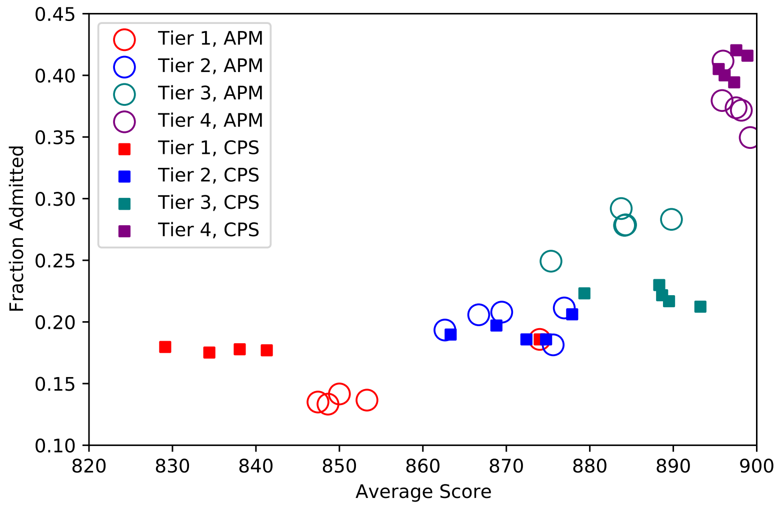
These estimates suggest that the gains from APM are economically meaningful. These gains stem from the variation across years in the joint distribution of student scores and tiers. This can be seen in Table 5.1, which shows the variability in the scores of the marginally admitted students from tiers 1, 2 and 3. More systematically, we visualize the difference in outcomes under CPS' reserves and the optimal APM by plotting the average scores and fraction admitted for each tier for each year under both mechanisms in Figure 5-5. There are two main differences between the allocations. First, the APM allocates systematically fewer tier 1 and tier 4 students and more tier 3 students. Second, the APM admits a greater fraction of students from each tier (especially tiers 1 and 3) in the years in which that tier scores well. This positive selection generates the welfare gains.

Robustness We now explore the robustness of APM to the three core assumptions of our analysis: (i) that CPS has the correct beliefs about the distribution of distributions of students, (ii) that CPS has preferences that lie in the assumed parametric family, and (iii) that CPS separately optimizes the sizes of all four tiers.

Our baseline analysis took the beliefs of CPS to be the true empirical distribution

[27] This is equivalent to increasing the percentage of students from tier 1 from 0.179 to 0.21 and decreasing the percentage of students from tier 4 from 0.407 to 0.378. This corresponds to swapping 8.7 students from the most overrepresented group (tier 4) for the most underrepresented group (tier 1) each year.

Figure 5-5: Comparing Admissions under the Optimal APM and the CPS Policy



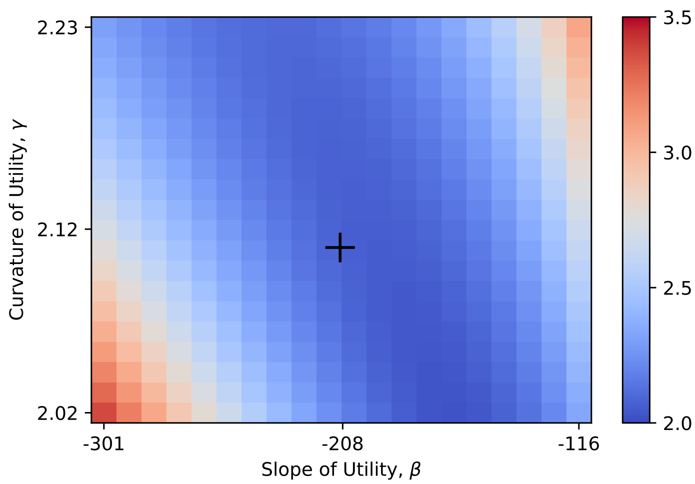
Notes: Each point corresponds to one of the four tiers of students in one of the five years under either the optimal APM or the CPS policy. The x-axis corresponds to the average score of those admitted from that tier in that year under that policy. The y-axis corresponds to the fraction of admitted students from that tier in that tier under that policy.

of student distributions over years. To test robustness to this assumption, we take $\hat{\Lambda}$ as a Dirac distribution on the realized distribution for each of the five years of our data and re-estimate the preference parameters. In Figure 5-6, we plot the difference in welfare under the optimal APM and CPS reserve policies over the full range of these re-estimated parameters (*i.e.,* we take the minimum and maximum of the estimated parameters across years as the ranges for the axes). We find that the gains from APM range from 2.0 to 3.5, while our baseline estimate was 2.1. Thus, the point estimate of our welfare gains from APM appears to be conservative by this metric.

To gauge robustness to the functional form we have assumed for CPS' preferences, in Appendix E.6.2, we estimate two different parametric specifications of utility. First, we consider a utility function that includes a loss term only for underrepresented tiers (and does not penalize overrepresentation of any tier). Second, we allow for CPS to care differentially about underrepresentation and overrepresentation by considering a utility function with separate coefficients for underrepresented and overrepresented tiers. We find that under these specifications, the improvement from APM corresponds to 9.7% and 8.7% of the loss from underrepresentation, which is attenuated relative to our baseline, but remains considerable.

To study the robustness of our findings to the assumption that CPS separately

Figure 5-6: Robustness of the Gains from APM



Notes: This chart plots the difference in empirical payoffs from the optimal APM and CPS reserve policy under alternative parameter values, with the shaded colors corresponding to the numerical value of the gains from APM, ranging from 2.0 to 3.5. The black '+' indicates our baseline parameter values. The ranges for the axes are obtained by estimating β and γ separately for each year of our data and separately taking the minimum and maximum estimated values of each set of estimated parameters.

optimizes the size of all four tiers, in Appendix E.6.1 we consider a setting where CPS sets a *single* reserve size for all tiers. As we now have only one moment condition, we vary γ over the interval [1,10], estimate $\beta^*(\gamma)$ as the exact solution to the moment condition, and compute the gains from APM as a function of γ. The *minimum* gain from APM over the estimated range is 1.98 points, which corresponds to 26.2% of the loss from underrepresentation under that parameterization. This is slightly smaller than our baseline estimate but still considerable.

Limitations Finally, we state some limitations of our analysis. First, even though we argue that our functional form assumptions are reasonable and parsimonious in this setting, there are possibly many other parametric utility functions that might represent the preferences of CPS. Unfortunately, with the available data, richer methods of preference estimation that allow for higher dimensionality of preference parameters are severely limited. This is because admissions rules are only set once and we can therefore use only six first-order conditions. Richer data in which we observe choices of students from different applicant sets (either in practice or in hypothetical choice settings) would allow for a more detailed analysis. Second, one of the main aims of the tier system employed by CPS is to increase racial diversity in the prestigious exam schools. Indeed, the pursued tier system is a race-neutral alternative that replaced the previous race-based system following two Supreme Court Rulings in 2003 and 2007 (see Ellison and Pathak, 2021, for a summary). Because of this, CPS uses tiers based on socioeconomic status instead of race and so we estimate their preferences over tiers. This notwithstanding, if admission rules could depend on race, then one could perform a similar analysis in which APM simply prioritize based on race rather than tier.

5.7 Conclusion

Motivated by the use of priority and quota policies in resource allocation settings with diversity concerns, we consider an authority that has separable preferences over scores and diversity. We introduce Adaptive Priority Mechanisms (APM) and characterize an APM that is both optimal and can be specified solely in terms of the preferences of the authority. We also study the priority and quota policies that are used in practice and show that they are optimal if and only if the authority is either risk-neutral or extremely risk-averse over diversity. Analyzing a setting with multiple authorities that dynamically admit agents, we show that the optimal APM is a dominant strategy. Thus, one could potentially advise authorities to follow an optimal APM with confidence (under our maintained assumptions on preferences) that they could do no better. Moreover, all authorities following the optimal APM implements the unique stable allocation. This notwithstanding, the stable allocation can fail to be efficient for the authorities. We propose a centralized adaptive priority mechanism with quotas to remedy this.

Our analysis has potential implications for improving the design of real-world allocation mechanisms. First, we show that while both priorities and quotas can be better than one another (depending on the risk preferences of the authority), they are generally suboptimal. Second, we show how to improve upon these mechanisms using APM that harness the strengths of these policies: APM benefit both from the guarantee effect of quotas in ensuring certain levels of admissions from various groups and the positive selection effect of priorities in expanding affirmative action when it is least costly. Our quantitative analysis using CPS data suggests that the use of APM can yield considerable welfare gains over the *status quo*. On the basis of our analysis, we conclude that APM may have a real-world use in delivering more desirable allocations of resources. Moreover, by virtue of their generality, APM could be applied in many settings, including the allocation of seats at schools, places at universities, and medical resources to patients.

Chapter 6

Nonlinear Pricing with Under-Utilization: A Theory of Multi-Part Tariffs

This chapter is jointly authored with Roberto Corrao and Karthik A. Sastry and has been published in the *American Economic Review* as Corrao, Flynn, and Sastry (2023).

6.1 Introduction

Digital goods often feature tiers of service within which marginal prices are zero. That is, they are sold according to multi-part tariffs. For example, making Google searches or browsing Facebook posts is always free; streaming certain movies from Amazon Prime's library is free after paying for a subscription, while additional movies can be streamed for a price; and reading *Wall Street Journal* articles is free on the margin for anyone up to a trial limit, and free for paid subscribers in unlimited quantities. A common theme in each of these cases is that the seller can monetize the buyer's time and attention via other channels, like serving advertisements, collecting valuable data, generating network effects, or addicting users. This indirect revenue is big business for the internet's largest players—for example, Google, Facebook, and Amazon respectively made 98%, 108%, and 60% of their net profit in 2020 from advertisement.[1]

But the fact that indirect revenue *can* make zero-marginal-price units profitable does not explain why they are optimal. The classical theory of nonlinear pricing as

[1] Net advertisement revenue figures are analyst estimates by eMarketer (eMarketer Insider Intelligence, 2020), and net income is from financial statements as collected by the *Wall Street Journal*.

screening (Mussa and Rosen, 1978; Maskin and Riley, 1984; Wilson, 1993) predicts that sellers should use smoothly varying marginal prices, instead of tiers of zero marginal prices, to extract maximal profits.[2] Understanding multi-part tariffs requires an alternative economic mechanism.

In this paper, we introduce a nonlinear pricing model with the following form of non-contractibility: once the buyer purchases the right to use a good, the seller cannot fully enforce the good's utilization. The scope for under-utilizing digital goods—and its potential influence on digital markets—is perhaps best illustrated by the historical failure of "pay-to-click" businesses that try to incentivize valuable usage (e.g., to pay people to see advertisements), only to be defrauded by users' simple cheating strategies (e.g., having a computer script click through a website).[3] Our model captures this key issue: providers cannot contract upon legitimate, valuable usage, or otherwise prevent fraudulent, valueless usage.

Our main result is that the optimal price schedule in the presence of under-utilization and usage-derived revenue is a multi-part tariff. Intuitively, non-contractibility of usage prevents sellers from charging negative marginal prices, which they would like to do to encourage valuable usage. Thus, they instead charge a price of zero on the margin. The remainder of our analysis explores the structure of multi-part tariffs and their welfare implications.

Model As in the classical nonlinear pricing framework, buyers differ in their demand for the product, represented by a scalar, privately known type, and have quasi-linear utility in money. Higher types correspond with higher demand for the product, embedded in the familiar assumptions that preferences are convex and satisfy strict single-crossing. The seller values transfers as well as usage-derived revenues from advertisement, data generation, network effects, and/or user addiction. To model non-contractibility of usage, we give buyers the ability to use less than they are allocated—for instance, if a buyer purchases the right to spend y hours on an online platform, they may choose to spend $x \leq y$ hours.

The seller chooses an arbitrary price schedule that assigns a price to each level of

[2] As we clarify in Section 6.3, pooling of many buyer types, as is standard under "ironed solutions," is unrelated to the issue of zero marginal pricing: under pooling, many buyers purchase the same amount of the good for the same price, but additional units of the good still have a strictly positive marginal price. Models with discrete buyer types or bang-bang solutions can also generate *weakly optimal* multi-part tariffs, but make the counterfactual prediction that no buyer ever consumes in the region with zero marginal prices.

[3] In Section 6.2.2, we discuss the case study of the pay-to-click AllAdvantage.com, and how modern legal infrastructure is designed to prevent platforms from compelling users to engage with advertisements.

purchases. Buyers first decide how much to purchase, and then what to use, within the scope of what is permitted by the contract. We study the problem of how to design the price schedule optimally, taking both of the buyers' decisions into account.

Optimal Pricing We characterize the seller's optimal pricing when buyers can freely dispose of what they purchase. The induced levels of buyer usage in the optimum cap the seller's preferred usage (i.e., what the seller would sell were usage fully contractible) with buyers' bliss points (i.e., what buyers would use were the product free).[4] The corresponding price schedule is flat whenever buyers consume their bliss points, since the marginal value of additional usage to the buyer is zero. Thus, sellers price according to *multi-part tariffs*.

We next explore the structure of multi-part tariffs. We show zero marginal pricing applies to more units of the good when there are greater marginal revenues from usage (e.g., from advertising) and smaller marginal information rents, or costs of screening. Intuitively, usage-derived revenue makes higher usage more attractive, while information rents distort down the amount of usage for all but the highest type.

The *shape* of these competing effects determines where in the price schedule zero-marginal-price regions emerge. Figure 6-1 previews four of the pricing schemes that our model can generate. *Free pricing*, in which all units have zero marginal price (e.g., Google search or Facebook), occurs when marginal revenues from usage globally dominate marginal information rents. *Freemium pricing* or *free trial pricing*, in which initial units of the good have zero marginal price and all subsequent "add-on" units have positive marginal price (e.g., the mobile game "Candy Crush Saga"), occurs when usage-derived revenues are highly concave (e.g., because unique users generate valuable data) and overwhelm information rents for only low-usage buyers. *Premium pricing*, in which initial units of the good are sold for a strictly positive marginal price and subsequent units are sold for zero marginal price (e.g., Amazon Prime Video), occurs when usage-derived revenue dominates information rents for only high-usage buyers. Indeed, as information rents vanish for the highest-usage buyers, globally positive marginal revenue from usage always generates a "premium tier." Thus, *hybrid pricing* schemes like the combination of a free trial and premium plan (e.g., *The Wall Street Journal*) can occur with sufficiently concave and increasing revenues from usage.

[4]Formally, we show this under the technical conditions that virtual surplus is strictly single-crossing in usage and agents' types, and that virtual surplus is strictly quasiconcave in usage. In Appendix F.2.1, we relax the first of these assumptions.

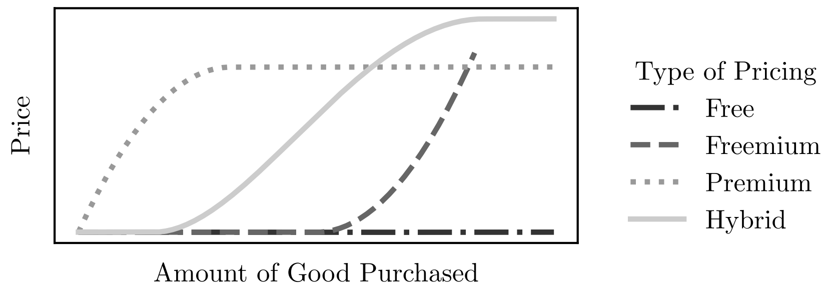


Figure 6-1: Example multi-part tariffs, which are derived in Section 6.4.

Welfare Implications We finally use our results to understand the effects on buyer welfare of both contractible usage and changes in the structure of demand and revenue. First, while non-contractibility improves buyer welfare under a fixed price schedule, all buyers would obtain greater welfare with perfectly contractible usage under the corresponding optimal price schedule. The intuition for this result is that the lack of contractibility of usage, which enables buyers to escape being forced to use the good, also prevents them from being paid compensating differentials for usage, which leads to forgone gains from trade. Second, in the absence of contractible usage, buyer welfare is less sensitive to changes in marginal usage-derived revenue. These results echo popular claims, made especially about social media products, that users are not fairly remunerated for "being the product, not the consumer."[5]

Related Literature The closest theoretical analysis to ours is by Grubb (2009), who demonstrates the optimality of three-part tariffs when selling to over-confident consumers who can freely dispose of the purchased good.[6] In Appendix F.3.3, we show how the setting of Grubb (2009) can be mapped to our framework, with overconfidence mapping to a particular kind of external revenue function. By considering a richer class of external revenue functions, our analysis endogenizes a richer set of pricing schemes and is applicable to a greater number of settings.

Our analysis relates to a literature on pricing under various constraints. Sundararajan (2004) studies non-linear pricing of information goods when the seller faces a "transaction cost" of measuring provision of the good, and derives an optimal tiered pricing schedule. Our analysis, by contrast, endogenizes multi-part tariffs as an op-

[5] As one example, Apple co-founder Steve Wozniak had the following to say about why he deleted his personal Facebook account: "[Facebook's] profits are all based on the user's info, but the users get none of the profits back [. . . .] As they say, with Facebook, you are the product" (Guynn and McCoy, 2018).

[6] Free disposal is also relevant for the sale of information goods, where agents may choose to optimally disregard information (see, e.g., Bergemann, Bonatti, and Smolin, 2018).

timal strategy in light of under-utilization. Amelio and Jullien (2012) and Choi and Jeon (2021) study markets with constraints for non-negative linear pricing and how bundling products across markets can effectively subvert such constraints. We complement this line of research by studying the *non-linear* pricing problem, albeit in a single market. Sartori (2021) studies the provision of goods that can be freely duplicated and damaged by the seller (as opposed to being freely disposed by the buyers). The optimal allocation in this setting exhibits inefficient bunching of types, without generating multi-part tariffs.

Our results fit into a theoretical literature on mechanism design with *ex post* moral hazard (e.g., Laffont and Tirole, 1986; Carbajal and Ely, 2013; Strausz, 2017; Gershkov, Moldovanu, Strack, and Zhang, 2021; Yang, 2022). However, given its focus on the possibility of under-utilization, our model of ex post moral hazard has a specific structure that admits tractable analysis and, at the same time, has previously not been analyzed.

Outline The rest of the paper proceeds as follows. Section 6.2 introduces our model. Section 6.3 solves for optimal contracts with free disposal. Section 6.4 studies the occurrence and structure of optimal multi-part tariffs. Section 6.5 studies the welfare implications of non-contractible usage. Section 6.6 concludes.

6.2 Model

6.2.1 Consumer Demand

There is a single good that can be bought and consumed in amounts $x \in X = [0, \overline{x}]$. There is a unit measure of consumers with privately known type $\theta \in \Theta = [0, 1]$ that parameterizes their demand. The type distribution $F \in \Delta(\Theta)$ admits a density f that is bounded away from zero on Θ. For example, x might be the time that an agent spends on an online platform, and θ shifts how much they enjoy this activity.

Consumers' type-specific preferences over consumption are represented by a twice continuously differentiable utility function $u : X \times \Theta \to \mathbb{R}$. We assume that higher types value consumption more and that all types have single-peaked preferences over consumption with the following three conditions: (i) u satisfies strict single-crossing in (x, θ);[7] (ii) for each $x \in X$, $u(x, \cdot)$ is monotone increasing over Θ; and (iii) for each $\theta \in \Theta$, $u(\cdot, \theta)$ is strictly quasiconcave over X. All consumer types value zero consumption the same as their outside option payoff, which we normalize to zero, or

[7] For avoidance of ambiguity, we mean that $u_{x\theta} > 0$, as per, for example, Nöldeke and Samuelson (2007).

$u(0, \theta) = 0$ for all types $\theta \in \Theta$. Agents have quasilinear preferences over consumption and money $t \in \mathbb{R}$, so their total payoff is $u(x, \theta) - t$.

6.2.2 Under-Utilization

The primary departure of our analysis from traditional non-linear pricing is the ability of consumers to under-utilize what they buy at zero cost. A consumer buying y can consume any $x \in [0, y]$. That is, they can *freely dispose* of purchased goods.

We argue that free disposal describes feasible contracting in digital goods markets, our primary application. For example, the *Wall Street Journal* can measure if a given consumer loads an article, but not if they actually read it. Likewise, Google can register that a search has been made, but not that this is done by a human as opposed to a bot.

Perhaps the most direct evidence for the non-contractibility of consumption is provided by the failure of "pay-to-click" internet businesses. One case study is the rise and fall of AllAdvantage.com, a venture that paid users to view a permanent banner ad when browsing the internet. *The New York Times*, who interviewed the company's founder as well as eager customers, asked "Can it Pay to Surf the Web?" in a July 1, 1999, headline (Guernsey, 1999). But AllAdvantage.com was quickly bogged down by users' finding simple ways to automate web surfing. This was concisely summarized in the headline of a *Wired* magazine article from July 10, 2000, which (unintentionally) answers the *Times*' original question: "It Pays to Cheat, Not to Surf" (Kang, 2000). In our model's language, consumers could purchase y hours of time browsing the internet with the AllAdvantage banner, but then under-utilize to consume $x \leq y$ hours, with the residual $y - x$ hours handled by bots.

6.2.3 Production and Revenues

The seller's revenue derives from two sources. The first is the total transfer from all buyers to the seller. The second is revenue that derives from consumers' usage of the good, net of production costs. This is represented by a continuously differentiable $\pi : X \times \Theta \to \mathbb{R}$.[8] The seller values zero consumption the same as their outside option revenue from not selling the product at all, which we normalize to zero, or $\pi(0, \theta) = 0$ for all $\theta \in \Theta$.

We have four primary justifications for usage-dependent revenue π for digital

[8]It may be reasonable to assume that production costs depend on the *allocation* rather than consumption of the good. But if costs are monotone, it is straightforward to argue that the seller will never produce more than is consumed and "waste" the product, leading to a representation of costs in terms of usage.

goods.

Advertisement Digital goods are commonly bundled with revenue-generating advertisements. For example, Google search results, Facebook social feeds, and *Wall Street Journal* articles all include advertisements. In these and other online settings, advertisers can directly measure both the number of times an advertisement is loaded (*impressions*) and the number of times an advertisement is clicked. Payments from the advertiser to the platform commonly depend *ex post* on both impressions and clicks per impression (*click through rate*).

We model these payments via our π as functions of platform consumption x, rather than purchases y—a *Wall Street Journal* user must load the article containing an advertisement, and perhaps click on it, to register a payment. We argue that the metaphor also applies when human and automated usage of online platforms may be substituted for one another. The aforementioned AllAdvantage failed because advertisers refused to pay out for inauthentic, bot-derived clicks. Modern advertisement contracts, having internalized the mistakes of the AllAdvantage era, tie payouts explicitly to "valid," human-derived clicks and impressions. As one example, the terms and conditions of Google AdSense, a popular service for adding advertisements to a webpage, define invalid activity as that "solicited or generated by payment of money, false representation, or requests for end-users to click on Ads or take other actions" (Google AdSense, 2020). For the erstwhile AllAdvantage.com, or any modern website monetized via AdSense, only human consumption x translates into revenue, while any bot-derived residual $y - x$ does not.

The function π is a possibly type-dependent mapping from usage to advertisement revenue. This subsumes details of consumer behavior and the advertisement contract, such as the rate with which consumers click advertisements and the payment per click.

Data Collection A trend in digital advertising over the last decade is the rise of *targeted advertisements* tuned toward individuals' interests as revealed by their online activity.[9] This phenomenon has helped open up a "data economy" in which producers profit from collecting information about consumers, either directly via selling data to marketing intermediaries or indirectly via applying data toward internal advertisement. As in the advertisement case, we use the function π to model the reduced-form "usage to revenue" schedule subsuming the translation of usage into data and the valuation of that data.

[9] A report by IHS Markit estimated that, in Europe, advertising that used behavioral data comprised 86% of all programmatic digital advertising (IHS Markit, 2017)

Network Effects Social media platforms, matching and networking services (e.g., Tinder and LinkedIn), online games (e.g., Fortnite and Candy Crush Saga), and content-streaming platforms with social rating systems (e.g., Netflix and Hulu) rely on active use to boost the appeal of their product. In Appendix F.3.1, we describe how a simple model in which platform externalities generate network effects can microfound an external revenue function π by affecting all agents' willingness to pay to participate in the platform. The framework can accommodate locally positive social externalities, as suggested by the previous examples, as well as negative externalities, due for instance to crowd-out or congestion.

Addiction Conventional wisdom and recent empirical evidence (Allcott, Gentzkow, and Song, 2022) suggest that addicted users are a major source of demand for phone apps and social media services. In particular, assume that the quantity $x' \in X$ purchased tomorrow by each consumer is increasing in their "addictive" consumption today $x \in X$, establishing an indirect link between current consumption x and future payments t'. Our revenue function π captures this effect in reduced form. In Appendix F.3.2, we illustrate how selling to myopic consumers with habit formation gives rise to an identical nonlinear pricing problem to the one we study, where π is the future revenue obtained by addicting agents today.[10]

6.2.4 The Nonlinear Pricing Problem

The seller's problem is to design a total revenue maximizing price schedule $T : X \to \overline{\mathbb{R}}$,[11] where $T(y)$ is the payment from a consumer purchasing $y \in X$. Following the choice of T, each type $\theta \in \Theta$ chooses whether to buy anything, how much to purchase $\xi(\theta) \in X$, and how much to ultimately consume $\phi(\theta) \in [0, \xi(\theta)]$. As is standard, we assume that the purchase and consumption functions $\xi : \Theta \to X$ and $\phi : \Theta \to X$ are the revenue-maximizing selections from the buyers' demand correspondence. Hence, the seller's problem can be formulated as:

$$
\begin{aligned}
\sup_{\phi,\xi,T} \quad & \int_{\Theta} \left(\pi(\phi(\theta),\theta) + T(\xi(\theta))\right) \mathrm{d}F(\theta) \\
\text{s.t.} \quad & \phi(\theta) \in \arg\max_{x \in [0,\xi(\theta)]} u(x,\theta) \quad \text{for all} \quad \theta \in \Theta && \text{(O)} \\
& \xi(\theta) \in \arg\max_{y \in X} \left\{ \max_{x \in [0,y]} u(x,\theta) - T(y) \right\} \quad \text{for all} \quad \theta \in \Theta && \text{(IC)} \\
& u(\phi(\theta),\theta) - T(\xi(\theta)) \geq 0 \quad \text{for all} \quad \theta \in \Theta && \text{(IR)}
\end{aligned}
\tag{198}
$$

[10] Our interpretations of our welfare results do not hold in this case, as π captures future consumption.

[11] We use $\overline{\mathbb{R}}$ to mean $\mathbb{R} \cup \{-\infty, \infty\}$.

The first constraint (O), or Obedience, establishes that each consumer θ chooses their optimal level of consumption $\phi(\theta)$ by optimally under-utilizing their initial purchase $\xi(\theta)$. The second constraint (IC), or Incentive Compatibility, embodies the consumers' optimal purchase $\xi(\theta)$, taking into account their subsequent ability to under-utilize. The final constraint (IR), or Individual Rationality, ensures that all consumers are willing to participate.[12]

6.3 Optimal Pricing

We first characterize the solutions of the seller's pricing problem (Equation 198). We then use a simple closed-form example, inspired by our digital goods applications, to both illustrate this result and show that it can imply the optimality of a multi-part tariff, a price schedule featuring units sold for a marginal price of zero.

6.3.1 Characterization of Optimal Pricing

We first define some important objects in which the optimal pricing schedule will be expressed. We define the consumer-optimal consumption of type θ as the function $\phi^A : \Theta \to X$:

$$\phi^A(\theta) = \arg\max_{x \in X} u(x, \theta) \tag{199}$$

This is unique and increasing because of the strict quasiconcavity of $u(\cdot, \theta)$ in x for all $\theta \in \Theta$ and strict single-crossing of u in (x, θ). In a digital-platform example, $\phi^A(\theta)$ is the amount of time that type θ would optimally spend on the platform were it freely available.

Second, we define the virtual surplus function $J : X \times \Theta \to \mathbb{R}$:

$$J(x, \theta) = \pi(x, \theta) + u(x, \theta) - \frac{1 - F(\theta)}{f(\theta)} u_\theta(x, \theta) \tag{200}$$

This is the total surplus $\pi + u$, net of the information rents required to ensure local incentive compatibility. For the remaining analysis, we will assume that the function J satisfies strict single-crossing in (x, θ) and is strictly quasiconcave in x. These standard technical assumptions guarantee that virtual surplus has a unique maximum and is maximized pointwise under the optimal contract, thereby ruling out cases with bunching, or multiple agent types' consuming the same bundle.[13] Therefore, there is

[12] Ensuring participation of all types is without loss of optimality for the seller owing to their ability to sell nothing and charge a price of zero, given the outside option of zero.

[13] In Appendix F.2.1, we relax the single-crossing assumption on J, strengthen strict quasiconcavity to strict concavity, and show how to adapt our analysis to settings which feature canonical bunching.

a unique, increasing producer-optimal consumption level $\phi^P : \Theta \to X$:

$$\phi^P(\theta) = \arg\max_{x \in X} J(x, \theta) \tag{201}$$

Continuing our example, a revenue-maximizing seller would induce a type-θ consumer to spend $\phi^P(\theta)$ time on the platform were usage perfectly contractible.

With these definitions in hand, we state our result describing optimal pricing:

Proposition 21 (Optimal Pricing)**.** *In any optimal contract, consumption is given by:*

$$\phi^* = \min\{\phi^P, \phi^A\} \tag{202}$$

The optimal price schedule for all $x \in [\phi^*(0), \phi^*(1)]$ *is uniquely given by:*[14]

$$T^*(x) = u(\phi^*(0), 0) + \int_{\phi^*(0)}^{x} u_x(z, {\phi^*}^{-1}(z))\,\mathrm{d}z \tag{203}$$

Moreover, purchases are part of an optimal contract if and only if they are a selection from the correspondence $\Xi_{\phi^*} : \Theta \rightrightarrows X$*:*[15]

$$\Xi_{\phi^*}(\theta) = \begin{cases} \{\phi^*(\theta)\} & \textit{if } \phi^*(\theta) < \phi^A(\theta) \\ \left[\phi^A(\theta), \inf_{\theta' \in [\theta, 1]} \left\{\phi^*(\theta') : \phi^*(\theta') < \phi^A(\theta')\right\}\right] & \textit{if } \phi^*(\theta) = \phi^A(\theta) \end{cases} \tag{204}$$

The proofs of this and all other results are in the Appendix. The first part of this result says that consumption in any optimal contract is the producer-optimal consumption level capped by the consumer-optimal consumption level. To understand this result, first observe that forcing a buyer to consume beyond their bliss-point level violates Obedience—they would avail of free disposal to consume less and reach their bliss point—and is therefore impossible. It is therefore *necessary* that $\phi \le \phi^A$. We moreover show that $\phi \le \phi^A$, combined with monotonicity of eventual consumption (also necessary for incentive compatibility by standard arguments), is *sufficient* for Obedience and Incentive Compatibility. Strict quasiconcavity of virtual surplus then implies that simply capping the optimum in the absence of free disposal with the bliss point is optimal.

[14]We define the inverse of a continuous consumption function $\phi : \Theta \to X$ on $[\phi^*(0), \phi^*(1)]$ as the standard inverse on $(\phi^*(0), \phi^*(1))$, extended to the boundaries with $\phi^{-1}(\phi^*(0)) = \lim_{x \to^+ \phi^*(0)} \phi^{-1}(x)$ and $\phi^{-1}(\phi^*(1)) = \lim_{x \to^- \phi^*(1)} \phi^{-1}(x)$. When $\phi^*(0) = \phi^*(1)$, we set $\phi^{-1}(\phi^*(0)) = 0$.

[15]With the convention that if the set over which the infimum is taken is empty, then the infimum is equal to the supremum of the codomain of the relevant objective function. For example, ∞ for $\mathbb{R}$, and 1 for $[0, 1]$.

Our argument generalizes and formalizes the following scenario. A digital platform like Twitter might want users to spend all 24 hours of the day reading content and seeing advertisements. But no contract could support this outcome—users would always prefer to use cheating software to mimic 24-hour usage, while reducing their actual consumption to their bliss point. This was exactly the problem that led to the downfall of AllAdvantage.com. A natural next choice is to design incentives to induce a feasible level of Twitter consumption (i.e., $\phi \leq \phi^A$) that maximizes revenue as well as possible given this constraint. When revenue is hump-shaped (i.e., strictly quasiconcave), this is achieved by enveloping over the bliss points ϕ^A and the revenue-maximizing points ϕ^P. This is what the optimal contract specifies.

The second part of the result derives an optimal price schedule that supports the optimum, using standard arguments that invoke the necessity of local Incentive Compatibility constraints. This price schedule is uniquely pinned down on the space of consumed amounts, $X^* = [\phi^*(0), \phi^*(1)]$, the image of Θ under the optimal consumption function. Away from this interval, there are many available options. For example, the seller can always offer a monotone price schedule by charging $T(\phi^*(0))$ for all $x < \phi^*(0)$ and setting an infinite price for all $x > \phi^*(1)$.[16] For the remainder of the analysis, we restrict attention to $T^* : X^* \to \mathbb{R}$ and study its properties in detail.

The third part of the result characterizes the set of type-specific purchases that support the optimum. This always includes simply setting purchases equal to consumption. However, when the constraint of free disposal binds, there are many possible levels of purchases which support the optimal allocation and yield the same total revenue for the seller. More precisely, for any interval of types for whom the constraint of under-utilization is binding, it is possible to have each type purchase any amount between their own bliss-point and the bliss-point of the highest type in that interval. An important implication of this multiplicity is that the seller may find it optimal not to offer every level of purchases. Concretely, a possible optimal solution for the seller is to offer only the maximum level of purchases for all types. Under this solution, the seller offers a tiered menu, which features discrete levels of provision, a property we explore in more detail in Section 6.4.

6.3.2 A Digital Pricing Example

We provide a more specialized intuition for the form of the optimal contract and foreshadow its pricing implications in a closed-form example.

[16] In Section 6.4.3, we also give sufficient conditions for optimal price schedules to be flat for $x \geq \phi^*(1)$, capturing the idea of unlimited subscriptions.

Example 4 (Digital Platform with Advertisement)**.** *A digital platform sells access time $x \in [0,1]$. Consumers have quadratic payoffs*

$$u(x,\theta) = \theta x - \frac{x^2}{2} \tag{205}$$

where θ is uniformly distributed on $[0,1]$. A consumer who spends time x on the platform clicks on $x(k-hx)$ advertisements, for $k>0$ and $h \in [0,k]$. The assumption $h \geq 0$ embodies user fatigue from seeing the same advertisements repeatedly, and $h \leq k$ ensures that total clicks remain positive. Each click yields revenue $p>0$ for the seller. Moreover, serving time x to the consumer has a constant marginal cost c. The seller's revenue function is therefore

$$\pi(x,\theta) = \left(\frac{\mathit{Revenue}}{\mathit{Click}} \cdot \frac{\mathit{Clicks}}{\mathit{Time}} - \frac{\mathit{Cost}}{\mathit{Time}}\right) \cdot \mathit{Time} = \alpha x - \frac{\beta}{2}x^2 \tag{206}$$

where we define $\alpha = pk - c$ and $\beta = 2ph$.

We can use Proposition 21 to solve for the seller's optimal pricing in this setting. The consumer-optimal and seller-optimal consumption levels, which respectively maximize payoffs u and virtual surplus J, are

$$\phi^A(\theta) = \theta \qquad \text{and} \qquad \phi^P(\theta) = \max\left\{0, \min\left\{1, \frac{\alpha + 2\theta - 1}{\beta + 1}\right\}\right\} \tag{207}$$

There are two possibilities for the shape of $\phi^* = \min\{\phi^A, \phi^P\}$. If $\beta \geq 1$, corresponding to high concavity of the revenue function (e.g., high user fatigue h), ϕ^A is optimal for sufficiently low types (if any). If $\beta < 1$, corresponding to low concavity of the revenue function (e.g., low user fatigue h), ϕ^A is optimal for sufficiently high types (if any). To most simply illustrate the result and its implications, we restrict attention to when $\alpha \leq 1$ and $\beta < 1$, in which case optimal consumption is

$$\phi^*(\theta) = \min\left\{\phi^A(\theta), \phi^P(\theta)\right\} = \begin{cases} \phi^P(\theta) & \text{if } \theta < \frac{1-\alpha}{1-\beta} \\ \phi^A(\theta) & \text{if } \theta \geq \frac{1-\alpha}{1-\beta} \end{cases} \tag{208}$$

To derive the optimal price schedule, we use the integral expression for prices (Equation 203) and simplify to obtain:

$$T^*(x) = \begin{cases} \frac{1-\alpha}{2}x - \frac{1-\beta}{4}x^2 & \text{if } x < \frac{1-\alpha}{1-\beta} \\ \frac{(1-\alpha)^2}{4(1-\beta)} & \text{if } x \geq \frac{1-\alpha}{1-\beta} \end{cases} \tag{209}$$

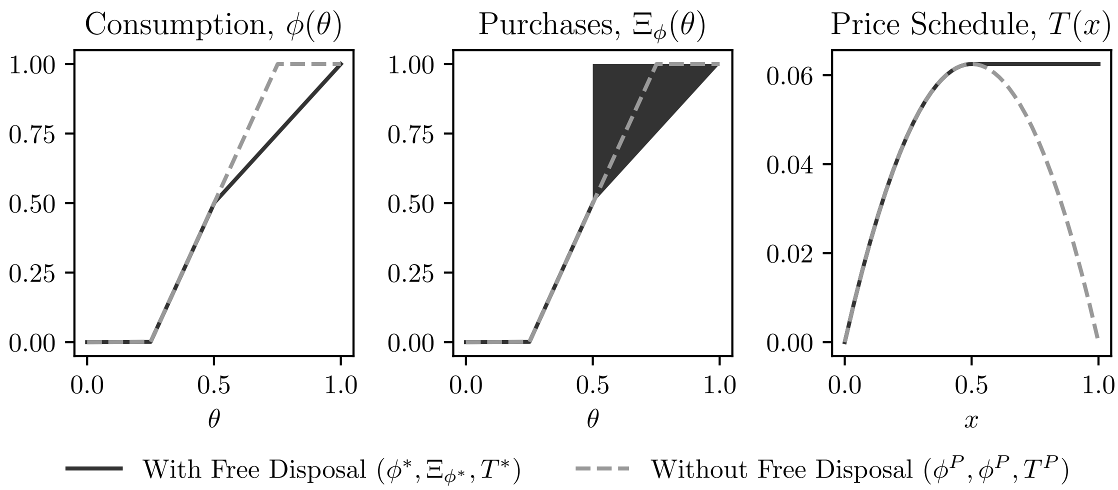


Figure 6-2: Optimal contracts with and without free disposal in Example 4.

We illustrate the properties of this optimal price schedule in Figure 6-2 when $\alpha = \frac{1}{2}$ *and* $\beta = 0$*. We show* $\phi^*(\theta)$ *in the leftmost panel and* $T^*(x)$ *in the right-most panel with a solid line. In the middle panel, we illustrate the purchase correspondence defined in Equation 204 as the shaded region. Under the optimal contract, the good is sold for a strictly positive marginal price until* $x = \frac{1}{2}$*, after which it is sold at a marginal price of zero. As a result, it is as if the seller offers buyers an unlimited subscription for the good at a fixed price, but allows them to buy less than this unlimited level for a discount.*

We now contrast these predictions with those in a variant setting with the same demand and external revenues, but no potential for free disposal (i.e., the legitimate time spent on the platform is perfectly contractible). The seller optimally sets $\phi^*(\theta) = \phi^P(\theta)$*. We illustrate, in the same case of* $\alpha = \frac{1}{2}$ *and* $\beta = 0$*, the consumption, purchases, and prices in the dashed lines of Figure 6-2. In this case, optimal prices are non-monotone. Starkly,* $x = 0$ *and* $x = 1$ *are both sold for free. This demonstrates the incentives for sellers in digital markets to use negative prices to induce valuable usage, which would be enforceable absent free disposal.*

6.4 The Occurrence and Structure of Multi-Part Tariffs

Having shown that multi-part tariffs can be optimal in an example, we now provide general conditions under which the optimal price schedule is a multi-part tariff. These conditions imply that the price schedule must be a multi-part tariff whenever marginal revenues from usage are strictly positive. We further provide sufficient conditions for optimal pricing to reduce to the freemium, premium and fixed cost pricing plans often observed in practice, and argue that these conditions realistically describe products with these pricing schemes.

6.4.1 The Occurrence of Multi-Part Tariffs

We first generally characterize when the optimal price schedule is a multi-part tariff. To this end, we now formally define flatness of price schedules and multi-part tariffs. A price schedule T is *flat* at x if there exists a neighborhood $O(x)$ such that T is constant on $O(x)$.[17] A price schedule T is a *multi-part tariff* if it is flat at some x. This definition builds on the conventional definition of a three-part tariff in which there is a fixed cost, an initial allotment of zero-marginal-price goods, and then additional units with positive marginal price.[18] Our more general definition can account for zero-marginal-pricing in multiple separate tiers and/or away from the "bottom."

To characterize the optimality of multi-part tariffs, we define the constrained marginal revenue function $H : X^* \to \mathbb{R}$ that maps outcomes to the marginal revenue of the seller for the type who most prefers that outcome:

$$H(x) = J_x\left(x, (\phi^A)^{-1}(x)\right) \tag{210}$$

To interpret H, imagine that the seller had to offer the product for free. The sign of $H(x)$ determines whether the seller would profit from more ($H(x) > 0$) or less ($H(x) < 0$) consumption from the type whose favorite consumption is x. The following result links flat pricing with this trade off.[19]

Proposition 22 (Multi-Part Tariffs)**.** *If $H(x) > 0$, then the optimal price schedule T^* is flat at x and, therefore, a multi-part tariff. Conversely, if the optimal price schedule T^* is a multi-part tariff that is flat at x, then $H(x) \geq 0$.*

To understand the intuition for this result, it is useful to re-write the sufficient

[17] Flatness of prices is more demanding than zero slope, or $T'(x) = 0$ wherever T' is defined, as it must apply on a neighborhood around x. Thus, zero marginal transfers or tariffs "at the top," meaning a specific maximal point in the type or action space, do not imply flatness by our definition—no unit of size $\varepsilon > 0$ is sold for zero price.

[18] A two-part tariff, via the conventional definition, combines fixed costs with positive marginal costs. This of course can be accommodated in the conventional non-linear pricing framework, with zero tiers, as the "intercept" of the tariff is a free parameter.

[19] Recall the definition of the inverse of a continuous consumption function from Footnote 14 in defining $\left(\phi^A\right)^{-1}$ on X^*. Moreover, it is not possible to strengthen the claim that if optimal price schedule T is flat at $x \in X^*$, then $H(x) \geq 0$ and make the inequality strict. The proof of Proposition 22 provides an explicit counterexample.

condition for flat pricing ($H(x) > 0$) as a comparison of two dueling economic forces:

$$\underbrace{f(\theta)\,\pi_x(x,\theta)}_{\text{marginal revenue}} > \underbrace{(1-F(\theta))\,u_{x\theta}(x,\theta)}_{\text{marginal information rent}} \tag{211}$$

where $\theta = (\phi^A)^{-1}(x)$. The first force (on the left-hand side) is the total marginal revenue from additional usage. Increasing π_x (e.g., from more intensive advertising) boosts this force. The second force (on the right-hand side) is the total marginal information rent paid to all higher-type consumers. Increasing complementarity or decreasing F in the hazard rate order enlarges these information rents.

When marginal revenue dominates marginal information rents, the seller would like to charge a negative marginal price to induce valuable usage. However, the constraint of free disposal makes this impossible. Thus, they do the next best thing and charge a price of zero on the margin. Conversely, under the opposite inequality, or $H(x) < 0$, marginal information rents dominate marginal revenues and the seller wishes for the buyer to consume less than their satiation point. Thus, free disposal poses no constraint, and they charge a positive marginal price to extract the buyer's (strictly positive) willingness-to-pay.

This result has two immediate implications for the possibility of multi-part tariffs. First, if marginal revenues from usage are everywhere negative ($\pi_x \leq 0$), multi-part tariffs are impossible. Thus, in situations with large production costs (e.g., physical goods), multi-part tariffs are unlikely to arise. Second, if marginal revenues from usage are everywhere strictly positive ($\pi_x > 0$), then the optimal price schedule *must* be a multi-part tariff. This claim follows because marginal information rents necessarily vanish for the highest types of agents, while marginal revenues (by assumption) are always strictly positive. Thus, in settings with low (marginal) production costs and valuable revenue from usage (e.g., digital goods), multi-part tariffs are likely to be optimal.

In the next section, we investigate the structure of multi-part tariffs. Before so doing, we first concretely illustrate why zero marginal pricing obtained in the case of Example 4 plotted in Figure 6-2. We also highlight how standard models cannot generate multi-part tariffs in two remarks.

Example 4 (continuing from p. 245)**.** *The constrained marginal revenue function is $H(x) = (\alpha - \beta x) - (1 - x)$, where the first term is the marginal revenue from additional usage and the second is the marginal increase of information rents. The presence of zero marginal pricing therefore relies on there being sufficiently high marginal adver-*

tising revenues (high α and low β), which derives in the underlying advertising model from a higher revenue per click or lower user fatigue. In the case plotted in Figure 6-2, we set $\alpha = \frac{1}{2}$ and $\beta = 0$. Thus, zero marginal pricing held when $H(x) > 0$ or $x > \frac{1}{2}$.

Remark 2 (Standard models do not generate multi-part tariffs)**.** Strictly optimal multi-part tariffs are not possible in the canonical Mussa and Rosen (1978) and Wilson (1993) screening models with a (convex) continuum of agent types.[20] A *weakly optimal* multi-part tariff is possible, for instance, in a specialization with a discrete number of types, by extending the domain of the offered menu with a constant price schedule. But we argue this is not economically meaningful, or relevant for our applications, for three reasons. First, no type would ever consume any outcome in the extended menu that does not lie in the initial menu, therefore ruling out variability of consumption within a pricing tier. This is clearly counterfactual—within the single tier of a zero-price product like Facebook, there is large variability in time spent on the platform. Second, there would be as many parts to the price schedule as unique consumer types, which is an arbitrary choice of the modeler. This prevents meaningful comparative statics for the number of observed tiers as a function of primitives. Third, there is no principled reason for arguing that the multi-part tariff is the "right" selection from the set of optimal price schedules which are extended off-menu.

Remark 3 (Bunching is unrelated to multi-part tariffs)**.** Optimal bunching in standard screening models is a different phenomenon from optimal multi-part tariffs. Intuitively, bunching is a feature that occurs in the type space Θ, where many different types buy the same amount. However, all units of the good are still sold at a strictly positive marginal price and so the optimal price schedule is never flat. In Appendix F.2.1, we solve for the optimal contract with bunching when J does not satisfy single-crossing in (x, θ) and is strictly concave in x by adapting the method of Nöldeke and Samuelson (2007). Example 4 provides an explicit illustration of how bunching is unrelated to the issue of zero marginal pricing: multiple buyers bunch on buying nothing, but there are still strictly positive marginal prices for the first marginal units of the good (see Figure 6-2).

6.4.2 Rationalizing Simple Pricing Schemes

We now leverage our characterization of multi-part tariffs to make theoretical predictions about the structure of pricing in various applications. We moreover argue that

[20] Even if the virtual surplus is not strictly concave and a multi-part tariff is weakly optimal, the same argument of this remark would apply.

these predictions line up with various forms of pricing that we observe in practice. We first define four common pricing schemes:

1. *Regular pricing*, in which all units of the good have strictly positive marginal price.

2. *Fixed pricing*, in which all units have the same price.

3. *Premium-tier pricing*, in which initial units have positive price until some point, after which all subsequent units have zero marginal price

4. *Introductory-offer pricing*, in which initial units have zero price and subsequent units have positive price

The last three of these are difficult to understand through the classical lens of non-linear pricing (see Remark 2). We now provide sufficient conditions, based solely on the case of our model where H crosses zero at most once, that delineate these pricing schemes in our model. They all immediately follow from Proposition 22.

Corollary 13 (Simple Pricing Schemes). *The following statements are true:*

1. *If $H(x) < 0$ for all $x \in X^*$, then optimal pricing is regular.*

2. *If $H(x) \geq 0$ for all $x \in X^*$, then optimal pricing is fixed.*

3. *If $H(x) < 0$ if and only if $x < \hat{x} \in int(X^*)$, then optimal pricing is of the premium-tier form, with threshold for strictly positive marginal prices given by $\hat{x}$.*

4. *If $H(x) \geq 0$ if and only if $x \leq \hat{x} \in int(X^*)$, then optimal pricing is of the introductory-offer form, with threshold for zero marginal prices given by $\hat{x}$.*

In several applications of this result, we now describe the model's theoretical predictions for pricing schemes based on plausible structures of external revenues. In each case, we argue that the model's predicted pricing scheme lines up with that which we observe in practice.

Case 1: Regular Pricing for Physical Goods Physical goods typically have significant production and/or transportation costs that swamp potential usage-based revenues (e.g., from word-of-mouth advertising). Seen through the lens of our model, this corresponds to a case with $\pi_x < 0$ and implies that $H < 0$. As per Corollary 13, our model therefore predicts that physical goods should feature regular pricing. This case is studied in the classical nonlinear pricing literature (Mussa and Rosen, 1978; Wilson, 1993) and of course matches the reality of how the majority of physical goods are priced.

Case 2: Fixed Pricing for Search Engines and Social Media Search engines and social media platforms, as we have motivated, derive large revenues from advertisement and data collection. Advertisement, in particular, often appears at a uniform rate during regular usage (e.g., sponsored search results in Google) and derives revenue per impression or click (i.e., marginal usage). Mapped to our model, $\pi_x > 0$ for all levels of usage and user types. Our model predicts that, if these revenues globally dominate information rents, then $H > 0$ and optimal pricing is fixed. Moreover, under the assumption that the lowest-type consumers would most prefer to consume nothing (e.g., make zero Google searches or spend zero minutes on Facebook), this fixed price is zero. This matches, of course, the observed pricing scheme of products like Google, Facebook, Twitter, and Instagram.

Case 3: Premium-Tier Pricing for Content-Streaming Content-streaming platforms, such as Amazon Prime, do not run external advertisements but instead use consumer data to fine-tune within-platform content recommendations. We might conjecture that the indirect revenues derived from this model are positive, but smaller than the sum of direct revenues earned by search engines and social media platforms from serving advertisements and selling data. Under our model, this corresponds to a case where $\pi_x > 0$, so $H(x) > 0$ for high levels of usage (where information rents are necessarily small, as we will formally clarify in Corollary 14), but $H(x) < 0$ for low levels of usage (where standard screening concerns dominate). Our model therefore predicts that marginal prices should initially be positive and free, in unlimited quantities, for high enough usage. This prediction matches Amazon Prime's streaming-library pricing, in which users can buy individual movies or shows for fixed prices or purchase the unlimited Prime subscription.

Case 4: Introductory-Offer Pricing for Cloud Storage, Mobile Games, and Networking Services A diverse class of goods derive a disproportionate amount of revenue from initial usage. For example, cloud storage platforms such as Dropbox and iCloud may benefit from a network externality, whereby all users find the product more valuable when they can share files with a larger number of unique people. A similar force is present for networking services platforms, such as LinkedIn. As a further example, mobile games (e.g., "Candy Crush Saga") can directly generate revenue from merely being opened, as they can scrape data points like the user's

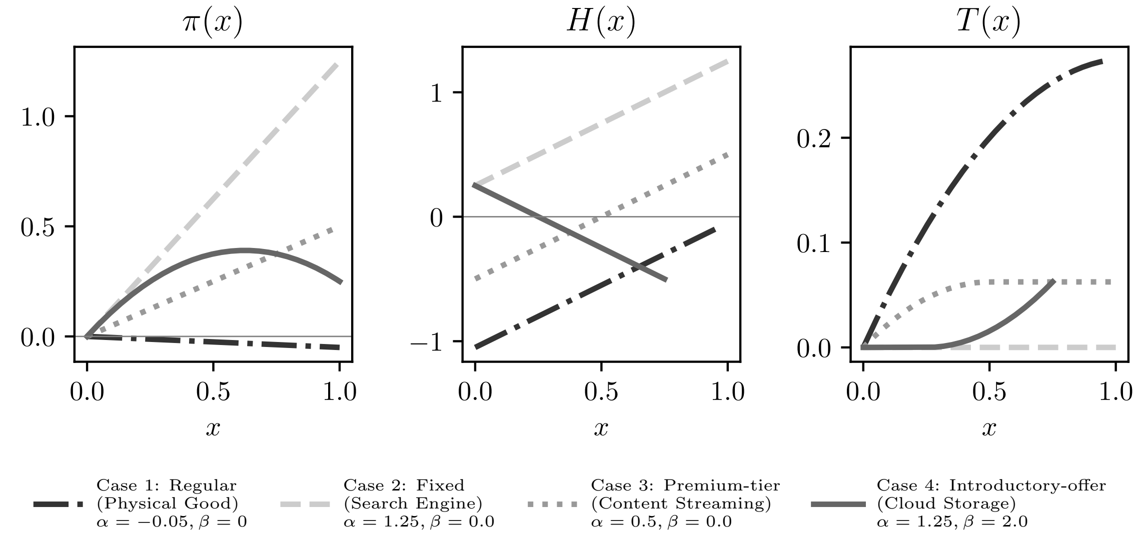


Figure 6-3: Cut-off price schedules in Example 4.

location and sell them to advertising firms. They are also likely to be addictive and feature diminishing marginal effects of past usage on future willingness-to-pay (i.e., the first minute of the game hooks the user more than the hundredth).[21] These settings can be represented in our model with marginal revenues that are diminishing ($\pi_{xx} < 0$), and potentially negative for high levels of usage. Because of this, $H(x) > 0$ is likely for low levels of usage (when marginal revenues are especially high) and $H(x) < 0$ is likely for high levels of usage (when marginal revenues are diminished). Thus, our model predicts an introductory offer of marginally free units. Moreover, if the lowest-type buyers would most prefer to consume nothing, then the introductory offer has no fixed cost either (i.e., it is a free trial). Dropbox features a free trial with limited space allotment, which can be increased with paid upgrades; LinkedIn has a free base version, which can be upgraded to a Premium mode with more features; and Candy Crush Saga gives the player a free allotment of lives, which can be replenished at a cost.

A Numerical Example We finally use Example 4 to illustrate these four pricing cases numerically, under different assumptions for the shape of the revenue function:

Example 4 (continuing from p. 245)**.** *The constrained marginal revenue function is* $H(x) = (\alpha - \beta x) - (1 - x)$*, which crosses zero at most once. The four cases of Corol-*

[21] Allcott, Gentzkow, and Song (2022) show that a model of total demand for popular phone apps, calibrated to empirical evidence, is consistent with habit formation, myopia, and diminishing returns.

lary 13 therefore summarize the possibilities for pricing. In Figure 6-3, we illustrate each of these cases. We plot the profit functions $\pi(x)$, constrained marginal revenue functions $H(x)$, and tariffs $T(x)$. Case 1 arises when marginal revenues from usage are sufficiently low or negative ($\alpha \leq 1$ and $\alpha \leq \beta$) as in our physical goods application. Case 2 occurs when marginal revenues from usage are always sufficiently high ($\alpha > 1$ and $\alpha > \beta$), as in our search engine and social media platforms applications. Case 3 obtains when marginal revenues from usage are intermediate and are dominated by information rents for low levels of usage ($\alpha \leq 1$ and $\alpha > \beta$), as in our content-streaming platforms applications. Case 4 occurs when there are initial high revenues from usage that then diminish sharply ($\alpha > 1$ and $\alpha \leq \beta$), as in our cloud storage, mobile games, and networking services applications.

6.4.3 More General Pricing Schemes: Unlimited Subscriptions, Trials, and Arbitrary-Part Tariffs

Our model also allows for richer pricing schemes that feature multiple regions of zero-marginal-pricing for various levels of usage. We first establish general sufficient conditions for the occurrence of zero-marginal-pricing at the bottom and top of the pricing schedule and link this to the co-occurrence of free trials and unlimited subscriptions for many goods. Finally, we illustrate in a stylized example how the model can be used to rationalize any number of regions of zero-marginal-pricing.

Unlimited Subscriptions and Trials A price schedule T features an *unlimited subscription* if it is flat at $\phi^*(1)$, the highest level of consumption under the optimal contract, and a *trial* if it is flat at $\phi^*(0)$, the lowest level of consumption under the optimal contract. These features were respectively present under the premium-tier pricing and introductory-offer pricing introduced in the previous subsection. We now derive general conditions under which unlimited subscriptions and trials are obtained in more general, possibly multi-tier pricing schedules:

Corollary 14 (Unlimited Subscriptions and Trials)**.** *The price schedule features an unlimited subscription if*

$$\pi_x(\phi^A(1), 1) > 0 \tag{212}$$

The price schedule features a trial if

$$f(0)\, \pi_x(\phi^A(0), 0) > u_{x\theta}(\phi^A(0), 0) \tag{213}$$

The first part shows that an unlimited subscription is optimal whenever marginal

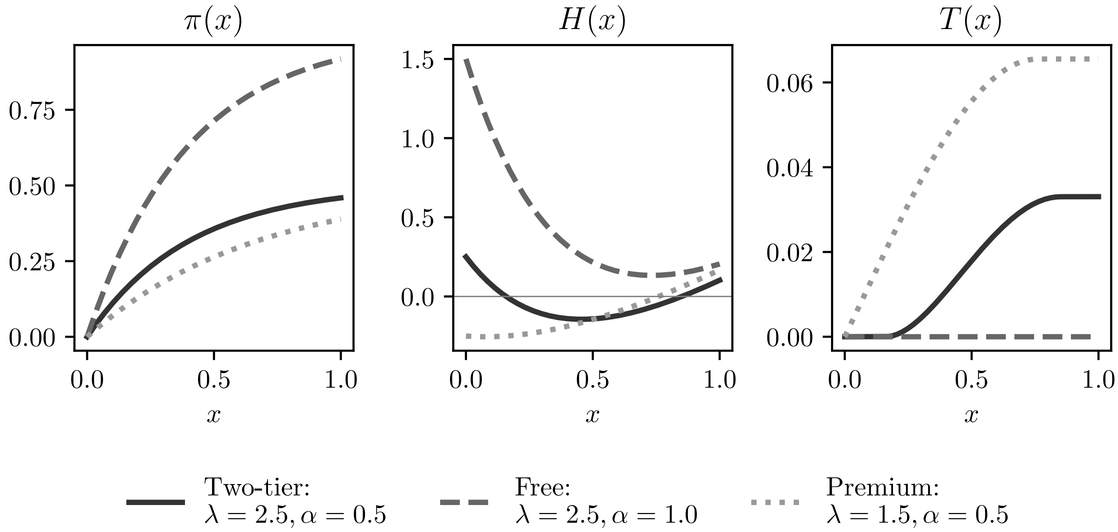


Figure 6-4: Hybrid price schedules in Example 5.

revenues from usage are positive when the highest type uses at their bliss-point level. As we previously discussed, a more demanding sufficient condition for this is that marginal revenues from usage are globally positive, as in our search engines, social media platforms, and content streaming applications. The lack of a countervailing force from marginal information rents reflects the fact that the seller does not distort allocations of the highest-type agents away from the first-best, surplus-maximizing allocation.

The second part shows that trial tiers are optimal whenever total marginal revenues from usage when the lowest type uses at their bliss-point level exceed the marginal information rent paid to *all* higher types. As we have highlighted, with zero or negative marginal revenues, this condition would never hold. Otherwise, it is more likely to hold in environments with high marginal revenues stemming from *low* levels of usage, as in our previous application to cloud storage, mobile games, and networking services.

These two conditions are mutually compatible. Thus, Corollary 14 opens the door to pricing schemes that feature both trials and unlimited subscriptions. We now provide a concrete example of this possibility and apply it to understand the pricing of online newspapers, such as the *Wall Street Journal* (wsj.com).

Example 5 (Optimal Pricing with Unlimited Subscriptions and Trials)**.** *Consumer preferences, the outcome space, and the type distribution are identical to those in Example 4. Advertisements are served at a constant rate, normalized to one. Con-*

sumers notice an advertisement according to an exponential process with hazard rate $\lambda > 0$, click on the first advertisement they notice, and ignore all subsequent advertisements. The seller earns α per click. The seller therefore receives the following payoff in expectation:

$$\pi(x,\theta) = \frac{\mathit{Revenue}}{\mathit{Click}} \cdot \mathit{Expected\ Clicks} = \alpha(1 - e^{-\lambda x}) \tag{214}$$

Thus, $H(x) = \lambda\alpha e^{-\lambda x} - (1-x)$, which has, depending on the values of (λ, α), either zero, one, or two tiers. The last case features both a trial and an unlimited subscription. In Figure 6-4, we illustrate all three possibilities.

The two-tier example, plotted with $\lambda = 2.5$ and $\alpha = 0.5$, satisfies both conditions in Corollary 14. First, the fact that marginal revenues remain positive for all levels of consumption guarantees an unlimited subscription, as this force dominates the vanishing information rents. Second, marginal revenues are high for low consumption because these agents are least likely to have already clicked an advertisement. In the two-tier calibration, this force dominates the marginal information rents that need to be paid to all higher-type agents. Moreover, because the lowest type most prefers zero consumption, the trial tier has zero price, and is therefore a "free trial."

This combination of a trial and an unlimited subscription is characteristic of online newspaper and content streaming platforms. For example, the *Wall Street Journal* features an initial allowance of free articles before charging a subscription fee for access to unlimited numbers of articles. This can be understood through the lens of our example as the optimal selling response to a situation in which readers' clicks generate revenue, but readers are unlikely to continue to click on advertisements multiple times.

Arbitrary-Part Tariffs In general, our results can be used to rationalize tariffs with an arbitrarily large number of flat regions. That is, if we define an $N+2$-*part tariff* as a price schedule with N flat-pricing intervals (in analogy with the definition of a three-part tariff as having one tier), our model can generate any $N+2$-part tariff for $N \in \mathbb{N}$.[22] To do this, one may simply construct a constrained marginal revenue function H that crosses zero the desired number of times. We show this constructively in the following (intentionally contrived) example:

Example 6 (Multi-Part Tariffs Can Have Arbitrarily Many Parts)**.** *Consumer preferences, the outcome space, and the type distribution are identical to those in Example*

[22] Formally speaking, we refer to a flat-pricing interval as a maximal interval $I \subset X^*$ such the price schedule is flat at all $x \in I$.

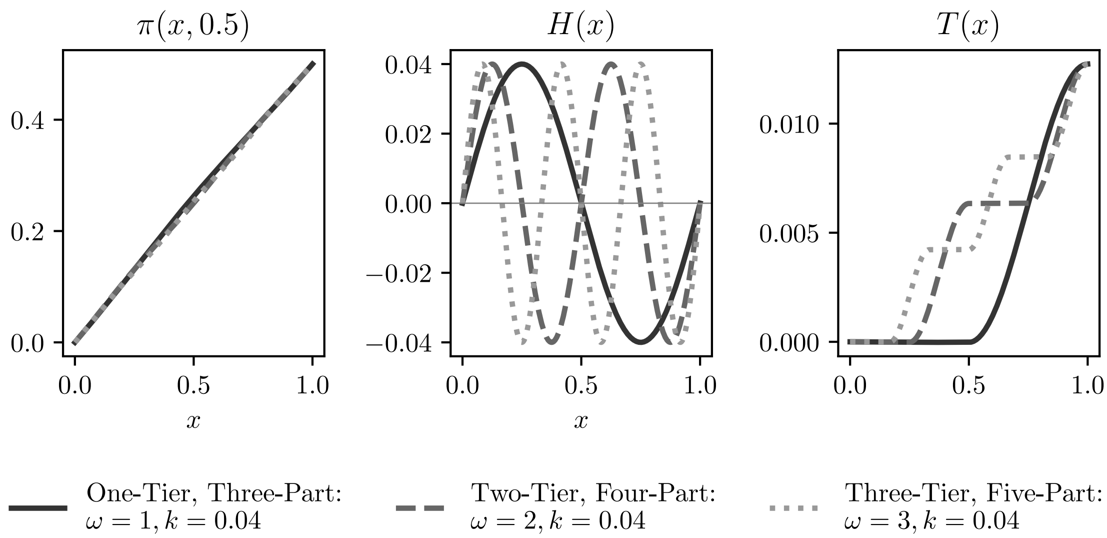


Figure 6-5: Arbitrary multi-part tariffs in Example 6.

4. The seller has a revenue function that increases in the usage of each type in proportion to how low that agent's type is (e.g., they derive particular value from targeting an advertisement at the most marginal users). Moreover, the value of attention paid by all types slightly oscillates over time. We capture this with the profit function:

$$\pi(x, \theta) = x(1 - \theta) - \frac{k}{2\pi\omega}(\cos(2\pi\omega x) - 1) \tag{215}$$

for some $\omega \in \mathbb{N}$ and $0 < k < \frac{1}{2\pi\omega}$.[23] *While the functional form of the cosine is intentionally ad hoc, observe in the left-most pane of Figure 6-5 that this revenue function only slightly deviates (for small k) from the linear baseline in inducing mild oscillations in revenue over time. The constrained marginal revenue function is $H(x) = k\sin(2\pi\omega x)$, which crosses zero from above ω times, generating an $\omega + 2$-part tariff. In Figure 6-5, we plot H and the optimal tariff T for $\omega \in \{1, 2, 3\}$ and $k = 0.04 < \frac{1}{6\pi}$. The tariffs respectively have one, two, and three plat-pricing intervals; these correspond to three-, four-, and five-part tariffs.*

6.5 Welfare Under Multi-Part Tariffs

Having studied the positive implications of under-utilization for pricing, we now study its welfare consequences. First, we study the effect of introducing under-utilization, or removing perfect contractibility, on consumer and producer welfare. Second, we

[23] Under this condition, it is immediate to verify that the induced virtual surplus function J satisfies our running assumptions: it is strictly concave and satisfies strict single-crossing.

study how the possibility for under-utilization mediates the welfare effects of changes in the structure of revenue. We apply these results to understand welfare in digital goods markets.

6.5.1 The Impact of Under-Utilization on Welfare

For an arbitrary price schedule T, we define *consumer welfare* under free disposal for each type θ as:

$$V(\theta;T) = \sup_{y\in X, x\in[0,y]} \{u(x,\theta) - T(y)\} \tag{216}$$

which is the payoff corresponding to an optimal purchase. We define *producer welfare with free disposal* for each purchasing type θ, $\Pi(\theta;T)$, as

$$\Pi(\theta,T) = \pi(\phi(\theta;T),\theta) + T(\xi(\theta;T)) \tag{217}$$

which is simply total revenue (i.e., the integrand in the objective in Problem 198), evaluated at a fixed T and the buyers' optimally chosen consumption $\phi(\theta;T)$ and purchases $\xi(\theta;T)$. We define $V^*(\theta)$ and $\Pi^*(\theta)$ as the corresponding equilibrium welfare quantities evaluated at the optimal price schedule, T^*. We finally define analogous quantities under no free disposal, or perfect contractibility, as $V_N(\theta;T)$, $\Pi_N(\theta;T)$, $V_N^*(\theta)$, and $\Pi_N^*(\theta)$.

The following result summarizes how consumer and producer surplus are affected by the possibility of under-utilization of consumption.

Proposition 23 (Contractibility and Welfare)**.** *For any price schedule T, $V(\theta;T) \geq V_N(\theta;T)$ for all $\theta \in \Theta$. However, under the optimal price schedules, $V^*(\theta) \leq V_N^*(\theta)$ and $\Pi^*(\theta) \leq \Pi_N^*(\theta)$ for all $\theta \in \Theta$.*

All consumers lose out from increased contractibility for any given pricing policy, as the scope for payoff-increasing actions declines; but all consumers *gain* from increased contractibility under the seller's reoptimized price schedule. Intuitively, consumers would like to commit *ex ante* to avoid the possibility of moral hazard and have the seller pay them to take certain actions; increasing contractibility provides exactly this commitment device. Sellers gain from increased contractibility (from selling to all types and, therefore, in total across types) for the simple reason that it increases the set of implementable allocations.

This result contextualizes the claim that users, particularly of social media platforms, are not fairly remunerated, given the fact that their time and data are the sources of these platforms' revenues. This idea is exemplified by Apple co-founder

Steve Wozniak's stated rationale for why he deleted his personal Facebook account (Guynn and McCoy, 2018):

> Users provide every detail of their life to Facebook and [...] Facebook makes a lot of advertising money off this. [...] The profits are all based on the user's info, but the users get none of the profits back. [...] As they say, with Facebook, you are the product.

This quote can be understood through the lens of our model. In light of the inherent non-contractibility of usage of social media platforms, consumers cannot be paid on the margin. In this context, Proposition 23 implies that consumers would be better off were their usage perfectly contractible and revenue-generative usage remunerated. However, our rationalization highlights that this issue is fundamentally technological: even if Facebook wished to pay users, they would not be able to do so without being exploited in the manner that led to the downfall of AllAdvantage.

6.5.2 Comparative Statics for Welfare

We next study how the presence of free disposal mediates the welfare effects of changes in usage-based revenues π and demand F. Changes in revenue may be driven by underlying changes in advertisement and data-collection technology while changes in demand may be driven by changes in demographics and the quality of goods.

We say that $\tilde{\pi}$ *exhibits more profitability of usage* than π if every unit of usage generates more revenue, that is, $\tilde{\pi}_x \geq \pi_x$. We say that $\tilde{F}$ *exhibits weaker demand* than F if F dominates $\tilde{F}$ in the hazard-rate order.[24] We write consumer and producer equilibrium welfare as functions of these arguments. We now show how these two comparative statics affect welfare, with and without free disposal:

Proposition 24 (Comparative Statics for Welfare)**.** *If $\tilde{\pi}$ exhibits more profitability of usage than π and $\tilde{F}$ exhibits weaker demand than F, then, for all $\theta \in \Theta$:*

$$0 \leq V^*(\theta; \tilde{\pi}, \tilde{F}) - V^*(\theta; \pi, F) \leq V_N^*(\theta; \tilde{\pi}, \tilde{F}) - V_N^*(\theta; \pi, F) \tag{218}$$

$$0 \leq \Pi^*(\theta; \tilde{\pi}, \tilde{F}) - \Pi^*(\theta; \pi, F) \leq \Pi_N^*(\theta; \tilde{\pi}, \tilde{F}) - \Pi_N^*(\theta; \pi, F) \tag{219}$$

This shows that both consumer and producer welfare increase, but less so than under perfect contractibility. In this sense, free disposal erodes the potential welfare

[24] Our nomenclature reflects the fact the hazard-rate order implies first-order stochastic dominance (See Theorem 1.B.1 in Shaked and Shanthikumar, 2007) and is therefore a (strong) notion of decreasing demand.

gains for both buyer and seller relative to a counterfactual perfect-contractibility world.

The intuition for this result is that, with higher marginal revenue from consumption for the seller or lower demand types (implying smaller distortions from information rents), the seller-optimal consumption is larger. This increases consumer and producer welfare, with and without free disposal. However, with free disposal, there are more types who consume their bliss point. Thus, both buyers and sellers have greater welfare gains under perfect contractibility of usage than under free disposal. This further underscores the sense in which perfect contractibility simulates commitment. Proposition 23 shows how this commitment increases welfare in *levels*; Proposition 24 shows how this commitment increases the *sensitivity* of welfare gains from changes in external revenue and demand.

This result allows us to shed light on the welfare effects of changes in advertising technology, such as the past decade's advent of more valuable targeted advertisements based on user data. We can capture this phenomenon as an increase in the profitability of usage for both platforms serving advertisements and platforms profiting from data sales. Through the lens of our model, any marginal increase in profitability, including the introduction of advertisement- or data-based revenue to a previously zero-revenue ($\pi = 0$) service, increases welfare for all consumers. These increases could be larger, however, if digital goods' usage were fully contractible. The reason is precisely the fact that users cannot be paid on the margin for their clicks and data. Moreover, even if only a subset of users are paid for their clicks and data, all users benefit due to changes in the overall price schedule.

This analysis has two potentially relevant implications for policy discussion about digital-goods regulation. First, government regulations that may diminish marginal advertising revenues, like the European Union's General Data Privacy Regulation, may weakly reduce consumer surplus. But the extent of these reductions, and their distributional consequences, depends on whether sellers are far on the interior of the zero marginal pricing constraint. That is, if marginal units are already free (e.g., for all units of Google or Facebook), then consumer surplus is unaffected and any (here, unmodeled) gains from the instrumental value of privacy lead to net consumer benefits.

Second, our focus on contractibility contrasts with an influential perspective in the literature that focuses instead on the lack of collective bargaining for users of online products (Posner and Weyl, 2018). Taken to the extreme, our results imply a thorny "privacy paradox" for consumers—the only way to properly reap the benefits of the surplus generated by targeted advertising is to surrender *additional* privacy by enabling tools to more precisely monitor usage and attention.

6.6 Conclusion and Summary of Extensions

We study optimal nonlinear pricing in environments with feasible under-utilization and usage-derived revenue, features that are ubiquitous in the digital goods context. We show how the combination of these two forces rationalizes the occurrence of multi-part tariffs, or price schedules that include at least one tier of zero marginal prices. The key mechanism is that sellers have an incentive to pay users on the margin, due to usage-derived revenue (e.g., from advertisement or data collection), but non-contractibility prevents such arrangements from being enforceable. More succinctly, zero marginal pricing is the sellers' constrained optimum in a world in which "pay to click" is impossible.

We apply these results to study positive and normative features of digital goods markets. We show how different structures of external revenue translate into different familiar pricing schemes like free trials, unlimited subscriptions, and free products. We moreover show that the scope for under-utilization reduces buyers' welfare and dampens the ability of buyers to reap the rewards from the revenue they generate.

We finally discuss additional analyses contained within the Appendix and ongoing work.

Optimal Bunching We assumed that virtual surplus was strictly single-crossing to rule out the possibility that multiple buyer types optimally bunch on the same level of consumption. In Appendix F.2.1, we relax this assumption, adapt the assignment approach of Nöldeke and Samuelson (2007), and characterize the optimal contract when bunching is a possibility. Analogously to Proposition 21, the optimal consumption function maximizes the suitable transformation of virtual surplus in this setting, subject to the constraint that no agent consumes more than their bliss point. Under this solution and as per our main analysis (Proposition 22), the price schedule is flat whenever this constraint binds and the optimal price schedule is a multi-part tariff.[25]

Under-Utilization with (Perfect) Competition It is of course natural to ask how our analysis in a monopoly setting would extend to markets with competition. In Appendix F.2.2, we solve for the equilibrium outcome of our screening model when

[25] In this setting, we cannot develop a similar characterization to that of Proposition 22 by comparing the marginal benefits and costs of additional consumption for the seller as the objective is no longer quasiconcave and global properties determine whether the constraint binds.

the monopolist faces a perfectly competitive fringe of potential entrants. Specifically, we model this free entry by adding an additional constraint imposing zero profits for the monopolist (i.e., total external revenues plus transfers). The equilibrium price schedule under perfect competition features zero marginal pricing more often than under monopoly pricing (Corollary 18 in Appendix F.2.2). The reason is that the perfect competition, modeled this way, leads firms to maximize total surplus instead of virtual surplus—put differently, competition ensures that all profits from screening and distorting down consumption are competed away. Total surplus is maximized by a higher level of consumption owing to the absence of information rents, and so the constraint of under-utilization is more often binding. As a result, our model implies that multi-part tariffs are likely to be more prevalent in scenarios with fierce competition between sellers than under monopoly.[26]

Pricing with Partially Contractible Usage In ongoing work, we study the effects of *partial* contractibility of usage on optimal nonlinear pricing. This allows us to capture further settings in which some levels of utilization are contractible (e.g., usage by high-end users) while others are not. In this model, agents switch endogenously from being consumers, who either pay for a good or receive it for free, to being workers, who are paid for their usage of a platform. This model can be used, for example, to capture the pricing of YouTube and TikTok, where low-end content creators use the platforms for free, while high-end content creators are paid.

[26]Another setting that features the same pricing implications as perfect competition, yet with a single monopolistic seller, is one in which buyers commit to participate before learning their types. This structure corresponds to the *ex ante* contracting setting of Grubb (2009).

Part III

Appendices

Appendix A

Appendix to The Macroeconomics of Narratives

A.1 Omitted Derivations and Proofs

A.1.1 Derivation of Equations 3 and 5

We first provide two assumptions under which Equation 3 holds as a linear approximation with a quadratic error bound. In what follows, we impose the technical requirements that $\mathcal{X}$ is convex, compact subset of $\mathbb{R}$, $\mathcal{Y}$ is a convex, compact subset of $\mathbb{R}^n$, Θ is a convex, compact subset of $\mathbb{R}^m$, and Ω is a convex, compact subset of $\mathbb{R}^r$. We first assume regularity conditions on payoffs to ensure the sufficiency of first-order conditions for optimality:

Assumption 12. *The utility function u is strictly concave and twice continuously differentiable. The conjectured aggregate outcome function $\hat{Y}$ is continuously differentiable.*

We next assume that agents' information about the random fundamentals is generated by location experiments that are conditionally independent of the fundamental, agents' preference shifters, and the narratives held by agents. We moreover assume that all narratives are equally sensitive to information.

Assumption 13. *The agents' information sets are generated by location experiments, i.e., $s_{it} = \theta_t + \nu_{it}$, where ν_{it} is a zero-mean random variable that is independent of θ_t, ω_i, and λ_{it}. Moreover, conditional on the signal, the conditional expectation of θ_t under each narrative k is given by $\mathbb{E}_k[\theta_t|s_{it}] = \kappa s_{it} + (1-\kappa)\mathbb{E}_k[\theta_t] + O(||s_{it}||^2)$.*

A sufficient condition for this assumption to hold with no approximation error is that all fundamentals and signals are Gaussian and the signal-to-noise ratio of the signal is constant across all narratives. Under these two assumptions, we can derive the form of the regression equation and that, modulo any misspecification error, the conditional expectation function is linear.

Proposition 25. *Under Assumptions 12 and 13, we have that:*

$$x_{it} = \gamma_i + \chi_t + \sum_{k \in \mathcal{K}} \delta_k \lambda_{k,it} + \varepsilon_{it} + O(||(x_{it}, Y_t, \theta_t, Q_t, \omega_i, \nu_{it}, \lambda_{it})||^2) \tag{220}$$

where ε_{it} is a zero mean random variable that is uncorrelated with γ_i, χ_t and λ_{it}. Thus, net of the misspecification error, the conditional expectation function is given by:

$$\mathbb{E}[x_{it}|i, t, \lambda_{it}] = \gamma_i + \chi_t + \sum_{k \in \mathcal{K}} \delta_k \lambda_{k,it} \tag{221}$$

Proof. By Assumption 12, from the agents' problems (Equation 2), their best replies must solve the following first-order condition (where we suppress all individual and time subscripts):

$$\mathbb{E}_{\pi_\lambda}\left[u_x(x, \hat{Y}(\theta), \theta, \omega)|s\right] = 0 \tag{222}$$

We linearize this first-order condition in (x, Y, θ, ω) around values $(\bar{x}, \bar{Y}, \bar{\theta}, \bar{\omega})$ which satisfy $u_x(\bar{x}, \bar{Y}, \bar{\theta}, \bar{\omega}) = 0$. This gives

$$\mathbb{E}_{\pi_\lambda}\left[u_{xx}(x - \bar{x}) + u'_{xY}(Y - \bar{Y}) + u'_{x\theta}(\theta - \bar{\theta}) + u'_{x\omega}(\omega - \bar{\omega})|s\right] + R = 0 \tag{223}$$

where $(u_{xx}, u_{xY}, u_{x,\theta}, u_{x\omega})$ are constants equal to the corresponding derivatives evaluated at $(\bar{x}, \bar{Y}, \bar{\theta}, \bar{\omega})$, the remainder R is $O(||(x, Y, \theta, \omega, \nu, \lambda)||^2)$. We can rearrange the above, and use the fact that ω is known to the agent, to write:

$$x = \bar{x} + \frac{1}{|u_{xx}|} u'_{x\omega}(\omega - \bar{\omega}) + \frac{1}{|u_{xx}|} \mathbb{E}_{\pi_\lambda}\left[u'_{xY}(Y - \bar{Y}) + u'_{x\theta}(\theta - \bar{\theta})|s\right] + \frac{1}{|u_{xx}|} R \tag{224}$$

Moreover, we know that $\hat{Y} = \hat{Y}(Q, \theta)$. Thus, assuming that $\hat{Y}$ is continuously differentiable, we may linearize $Y = \bar{Y} + Y'_Q(Q - \bar{Q}) + Y'_\theta(\theta - \bar{\theta}) + \hat{R}$, where $\bar{Y} = \hat{Y}(\bar{Q}, \bar{\theta})$ and $\hat{R}$ is the error induced by the approximation of $\hat{Y}$, which is $O(||Y, Q, \theta||^2)$. Substituting this approximation into Equation 224 gives

$$
\begin{aligned}
x &= \bar{x} + \frac{1}{|u_{xx}|}u'_{x\omega}(\omega - \bar{\omega}) + \frac{1}{|u_{xx}|}\mathbb{E}_{\pi_\lambda}\left[u'_{xY}(Y'_Q(Q-\bar{Q}) + Y'_\theta(\theta - \bar{\theta})) + u'_{x\theta}(\theta - \bar{\theta})|s\right] + \tilde{R} \\
&= \gamma + \tilde{\chi} + \mathbb{E}_{\pi_\lambda}\left[\tilde{\theta}|s\right] + \tilde{R}
\end{aligned}
\tag{225}
$$

where $\gamma = \bar{x} + \frac{1}{|u_{xx}|}u'_{x\omega}(\omega - \bar{\omega}) - \frac{1}{|u_{xx}|}u'_{xY}(Y'_Q\bar{Q} + Y'_\theta\bar{\theta}) - \frac{1}{|u_{xx}|}u'_{x\theta}\bar{\theta}$, $\tilde{\chi} = \frac{1}{|u_{xx}|}u'_{xY}Y'_Q Q$, and $\tilde{\theta} = \frac{1}{|u_{xx}|}\left(u'_{xY}Y'_\theta + u'_{x\theta}\right)\theta$, $\tilde{R} = \frac{1}{|u_{xx}|}R + \hat{R}$.

We next re-write the conditional expectation of $\tilde{\theta}$ as linear in two arguments, the signal s and prior mean $\mathbb{E}_{\pi_\lambda}[\tilde{\theta}]$. Using first the linearity of the expectation operator and the linearity of forming beliefs from narratives, and second Assumption 13 to re-write each narrative-specific conditional expectation in terms of the signal and the prior, we write

$$
\mathbb{E}_{\pi_\lambda}\left[\tilde{\theta}|s\right] = \sum_{k\in\mathcal{K}}\lambda_k\mathbb{E}_k[\tilde{\theta}|s] = \kappa\tilde{s} + (1-\kappa)\mathbb{E}_{\pi_\lambda}[\tilde{\theta}] + \check{R} \tag{226}
$$

where $\mathbb{E}_{\pi_\lambda}[\theta] = \sum_{k\in\mathcal{K}}\lambda_k\mathbb{E}_k[\theta]$ is the average prior mean across narratives, $\check{R}$ is the error induced by the approximation and is $O(||(\theta,\nu,\lambda)||^2)$, and the transformed signal is $\tilde{s} = \frac{1}{|u_{xx}|}\left(u'_{xY}Y'_\theta + u'_{x\theta}\right)s = \tilde{\theta} + \tilde{\nu}$, with $\tilde{\nu} = \frac{1}{|u_{xx}|}\left(u'_{xY}Y'_\theta + u'_{x\theta}\right)\nu$ independent of $\tilde{\theta}$ and of mean zero by Assumption 13. Defining $\chi = \tilde{\chi} + \kappa\tilde{\theta}$, $\varepsilon = \kappa\tilde{\nu}$ and $\delta_k = (1-\alpha)\mathbb{E}_k[\tilde{\theta}]$, we may write:

$$
x = \gamma + \chi + \sum_{k\in\mathcal{K}}\delta_k\lambda_k + \varepsilon + \bar{R} \tag{227}
$$

where $\bar{R} = \tilde{R} + \check{R} = O(||(x,Y,\theta,Q,\omega,\nu,\lambda)||^2)$. Re-introducing subscripts, we have ω_i, Q_t, $\tilde{\theta}_t$, $\lambda_{k,it}$ and ε_{it}. Thus, we have the claimed regression equation:

$$
x_{it} = \gamma_i + \chi_t + \sum_{k\in\mathcal{K}}\delta_k\lambda_{k,it} + \varepsilon_{it} + O(||(x_{it},Y_t,\theta_t,Q_t,\omega_i,\nu_{it},\lambda_{it})||^2) \tag{228}
$$

As ε_{it} has zero mean and is uncorrelated with γ_i, χ_t and λ_{it}, the claimed formula for the conditional expectation function follows. □

We now turn to the narrative updating rule. We impose the following assumption:

Assumption 14. *The updating rule P is continuously differentiable.*

We finally derive Equation 5 under this condition:

Proposition 26. *Under Assumption 14, we have that:*

$$\mathbb{P}[\lambda_{it} = \lambda | \lambda_{i,t-1}, Y_{t-1}, Q_{t-1}] = \zeta_\lambda + u'_\lambda \lambda_{i,t-1} + r'_\lambda Y_{t-1} + s'_\lambda Q_{t-1} + O(||(\lambda_{i,t-1}, Y_{t-1}, Q_{t-1})||^2) \tag{229}$$

Proof. By definition we have that $\mathbb{P}[\lambda_{it} = \lambda | \lambda_{i,t-1}, Y_{t-1}, Q_{t-1}] = P_\lambda(\lambda_{i,t-1}, Y_{t-1}, Q_{t-1})$. Linearizing this expression under Assumption 14, we immediately have:

$$\mathbb{P}[\lambda_{it} = \lambda | \lambda_{i,t-1} = \lambda', Y_{t-1}, Q_{t-1}] = \zeta_\lambda + u_{\lambda,\lambda'} + r'_\lambda Y_{t-1} + s'_\lambda Q_{t-1} + O(||(\lambda_{i,t-1}, Y_{t-1}, Q_{t-1})||^2) \tag{230}$$

Completing the proof. □

A.1.2 Proof of Proposition 1

Proof. We guess and verify that there exists a unique quasi-linear equilibrium. That is, there exists a unique equilibrium of the following form for some parameters $a_0, a_1 \in \mathbb{R}$ and function $f : [0,1] \to \mathbb{R}$:

$$\log Y(\theta, Q) = a_0 + a_1 \log \theta + f(Q) \tag{231}$$

To verify this conjecture, we need to compute best replies under this conjecture and show that when we aggregate these best replies that the conjecture is consistent and, moreover, that it is consistent for a unique triple (a_0, a_1, f).

From the arguments in the main text, we need to compute two objects: $\log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right]$ and $\log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right]$. We can compute the first object directly. Conditional on a signal s_{it} and a narrative weight λ_{it}, we have that the distribution of the aggregate component of productivity is:

$$\log \theta_t | s_{it}, \lambda_{it} \sim N\left(\kappa s_{it} + (1-\kappa)(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P), \sigma^2_{\theta|s}\right) \tag{232}$$

by the standard formula for the conditional distribution of jointly normal random variables, where:

$$\kappa = \frac{1}{1 + \frac{\sigma_e^2}{\sigma_\theta^2}} \quad \text{and} \quad \sigma^2_{\theta|s} = \frac{1}{\frac{1}{\sigma_\theta^2} + \frac{1}{\sigma_e^2}} \tag{233}$$

with κ being the signal-to-noise ratio and $\sigma^2_{\theta|s}$ the variance of fundamentals conditional on the signal. Thus, idiosyncratic productivity has conditional distribution given by:

$$\log \theta_{it} | s_{it}, \lambda_{it} \sim N\left(\log \gamma_i + \kappa s_{it} + (1-\kappa)(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P), \sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) \tag{234}$$

where we will denote the above mean by μ_{it} and variance by η^2. Hence, rewriting and using the moment generating function of a normal random variable, we have that:

$$\begin{aligned} \log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right] &= \log \mathbb{E}_{it}\left[\exp\left\{-\frac{1+\psi}{\alpha}\log\theta_{it}\right\}\right] \\ &= -\frac{1+\psi}{\alpha}\mu_{it} + \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\eta^2 \end{aligned} \tag{235}$$

Under our conjecture (Equation 231), we can moreover compute:

$$\begin{aligned} \log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right] &= \log \mathbb{E}_{it}\left[\exp\left\{\left(\frac{1}{\epsilon}-\gamma\right)(a_0 + a_1\log\theta_t + f(Q_t))\right\}\right] \\ &= \left(\frac{1}{\epsilon}-\gamma\right)[a_0 + a_1(\mu_{it} - \log\gamma_i) + f(Q_t)] + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\left[\eta^2 - \sigma_{\tilde{\theta}}^2\right] \end{aligned} \tag{236}$$

Thus, we have that best replies under our conjecture are given by:

$$\begin{aligned} \log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\mu_{it} - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\eta^2 \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)[a_0 + a_1(\mu_{it}-\log\gamma_i) + f(Q_t)] + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\left[\eta^2-\sigma_{\tilde{\theta}}^2\right]\Bigg] \end{aligned} \tag{237}$$

To confirm the conjecture, we must now aggregate these levels of production and show that they are consistent with the conjecture. Performing this aggregation we have that:

$$\begin{aligned} \log Y_t &= \log\left[\left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}}\right] \\ &= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}\right] \\ &= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right] \end{aligned} \tag{238}$$

Moreover, expanding the terms in Equation 237, we have that:

$$\begin{aligned}\log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\ &+ \frac{1+\psi}{\alpha}\left[\log\gamma_i + \kappa s_{it} + (1-\kappa)\left[\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right]\right] \\ &- \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}}\right) \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\left[\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right]\right) + f(Q_t)\right] \\ &+ \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}\Bigg]\end{aligned} \tag{239}$$

which is, conditional on λ_{it}, normally distributed as both $\log\gamma_i$ and s_{it} are both normal. Hence, we write $\log x_{it}|\lambda_{it} \sim N(\delta_t(\lambda_{it}), \hat{\sigma}^2)$, where:

$$\begin{aligned}\delta_t(\lambda_{it}) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) \\ &+ \frac{1+\psi}{\alpha}\left[\mu_\gamma + \kappa\log\theta_t + (1-\kappa)\left[\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right]\right] \\ &- \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}}\right) \\ &+ \left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1\left(\kappa\log\theta_t + (1-\kappa)\left[\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right]\right) + f(Q_t)\right] \\ &+ \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}\Bigg]\end{aligned} \tag{240}$$

and:

$$\hat{\sigma}^2 = \left(\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\right)^2\left[\left(\frac{1+\psi}{\alpha}\right)^2\sigma^2_\gamma + \kappa^2\left[\frac{1+\psi}{\alpha} + a_1\left(\frac{1}{\epsilon}-\gamma\right)\right]^2\sigma^2_e\right] \tag{241}$$

Thus, we have that:

$$\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right] = \exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(\lambda_{it}) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} \tag{242}$$

and so:

$$\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right] = Q_t\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(1)+\frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}$$
$$+(1-Q_t)\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(0)+\frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}$$
$$=\left[Q_t\exp\left\{\frac{\epsilon-1}{\epsilon}(\delta_t(1)-\delta_t(0))\right\}+(1-Q_t)\right]\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_t(0)+\frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\} \tag{243}$$

Yielding:

$$\log Y_t = \delta_t(0)+\frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2+\frac{\epsilon}{\epsilon-1}\log\left(Q_t\exp\left\{\frac{\epsilon-1}{\epsilon}(\delta_t(1)-\delta_t(0))\right\}+(1-Q_t)\right) \tag{244}$$

where we define $\alpha\delta^{OP}=\delta_t(1)-\delta_t(0)$ and compute:

$$\delta_t(1)-\delta_t(0)=\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\left(\frac{1+\psi}{\alpha}+a_1\left(\frac{1}{\epsilon}-\gamma\right)\right)(1-\kappa)(\mu_O-\mu_P)=\alpha\delta^{OP} \tag{245}$$

and note that this is a constant. Finally, we see that $\delta_t(0)$ is given by:

$$\delta_t(0)=\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right)+\frac{1+\psi}{\alpha}(\mu_\gamma+(1-\kappa)\mu_P)-\frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}})$$
$$+\left(\frac{1}{\epsilon}-\gamma\right)(a_0+a_1(1-\kappa)\mu_P)+\frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}$$
$$+\left[\frac{1+\psi}{\alpha}+a_1\left(\frac{1}{\epsilon}-\gamma\right)\right]\kappa\log\theta_t+\left(\frac{1}{\epsilon}-\gamma\right)f(Q_t)\Bigg] \tag{246}$$

By matching coefficients between Equations 244 and Equation 231, we obtain a_0, a_1, and f.

We first match coefficients on $\log\theta_t$ to obtain an equation for a_1:

$$a_1=\frac{\left[\frac{1+\psi}{\alpha}+a_1\left(\frac{1}{\epsilon}-\gamma\right)\right]\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \tag{247}$$

Under our maintained assumption that $\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\in[0,1)$, as $\kappa\in[0,1]$, we have that this has a unique solution:

$$a_1 = \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1-\frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \tag{248}$$

It is moreover positive.

Second, by collecting terms with Q_t we obtain an equation for f:

$$f(Q) = \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} f(Q) + \frac{\epsilon}{\epsilon-1}\log\left(1+Q\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\right\}-1\right]\right) \tag{249}$$

which has a unique solution as $\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \in [0,1)$ and can be solved to yield:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(1+Q\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\right\}-1\right]\right) \tag{250}$$

where we observe that δ^{OP} depends only on primitive parameters and a_1, for which we have already solved. Finally, by collecting constants, we obtain an equation for a_0:

$$\begin{aligned} a_0 = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left(\mu_\gamma+(1-\kappa)\mu_P\right) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}}\right) \\ + \left(\frac{1}{\epsilon}-\gamma\right)\left(a_0+a_1(1-\kappa)\mu_P\right) + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}\Bigg] + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 \end{aligned} \tag{251}$$

Solving this equation yields:

$$\begin{aligned} a_0 = \frac{1}{1-\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\Bigg[\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left(\mu_\gamma+(1-\kappa)\mu_P\right) - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}+\sigma^2_{\tilde{\theta}}\right) \\ + \left(\frac{1}{\epsilon}-\gamma\right)a_1(1-\kappa)\mu_P + \frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}\Bigg] + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2\Bigg] \end{aligned} \tag{252}$$

which we observe depends only on parameters, a_1, and $\hat{\sigma}^2$. Moreover, $\hat{\sigma}^2$ depends only on parameters and a_1. Thus, given that we have solved for a_1, we have now recovered a_0, a_1, and f uniquely and verified that there exists a unique quasi-linear equilibrium. Finally, to obtain the formula for the best reply of agents, simply substitute a_0, a_1, and f into Equation 239 and label the coefficients as in the claim. □

A.1.3 Proof of the Claims in Remark 1

We now prove the claims made in Remark 1. We have already shown that there exists a unique quasi-linear equilibrium. More generally, we seek to rule out an equilibrium of any other form. To do so, we show that there is a unique equilibrium when fundamentals are bounded by some $M \in \mathbb{R}$, $\log \theta_t \in [-M, M]$, $\log \gamma_i \in [-M, M]$, $\log \tilde{\theta}_{it} \in [-M, M]$, and $e_{it} \in [-M, M]$.

Lemma 7. *When fundamentals are bounded, there exists a unique equilibrium*

Proof. To this end, we can recast any equilibrium function $\log Y(\theta, q)$ as one that solves the fixed point in Equation 37. In the case where fundamentals are bounded, this can be accomplished by demonstrating that the implied fixed-point operator is a contraction by verifying Blackwell's sufficient conditions. More formally, consider the space of bounded, real-valued functions $\mathcal{C}$ under the L^∞-norm and consider the operator $V_M : \mathcal{C} \to \mathcal{C}$ given by:

$$V_M(g)(\theta, Q) = \frac{\epsilon}{\epsilon - 1} \log \mathbb{E}_{(\theta,Q)} \left[\exp \left\{ \frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\log \left(\frac{1 - \frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}} \right) \right.\right.\right. \\ \left.\left.\left. - \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ -\frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] + \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \right) \right\} \right] \tag{253}$$

The following two conditions are sufficient for this operator to be a contraction: (i) monotonicity: for all $g, h \in \mathcal{C}$ such that $g \geq h$, we have that $V_M(g) \geq V_M(h)$ (ii) discounting: there exists a parameter $c \in [0, 1)$ such that for all $g \in \mathcal{C}$ and $a \in \mathbb{R}_+$ and $V_M(g + a) \leq V_M(g) + ca$. Thus, as the space of bounded functions under the L^∞-norm is a complete metric space, if Blackwell's conditions hold, then by the Banach fixed-point theorem, there exists a unique fixed point of the operator V_M.

To complete this argument, we now verify (i) and (ii). To show monotonicity, observe that $\frac{1}{\epsilon} - \gamma \geq 0$ as $\omega \geq 0$ and recall that $\epsilon > 1$. Thus, we have that:

$$\log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] \geq \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) h \right\} \right] \tag{254}$$

for all (s, Q). And so $V_M(g)(\theta, Q) \geq V_M(h)(\theta, Q)$ for all (θ, Q). To show discounting, observe that:

$$\log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) (g+a) \right\} \right] = \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] + \left(\frac{1}{\epsilon} - \gamma \right) a \tag{255}$$

And so:

$$\begin{aligned} V_M(g+a)(\theta,Q) &= \frac{\epsilon}{\epsilon-1} \log \mathbb{E}_{(\theta,Q)} \left[\exp \left\{ \frac{\frac{\epsilon-1}{\epsilon}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\log \left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}} \right) \right. \right. \right. \\ &- \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ -\frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] + \log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right] + \left. \left. \left. \left(\frac{1}{\epsilon} - \gamma \right) a \right) \right\} \right] \\ &= V_M(g)(\theta,Q) + \omega a \end{aligned} \tag{256}$$

where $\omega \in [0,1)$ by assumption. Note that the modulus of contraction ω is precisely the claimed strategic complementarity parameter in Equation 24. This verifies equilibrium uniqueness. $\square$

Away from the case with bounded fundamentals, the above strategy cannot be used to demonstrate uniqueness. Even though the fixed-point operator still satisfies Blackwell's conditions, the relevant function space now becomes any L^p-space for $p \in (1,\infty)$ and the sup-norm over such spaces can be infinite, making Blackwell's conditions insufficient for V to be a contraction. In this case, we show that the unique quasi-linear equilibrium in the unbounded fundamentals case is an appropriately-defined ε-equilibrium for any $\varepsilon > 0$. Let the unique quasi-linear equilibrium we have guessed and verified be $\log Y^*$. We say that g is a ε-equilibrium if

$$||g - V_M(g)||_p < \varepsilon \tag{257}$$

where $||\cdot||_p$ is the L^p-norm. In words, g is a ε-equilibrium if its distance from being a fixed point is at most ε. The following Lemma establishes that Y^* is a ε-equilibrium for bounded fundamentals for any $\varepsilon > 0$ for some bound M:

Lemma 8. *For every $\varepsilon > 0$, there exists an $M \in \mathbb{N}$ such that $\log Y^*$ is a ε-equilibrium.*

Proof. Now extend from $\mathcal{C}$, $V_M : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ as in Equation 253. We observe that V_M is continuous in the limit in M in the sense that $V_M(g) \to V(g)$ as $M \to \infty$ for all $g \in L^p(\mathbb{R})$. This observation follows from noting that both $\log \mathbb{E}_{(s,Q)} \left[\exp \left\{ -\frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right]$ and $\log \mathbb{E}_{(s,Q)} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) g \right\} \right]$ are convergent pointwise for $M \to \infty$ for all (s,Q). In Proposition 1, we showed that $V(\log Y^*) = \log Y^*$.

Thus, we have that: $V_M(\log Y^*) \to V(\log Y^*) = \log Y^*$, which implies that:

$$\lim_{M\to\infty} || \log Y^* - V_M(\log Y^*)||_p = 0 \tag{258}$$

which implies that for every $\varepsilon > 0$, there exists a $\bar{M} \in \mathbb{N}$ such that:

$$|| \log Y^* - V_M(\log Y^*)||_p < \varepsilon \quad \forall M \in \mathbb{N} : M > \bar{M} \tag{259}$$

Completing the proof. □

A.1.4 Proof of Corollary 1

Proof. From Equation 239, we may express:

$$\begin{aligned} \log x_{it} = \text{cons} + b_3 f(Q_t) + \frac{1+\psi}{\alpha}(\log \gamma_i + \kappa s_{it}) + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \kappa s_{it} \\ \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\right] (1-\kappa)(\mu_O - \mu_P)\lambda_{it} \end{aligned} \tag{260}$$

We substitute this expression into $\log L_{it} = \frac{1}{\alpha}(\log x_{it} - \log \theta_{it})$ to write

$$\begin{aligned} \log L_{it} = & -\frac{1}{\alpha}\log \theta_{it} + \text{cons} + c_4 f(Q_t) + \text{cons}_i \\ & + \frac{1}{\alpha}\left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\right] \kappa \log \theta_t \\ & + \frac{1}{\alpha}\frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma\right) \frac{\frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}{1 - \frac{\left(\frac{1}{\epsilon}-\gamma\right)\kappa}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\right] (1-\kappa)(\mu_O - \mu_P)\lambda_{it} \\ & + \xi'_{it} \end{aligned} \tag{261}$$

where $\xi'_{it} \sim N(0, \sigma_\xi^2)$ and IID. Comparing the above with the definition of $\alpha\delta^{OP}$ in Equation 245, we see that the coefficient on λ_{it} in the above expression is δ^{OP}. Subtracting lagged labor from both sides gives the claimed regression equation. We finally observe from Equation 250 that $f(Q)$ depends on $(\epsilon, \gamma, \psi, \alpha)$ and δ^{OP}. Hence, given $(\epsilon, \gamma, \psi, \alpha)$, f is identified uniquely from the studied regression estimate. □

A.1.5 Proof of Corollary 2

Proof. This is immediate by substituting Equation 38 into Equation 28. □

A.1.6 Proof of Proposition 2

Proof. We prove the three claims in sequence.

(1) The map $T_\theta : [0,1] \to [0,1]$ is continuous for all $\theta \in \Theta$ as f, P_O and P_P are continuous functions. Moreover, it maps a convex and compact set to itself. Thus, by Brouwer's fixed point theorem, there exists a Q^*_θ such that $Q^*_\theta = T_\theta(Q^*_\theta)$ for all $\theta \in \Theta$.

(2) To characterize the existence of extremal steady states, observe that $Q = 1$ is a steady state for θ if and only if $T_\theta(1) = P_O(a_o + a_1 \log\theta + f(1), 1, 0) = 1$ and $Q = 0$ is a steady state for θ if and only if $T_\theta(0) = P_P(a_0 + a_1 \log\theta, 0, 0) = 0$. Thus, $Q = 1$ is a steady state if and only if $P_O^{-1}(1; 1) \leq a_0 + a_1 \log\theta + f(1)$ and $Q = 0$ is a steady state if and only if $P_P^{-1}(0; 0) \geq a_0 + a_1 \log\theta$. To obtain the result as stated, we re-arrange these inequalities in terms of $\log\theta$ and exponentiate.

(3) To analyze the stability of the extremal steady states, observe that if $T'_\theta(Q^*) < 1$ at a steady state Q^*, then Q^* is stable. When it exists (which it does almost everywhere), we have that:

$$
\begin{aligned}
T'_\theta(Q) &= P_O(a_0 + a_1 \log\theta + f(Q), Q, 0) - P_P(a_0 + a_1 \log\theta + f(Q), Q, 0) \\
&\quad + Q\frac{\mathrm{d}}{\mathrm{d}Q}P_O(a_0 + a_1 \log\theta + f(Q), Q, 0) + (1-Q)\frac{\mathrm{d}}{\mathrm{d}Q}P_P(a_0 + a_1 \log\theta + f(Q), Q, 0)
\end{aligned} \tag{262}
$$

Thus, for $\theta < \theta_P$ and $Q = 0$:

$$
\begin{aligned}
T'_\theta(0) &= P_O(a_0 + a_1 \log\theta, 0, 0) - P_P(a_0 + a_1 \log\theta, 0, 0) \\
&\quad + \frac{\mathrm{d}}{\mathrm{d}Q}P_P(a_0 + a_1 \log\theta + f(Q), Q, 0) \mid_{Q=0} \\
&= P_O(a_0 + a_1 \log\theta, 0, 0)
\end{aligned} \tag{263}
$$

where the second equality follows by observing that all of P_P, $\frac{\partial P_P}{\partial \log Y}$, and $\frac{\partial P_P}{\partial Q}$ are zero for $\theta < \theta_P$. Thus, we have that $T'_\theta(0) < 1$ when $P_O(a_0 + a_1 \log\theta, 0, 0) < 1$. Moreover, for $\theta < \theta_P$, we have that: $P_O(a_0 + a_1 \log\theta, 0, 0) \leq P_O(a_0 + a_1 \log\theta_P, 0, 0) = P_O(P_P^{-1}(0; 0), 0, 0)$. Thus, a sufficient condition for $T'_\theta(0) < 1$ for $\theta < \theta_P$ is that $P_O(P_P^{-1}(0; 0), 0, 0) < 1$.

For $\theta > \theta_O$ and $Q = 1$, we have that:

$$\begin{aligned}T_\theta'(1) &= P_O(a_0 + a_1 \log\theta + f(1), 1, 0) - P_P(a_0 + a_1 \log\theta + f(1), 1, 0) \\ &\quad + \frac{\mathrm{d}}{\mathrm{d}Q} P_O(a_0 + a_1 \log\theta + f(1), 1, 0) \mid_{Q=1} \\ &= 1 - P_P(a_0 + a_1 \log\theta + f(1), 1, 0)\end{aligned} \tag{264}$$

where the second equality follows again by observing that $P_O = 1$ and both $\frac{\partial P_O}{\partial \log Y}$ $\frac{\partial P_O}{\partial Q}$ are zero for $\theta > \theta_O$. Hence, we have that $T_\theta'(1) < 1$ when $P_P(a_0+a_1\log\theta+f(1), 1, 0) > 0$. For $\theta > \theta_O$ we have that $P_P(a_0+a_1\log\theta+f(1), 1, 0) \geq P_P(a_0+a_1\log\theta_O+f(1), 1) = P_P(P_O^{-1}(1, 1), 1, 0)$. Thus, a sufficient condition for $T_\theta'(1) < 1$ for $\theta > \theta_O$ is that $P_P(P_O^{-1}(1, 1), 1, 0) > 0$. □

A.1.7 Proof of Corollary 3

Proof. By Proposition 2, the extremal steady states coexist if and only if $\theta \in [\theta_O, \theta_P]$, which is non-empty if and only if $\theta_O \leq \theta_P$ which is equivalent to $P_O^{-1}(1; 1) - P_P^{-1}(0; 0) \leq f(1)$. □

A.1.8 Proof of Lemma 1

Proof. Fix $\theta \in \Theta$. We first study the SSC-A case. By SSC-A of T we have that there exists $\hat{Q}_\theta \in [0, 1]$ such that $T_\theta(Q) > Q$ for all $Q \in (0, \hat{Q}_\theta)$ and $T_\theta(Q) < Q$ for all $Q \in (\hat{Q}_\theta, 1)$. As T_θ is continuous we have that $T_\theta(\hat{Q}_\theta) = \hat{Q}_\theta$. Consider now some $Q_0 \in (0, 1)$ such that $Q_0 \neq \hat{Q}_\theta$. We have that $T_\theta(Q_0) > \hat{Q}_\theta$ if $Q_0 < \hat{Q}_\theta$ and $T_\theta(Q_0) < \hat{Q}_\theta$ if $Q_0 > \hat{Q}_\theta$. Hence, there exists at most one $Q^* \in (0, 1)$ such that $T_\theta(Q^*) = Q^*$. Thus, there exist at most three steady states $Q^* = 0$, $Q^* = \hat{Q}_\theta$, and $Q^* = 1$.

To find the basins of attraction of these steady states, fix $Q_0 \in (0, 1)$ and consider the sequence $\{T_\theta^n(Q_0)\}_{n\in\mathbb{N}}$. For a steady state Q^*, its basin of attraction is:

$$\mathcal{B}_\theta(Q^*) = \left\{Q_0 \in [0, 1] : \lim_{n\to\infty} T_\theta^n(Q_0) = Q^*\right\} \tag{265}$$

First, consider $Q_0 \in (0, \hat{Q}_\theta)$. We now show by induction that $T_\theta^n(Q_0) \geq T_\theta^{n-1}(Q_0)$ for all $n \in \mathbb{N}$. Consider $n = 1$. We have that $T_\theta(Q_0) > Q_0$ as T is SSC-A and $Q_0 < \hat{Q}_\theta$. Suppose now that $T_\theta^n(Q_0) \geq T_\theta^{n-1}(Q_0)$. We have that:

$$T_\theta^{n+1}(Q_0) = T_\theta \circ T_\theta^n(Q_0) \geq T_\theta \circ T_\theta^{n-1}(Q_0) = T_\theta^n(Q_0) \tag{266}$$

by monotonicity of T_θ, which proves the inductive hypothesis. Observe moreover that the sequence $\{T_\theta^n(Q_0)\}_{n\in\mathbb{N}}$ is bounded as $T_\theta^n(Q_0) \in [0,1]$ for all $n \in \mathbb{N}$. Hence, by the monotone convergence theorem, $\lim_{n\to\infty} T_\theta^n(Q_0)$ exists. Toward a contradiction, suppose that $Q_0^\infty = \lim_{n\to\infty} T_\theta^n(Q_0) > \hat{Q}_\theta$. By SSC-A of T we have that $T_\theta(Q_0^\infty) > Q_0^\infty$, but this contradicts that $Q_0^\infty = \lim_{n\to\infty} T_\theta^n(Q_0)$. Thus, we have that $Q_0^\infty = \hat{Q}_\theta$. Hence, $(0, \hat{Q}_\theta) \subseteq \mathcal{B}_\theta(\hat{Q}_\theta)$. Second, consider $Q_0 = \hat{Q}_\theta$. We have that $T_\theta(\hat{Q}_\theta) = \hat{Q}_\theta$. Thus, $Q_0^\infty = \hat{Q}_\theta$. Hence, $\hat{Q}_\theta \in \mathcal{B}_\theta(\hat{Q}_\theta)$. Third, consider $Q_0 \in (\hat{Q}_\theta, 1)$. Following the arguments of the first part, we have that $(\hat{Q}_\theta, 1) \subseteq \mathcal{B}_\theta(\hat{Q}_\theta)$. Thus, $(0,1) \subseteq \mathcal{B}_\theta(\hat{Q}_\theta)$. Moreover, if $Q = 0$ or $Q = 1$ are steady states, they can only have basins of attraction in $[0,1] \setminus \mathcal{B}_\theta(\hat{Q}_\theta)$, which implies that they are unstable and can only have basins of attraction $\{0\}$ and $\{1\}$.

The analysis of the SSC-B case follows similarly. By SSC-B of T we have that there exists $\hat{Q}_\theta \in [0,1]$ such that $T_\theta(Q) > Q$ for all $Q \in (\hat{Q}_\theta, 1)$ and $T_\theta(Q) < Q$ for all $Q \in (0, \hat{Q}_\theta)$. As T_θ is continuous, we have that $T_\theta(\hat{Q}_\theta) = \hat{Q}_\theta$. Consider now some $Q_0 \in (0,1)$ such that $Q_0 \neq \hat{Q}_\theta$. Observe that $T_\theta(Q_0) < \hat{Q}_\theta$ if $Q_0 < \hat{Q}_\theta$ and $T_\theta(Q_0) > \hat{Q}_\theta$ if $Q_0 > \hat{Q}_\theta$. Hence, there exists at most one $Q^* \in (0,1)$ such that $T_\theta(Q^*) = Q^*$. Thus, there exist at most three steady states $Q^* = 0$, $Q^* = \hat{Q}_\theta$, and $Q^* = 1$.

To find the basins of attraction of these steady states, first consider $Q_0 \in (0, \hat{Q}_\theta)$. We now show by induction that $T_\theta^n(Q_0) \leq T_\theta^{n-1}(Q_0)$ for all $n \in \mathbb{N}$. Consider $n = 1$. We have that $T_\theta(Q_0) < Q_0$ as T is SSC-B and $Q_0 < \hat{Q}_\theta$. Suppose now that $T_\theta^n(Q_0) \leq T_\theta^{n-1}(Q_0)$. We have that:

$$T_\theta^{n+1}(Q_0) = T_\theta \circ T_\theta^n(Q_0) \leq T_\theta \circ T_\theta^{n-1}(Q_0) = T_\theta^n(Q_0) \tag{267}$$

by monotonicity of T_θ, which proves the inductive hypothesis. Observe moreover that the sequence $\{T_\theta^n(Q_0)\}_{n\in\mathbb{N}}$ is bounded as $T_\theta^n(Q_0) \in [0,1]$ for all $n \in \mathbb{N}$. Hence, by the monotone convergence theorem, $\lim_{n\to\infty} T_\theta^n(Q_0)$ exists. Finally, toward a contradiction, suppose that $Q_0^\infty = \lim_{n\to\infty} T_\theta^n(Q_0) > 0$. By SSC-B of T we have that $T_\theta(Q_0^\infty) < Q_0^\infty$, but this contradicts that $Q_0^\infty = \lim_{n\to\infty} T_\theta^n(Q_0)$. Thus, we have that $Q_0^\infty = 0$. Hence, $[0, \hat{Q}_\theta) \subseteq \mathcal{B}_\theta(0)$. Second, consider $Q_0 = \hat{Q}_\theta$. We have that $T_\theta(\hat{Q}_\theta) = \hat{Q}_\theta$. Thus, $Q_0^\infty = \hat{Q}_\theta$. Hence $\hat{Q}_\theta \in \mathcal{B}_\theta(\hat{Q}_\theta)$. Third, consider $Q_0 \in (\hat{Q}_\theta, 1]$. By the exact arguments of the first part, we have that $(\hat{Q}_\theta, 1] \subseteq \mathcal{B}_\theta(1)$. Observing $\mathcal{B}_\theta(0)$, $\mathcal{B}_\theta(\hat{Q}_\theta)$, and $\mathcal{B}_\theta(1)$ are disjoint completes the proof. $\square$

A.1.9 Proof of Proposition 3

Proof. By Proposition 1 and substituting the form of the shock process from Equation 47, we obtain the formula for the output IRF. For the fraction of optimists, we see that:

$$\begin{aligned} Q_2 &= \hat{Q}_1 P_O(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) + (1 - \hat{Q}_1) P_P(a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1), \hat{Q}_1, 0) \\ &> \hat{Q}_1 P_O(a_0 + f(\hat{Q}_1), \hat{Q}_1, 0) + (1 - \hat{Q}_1) P_P(a_0 + f(\hat{Q}_1), \hat{Q}_1, 0) = \hat{Q}_1 \end{aligned} \tag{268}$$

and $Q_t = T_1(\log Y_{t-1}, Q_{t-1})$ for $t \geq 3$ by iterating forward. That Q_t monotonically declines to $\hat{Q}_1$ follows from Lemma 1 as we are in the SSC-A case. The hump shape is obtained if $\log Y_1 \leq \log Y_2$. This corresponds to

$$\log Y_1 = a_0 + a_1 \log \hat{\theta} + f(\hat{Q}_1) \leq a_0 + f(Q_2) = \log Y_2 \tag{269}$$

which rearranges to the desired expression. □

A.1.10 Proof of Proposition 4

Proof. We first derive the IRF functions. The formula for the output IRF follows Proposition 3. For the IRF for the fraction of optimists, we simply observe that $Q_0 = Q_1 = 0$ and $Q_2 = P_P(a_0 + a_1 \log \hat{\theta}, 0, 0)$, and that $Q_t = T_1(Q_{t-1})$ for $t \geq 3$ by iterating forward.

We now describe the properties of the IRFs as a function of the size of the initial shock $\hat{\theta}$. First, observe that $Q_2 = P_P(a_0 + a_1 \log \hat{\theta}, 0, 0)$. Thus, we have that $Q_2 = 0$ if and only if $P_P^{-1}(0; 0) \geq a_0 + a_1 \log \hat{\theta}$ which holds if and only if $\hat{\theta} \leq \theta_P$. For any $\hat{\theta} > \theta_P$ it follows that $Q_2 > 0$. As we lie in the SSC class, by Lemma 1, we have that the steady states $Q = 0$, $Q = 1$, and $Q = \hat{Q}_1$ have basins of attraction given by $[0, \hat{Q}_1)$, $(\hat{Q}_1, 1]$, $\{\hat{Q}_1\}$. Thus, if $Q_2 < \hat{Q}_1$, we have monotone convergence of Q_t to 0. If $Q_2 = \hat{Q}_1$, then $Q_t = \hat{Q}_t$ for all $t \in \mathbb{N}$. If $Q_2 > \hat{Q}_1$, we have monotone convergence of Q_t to 1. Moreover, the threshold for $\hat{\theta}$ such that $Q_2 = \hat{Q}^*$ is $\exp\left\{\frac{P_P^{-1}(\hat{Q}_1; 0) - a_0}{a_1}\right\}$.

Finally, to find the condition such that the IRF is hump-shaped, we observe that this occurs if and only if $f(Q_2) > a_1 \log \hat{\theta}$ as Q_t is monotonically decreasing for $t \geq 2$, which is precisely the claimed condition. □

A.1.11 Proof of Proposition 5

Proof. We prove this result by first constructing fictitious processes for optimism that bound above and below the true optimism process for all realizations of $\{\theta_t\}_{t \in \mathbb{N}}$

before the stopping time. We can then use this to bound the stopping times' distributions in the sense of first-order stochastic dominance and use this fact to bound the expectations.

First, consider the case where we seek to bound $\tau_{PO} = \min\{t \in \mathbb{N} : Q_t = 1, Q_0 = 0\}$. In the model, we have that $Q_{t+1} = T(Q_t, \nu_t)$. Fix a path of fundamentals and narrative shocks $\{\nu_t\}_{t\in\mathbb{N}} = \{\theta_t, \varepsilon_t\}_{t\in\mathbb{N}}$ and define the fictitious $\overline{Q}$ process as:

$$\overline{Q}_{t+1} = \mathbb{I}[T(\overline{Q}_t, \nu_t) = 1] \tag{270}$$

with $\overline{Q}_0 = 0$. We prove by induction that $\overline{Q}_t \leq Q_t$ for all $t \in \mathbb{N}$. Consider first the base case that $t = 1$:

$$\overline{Q}_1 = \mathbb{I}[T(0, \nu_0) = 1] \leq T(0, \nu_0) = Q_1 \tag{271}$$

Toward the inductive hypothesis, suppose that $\overline{Q}_{t-1} \leq Q_{t-1}$. Then we have that:

$$\overline{Q}_t = \mathbb{I}[T(\overline{Q}_{t-1}, \nu_{t-1}) = 1] \leq \mathbb{I}[T(Q_{t-1}, \nu_{t-1}) = 1] \leq T(Q_{t-1}, \nu_{t-1}) = Q_t \tag{272}$$

where the first inequality follows by the property that $T(\cdot, \nu)$ is a monotone increasing function.

As $\overline{Q}_t \leq Q_t$ for all $t \in \mathbb{N}$, we have that:

$$\overline{\tau}_{PO} = \min\{t \in \mathbb{N} : \overline{Q}_t = 1, \overline{Q}_0 = 0\} \geq \min\{t \in \mathbb{N} : Q_t = 1, Q_0 = 0\} = \tau_{PO} \tag{273}$$

Else, we would have at $\overline{\tau}_{PO}$ that $Q_{\overline{\tau}_{PO}} < \overline{Q}_{\overline{\tau}_{PO}}$, which is a contradiction.

We now have a pathwise upper bound on τ_{PO}. We now characterize the distribution of the bound. Observe that the possible sample paths for $\{\overline{Q}_t\}_{t\in\mathbb{N}}$ until stopping are given by the set:

$$\mathcal{G}_{PO} = \{(0^{(n-1)}, 1)\} : n \geq 1\} \tag{274}$$

Moreover, conditional on $\overline{Q}_{t-1} = 0$, the distribution of $\overline{Q}_t$ is independent of $\{\nu_s\}_{s\leq t-1}$. Thus, the fictitious stopping time $\overline{\tau}_{PO}$ has a geometric distribution with parameter given by $\mathbb{P}[Q_{t+1} = 1 | Q_t = 0]$. This parameter is given by:

$$\begin{aligned}\mathbb{P}[Q_{t+1} = 1 | Q_t = 0] &= \mathbb{P}\left[P_P(a_0 + a_1 \log \theta_t, 0, \varepsilon_t) = 1\right] \\ &= \mathbb{P}\left[\theta_t \geq \exp\left\{\frac{P_P^{\dagger}(1; 0, \varepsilon_t) - a_0}{a_1}\right\}\right] \\ &= 1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^{\dagger}(1; 0, \varepsilon) - a_0}{a_1}\right\}\right)\right]\end{aligned} \tag{275}$$

Thus, we have established a stronger result and provided a distributional bound on the stopping time:

$$\tau_{PO} \prec_{FOSD} \overline{\tau}_{PO} \sim \text{Geo}\left(1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^{\dagger}(1; 0, \varepsilon) - a_0}{a_1}\right\}\right)\right]\right) \tag{276}$$

An immediate corollary is that:

$$T_{PO} = \mathbb{E}[\tau_{PO}] \leq \mathbb{E}[\overline{\tau}_{PO}] = \frac{1}{1 - \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_P^{\dagger}(1;0,\varepsilon) - a_0}{a_1}\right\}\right)\right]} \tag{277}$$

We can apply appropriately adapted arguments for the other case, where we now define:

$$\underline{Q}_{t+1} = \mathbb{I}[T(\underline{Q}_t, \nu_t) \neq 0] \tag{278}$$

with $\underline{Q}_0 = 1$. In this case, by an analogous induction have that $\underline{Q}_t \geq Q_t$ for all $t \in \mathbb{N}$ for all sequences $\{\nu_t\}_{t \in \mathbb{N}}$. And so, we have that if $\underline{Q}_t$ has reached 0 then so too has Q_t. The possible sample paths in this case are:

$$\mathcal{G}_{OP} = \{(1^{(n-1)}, 0)\} : n \geq 1\} \tag{279}$$

So again the stopping time has a geometric distribution, this time with parameter:

$$\begin{aligned}\mathbb{P}[Q_{t+1} = 0 | Q_t = 1] &= \mathbb{P}\left[\theta_t \leq \exp\left\{\frac{P_O^{\dagger}(0; 1, \varepsilon_t) - a_0 - f(1)}{a_1}\right\}\right] \\ &= \mathbb{E}_G\left[H\left(\exp\left\{\frac{P_O^{\dagger}(0; 1, \varepsilon) - a_0 - f(1)}{a_1}\right\}\right)\right]\end{aligned} \tag{280}$$

And so we have:

$$T_{OP} \leq \frac{1}{\mathbb{E}_G\left[H\left(\exp\left\{\frac{P_O^{\dagger}(0;1,\varepsilon) - a_0 - f(1)}{a_1}\right\}\right)\right]} \tag{281}$$

It remains to show that these bounds are tight. To do so, we derive a law H such that $Q_t = \overline{Q}_t = \underline{Q}_t$ for all $t \in \mathbb{N}$. Concretely, define the set:

$$\Theta^* = \left(-\infty, \exp\left\{\frac{P_O^\dagger(0;1,0) - a_0 - f(1)}{a_1}\right\}\right] \cup \left[\exp\left\{\frac{P_P^\dagger(1;0,0) - a_0}{a_1}\right\}, \infty\right) \quad (282)$$

and suppose that θ takes values only in this set, where the two sub-intervals are disjoint as $P_O^\dagger(0;1,0) - P_P^\dagger(1;0,0) \leq f(1)$. Moreover, suppose that narrative shocks equal zero with probability one. In this case, starting from $Q_t = 1$, the only possible values for Q_{t+1} are zero and one. Moreover, starting from $Q_t = 0$, the only possible values for Q_{t+1} are zero and one. Thus, in either case, $Q_t = \overline{Q}_t = \underline{Q}_t$ pathwise and $T_{OP} = T^*_{OP}$ and $T_{PO} = T^*_{PO}$. It is worth noting that such a distribution can be obtained by considering a limit of normal-mixture distributions. Concretely, suppose that H is derived as a mixture of two normal distributions $N(\mu_A, \sigma^2)$ and $N(\mu_B, \sigma^2)$ for $\mu_A < \exp\left\{\frac{P_O^\dagger(0;1,0) - a_0 - f(1)}{a_1}\right\}$ and $\mu_B > \exp\left\{\frac{P_P^\dagger(1;0,0) - a_0}{a_1}\right\}$. Taking the limit as $\sigma \to 0$, the support of H converges to being contained within Θ^*. $\square$

A.2 Model Extensions

This appendix covers several model extensions. First, we study equilibrium dynamics under a benchmark model of Bayesian model updating and contrast these predictions with those obtained in our main analysis (A.2.1). Second, we theoretically characterize and quantify the normative implications of narrative fluctuations (A.2.2). Third, fourth, fifth, and sixth we extend the baseline model to respectively incorporate a continuum of different levels of optimism (A.2.3), narratives about idiosyncratic fundamentals (A.2.4), multi-dimensional narratives and persistent fundamentals (A.2.5), and narrative updating that depends on idiosyncratic fundamentals (A.2.6). In each case, we characterize equilibrium dynamics and show how our main theoretical insights extend. Seventh, we show how endogenous cycles and chaotic dynamics can obtain when agents are contrarian and implement an empirical test for their presence (A.2.7). Eighth, we highlight the role of higher-order beliefs and show how our analysis could generalize to other settings by deriving a similar law of motion for optimism in abstract, linear beauty contest games à la Morris and Shin (2002) (A.2.8). Finally, we sketch an extension of our abstract framework to allow for persistent idiosyncratic states and adjustment costs and discuss the implications for our measurement (A.2.9).

A.2.1 Comparison to the Bayesian Benchmark

Consider an alternative model in which each agent i initially believes the optimistic model is correct with probability $\lambda_{i0} \in (0,1)$, and subsequently updates this probability by observing aggregate output and aggregate optimism and applying Bayes' rule under rational expectations. Formally, this corresponds to the following law of motion for Q_t:

$$Q_{t+1} = \int_{[0,1]} \mathbb{P}_i[\mu = \mu_O | \{\log Y_j, Q_j\}_{j=0}^t] \, \mathrm{d}i \tag{283}$$

where $\mathbb{P}_i[\mu = \mu_0|\emptyset] = \lambda_{i0}$ for some $\lambda_{i0} \in (0,1)$ for all $i \in [0,1]$, and conditional probabilities are computed under rational expectations with knowledge of $\{\lambda_{i0}\}_{i\in[0,1]}$. We define the log-odds ratio of an agent's belief as $\Omega_{it} = \log \frac{\lambda_{it}}{1-\lambda_{it}}$. The following Proposition characterizes the dynamics of agents' subjective models under the Bayesian benchmark:

Proposition 27 (Dynamics under the Bayesian Benchmark)**.** *Each agent's log-odds ratio follows a random walk with drift, or* $\Omega_{i,t+1} = \Omega_{it} + a + \xi_t$*, where* $a = \mathbb{E}_H\left[\frac{(\log\theta_t - \mu_P)^2 - (\log\theta_t - \mu_O)^2}{\sigma^2}\right]$ *and* ξ_t *is an IID, mean-zero random variable. The economy converges almost surely to either extreme optimism* ($a > 0$) *or extreme pessimism* ($a < 0$)*. The dynamics of* *the economy are asymptotically described by:*

$$\log Y_t = \begin{cases} a_0 + a_1 \log\theta_t & \textit{if } a < 0, \\ a_0 + a_1 \log\theta_t + f(1) & \textit{if } a > 0. \end{cases} \tag{284}$$

Thus, the economy does not feature steady state multiplicity, hump-shaped or discontinuous IRFs, or the possibility for boom-bust cycles.

Proof. The equilibrium Characterization of Proposition 1 still holds. Moreover, Q_0 is known to all agents. Thus, they can identify θ_0 as:

$$\theta_0 = \frac{\log Y_0 - a_0 - f(Q_0)}{a_1} \tag{285}$$

Thus, we have that $\lambda_{i1} = \mathbb{P}[\mu = \mu_O | \theta_0, \lambda_{i0}]$. Moreover, all agents know that $Q_1 = \int_{[0,1]} \lambda_{i1}\, \mathrm{d}i$. Thus, agents can sequentially identify θ_t by observing only $\{Y_j\}_{j \leq t}$ (and not $\{Q_j\}_{j\leq t}$) by computing:

$$\theta_t = \frac{\log Y_t - a_0 - f(Q_t)}{a_1} \tag{286}$$

Thus, we can describe the evolution of agents' beliefs by computing:

$$\lambda_{i,t+1} = \mathbb{P}_i[\mu = \mu_O | \{\theta_j\}_{j=1}^t] = \lambda_{i,t+1} = \mathbb{P}_i[\mu = \mu_O | \{Y_j\}_{j=1}^t] \tag{287}$$

By application of Bayes rule, we obtain:

$$\lambda_{i,t+1} = \mathbb{P}[\mu = \mu_O | \theta_t, \lambda_{i,t}] = \frac{f_O(\theta_t)\lambda_{i,t}}{f_O(\theta_t)\lambda_{i,t} + f_P(\theta_t)(1-\lambda_{i,t})} \tag{288}$$

which implies that:

$$\begin{aligned} \frac{\lambda_{i,t+1}}{1-\lambda_{i,t+1}} &= \frac{f(\log\theta_t | \mu = \mu_O)}{f(\log\theta_t|\mu = \mu_P)} \frac{\lambda_{i,t}}{1-\lambda_{i,t}} \\ &= \exp\left\{ \frac{(\log\theta_t - \mu_P)^2 - (\log\theta_t - \mu_O)^2}{\sigma^2} \right\} \frac{\lambda_{i,t}}{1-\lambda_{i,t}} \end{aligned} \tag{289}$$

Defining $\Omega_{it} = \log\frac{\lambda_{i,t}}{1-\lambda_{i,t}}$ and $a = \mathbb{E}_H\left[\frac{(\log\theta_t - \mu_P)^2 - (\log\theta_t-\mu_O)^2}{\sigma^2}\right]$ and $\xi_t = \frac{(\log\theta_t-\mu_P)^2 - (\log\theta_t - \mu_O)^2}{\sigma^2} -$

a, we then have that:

$$\begin{aligned}\Omega_{i,t+1} &= \Omega_{i,t} + \frac{(\log\theta_t - \mu_P)^2 - (\log\theta_t - \mu_O)^2}{\sigma^2} \\ &= \Omega_{it} + a + \xi_t\end{aligned} \tag{290}$$

which is a random walk with drift, with the drift and stochastic increment claimed in the statement. Iterating, dividing by t, and applying the law of large numbers, we obtain:

$$\frac{\Omega_{i,t}}{t} = \frac{1}{t}\Omega_{i,0} + \frac{t-1}{t}a + \frac{1}{t}\sum_{i=1}^{t}\xi_i \to^{a.s.} a \tag{291}$$

Hence, almost surely, we have that $Q_t \to 1$ if $a > 0$ and $Q_t \to 0$ if $a < 0$.

Hence, the dynamics are asymptotically described by Proposition 1 with $Q_t = 1$ if $a > 0$ and $Q_t = 0$ if $a < 0$. The resulting properties for output follow immediately from combining this characterization for Q_t with the characterization in our main analysis of equilibrium output conditional on optimism and fundamentals (Proposition 1), which continues to hold in the model of this appendix. □

The optimist fraction Q converges to either 0 or 1 in the long run because one model is unambiguously better-fitting, and this will be revealed with infinite data. Moreover, the log-odds ratio converges linearly and so the odds ratio in favor of the better fitting model converges exponentially quickly. Thus the Bayesian benchmark model makes a prediction that is at odds with our finding of cyclical dynamics for aggregate optimism (Figure 1-1), and moreover, in the long run, rules out the features of macroeconomic dynamics that we derive in Section 1.6 as consequences of the endogenous evolution of narrative optimism.

A.2.2 Welfare Implications

In this appendix, we derive the normative implications of narratives for the economy.

Theory The following result characterizes welfare along any path for the fraction of optimists in the population and the conditions under which a steady state of extreme optimism is preferred to one of extreme pessimism:

Proposition 28 (Narratives and Welfare)**.** *For any path of aggregate optimism* $\mathbf{Q} =$

$\{Q_t\}_{t=0}^{\infty}$, *aggregate welfare is given by*

$$\begin{aligned}\mathcal{U}(\mathbf{Q}) = U_C^* \sum_{t=0}^{\infty} \beta^t \exp\left\{(1-\gamma) f(Q_t)\right\} \\ - U_L^* \sum_{t=0}^{\infty} \beta^t \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t)\right) \exp\left\{(1+\psi) d_3 f(Q_t)\right\}\end{aligned} \tag{292}$$

for some positive constants U_C^*, U_L^*, d_2 and d_3 that are provided in the proof of the result. Thus, there is higher welfare in an optimistic steady state than in a pessimistic steady state if and only if

$$\frac{U_C^*}{U_L^*} \times \frac{\exp\left\{(1-\gamma) f(1)\right\} - 1}{\exp\left\{(1+\psi)(d_2 + d_3 f(1))\right\} - 1} > 1 \tag{293}$$

Moreover, when the pessimistic narrative is correctly specified, extreme optimism is welfare-equivalent to an ad valorem price subsidy for intermediate goods producers of:

$$\tau^* = \exp\left\{(1-\omega)\left(\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}\right) f(1)\right\} - 1 \tag{294}$$

Proof. We have that welfare for any path of optimism $\mathbf{Q} = \{Q_t\}_{t \in \mathbb{N}}$ is given by:

$$\mathcal{U}(\mathbf{Q}) = \sum_{t=0}^{\infty} \beta^t \left(\mathbb{E}_H \left[\frac{C_t(Q_t, \theta_t)^{1-\gamma}}{1-\gamma} \right] - \mathbb{E}_H \left[\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1+\psi} \mathrm{d}i \right] \right) \tag{295}$$

By market clearing, we have that $C_t = Y_t$ for all t. Thus, using the formula for equilibrium aggregate output from Proposition 1 and our assumption that $\log \theta_t$ is Gaussian under H, we have that the consumption component of welfare is given by:

$$\begin{aligned}\mathbb{E}_H \left[\frac{C_t^{1-\gamma}(Q_t, \theta_t)}{1-\gamma} \right] &= \mathbb{E}_H \left[\frac{1}{1-\gamma} \exp\left\{(1-\gamma) \log Y(Q_t, \theta)\right\} \right] \\ &= \mathbb{E}_H \left[\frac{1}{1-\gamma} \exp\left\{(1-\gamma) \left(a_0 + a_1 \log \theta + f(Q_t)\right)\right\} \right] \\ &= \frac{1}{1-\gamma} \exp\left\{(1-\gamma)\left(a_0 + a_1 \mu_H + f(Q_t)\right) + \frac{1}{2} a_1^2 \sigma_H^2 \right\} \\ &= \frac{1}{1-\gamma} \exp\left\{(1-\gamma)\left(a_0 + a_1 \mu_H\right) + \frac{1}{2} a_1^2 \sigma_H^2 \right\} \exp\left\{(1-\gamma) f(Q_t)\right\} \\ &= U_C^* \exp\left\{(1-\gamma) f(Q_t)\right\}\end{aligned} \tag{296}$$

From Proposition 1, we moreover have that labor employed by each firm can be written as:

$$L_{it} = d_1 \log \theta_t + d_2 \lambda_{it} + d_3 f(Q_t) + v_{it} \tag{297}$$

where v_{it} is Gaussian and IID over i. Hence given θ and Q_t:

$$\begin{aligned} &\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1+\psi} \mathrm{d}i \\ &= \frac{1}{1+\psi} \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t) \right) \\ &\quad \times \exp\left\{ (1+\psi)(d_1 \log \theta + \mu_v + d_3 f(Q_t)) + \frac{1}{2}(1+\psi)^2 \sigma_v^2 \right\} \end{aligned} \tag{298}$$

Hence, the expectation over θ is given by:

$$\begin{aligned} &\mathbb{E}_H \left[\int_{[0,1]} \frac{L_{it}(\gamma_i, s_{it}, Q_t)^{1+\psi}}{1+\psi} \mathrm{d}i \right] \\ &= \frac{1}{1+\psi} \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t) \right) \\ &\quad \times \exp\left\{ (1+\psi) d_3 f(Q_t) \right\} \exp\left\{ (1+\psi)(d_1 \mu_H + \mu_v) + \frac{1}{2}(1+\psi)^2 (\sigma_v^2 + d_1^2 \sigma_H^2) \right\} \\ &= U_L^* \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t) \right) \exp\left\{ (1+\psi) d_3 f(Q_t) \right\} \end{aligned} \tag{299}$$

And so total welfare under narrative path $\mathbf{Q}$ is given by:

$$\begin{aligned} \mathcal{U}(\mathbf{Q}) = U_C^* &\sum_{t=0}^{\infty} \beta^t \exp\left\{ (1-\gamma) f(Q_t) \right\} \\ &- U_L^* \sum_{t=0}^{\infty} \beta^t \left(Q_t \exp\{(1+\psi)d_2\} + (1-Q_t) \right) \exp\left\{ (1+\psi) d_3 f(Q_t) \right\} \end{aligned} \tag{300}$$

The final inequality follows by noting that $f(0) = 0$ and rearranging this expression.

Now consider the benchmark model but where, without loss of generality, all agents are pessimistic $Q_t = 0$ and a planner levies an *ad valorem* subsidy. That is, when the consumer price is $p_{it}^C = Y_t^{\frac{1}{\varepsilon}} x_{it}^{-\frac{1}{\varepsilon}}$, the price received by the producer is

$p_{it}^P = (1+\tau)p_{it}^C$. Under this subsidy, each producer's first-order condition is:

$$\begin{aligned} \log x_{it} =& \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\log \left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}} \right) - \log \mathbb{E}_{it} \left[\exp \left\{ -\frac{1+\psi}{\alpha} \log \theta_{it} \right\} \right] \right. \\ & \left. + \log \mathbb{E}_{it} \left[\exp \left\{ \left(\frac{1}{\epsilon} - \gamma \right) \log Y_t \right\} \right] \right) + \Xi(\tau) \end{aligned} \tag{301}$$

where $\Xi(\tau) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}} \log(1+\tau)$. By identical arguments to Proposition 1, we have that there is a unique quasi-linear equilibrium, where:

$$\log Y(\theta, \tau) = a_0 + a_1 \log \theta + \frac{1}{1-\omega} \Xi(\tau) \tag{302}$$

and a_0 and a_1 are as in Proposition 1. Hence, in this equilibrium we have that:

$$\log x_{it}(\tau) = \log x_{it}(0) + \frac{1}{1-\omega} \Xi(\tau) \tag{303}$$

Which implies that:

$$\log L_{it}(\tau) = \log L_{it}(0) + \frac{1}{\alpha} \frac{1}{1-\omega} \Xi(\tau) \tag{304}$$

And so, welfare under the subsidy τ is given by:

$$\begin{aligned} \mathcal{U}(\tau) = U_C^* \sum_{t=0}^{\infty} \beta^t \exp \left\{ (1-\gamma) \frac{1}{1-\omega} \Xi(\tau) \right\} \\ - U_L^* \sum_{t=0}^{\infty} \beta^t \exp \left\{ (1+\psi) d_3 \frac{1}{1-\omega} \Xi(\tau) \right\} \end{aligned} \tag{305}$$

as $d_3 = \frac{1}{\alpha}$. Hence:

$$\mathcal{U}(1) = \mathcal{U}(\tau^*) \tag{306}$$

where τ^* is such that $\frac{1}{1-\omega} \Xi(\tau^*) = f(1)$. Hence:

$$\tau^* = \exp \left\{ (1-\omega) \left(\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon} \right) f(1) \right\} - 1 \tag{307}$$

Completing the proof. □

This result sheds light on the potential for non-fundamental optimism to increase aggregate welfare. In the presence of the product market monopoly and labor market monopsony distortions, intermediates goods firms under-hire labor and under-produce goods. As a result, if irrational optimism causes them to produce more output, but not so much that the household over-supplies labor, then it has the potential to be welfare improving. The final part of the proposition then reduces this question to assessing if the implied optimism-equivalent subsidy is less than the welfare-optimal subsidy. Thus, optimism in the economy can serve the role of undoing monopoly frictions and thereby has the potential to be welfare-improving, even when misspecified.

Quantification Proposition 28 can be directly applied in our numerical calibration from Section 1.7 to calculate the welfare effects of narrative optimism without approximation. We calculate the average payoff of the representative household under three scenarios. The first corresponds to the calibrated narrative dynamics in simulation, under the assumption that the pessimistic model is correctly specified.[1] The second is a counterfactual scenario with permanent extreme optimism, or $Q_t \equiv 1$ for all t. The third is a counterfactual scenario with permanent extreme pessimism, or $Q_t \equiv 0$ for all t, and an *ad valorem* subsidy of τ to all producers. We use the third scenario to translate the first and second into payoff-equivalent subsidies. We find that both contagious and extreme optimism are welfare-increasing relative to extreme pessimism in autarky (*i.e,* $\tau = 0$). In payoff units, they correspond respectively to equivalent subsidies of 1.33% and 2.59%. Our finding of an overall positive welfare effect for contagious optimism suggests that, in our macroeconomic calibration, losses from inducing misallocation are more than compensated by level increases in output.

A.2.3 Continuous Narratives

Our main analysis featured two levels of optimism. However, much of our analysis generalizes to a setting with a continuum of levels of optimism. For simplicity, in this section we abstract from optimism shocks. The model is as in Section 1.5, but now $\mu \in [\mu_P, \mu_O]$ and the distribution of narratives is given by $Q_t \in \Delta([\mu_P, \mu_O])$. The probabilistic transition between models is now given by a Markov kernel $P : [\mu_P, \mu_O] \times \mathcal{Y} \times \Delta^2([\mu_P, \mu_O]) \to \Delta([\mu_P, \mu_O])$ where $P_{\mu'}(\mu, \log Y, Q)$ is the density of agents who have model μ who switch to μ' when aggregate output is Y and the distribution of narratives is Q.

Characterizing Equilibrium Output By modifying the guess-and-verify arguments that underlie Proposition 1, we can obtain an almost identical representation of equilibrium aggregate output:

[1] Relative to the positive analysis, the normative analysis requires two additional model parameters. We set the idiosyncratic component of productivity to have unit mean and zero variance.

Proposition 29 (Equilibrium Characterization with Continuous Narratives)**.** *There exists a quasi-linear equilibrium:*

$$\log Y(\log\theta_t, Q_t) = a_0 + a_1 \log\theta_t + f(Q_t) \tag{308}$$

Moreover, the density of narratives evolves according to the following difference equation:

$$\mathrm{d}Q_{t+1}(\mu') = \int_{\mu_P}^{\mu_O} P_{\mu'}(\mu, a_0 + a_1 \log\theta_t + f(Q_t), Q_t)\mathrm{d}Q_t(\mu) \tag{309}$$

Proof. By appropriately modifying the steps of the proof of Proposition 1, the result follows. Throughout, simply replace $\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P$ with $\tilde{\mu}_{it} \sim Q_t$ and λ_{it} with $\tilde{\mu}_{it}$ as appropriate. The proof follows as written until the aggregation step. At this point, we instead obtain:

$$\log Y_t = \delta_t(\mu_P) + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 + \frac{\epsilon}{\epsilon-1}\log\left(\int_{\mu_P}^{\mu_O} \exp\left\{\frac{\epsilon-1}{\epsilon}(\delta_t(\tilde{\mu}) - \delta_t(\mu_P))\right\}\mathrm{d}Q_t(\tilde{\mu})\right) \tag{310}$$

where $\delta_t(\mu_P) = \delta_t(0)$ and $\delta_t(\tilde{\mu}) - \delta_t(\mu_P) = \alpha\delta^{OP}\frac{\tilde{\mu}-\mu_P}{\mu_O-\mu_P}$. Hence, we have that a_0 and a_1 are as in Proposition 1 and f is instead given by:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(\int_{\mu_P}^{\mu_O}\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{\tilde{\mu}-\mu_P}{\mu_O-\mu_P}\right\}\mathrm{d}Q(\tilde{\mu})\right) \tag{311}$$

Completing the proof. □

Importantly, observe that we still obtain a marginal representation in terms of the partial equilibrium effect of going from full pessimism to full optimism on hiring δ^{OP}, as we have empirically estimated.

Equilibrium Dynamics We have seen that a continuum of models poses no difficulty for the static analysis. The challenge for the dynamic analysis is that the state variable, the evolution of which is fully characterized by Proposition 29, is now infinite-dimensional. This notwithstanding, by use of approximation arguments, we can reduce the dynamics to an essentially identical form to that which we have studied in the main text.

To this end, define the cumulant generating function (CGF) of the cross-sectional distribution of narratives as:

$$K_Q(\tau) = \log\left(\mathbb{E}_Q[\exp\{\tau\tilde{\mu}\}]\right) \tag{312}$$

We therefore have that $\log\left(\mathbb{E}_Q[\exp\{\tau(\tilde{\mu}-z)\}]\right) = K_Q(\tau) - \tau z$. It follows by Equation 311 that:

$$f(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\omega}\left[K_Q\left(\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{1}{\mu_O-\mu_P}\right) - \frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{\mu_P}{\mu_O-\mu_P}\right] \tag{313}$$

By Maclaurin series expansion, we can express the CGF to first-order as:

$$K_Q(\tau) = \mu_Q\tau + O(\tau^2) \tag{314}$$

We therefore have that:

$$f(Q) = \frac{1}{1-\omega}\alpha\delta^{OP}\frac{\mu_Q-\mu_P}{\mu_O-\mu_P} + O\left(\left(\frac{\epsilon-1}{\epsilon}\alpha\delta^{OP}\frac{1}{\mu_O-\mu_P}\right)^2\right) \tag{315}$$

We now can express the static, general equilibrium effects in terms of mean of the narrative distribution. With some abuse of notation, we now write $f(\mu_Q) = f(Q)$. Of course, this CGF-based approach would allow one to consider higher-order effects through the variance, skewness, kurtosis, and higher cumulants as desired.

In the next steps, we provide conditions on updating that allow us to express the dynamics solely in terms of the mean of the narrative distribution. To do this, we assume that $P_{\mu'}(\mu, \log Y, Q) = P_{\mu'}(\mu'', \log Y, \mu_Q)$ for all $Q \in \Delta^2([\mu_P, \mu_O])$ and all $\mu, \mu', \mu'' \in [\mu_P, \mu_O]$. This is tantamount to assuming no stubbornness (all agents update the same regardless of the model they start with) and that contagiousness only matters via the mean. Under this assumption, we can write $P_{\mu'}(\log Y(\log\theta, \mu_Q), \mu_Q)$ and express the difference equation as:

$$\begin{aligned} \mathrm{d}Q_{t+1}(\mu') &= \int_{\mu_P}^{\mu_O} P_{\mu'}(a_0 + a_1\log\theta_t + f(\mu_{Q,t}), \mu_{Q,t})\mathrm{d}Q_t(\mu) \\ &= P_{\mu'}(a_0 + a_1\log\theta_t + f(\mu_{Q,t}), \mu_{Q,t}) \end{aligned} \tag{316}$$

It then suffices to take the mean of Q_{t+1} to express the system in terms of the one-dimensional state variable $\mu_{Q,t}$:

$$\mu_{Q,t+1} = T(\mu_{Q,t}, \theta_t) = \int_{\mu_P}^{\mu_O} \mu' P_{\mu'}(a_0 + a_1\log\theta_t + f(\mu_{Q,t}), \mu_{Q,t})\mathrm{d}\mu' \tag{317}$$

Which is simply a continuous state analog of the difference equation expressed in Corollary 2 expressed in terms of average beliefs.

Steady State Multiplicity We now obtain the analogous characterization of extremal steady state multiplicity in this setting, *i.e.,* when it is possible that all agents being maximally pessimistic and all agents being maximally optimistic are simultaneously deterministic steady states. To this end, define the following two inverses:

$$\begin{aligned} \hat{P}^{-1}(x;\mu_Q) &= \sup\{Y : P(Y,Q) = \delta_x\} \\ \check{P}^{-1}(x;\mu_Q) &= \inf\{Y : P(Y,Q) = \delta_x\} \end{aligned} \tag{318}$$

where δ_x denotes the Dirac delta function on x. We define analogous objects to the previous θ_O and θ_P:

$$\theta_O = \exp\left\{\frac{\check{P}^{-1}(\mu_O;\mu_O) - a_0 - f(1)}{a_1}\right\},\ \theta_P = \exp\left\{\frac{\hat{P}^{-1}(\mu_P;\mu_P) - a_0 - f(1)}{a_1}\right\} \tag{319}$$

The following result establishes that these thresholds characterize extremal multiplicity:

Proposition 30 (Steady State Multiplicity with Continuous States)**.** *Extreme optimism and pessimism are simultaneously deterministic steady states for θ if and only if $\theta \in [\theta_O, \theta_P]$, which is non-empty if and only if*

$$\check{P}^{-1}(\mu_O;\mu_O) - \hat{P}^{-1}(\mu_P;\mu_P) \leq f(1) \tag{320}$$

Proof. This follows exactly the same steps as the proofs of Proposition 2 and Corollary 3, replacing the appropriate inverses defined above. □

Thus, the same conditions that give rise to multiplicity with binary narratives obtain with a continuum of levels of optimism. Indeed, observe that restricting to first-order approximations above was unnecessary. We could have considered an arbitrary order, say k, of approximation of the CGF and obtained a system of difference equations for the first k cumulants. Proposition 30 would still hold as written, as under the extremal steady states, all higher cumulants are identically zero and remain so under the provided condition. Naturally, however, the general dynamics only reduce to those resembling the simple model under the first-order approximation. Nevertheless, we observe that this is a first-order approximation to the exact equilibrium dynamics and not simply an approximation of the dynamics of an approximate equilibrium.

A.2.4 Narratives About Idiosyncratic Fundamentals

In the main analysis, we assumed that narratives described properties of aggregate fundamentals. In this section, we characterize equilibrium dynamics when narratives describe properties of idiosyncratic fundamentals. Concretely, we now instead suppose that all agents believe that $\log\theta_t \sim N(0,\sigma^2)$, or agree about the distribution of aggregate productivity. Moreover, as in the baseline, all agents believe that others' idiosyncratic productivity follows $\log\tilde{\theta}_{jt} \sim N(0,\sigma^2_{\tilde{\theta}})$ for all $j \neq i$. However, agents disagree about the mean of their own idiosyncratic productivity: optimistic agents believe that $\log\tilde{\theta}_{it} \sim N(\mu_O,\sigma^2_{\tilde{\theta}})$ while pessimistic agents believe that $\log\tilde{\theta}_{it} \sim N(\mu_P,\sigma^2_{\tilde{\theta}})$. The rest of the model is identical.

In this context, dynamics are identical conditional on the static relationship between output and narratives. Moreover, the static relationship between output and narratives is now identical (up to a constant) conditional on estimating the partial equilibrium effect of optimism on hiring. This is formalized by the following result:

Proposition 31 (Equilibrium Characterization with Narratives About Idiosyncratic Fundamentals)**.** *There exists a unique equilibrium such that:*

$$\log Y(\log\theta_t, Q_t) = \tilde{a}_0 + a_1\log\theta_t + \tilde{f}(Q_t) \tag{321}$$

for coefficients $\tilde{a}_0$ and $a_1 > 0$, and a strictly increasing function f, where a_1 is identical to that from Proposition 1 and

$$\tilde{f}(Q) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(1+Q\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\tilde{\delta}^{OP}\right\}-1\right]\right) \tag{322}$$

where $\tilde{\delta}^{OP}$ is defined in Equation 323.

Proof. The proof follows exactly the steps of the proof of Proposition 1 where the aggregate narrative is replaced with an idiosyncratic one. To be concrete, the computation of $\log\mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right]$ and the method of aggregation are identical to those in the proof of Proposition 1. The only difference is in the computation of $\log\mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right]$. Now, Equation 236 differs in that $\mu_{it} = \log\gamma_i + \kappa s_{it}$. Tracking this through to Equation 240, lines 1, 2, 3, and 5 are identical and line 4 differs only in that the term $(1-\kappa)[\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P]$ is now set equal to zero. The analysis then follows up to Equation 245, at which point we have that the exact formula for δ^{OP} changes and is now given by:

$$\alpha\tilde{\delta}^{OP} = \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}}(1-\kappa)(\mu_O - \mu_P) \tag{323}$$

The formula for $\delta_t(0)$ is identical except for in the second line where the term $a_1(1-\kappa)\mu_P$ is now equal to zero. The formula for a_1 remains the same. Conditional on $\tilde{\delta}^{OP}$, the formula for f remains the same. The formula for a_0 is identical except for the second line where the term $(1/\epsilon - \gamma)a_1(1-\kappa)\mu_P$ is now equal to zero. □

This Proposition makes clear that output differs in this case only up to an intercept and in changing the mapping from structural parameters to the partial-equilibrium effect of optimism on hiring. Nonetheless, interpreted via the model above, our empirical exercise directly identifies the now-relevant parameter $\tilde{\delta}^{OP}$. As a result, neither our theoretical nor quantitative analysis is sensitive to making narratives be about idiosyncratic conditions. The only difference is that the point calibrations for κ and $(\mu_O - \mu_P)$ would change, while the aggregate dynamics would remain identical.

A.2.5 Multi-Dimensional Narratives and Persistent States

Our baseline model featured two narratives regarding the mean of fundamentals and transitory fundamentals, but we live in a world of many competing narratives regarding many aspects of reality and potentially persistent fundamentals. In this extension, we broaden our analysis to study a class of three-dimensional narratives, which is essentially exhaustive within the Gaussian class. For simplicity, we abstract from narrative shocks in this analysis. Concretely, suppose that agents believe that the aggregate component of fundamentals follows:

$$\log\theta_t = (1-\rho)\mu + \rho\log\theta_{t-1} + \sigma\nu_t \tag{324}$$

with $\nu_t \sim N(0,1)$ and IID. Narratives now correspond to a vector of (μ,ρ,σ), indexing the mean, persistence and variance of the process for fundamentals. The set of narratives can therefore be represented by $\{(\mu_k,\rho_k,\sigma_k)\}_{k\in\mathcal{K}}$. We restrict that agents place Dirac weights on this set, so that they only ever believe one narrative at a time, and let $Q_{t,k}$ be the fraction of agents who believe narrative (μ_k,ρ_k,σ_k) at time t. Finally, we assume that agents face the same signal-to-noise ratio κ, regardless of the narrative that they hold.[2] Together, these assumptions ensure that agents' posteriors are normal and place a common weight on narratives when agents form

[2]Formally, this means that the variance of the noise in agents' signals satisfies $\sigma^2_{\varepsilon,k} \propto \sigma^2_k$ across narratives.

their expectations of fundamentals.

By modifying the functional guess-and-verify arguments from Proposition 1, we characterize equilibrium output in this setting in the following result:

Proposition 32 (Equilibrium Characterization with Multi-Dimensional Narratives and Persistence)**.** *There exists a quasi-linear equilibrium:*

$$\log Y(\log\theta_t, \log\theta_{t-1}, Q_t) = a_0 + a_1 \log\theta_t + a_2 \log\theta_{t-1} + f(Q_t, \theta_{t-1}) \tag{325}$$

for some $a_1 > 0$, $a_2 \geq 0$, and f. In this equilibrium, the distribution of narratives in the population evolves according to:

$$Q_{t+1,k} = \sum_{k' \in \mathcal{K}} Q_{t,k'} P_{k'}(k, a_0 + a_1 \log\theta_t + a_2 \log\theta_{t-1} + f(Q_t, \theta_{t-1}), Q_t) \tag{326}$$

Proof. We follow the same steps as in the proof of Proposition 1, appropriately adapted to this richer setting. First, we guess an equilibrium of the form:

$$\log Y(\log\theta_t, \log\theta_{t-1}, Q_t) = a_0 + a_1 \log\theta_t + a_2 \log\theta_{t-1} + f(Q_t, \theta_{t-1}) \tag{327}$$

To verify that this is an equilibrium, we need to compute agents' best replies under this conjecture, aggregate them, and show that they are consistent with this guess once aggregated.

We first find agents' posterior beliefs given narrative weights. Let E denote the standard basis for $\mathbb{R}^K$ with k-th basis vector denoted by

$$e_k = \{\underbrace{0, \ldots, 0}_{k-1}, 1, \underbrace{0, \ldots, 0}_{K-k}\} \tag{328}$$

We have that $\lambda_{it} = e_k$ for some $k \leq K$. Under this narrative loading, we have that agent's posteriors are given by:

$$\log\theta_{it} | \lambda_{it}, s_{it} \sim N\left(\log\gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \tag{329}$$

with:

$$\begin{aligned} \mu(e_k, \theta_{t-1}) &= (1-\rho_k)\mu_k + \rho_k \log\theta_{t-1} \\ \sigma^2_{\theta|s}(e_k) &= \frac{1}{\frac{1}{\sigma_k^2} + \frac{1}{\sigma^2_{\epsilon,k}}} \quad \kappa = \frac{1}{1 + \frac{\sigma^2_{\varepsilon,k}}{\sigma^2_k}} \end{aligned} \tag{330}$$

for all $k \leq K$,where κ does not depend on k as $\sigma^2_{\varepsilon,k} \propto \sigma^2_k$. Hence, we can compute agents' best replies by evaluating:

$$\log \mathbb{E}_{it}\left[\theta_{it}^{-\frac{1+\psi}{\alpha}}\right] = -\frac{1+\psi}{\alpha}\left(\log\gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right) \tag{331}$$

$$\begin{aligned}\log \mathbb{E}_{it}\left[Y_t^{\frac{1}{\epsilon}-\gamma}\right] &= \left(\frac{1}{\epsilon}-\gamma\right)\left(a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1})\right) + a_2\log\theta_{t-1} + f(Q_t, \theta_{t-1})\right)\\ &+ \frac{1}{2}\left(\frac{1}{\epsilon}-\gamma\right)^2 a_1^2\sigma^2_{\theta|s}(\lambda_{it})\end{aligned} \tag{332}$$

By substituting this into agents' best replies, we obtain:

$$\begin{aligned}\log x_{it} = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left[\log\gamma_i + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it},\theta_{t-1})\right]\\ &-\frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\tilde{\theta}}\right)\\ &+\left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1\left(\kappa s_{it} + (1-\kappa)\mu(\lambda_{it},\theta_{t-1})\right) + a_2\log\theta_{t-1} + f(Q_t,\theta_{t-1})\right]\\ &+\frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}(\lambda_{it})\Bigg]\end{aligned} \tag{333}$$

which we observe is conditional normally distributed as $\log x_{it}|\lambda_{it} \sim N(\delta_t(\lambda_{it}), \hat{\sigma}^2)$ with $\hat{\sigma}^2$ as in Equation 241 and:

$$\begin{aligned}\delta_t(e_k) = \frac{1}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\Bigg[&\log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right)\\ &+\frac{1+\psi}{\alpha}\left[\log\gamma_i + \kappa\log\theta_t + (1-\kappa)\mu(e_k,\theta_{t-1})\right]\\ &-\frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2\left(\sigma^2_{\theta|s}(e_k) + \sigma^2_{\tilde{\theta}}\right)\\ &+\left(\frac{1}{\epsilon}-\gamma\right)\left[a_0 + a_1\left(\kappa\log\theta_t + (1-\kappa)\mu(e_k,\theta_{t-1})\right) + a_2\log\theta_{t-1} + f(Q_t,\theta_{t-1})\right]\\ &+\frac{1}{2}a_1^2\left(\frac{1}{\epsilon}-\gamma\right)^2\sigma^2_{\theta|s}(e_k)\Bigg]\end{aligned} \tag{334}$$

for all $k \leq K$. Aggregating these best replies, using Equation 242, we obtain that:

$$\begin{aligned}
\log Y_t &= \frac{\epsilon}{\epsilon-1}\log \mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\lambda_{it}\right]\right] \\
&= \frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(e_k)+\frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}\right) \\
&= \delta_t(e_1)+\frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2+\frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\left(\delta_t(e_k)-\delta_t(e_1)\right)\right\}\right)
\end{aligned} \tag{335}$$

where $\hat{\sigma}^2$ is a constant, $\delta_t(e_1)$ depends linearly on $\log\theta_t$ and $\log\theta_{t-1}$ and $\delta_t(e_k)-\delta_t(e_1)$ does not depend on $\log\theta_t$ for all $k \leq K$ and can therefore be written as $\delta_{k1}(\theta_{t-1})$. Moreover, by matching coefficients, we obtain that a_1 is the same as in the proof of Proposition 1. And we find that f must satisfy:

$$f(Q,\theta_{t-1}) = \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}f(Q,\theta_{t-1})+\frac{\epsilon}{\epsilon-1}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1})\right\}\right) \tag{336}$$

and so:

$$f(Q,\theta_{t-1}) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}}\log\left(\sum_k Q_{t,k}\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1})\right\}\right) \tag{337}$$

Completing the proof. □

In the multidimensional narrative case with persistence, the past value of fundamentals interacts non-linearly with the cross-sectional narrative distribution in affecting aggregate output. However, without more structure, the properties of the dynamics generated by this multi-dimensional system are essentially unrestricted.

A.2.6 Persistent Idiosyncratic Shocks and Belief Updating

We now extend the analysis from Section A.2.5 to the case where agents' idiosyncratic states drive narrative updating and are persistent. Concretely, in that setting, we let $P_{k'}$ depend on $(Y_t, Q_t, \tilde{\theta}_{it})$ and idiosyncratic productivity shocks evolve according to an AR(1) process:

$$\log\tilde{\theta}_{it} = \rho_{\tilde{\theta}}\log\tilde{\theta}_{i,t-1}+\zeta_{it} \tag{338}$$

where $0<\rho_{\tilde{\theta}}<1$ and $\zeta_{it}\sim N(0,\sigma_\zeta^2)$. We let $F_{\tilde{\theta}}$ denote the stationary distribution of $\tilde{\theta}_{it}$, which coincides with the cross-sectional marginal distribution of $\tilde{\theta}_{it}$ for all $t\in\mathbb{N}$.

The additional theoretical complication these two changes induce is that the marginal distribution of narratives Q_t is now insufficient for describing aggregate output. This is because narratives λ_{it} and idiosyncratic fundamentals $\tilde{\theta}_{it}$ are no longer independent as λ_{it} and $\tilde{\theta}_{it}$ both depend on $\tilde{\theta}_{it-1}$. The relevant state variable is now the joint distribution of narratives and idiosyncratic productivity $\check{Q}_t \in \Delta(\Lambda \times \mathbb{R})$. We denote the marginals as Q_t and $F_{\tilde{\theta}}$, and the conditional distribution of narratives given $\tilde{\theta}$ as $\check{Q}_{t,k|\tilde{\theta}} = \frac{\check{Q}_{t,k}(\tilde{\theta})}{f_{\tilde{\theta}}(\tilde{\theta})}$.

Proposition 33 (Equilibrium Characterization with Multi-Dimensional Narratives, Aggregate and Idiosyncratic Persistence, and Idiosyncratic Narrative Updating)**.** *There exists a quasi-linear equilibrium:*

$$\log Y(\log\theta_t, \log\theta_{t-1}, \check{Q}_t) = a_0 + a_1 \log\theta_t + a_2 \log\theta_{t-1} + f(\check{Q}_t, \theta_{t-1}) \tag{339}$$

for some $a_1 > 0$, $a_2 \geq 0$, and f.

Proof. This proof follows closely that of Proposition 32. Under narrative loading λ_{it}, we have that the agent's posterior regarding $\log\theta_{it}$ is given by:

$$\log\theta_{it}|\tilde{\theta}_{it-1}, \lambda_{it}, s_{it} \sim N\left(\log\gamma_i + \rho_{\tilde{\theta}} \log\tilde{\theta}_{it-1} + \kappa s_{it} + (1-\kappa)\mu(\lambda_{it}, \theta_{t-1}), \sigma^2_{\theta|s}(\lambda_{it}) + \sigma^2_{\xi}\right) \tag{340}$$

where $\mu(\lambda_{it}, \theta_{t-1})$, κ, and $\sigma^2_{\theta|s}(\lambda_{it})$ are as in Proposition 32. Then substitute $\log\gamma_i + \rho_{\tilde{\theta}}\tilde{\theta}_{it-1}$ for $\log\gamma_i$ and follow the Proof of Proposition 32 until the aggregation step (Equation 335). We now instead have that:

$$\begin{aligned}
\log Y_t &= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\log x_{it}\right\}|\tilde{\theta}_{it-1}, \lambda_{it}\right]\right] \\
&= \frac{\epsilon}{\epsilon-1}\log\mathbb{E}_t\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\delta_t(e_k, \tilde{\theta}_{it-1}) + \frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\right)^2\hat{\sigma}^2\right\}\right] \\
&= \frac{\epsilon}{\epsilon-1}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\varepsilon}\delta_t(e_k, \tilde{\theta}) + \frac{1}{2}\left(\frac{\epsilon-1}{\varepsilon}\right)^2\hat{\sigma}^2\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right) \\
&= \delta_t(e_1, 1) + \frac{1}{2}\frac{\epsilon-1}{\epsilon}\hat{\sigma}^2 \\
&\quad + \frac{\epsilon}{\epsilon-1}\log\left(\int\sum_k \check{Q}_{t,k|\tilde{\theta}}\exp\left\{\frac{\epsilon-1}{\epsilon}\left(\delta_t(e_k, \tilde{\theta}) - \delta_t(e_1, 1)\right)\right\}\mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right)
\end{aligned} \tag{341}$$

Again, $\hat{\sigma}^2$ is a constant and $\delta_t(e_1, 0)$ depends linearly on $\log\theta_t$ and $\log\theta_{t-1}$ and $\delta_t(e_k, \tilde{\theta}) - \delta_t(e_1, 1)$ does not depend on $\log\theta_t$ for all $k \leq K$. Thus, we may write it as $\delta_{k1}(\theta_{t-1}, \tilde{\theta})$. Again, a_1 is the same as in Proposition 1. By the same steps as in Proposition 32, we then have that:

$$f(\check{Q}, \theta_{t-1}) = \frac{\frac{\epsilon}{\epsilon-1}}{1 - \frac{\frac{1}{\epsilon}-\gamma}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}} \log\left(\int \sum_k \check{Q}_{t,k|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{k1}(\theta_{t-1}, \tilde{\theta})\right\} \mathrm{d}F_{\tilde{\theta}}(\tilde{\theta})\right) \tag{342}$$

Completing the proof. □

We can use this result to study the additional effects induced by persistent idiosyncratic fundamentals. To do this, we restrict to the case of our main analysis with optimism and pessimism. In this context, we have that:

$$f(\check{Q}) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\omega} \log\left(\mathbb{E}_{\tilde{\theta}}\left[\check{Q}_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\} + (1-\check{Q}_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\}\right]\right) \tag{343}$$

where:

$$\begin{aligned} \delta_{OP}(\tilde{\theta}) &= \alpha\delta_{OP} + \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}} \log\tilde{\theta} \\ \delta_{PP}(\tilde{\theta}) &= \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}} \log\tilde{\theta} \end{aligned} \tag{344}$$

We define $\xi = \frac{\frac{1+\psi}{\alpha}}{\frac{1+\psi-\alpha}{\alpha}+\frac{1}{\epsilon}}\rho_{\tilde{\theta}}$ and observe that we can write:

$$\begin{aligned} &\check{Q}_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\} + (1-\check{Q}_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\} \\ &= Q_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\left(\alpha\delta_{OP} + \xi\log\tilde{\theta}\right)\right\} + (1-Q_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\} \\ &= Q_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\}\left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\} - 1\right] + \exp\left\{\frac{\epsilon-1}{\epsilon}\xi\log\tilde{\theta}\right\} \end{aligned} \tag{345}$$

Taking the expectation of the relevant terms, we obtain:

$$\begin{aligned} &\mathbb{E}_{\tilde{\theta}}\left[\check{Q}_{t|\tilde{\theta}} \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{OP}(\tilde{\theta})\right\} + (1-\check{Q}_{t|\tilde{\theta}}) \exp\left\{\frac{\epsilon-1}{\epsilon}\delta_{PP}(\tilde{\theta})\right\}\right] \\ &= \left[\exp\left\{\frac{\epsilon-1}{\epsilon}\alpha\delta_{OP}\right\} - 1\right] \exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2 \frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\} Q_t \\ &\quad + \mathrm{Cov}_t\left(Q_{t|\tilde{\theta}}, \tilde{\theta}^{\frac{\epsilon-1}{\epsilon}\xi}\right) + \exp\left\{\frac{1}{2}\left(\frac{\epsilon-1}{\epsilon}\xi\right)^2 \frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2}\right\} \end{aligned} \tag{346}$$

Thus, we have that the contribution of optimism to output is given by:

$$
\begin{aligned}
f(\check{Q}_t) = \frac{\frac{\epsilon}{\epsilon-1}}{1-\omega} \log \Bigg(& \left[\exp\left\{ \frac{\epsilon-1}{\epsilon} \alpha \delta_{OP} \right\} - 1 \right] \exp\left\{ \frac{1}{2} \left(\frac{\epsilon-1}{\epsilon} \xi \right)^2 \frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2} \right\} Q_t \\
& + \mathrm{Cov}_t \left(Q_{t|\tilde{\theta}}, \tilde{\theta}^{\frac{\epsilon-1}{\epsilon}\xi} \right) + \exp\left\{ \frac{1}{2} \left(\frac{\epsilon-1}{\epsilon} \xi \right)^2 \frac{\sigma_\zeta^2}{1-\rho_{\tilde{\theta}}^2} \right\} \Bigg)
\end{aligned} \tag{347}
$$

We observe that the first term is almost identical to that in our main analysis. This term is now intermediated by the effect of heterogeneity in previous productivity (to see this, observe that this vanishes when $\rho_{\tilde{\theta}} = 0$). Second, there is a new effect stemming from the covariance of optimism and productivity. Intuitively, when more optimistic firms are also more productive, they increase their production by more and this increases output. Finally, there is a level effect of heterogeneous productivity.

Thus, the sole new qualitative force is the covariance effect. To the extent that this does not vary with time, it can have no effect on dynamics. We investigate this in the data by estimating the regression model

$$
\log \hat{\theta}_{it} = \sum_{\tau=1995}^{2019} \beta_\tau \cdot (\mathrm{opt}_{i\tau} \cdot \mathbb{I}[\tau = t]) + \chi_{j(i),t} + \gamma_i + \varepsilon_{it} \tag{348}
$$

where $(\chi_{j(i),t}, \gamma_i)$ are industry-by-time and firm fixed effects, and β_s measures the (within-industry, within-firm) difference in mean log TFP for optimistic and pessimistic firms in each year. If the β_s vary systematically with the business cycle, then the shifting productivity composition of optimists over the business cycle is an important component of business-cycle dynamics.

We plot our coefficient estimates β_τ in Figure A-15. The estimates are generally positive, but economically small relative to the large observed variation in TFP, $\log \theta_{it}$, which has an in-sample standard deviation of 0.84. Outside of the first two years and last year of the sample, we find limited evidence of time variation. Moreover, the variation that exists is not obviously correlated with the business cycle. This suggests that the compositional effect for optimists driven by narrative updating in response to idiosyncratic conditions is not, at least in our data, quantitatively significant.

A.2.7 Contrarianism, Endogenous Cycles, and Chaos

The baseline model can generate neither endogenous cycles nor chaotic dynamics without extrinsic shocks to fundamentals (as made formal by Lemma 1). This is because the probability that agents become optimistic is always increasing in the fraction of optimists in equilibrium.

In this appendix, we relax this assumption and delineate precise, testable conditions under which cyclical and chaotic dynamics occur in the absence of fundamental and aggregate shocks. We do so in a model with "contrarian" agents whose updating contradicts recent data and/or consensus. Our analysis of endogenous narratives with contrarianism therefore complements the literature on endogenous cycles in macroeconomic models (see, *e.g.*, Boldrin and Woodford, 1990; Beaudry, Galizia, and Portier, 2020) by providing a further potential micro-foundation for the existence of endogenous cycles.

We begin by defining cycles and chaos. There exists a cycle of period $k \in \mathbb{N}$ if $Q = T^k(Q)$ and all elements of $\{Q, T(Q), \ldots, T^{k-1}(Q)\}$ are non-equal. We will say that there are *chaotic dynamics* if there exists an uncountable set of points $S \subset [0,1]$ such that (i) for every $Q, Q' \in S$ such that $Q \neq Q'$, we have that $\limsup_{t\to\infty} |T^t(Q) - T^t(Q')| > 0$ and $\liminf_{t\to\infty} |T^t(Q) - T^t(Q')| = 0$ and (ii) for every $Q \in S$ and periodic point $Q' \in [0,1]$, $\limsup_{t\to\infty} |T^t(Q) - T^t(Q')| > 0$. This definition of chaos is due to Li and Yorke (1975) and can be understood as saying that there is a large set of points such that the iterated dynamics starting from any two points in this set get both far apart and vanishingly close.

A Variant Model with the Potential for Cycles and Chaos We will study the issue of cycles and chaos under the simplifying assumption that,[3] in equilibrium, the induced probabilities that optimists and pessimists respectively become optimists are quadratic and given by:[4]

$$\tilde{P}_O(Q) = a_O + b_O Q - cQ^2\,,\; \tilde{P}_P(Q) = a_P + b_P Q - cQ^2 \tag{349}$$

with parameters $(a_O, a_P, b_O, b_P, c) \in \mathbb{R}^5$ such that $P_O([0,1]), P_P([0,1]) \subseteq [0,1]$. The parameters a_O and a_P index stubbornness, b_O and b_P capture both contagiousness and associativeness (through the subsumed equilibrium map), and c captures any non-linearity.

The following result describes the potential dynamics:

[3] This simplifying assumption is without any qualitative loss as this model can demonstrate the full range of potential cyclical and chaotic dynamics.

[4] This can be microfounded in a generalization our earlier LAC model by taking $P_i(\log Y, Q) = u_i + r_i \log Y + s_i Q - cQ^2$ for $i \in \{O, P\}$ and approximating $f(Q) \approx \frac{\alpha\delta^{OP}}{1-\omega} Q$. In this case:

$$\tilde{P}_i(Q) = (u_i + r_i a_0 + r_i a_1 \log\theta) + \left(r_i \frac{\alpha\delta^{OP}}{1-\omega} + s_i\right) Q - cQ^2$$

Proposition 34. *The following statements are true:*

1. *When $\tilde{P}_O \geq \tilde{P}_P$ and both are monotone, there are neither cycles of any period nor chaotic dynamics.*

2. *When $\tilde{P}_O$ and $\tilde{P}_P$ are linear, cycles of period 2 are possible, cycles of any period $k > 2$ are not possible, and chaotic dynamics are not possible.*

3. *Without further restrictions on $\tilde{P}_O$ and $\tilde{P}_P$, cycles of any period $k \in \mathbb{N}$ and chaotic dynamics are possible.*

Proof. The dynamics of optimism are characterized by the transition map

$$\begin{aligned} T(Q) &= Q(a_O + b_O Q - cQ^2) + (1-Q)(a_P + b_P Q - cQ^2) \\ &= a_P + (a_O - a_P + b_P)Q - (c + b_P - b_O)Q^2 \end{aligned} \tag{350}$$

where we define $\omega_0 = a_P$, $\omega_1 = (a_O - a_P + b_P)$, $\omega_2 = (c + b_P - b_O)$ for simplicity. We first show that the dynamics described by T are topologically conjugate to those of the logistic map $\check{T}(x) = \eta x(1-x)$ with

$$\eta = 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)} \tag{351}$$

Two maps $T : [0,1] \to [0,1]$ and $T' : [0,1] \to [0,1]$ are topologically conjugate if there exists a continuous, invertible function $h : [0,1] \to [0,1]$ such that $T' \circ h = h \circ T$. If T is topologically conjugate to T' and we know the orbit of T', we can compute the orbit of T via the formula:

$$T^k(Q) = \left(h^{-1} \circ T'^k \circ h\right)(Q) \tag{352}$$

Hence, we can prove the properties of interest using known properties of the map $\check{T}$ as well as the mapping from the deeper parameters of T to the parameters of $\check{T}$.

To show the topological conjugacy of T and $\check{T}$, we proceed in three steps:

1. *T is topically topologically conjugate to the quadratic map $\hat{T}(Q) = Q^2 + k$ for appropriate choice of k.* We guess the following homeomorphism $\hat{h}(Q) = \hat{\alpha} + \hat{\beta} Q$. Plugging $\hat{h}$ in $\hat{T}$, we have that:

$$\hat{T}(\hat{h}(Q)) = (k + \hat{\alpha}^2) + 2\hat{\alpha}\hat{\beta}Q + \hat{\beta}^2 Q^2 \tag{353}$$

Inverting $\hat{h}$ and applying it to this expression yields:

$$\hat{h}^{-1}(\hat{T}(\hat{h}(Q))) = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}} + 2\hat{\alpha}Q + \hat{\beta}Q^2 \tag{354}$$

To verify topological conjugacy, we need to show that $T(Q) = \hat{h}^{-1}(\hat{T}(\hat{h}(Q)))$. Matching coefficients, this is the case if and only if:

$$\omega_0 = \frac{k + \hat{\alpha}(\hat{\alpha} - 1)}{\hat{\beta}},\ \omega_1 = 2\hat{\alpha},\ \omega_2 = -\hat{\beta} \tag{355}$$

We therefore have that:

$$k = \hat{\beta}\omega_0 + \hat{\alpha}(1 - \hat{\alpha}) = -\omega_2\omega_0 + \frac{\omega_1}{2}\left(1 - \frac{\omega_1}{2}\right) \tag{356}$$

with $\hat{h}(Q) = \frac{\omega_1}{2} - \omega_2 Q$.

2. *$\hat{T}$ is topologically conjugate to $\check{T}$ for appropriate choice of η.* We guess the following homeomorphism $\check{h}(Q) = \check{\alpha} + \check{\beta}Q$. Plugging $\check{h}$ in $\check{T}$, we obtain:

$$\check{T}(\check{h}(Q)) = \eta\left(\check{\alpha}(1 - \check{\alpha}) + \check{\beta}(1 - 2\check{\alpha})Q - \check{\beta}^2 Q^2\right) \tag{357}$$

Inverting $\check{h}$ and applying it, we obtain:

$$\check{h}^{-1}(\check{T}(\check{h}(Q))) = \frac{\eta\check{\alpha}(1 - \check{\alpha}) - \check{\alpha}}{\check{\beta}} + \eta(1 - 2\check{\alpha})Q - \eta\check{\beta}Q^2 \tag{358}$$

Matching coefficients, we find:

$$k = \frac{\eta\check{\alpha}(1 - \check{\alpha}) - \check{\alpha}}{\check{\beta}},\ 0 = \eta(1 - 2\check{\alpha}),\ 1 = -\eta\check{\beta} \tag{359}$$

We therefore obtain that:

$$k = \eta(\check{\alpha} - \eta(1 - \check{\alpha})) = \frac{\eta}{2}\left(1 - \frac{\eta}{2}\right) \tag{360}$$

which implies that $\eta = 1 + \sqrt{1 - 4k}$ with $\check{h}(Q) = \frac{1}{2} - \frac{1}{1+\sqrt{1-4k}}Q$.

3. *T is topologically conjugate to $\check{T}$ for appropriate choice of η.* We now compose the mappings proved in steps 1 and 2 to show

$$T = \hat{h}^{-1} \circ \check{h}^{-1} \circ \check{T} \circ \check{h} \circ \hat{h} \tag{361}$$

with

$$\eta = 1 + \sqrt{1 - 4\left(-\omega_2\omega_0 + \frac{\omega_1}{2}\left(1 - \frac{\omega_1}{2}\right)\right)} = 1 + \sqrt{(\omega_1 - 1)^2 + 4\omega_2\omega_0}$$
$$= 1 + \sqrt{(a_O - a_P + b_P - 1)^2 + 4a_P(c + b_P - b_O)} \tag{362}$$

and therefore that T is topologically conjugate to $\check{T}$.

Having shown the conjugacy of T to $\check{T}$, we now find bounds on η implied by each case and use this conjugacy to derive the implications for possible dynamics. The following points prove each claim 1-3 in the original Proposition.

1. $\tilde{P}_O \geq \tilde{P}_P$ and both are monotone. Thus, T is increasing and there cannot be cycles or chaos. This implies that $\eta < 3$ (see Weisstein, 2001, for reference).

2. $\tilde{P}_O$ and $\tilde{P}_P$ are linear. It suffices to show that we can attain $\eta > 3$ but that η must be less than $1 + \sqrt{6}$ (see Weisstein, 2001, for reference). In this case, $c = 0$. This is in addition to the requirements that $\max_{Q\in[0,1]} \tilde{P}_i(Q) \leq 1$ and $\min_{Q\in[0,1]} \tilde{P}_i(Q) \geq 0$ for $i \in \{O, P\}$, which can be expressed as:

$$\max_{Q\in[0,1]} \tilde{P}_i(Q) = \max\left\{a_i, a_i + b_i - c, \left(a_i + \frac{b_i^2}{4c}\right)\mathbb{I}[0 \leq b_i \leq 2c]\right\} \leq 1$$
$$\min_{Q\in[0,1]} \tilde{P}_i(Q) = \min\{a_i, a_i + b_i - c\} \geq 0 \tag{363}$$

   The maximal value of η consistent with these restrictions can therefore be obtained by solving the following program:

$$\max_{(a_O,a_P,b_O,b_P)\in\mathbb{R}^4} (a_O - a_P + b_P - 1)^2 + 4a_P(b_P - b_O)$$
$$\text{s.t.} \max\{a_O, a_O + b_O\} \leq 1, \max\{a_P, a_P + b_P\} \leq 1 \tag{364}$$
$$\min\{a_O, a_O + b_O\} \geq 0, \min\{a_P, a_P + b_P\} \geq 0$$

   Exact solution of this program via Mathematica yields that the maximum value is 5. This implies that the maximum value of η is $1 + \sqrt{5} \approx 3.23$, which is greater than 3 but less than $1 + \sqrt{6}$. Moreover, this maximum is attained at $a_O = 0, a_P = 1, b_O = 0, b_P = -1$.

3. No further restrictions on $\tilde{P}_O$ and $\tilde{P}_P$. We can attain $\eta = 4$ by setting $a_0 = a_P = 0$, $b_O = b_P = 4$, $c = 4$. Thus, cycles of any period $k \in \mathbb{N}$ and chaotic dynamics can occur (see Weisstein, 2001, for reference).

□

The proof of this result follows a classic approach of recasting a quadratic difference equation as a logistic difference equation via topological conjugacy (see, *e.g.,* Battaglini, 2021; Deng, Khan, and Mitra, 2022). The restrictions on structural parameters implied by the hypotheses of the proposition then yield upper bounds on the possible logistic maps and allow us to characterize the possible dynamics using known results.

To understand this result, observe in our baseline case in which T is monotone that cycles and chaos are not possible. This is because there is no potential for optimism to sufficiently overshoot its steady state. By contrast, when $\tilde{P}_O$ and $\tilde{P}_P$ are either non-monotone or non-ranked, two-period cycles can take place where the economy undergoes endogenous boom-bust cycles with periods of high optimism and high output ushering in periods of low optimism and low output (and *vice versa*) as contrarians switch positions and consistently overshoot the (unstable) steady state. Finally, when $\tilde{P}_O$ and $\tilde{P}_P$ are non-linear and non-monotone, essentially any richness of dynamics can be achieved via erratic movements in optimism that are extremely sensitive to initial conditions.

An Empirical Test for Cycles and Chaos Proposition 34 shows how to translate an updating rule of the form of Equation 349 into predictions about the potential for cycles and chaos. We now estimate this updating rule in the data to test these predictions empirically. Concretely, in our panel dataset of firms, we estimate the regression model

$$\begin{aligned}\text{opt}_{it} = \alpha_1\, \text{opt}_{i,t-1} + \beta_1 \text{opt}_{i,t-1} \cdot \overline{\text{opt}}_{i,t-1} + \\ \beta_2(1 - \text{opt}_{i,t-1}) \cdot \overline{\text{opt}}_{i,t-1} + \tau\, (\overline{\text{opt}}_{i,t-1})^2 + \gamma_i + \varepsilon_{it}\end{aligned} \tag{365}$$

where γ_i is a firm fixed effect. This model allows the effects of contagiousness to depend on agents' previous state. In the mapping to Equation 349, $\alpha = a_P$, $\alpha_1 = a_O - a_P$, $\beta_1 = b_O$, $\beta_2 = b_P$, and $\tau = c$. With estimates of each regression parameter, denoted by a hat, we also obtain an estimate of the logistic map parameter η defined in Equation 351:

$$\hat{\eta} = 1 + \sqrt{(\hat{\alpha}_1 + \hat{\beta}_2 - 1)^2 + 4\hat{\alpha}_1(\hat{\tau} + \hat{\beta}_2 - \hat{\beta}_1)} \tag{366}$$

Since $\hat{\eta}$ is a nonlinear function of estimated parameters in the regression, we can conduct inference on $\hat{\eta}$ using the delta method. Moreover, this constitutes a test for the possibility of cycles and chaos in the model by the logic of Proposition 34. Specifically, as described in the proof of that result, there are two main cases. First, if $\eta < 3$, then case 1 of the result obtains: there are neither cycles of any period nor chaotic dynamics. Second, if $\eta \geq 3$, there can be cycles of period 2 or more and/or chaos. Moreover, if $\eta > 3.57$, chaotic dynamic obtain.

Our estimates are presented in Table A.20. Our point estimate of η is 1.443 and the 95% confidence interval is $(0.076, 2.810)$. This rules out, at the 5% level, the presence of cycles and/or chaos. The 99% confidence interval is $(-0.354, 3.240)$, which does not rule out cycles. The p-value for the chaotic dynamics threshold is 0.001. Thus, our results provide strong evidence against the possibility of chaos due to contagious optimism, and marginally weaker evidence against the possibility of cycles. This test complements the literature on endogenous cycles in macroeconomic models (see, *e.g.*, Boldrin and Woodford, 1990; Beaudry, Galizia, and Portier, 2020) by providing a micro-founded test within a structural economic model, which may ameliorate challenges associated with interpreting pure time-series evidence (see, e.g., Werning, 2017).

A.2.8 Narratives in Games and the Role of Higher-Order Beliefs

We have studied a micro-founded business-cycle model, but the basic insights extend much more generally to abstract, linear beauty contest games. Importantly, these settings provide us with an ability to disentangle the dual roles of narratives in affecting both agents' first-order and higher-order beliefs about fundamentals.

Concretely, suppose that agents' best replies are given by the following beauty contest form (see, *e.g.*, Morris and Shin, 2002):

$$x_{it} = \alpha \mathbb{E}_{it}[\theta_t] + \beta \mathbb{E}_{it}[Y_t] \tag{367}$$

where $\alpha > 0$ and $\beta \in [0, 1)$. This linear form for best replies is commonly justified by (log-)linearization of some underlying best response function (see, *e.g.*, Angeletos and Pavan, 2007). For example, log-linearization of the agents' best replies in the baseline model of this section yields such an equation with $\beta = \omega$ and all variables above standing in for their log-counterparts. Moreover, suppose that aggregation is linear so that $Y_t = \int_{[0,1]} x_{it} \mathrm{d}i$. This can similarly be justified via an appropriate first-order expansion of non-linear aggregators. Finally, we let the structure of narratives be as before.

Toward characterizing equilibrium, we define the average expectations operator:

$$\overline{\mathbb{E}}_t[\theta_t] = \int_{[0,1]} \mathbb{E}_{it}\left[\theta_t\right] \mathrm{d}i \tag{368}$$

and the higher-order average expectations operator for $k \in \mathbb{N}$ as:

$$\overline{\mathbb{E}}_t^k[\theta_t] = \int_{[0,1]} \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^{k-1}[\theta_t]\right] \mathrm{d}i \tag{369}$$

Moreover, we observe by recursive substitution that equilibrium aggregate output is given by:

$$Y_t = \alpha \sum_{k=1}^{\infty} \beta^{k-1} \overline{\mathbb{E}}_t^k[\theta_t] \tag{370}$$

We can therefore solve for the unique equilibrium by computing the hierarchy of higher-order expectations. We can do this in closed-form by observing that agents' idiosyncratic first-order beliefs are given by:

$$\mathbb{E}_t[\theta_t|s_{it}, \lambda_{it}] = \kappa s_{it} + (1-\kappa)\left(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right) \tag{371}$$

which allows us to compute average first-order expectations of fundamentals as:

$$\overline{\mathbb{E}}_t[\theta_t] = \kappa\theta_t + (1-\kappa)(Q_t\mu_O + (1-Q_t)\mu_P) \tag{372}$$

which is a weighted average between true fundamentals and the average impact of narratives on agents' priors. By taking agents' expectations over this object and averaging, we compute higher-order average expectations as:

$$\overline{\mathbb{E}}_t^k[\theta_t] = \kappa^k\theta_t + (1-\kappa^k)(Q_t\mu_O + (1-Q_t)\mu_P) \tag{373}$$

which is again a weighted average between the state and agents' priors, but now with a geometrically increasing weight on narratives as we consider higher-order average beliefs.

The following result characterizes aggregate output and agents' best replies in the unique equilibrium:

Proposition 35 (Narratives and Higher-Order Beliefs)**.** *There exists a unique equi-*

librium. In this unique equilibrium, aggregate output is given by:

$$Y_t = \frac{\alpha}{1-\beta}\left(\frac{(1-\beta)\kappa}{1-\beta\kappa}\theta_t + \frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right)\right) \tag{374}$$

Moreover, agents' actions follow:

$$\begin{aligned} x_{it} &= \alpha\frac{1}{1-\beta\kappa}\left[\kappa\theta_t + \kappa e_{it} + (1-\kappa)\left(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right)\right] \\ &\quad + \beta\frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right) \end{aligned} \tag{375}$$

Proof. To substantiate the arguments in the main text, by aggregating Equation 367, we obtain that:

$$Y_t = \alpha\overline{E}_t[\theta_t] + \beta\overline{E}_t[Y_t] \tag{376}$$

Thus, by recursive substitution k times we obtain that:

$$Y_t = \alpha\sum_{j=1}^{k}\beta^{j-1}\overline{E}_t^j[\theta_t] + \beta^k\overline{E}_t^k[Y_t] \tag{377}$$

Moreover, we have that:

$$\overline{\mathbb{E}}_t^j[\theta_t] = \kappa^j\theta_t + (1-\kappa^j)(Q_t\mu_O + (1-Q_t)\mu_P) \tag{378}$$

and thus that:

$$\begin{aligned} \alpha\sum_{j=1}^{k}\beta^{j-1}\overline{E}_t^j[\theta_t] &= \alpha\sum_{j=1}^{k}\beta^{j-1}\left(\kappa^j\theta_t + (1-\kappa^j)(Q_t\mu_O + (1-Q_t)\mu_P)\right) \\ = \alpha\sum_{j=1}^{k}\beta^{j-1}(Q_t\mu_O + (1-Q_t)\mu_P) &+ \alpha\beta^{-1}\sum_{j=1}^{k}(\beta\kappa)^j\left[\theta_t - (Q_t\mu_O + (1-Q_t)\mu_P)\right] \end{aligned} \tag{379}$$

Hence:

$$\begin{aligned} \lim_{k\to\infty}\alpha\sum_{j=1}^{k}\beta^{j-1}\overline{E}_t^j[\theta_t] = &\frac{\alpha}{1-\beta}(Q_t\mu_O + (1-Q_t)\mu_P) + \\ &\frac{\alpha\kappa}{1-\beta\kappa}\left[\theta_t - (Q_t\mu_O + (1-Q_t)\mu_P)\right] \end{aligned} \tag{380}$$

We therefore have that there is a unique equilibrium if $\lim_{k\to\infty}\beta^k\overline{E}_t^k[Y_t] = 0$. Hellwig and Veldkamp (2009) show in Proposition 1 of their supplementary material that all equilibria differ on a most a measure zero set of fundamentals. In this setting, this implies that $\lim_{k\to\infty}\beta^k \overline{E}_t^k[Y_t] = c$ for some $c \in \mathbb{R}$ for almost all $\theta \in \Theta$. Hence, the equilibrium is given by:

$$\begin{aligned} Y_t &= \frac{\alpha}{1-\beta}(Q_t\mu_O + (1-Q_t)\mu_P) + \frac{\alpha\kappa}{1-\beta\kappa}\left[\theta_t - (Q_t\mu_O + (1-Q_t)\mu_P)\right] + c \\ &= \frac{\alpha}{1-\beta}\left(\frac{(1-\beta)\kappa}{1-\beta\kappa}\theta_t + \frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right)\right) + c \end{aligned} \tag{381}$$

But then we have that $c = 0$ by computing $\lim_{k\to\infty}\beta^k \overline{E}_t^k[Y_t] = 0$ under this equilibrium.

Finally, to solve for individual actions under this equilibrium, we compute:

$$\begin{aligned} x_{it} &= \alpha\mathbb{E}_{it}[\theta_t] + \beta\mathbb{E}_{it}[Y_t] \\ &= \alpha\mathbb{E}_{it}[\theta_t] + \beta\mathbb{E}_{it}\left[\frac{\alpha}{1-\beta}\left(\frac{(1-\beta)\kappa}{1-\beta\kappa}\theta_t + \frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right)\right)\right] \\ &= \left(\alpha + \beta\frac{\alpha}{1-\beta}\frac{(1-\beta)\kappa}{1-\beta\kappa}\right)\mathbb{E}_{it}[\theta_t] + \beta\frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right) \\ &= \alpha\frac{1}{1-\beta\kappa}\left(\kappa s_{it} + (1-\kappa)\left(\lambda_{it}\mu_O + (1-\lambda_{it})\mu_P\right)\right) \\ &\quad + \beta\frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}\left(Q_t\mu_O + (1-Q_t)\mu_P\right) \end{aligned} \tag{382}$$

Completing the proof. □

This result allows us to see how narratives affect output by propagating up through the hierarchy of higher-order beliefs. Concretely, we have that the static impulse response of output to a contemporaneous shock to the fraction of optimists in the population is given by:

$$\frac{\partial Y_t}{\partial Q_t} = \frac{\alpha}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}(\mu_O - \mu_P) = \alpha\sum_{j=1}^{\infty}\beta^{j-1}(1-\kappa^j)(\mu_O - \mu_P) \tag{383}$$

The first expression is composed of the relative importance of fundamentals $\frac{\alpha}{1-\beta}$, the impact of prior beliefs on the entire hierarchy of higher-order beliefs about exogenous and endogenous outcomes $\frac{1-\kappa}{1-\beta\kappa}$ and the difference between the two narratives $\mu_O - \mu_P$. The second expression re-expands the heirarchy of beliefs, to highlight how fraction

$$\frac{\beta^{j-1}(1-\kappa^j)}{\frac{1}{1-\beta}\frac{1-\kappa}{1-\beta\kappa}} \tag{384}$$

of the total effect is driven by beliefs of order j. These weights decline more slowly if complementarity β or prior weights $1-\kappa$ are high.

Finally, our result shows how the regression equation relating individual actions with narrative weights, estimated in our main analysis, holds in equilibrium in the linearized beauty contest. Thus, our empirical strategy is compatible with the interpretation that the macroeconomy is best described by a linear beauty contest, and moreover can be ported to other settings where this modeling assumption may be appropriate, such as that of financial speculation (see *e.g.,* Allen, Morris, and Shin, 2006).

A.2.9 Model with Firm Dynamics

We now sketch an augmentation of our baseline conceptual model of the firm from which we derived our earlier estimating equations (see Appendix A.1.1) to allow for persistent idiosyncratic states and adjustment costs. This allows us to more formally justify why controlling for firm productivity and lagged labor is sufficient to account for the presence of adjustment costs to first-order.

In every period t, each firm i still takes an action $x_{it} \in \mathcal{X}$. Their objective function still takes as an input their action, aggregate outcomes $Y_t \in \mathcal{Y}$, and aggregate fundamentals θ_t (which in analogy to the previous appendix sections, we allow to follow a first-order (continuous) Markov process). However, they now have idiosyncratic fundamentals $\tilde{\theta}_{it}$, which follow a first-order (continuous) Markov process. Moreover, their actions are subject to adjustment costs $\Phi : \mathbb{R} \to \mathbb{R}_+$ equal to $\Phi(x - x_{-1})$ when their last action was x_{-1}. Thus, we let their flow utility be $u(x, Y, \theta, \tilde{\theta}) - \Phi(x - x_{-1})$. The firm discounts the future at rate $\beta_i \in [0, 1)$. The aggregate state variables in period t are the distribution of x_{it-1} in the population F^x_{t-1}, the distribution of narratives in the population Q_t, and the level of current and past aggregate fundamentals θ_t and θ_{t-1}. Thus, equilibrium aggregate output is described by some function $\hat{Y}(F^x_{t-1}, Q_t, \theta_t, \theta_{t-1})$. Moreover, observe at time t that the following are the state variables for a firm: (i) the level of idiosyncratic productivity in the previous period $\tilde{\theta}_{it-1}$ (ii) the level of aggregate productivity in the previous period θ_{t-1} (iii) the firm's action in the previous period x_{it-1} (iv) the narrative entertained by the agent λ_{it} (v) their current signal about fundamentals s_{it}, and (vi) the additional aggregate states (F^x_{t-1}, Q_t).

We can therefore represent any firm policy function as:

$$x_{it} = g(x_{it-1}, \theta_{t-1}, \tilde{\theta}_{it-1}, F^x_{t-1}, Q_t, \lambda_{it}, s_{it}) \tag{385}$$

If this is differentiable, we may linearize it to obtain:

$$x_{it} \approx \gamma_i + \chi_t + \sum_{k=1}^{K} \delta_k \lambda_{k,it} + \gamma\theta_{it-1} + \omega x_{it-1} + \varepsilon_{it} \tag{386}$$

where the aggregate fixed effect now absorbs $(F^x_{t-1}, Q_t, \theta_t, \theta_{t-1})$, θ_{it-1} capture agents' idiosyncratic expectations of future fundamentals, and x_{it-1} captures their adjustment costs.

A.3 Additional Details on Textual Data

A.3.1 Obtaining and Processing 10-Ks

Here, we describe our methodology for obtaining and processing raw data on 10-K filings. We start with raw `html` files downloaded directly from the SEC's EDGAR (Electronic Data Gathering, Analysis, and Retrieval) system. Each of these files corresponds to a single 10-K filing. Each file is identified by its unique accession number. In its heading, each file also contains the end-date for the period the report concerns (*e.g.*, 12/31/2018 for a FY 2018 ending in December), and a CIK (Central Index Key) firm identifier from the SEC. We use standard linking software provided by Wharton Research Data Services (WRDS) to link CIK numbers and fiscal years to the alternative firm identifiers used in data on firm fundamentals and stock prices. We have, in our original dataset, 182,259 files.

We follow the following steps to turn each document, now identified by firm and year, into a bag-of-words representation:

1. *Cleaning raw text.* We first translate the document into unformatted text. Specifically, we follow the following steps in order:

    (a) Removing hyperlinks and other web addresses

    (b) Removing html formatting tags encased in the brackets <>

    (c) Making all text lowercase

    (d) Removing extra spaces, tabs, and new lines.

    (e) Removing punctuation

    (f) Removing non-alphabetical characters

2. *Removing stop words.* Following standard practice, we remove "stop words" which are common in English but do not convey specific meaning in our analysis. We use the default English stop word list in the `nltk` Python package. Example stopwords include articles ("a","the"), pronouns ("I","my"), prepositions ("in","on"), and conjunctions ("and","while").

3. *Lemmatizing documents.* Again following standard practice, we use lemmatization software to reduce words to their common roots. We use the default English-language lemmatizer of the `spacy` Python package. The lemmatizer uses both the word's identity and its content to transform sentences. For instance, when each is used as a verb, "meet," "met," and "meeting" are commonly

lemmatized to "meet." But if the software predicts that "meeting" is used as a noun, it will be lemmatized as the noun "meeting."

4. *Estimating a bigram model.* We estimate a bigram model to group together commonly co-occurring words as single two-word phrases. We use the `phrases` function of the `gensim` package. The bigram modeler groups together words that are almost always used together. For instance, if our original text data set were the 10-Ks of public firms Nestlé and General Mills, the model may determine that "ice" and "cream," which almost always appear together, are part of a bigram "ice_cream."

5. *Computing the bag of words representation.* Having now expressed each document as a vector of clean words (*i.e.,* single words and bigrams), we simply collapse these data to frequencies.

Finally, note that our procedure uses *all* of the non-formatting text in the 10K. This includes all sections of the documents, and does not limit to the Management Discussion and Analysis (MD&A) section. This is motivated by the fact that management's discussion is not limited to one section SEC (2011). Moreover, prior literature has found that textual analysis of the entire 10-K versus the MD&A section tends to closely agree, and that limiting scope to the MD&A section has limited practical benefits due to the trade-off of limiting the amount of text per document (Loughran and McDonald, 2011a).

A.3.2 Obtaining and Processing Conference Call Text

We obtain the full text of sales and earnings conference calls from 2002 to 2014 from the Fair Disclosure (FD) Wire service. The original sample includes 261,034 documents, formatted as raw text. We next subset to documents that have reported firm names and stock tickers, which are automatically associated with documents by Lexis Nexis. When matches are probabilistic, we use the first (highest probability) match.[5] We finally restrict to firms that are listed on one of three US stock exchanges: the NYSE, the NASDAQ, or the NYSE-MKT (Small Cap). We finally connect tickers to the firm identifiers in our fundamentals data using the master cross-walk available on Wharton Research Data Services (WRDS). These operations together reduce the sample size to 158,810 calls. We clean these data by conducting steps 1-3 described

[5] In the essentially zero-measure cases in which there is a tie, we take the alphabetically first ticker.

above in Appendix A.3.1. We then calculate positive word counts, negative word counts, and optimism exactly as described in the main text for the 10-K data.

A.3.3 Measuring Positive and Negative Words

To calculate sets of positive and negative 10K words, we use the updated dictionary available online at McDonald (2021) as of June 2020. This dictionary includes substantial updates relative to the dictionaries associated with the original Loughran and McDonald (2011a) publication. These changes are reviewed in the *Documentation* available at McDonald (2021).

The Loughran-McDonald dictionary includes 2345 negative words and 347 positive words. The dictionary is constructed to include multiple forms of each relevant word. For instance, the first negative root "abandon" is listed as: "abandon," "abandoned," "abandoning," "abandonment," "abandonments," and "abandons." To ensure consistency with our own lemmatization procedure, we first map each unique word to all of its possible lemmas using the `getAllLemmas` function of the `lemminflect` Python package, which is an extension to the `spacy` package we use for lemmatization. We then construct a new list of negative words by combining the original list of negative words with all new, unique lemmas to which a negative word mapped (and similarly for positive words). This procedure results in new lists of 2411 negative words and 366 positive words, which map exactly to the words that appear in our cleaned bag of words representation. We list the top ten most common positive and negative words from this cleaned set in Table A.1. In particular, to make the table most legible, we first associate words with their lemmas, then count the sum of document frequencies for each associated word (which may exceed one), and then print the most common word associated with the lemma.

A.4 Additional Details on Firm Fundamentals Data

A.4.1 Compustat: Data Selection

Our data selection criteria and variable definitions are identical to those used in Flynn and Sastry (2022a). In this Appendix, we review essential points. We refer the reader to the Appendix material of Flynn and Sastry (2022a) for certain details.

Our dataset is Compustat Annual Fundamentals. Our main variables of interest are defined in Appendix Table A.21. We restrict the sample to firms based in the United States, reporting statistics in US Dollars, and present in the "Industrial" dataset. We exclude firms whose 2-digit NAICS is 52 (Finance and Insurance) or 22 (Utilities). This filter eliminates firms in two industries that, respectively, may have highly non-standard production technology and non-standard market structure.

We summarize our definitions of major "input and output" variables in Appendix Table A.21. For labor choice, we measure the number of employees. For materials expenditure, we measure the sum of reported variable costs (`cogs`) and sales and administrative expense (`xsga`) net of depreciation (`dp`).[6] As in Ottonello and Winberry (2020) and Flynn and Sastry (2022a), we use a perpetual inventory method to calculate the value of the capital stock. We start with the first reported observation of gross value of plant, property, and equipment and add net investment or the differences in net value of plant, property, and equipment. Note that, because all subsequent analysis is conditional on industry-by-time fixed effects, it is redundant at this stage to deflate materials and capital expenditures by industry-specific deflators.

We categorize the data into 44 sectors. These are defined at the 2-digit NAICS level, but for the Manufacturing (31-33) and Information (51) sectors, which we classify at the 3-digit level to achieve a better balance of sector size. More summary information about these industries is provided in Appendix F of Flynn and Sastry (2022a).

A.4.2 Compustat: Calculation of TFP

When calculating firms' Total Factor Productivity, we restrict attention to a subset of our sample that fulfils the following inclusion criteria:

1. Sales, material expenditures, and capital stock are strictly positive;

2. Employees exceed 10;

[6] A small difference from Flynn and Sastry (2022a) is that, in assessing the firms' costs and later calculating TFP, we do not "unbundle" materials expenditures on labor and non-labor inputs using supplemental data on annual wages.

3. Acquisitions as a proportion of assets (`aqc` over `at`) does not exceed 0.05.

The first ensures that all companies meaningfully report all variables of interest for our production function estimation; the second applies a stricter cut-off to eliminate firms that are very small, and lead to outlier estimates of productivity and choices. The third is a simple screening device for large acquisitions which may spuriously show up as large innovations in firm choices and/or productivity.

Our method for recovering total factor productivity is based on cost shares. In brief, we use cost shares for materials to back out production elasticities, and treat the elasticity of capital as the implied "residual" given an assumed mark-up $\mu > 1$ (in our baseline, $\mu = 4/3$) and constant physical returns-to-scale. The exact procedure is the following:

1. For all firms in industry j, calculate the estimated materials share:

$$\text{Share}_{M,j'} = \frac{\sum_{i:j(i)=j'} \sum_t \text{MaterialExpenditure}_{it}}{\sum_{i:j(i)=j'} \sum_t \text{Sales}_{it}} \tag{387}$$

2. If $\text{Share}_{M,j'} \leq \mu^{-1}$, then set

$$\begin{aligned} \alpha_{M,j'} &= \mu \cdot \text{Share}_{M,j'} \\ \alpha_{K,j'} &= 1 - \alpha_{M,j'} \end{aligned} \tag{388}$$

3. Otherwise, adjust shares to match the assumed returns-to-scale, or set

$$\begin{aligned} \alpha_{M,j'} &= 1 \\ \alpha_{K,j'} &= 0 \end{aligned} \tag{389}$$

To translate our production function estimates into productivity, we calculate a "Sales Solow Residual" $\tilde{\theta}_{it}$ of the following form:

$$\log \tilde{\theta}_{it} = \log \text{Sales}_{it} - \frac{1}{\mu} \left(\alpha_{M,j(i)} \cdot \log \text{MatExp}_{it} + \alpha_{K,j(i)} \cdot \log \text{CapStock}_{it} \right) \tag{390}$$

We finally define our estimate $\log \hat{\theta}$ as the previous net of industry-by-time fixed effects

$$\log \hat{\theta}_{it} = \log \tilde{\theta}_{it} - \chi_{j(i),t} \tag{391}$$

Theoretical Interpretation The aforementioned method recovers physical productivity ("TFPQ") under the assumptions, consistent with our quantitative model, that firms operate constant returns-to-scale technology and face an isoleastic, downward-sloping demand curve of *known* elasticity (equivalently, they charge a known markup). The idea is that, given the known markup, we can impute firms' (model-consistent) costs as a fixed fraction of sales and then calculate the theoretically desired cost shares. Here, we describe the simple mathematics.

There is a single firm i operating in industry j with technology

$$Y_i = \theta_i M_i^{\alpha_j} K_i^{1-\alpha_j} \tag{392}$$

They act as a monopolist facing the demand curve

$$p_i = Y_i^{-\frac{1}{\epsilon}} \tag{393}$$

for some inverse elasticity $\epsilon > 1$. Observe that this is, up to scale, the demand function faced by monopolistically competitive intermediate goods producers in our model. The firm's revenue is therefore $p_i Y_i = Y_i^{1-\frac{1}{\epsilon}}$. Finally, the firm can buy materials at industry-specific price q_j and rent capital at rate r_j. The firm's program for profit maximization is therefore

$$\max_{M_i, K_i} \left\{ (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} - q_j M_i - r_j K_i \right\} \tag{394}$$

We first justify our formulas for the input shares (Equation 388). To do this, we solve for the firm's optimal input choices. This is a concave problem, in which first-order conditions are necessary and sufficient. These conditions are

$$\begin{aligned} q_j &= M_i^{-1} \alpha_j \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \\ r_j &= K_i^{-1} (1-\alpha_j) \left(1 - \frac{1}{\epsilon}\right) (\theta_i M_i^{\alpha_j} K_i^{1-\alpha_j})^{1-\frac{1}{\epsilon}} \end{aligned} \tag{395}$$

Re-arranging, and substituting in $p_i = Y_i^{-\frac{1}{\epsilon}}$, we derive

$$\begin{aligned} \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{q_j M_i}{p_i Y_i} \\ 1 - \alpha_j &= \frac{\epsilon}{\epsilon - 1} \frac{r_j K_i}{p_i Y_i} \end{aligned} \tag{396}$$

Or, in words, that the materials elasticity is $\frac{\epsilon}{\epsilon-1}$ times the ratio of materials input expenditures to sales. Observe also that, by re-arranging the two first-order conditions, we can write expressions for production and the price

$$Y = \left(\left(\frac{\epsilon - 1}{\epsilon} \right) \theta_i \left(\frac{\alpha_j}{q_j} \right)^{\alpha} \left(\frac{1 - \alpha_j}{r_j} \right)^{1-\alpha_j} \right)^{\epsilon} \Rightarrow p = \left(\frac{\epsilon}{\epsilon - 1} \right) \theta_i^{-1} \left(\frac{q_j}{\alpha_j} \right)^{\alpha_j} \left(\frac{r_j}{1 - \alpha_j} \right)^{1-\alpha_j} \tag{397}$$

and observe that $\theta_i^{-1} \left(\frac{q_j}{\alpha_j} \right)^{\alpha_j} \left(\frac{r_j}{1-\alpha_j} \right)^{1-\alpha_j}$ is the firm's marginal cost. Hence, we can define $\mu = \frac{\epsilon}{\epsilon-1} > 1$ as the firm's markup and write the shares as required:

$$\alpha = \mu \frac{q_j M_i}{p_i Y_i} \tag{398}$$

Finally, we now apply Equations 390 and 391 to calculate productivity. Assume that we observe materials expenditure $q_j M_i$ and capital value $p_{K,j} K_i$, where $p_{K,j}$ is an (unobserved) price of capital. We find

$$\log \tilde{\theta}_i = \left(1 - \frac{1}{\epsilon} \right) (\log \theta_i - \alpha \log q_j - (1 - \alpha) \log p_{K,j}) \tag{399}$$

We finally observe that the industry-level means are

$$\chi_j = \left(1 - \frac{1}{\epsilon} \right) \left(\log \bar{\theta}_j - \alpha \log q_j - (1 - \alpha) \log p_{K,j} \right) \tag{400}$$

where $\log \bar{\theta}_j$ is the mean of $\log \theta_i$ over the industry. Hence,

$$\log \hat{\theta}_i = \left(1 - \frac{1}{\epsilon} \right) (\log \theta_i) \tag{401}$$

or our measurement captures physical TFP, up to scale.

A.5 Additional Empirical Results

A.5.1 A Test for Coefficient Stability

Here, we study the bias that may arise from omitted variables in our estimation of the effect of narrative optimism on hiring, or δ^{OP} in Section 1.4.1, Equation 12, and Table 1.1. In particular, we apply the method of Oster (2019) to bound bias in the estimate of δ^{OP} under external assumptions about selection on unobservable variables and to calculate an extent of unobservable selection that could be consistent with a point estimate $\delta^{OP} = 0$ that corresponds to our null hypothesis (*i.e.,* "narrative optimism is irrelevant for hiring"). We find that our results are highly robust by this criterion.

Set-up and Review of Methods To review, our estimating equation is

$$\Delta \log L_{it} = \delta^{OP} \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it} \tag{402}$$

Hiring and optimism are constructed as described in Section 1.3, at the level of firms and fiscal years. We treat firm and industry-by-time fixed effects as baseline controls that are necessary for interpreting the regression.[7] As our main "discretionary" controls, we consider current and past TFP and lagged labor—that is, $X_{it} = \{\log \hat{\theta}_{it}, \log \hat{\theta}_{i,t-1}, \log L_{i,t-1}\}$. Under our baseline model, these controls help increase precision, as they are in principle observable variables that explain hiring (Corollary 1). Thus, in this Appendix, we will study the regression model in which the fixed effects are partialed out of both the outcome, main regression, and controls, as indicated below with the $\perp$ superscript:

$$\Delta \log L_{it}^{\perp} = \delta^{OP} \text{opt}_{it}^{\perp} + \tau' X_{it}^{\perp} + \varepsilon_{it}^{\perp} \tag{403}$$

The essence of the method proposed by Oster (2019), who builds on the approach of Altonji, Elder, and Taber (2005), is to extrapolate the change in the coefficient in interest upon the addition of control variables, taking into account the better fit (*i.e.,* additional R^2) from adding the new regressors. To exemplify the logic, consider a case in which we first estimated Equation 403 without controls, obtaining a coefficient estimate of $\hat{\delta}_{NC}^{OP}$ and an R^2 of $\hat{R}_{NC}^2$, and then estimated the same equation with controls, obtaining a coefficient estimate of $\hat{\delta}_{C}^{OP}$ and an R^2 of $\hat{R}_{C}^2$. Both estimates are restricted to a common sample, for comparability. If $\hat{R}_{C}^2 = 1$, then (up to estimation

[7] The latter, in particular, controls for the effect of fundamentals on hiring in our macroeconomic model. We leverage this interpretation of the *biased* estimate of δ^{OP} from a regression lacking this fixed effect in Appendix A.6.3.

error) we might presume that $\hat{\delta}_C^{OP} - \hat{\delta}_{NC}^{OP}$ estimates the entirety of the theoretically possible omitted variables bias, as there is no remaining unmodeled variation in hiring. If $\hat{R}_C^2 < 1$ and $\hat{R}_C^2 - \hat{R}_{NC}^2$ is small (*i.e.,* the controls did not greatly improve fit), then we might presume that the residual still contains unobserved variables that could contribute toward more bias—in other words, the observed omitted variables bias $\hat{\delta}_C^{OP} - \hat{\delta}_{NC}^{OP}$ is only a small fraction of what is possible.

To formalize this idea, Oster (2019) introduces two auxiliary parameters: λ (the *proportional degree of selection*, called δ in the original paper), which controls the relative effect of observed and unobserved controls on the outcome, and $\bar{R}^2$, which is the maximum achievable fit of the regression with all (possibly bias-inducing) controls, presumed in the example above to be 1. Conditional on $\bar{R}^2$, Oster (2019) proposes an intuitively reasonable (and, in special cases and under specific asymptotic arguments, consistent) estimator for the degree of selection required to induce a zero coefficient, $\hat{\lambda}^*$. Conditional on both $\bar{R}^2$ and λ, Oster (2019) also proposes a bias-corrected coefficient estimator, which is $\hat{\delta}_{OP}^*$ in our language.

The key parameter that the researcher has to specify for the first calculation is $\bar{R}^2$: the proportion of variance in the outcome variable (hiring, net of firm and sector-by-time fixed effects) that can be explained by factors that correlate with the variable of interest (optimism) and explain the outcome variable. As the main source of omitted variation that could influence optimism and hiring is news about fundamentals, we benchmark $\hat{\bar{R}}^2$ by estimating a regression in which we include our base control set $X_{it} = \{\log \hat{\theta}_{it}, \log \hat{\theta}_{i,t-1}, \log L_{i,t-1}\}$ and control for two years of future fundamentals and labor choice, or

$$Z_{it} = \{\log \hat{\theta}_{i,t+1}, \log \hat{\theta}_{i,t+2}, \log L_{i,t+1}, \log L_{i,t+2}\}$$

This yields $\hat{\bar{R}}^2 = 0.459$. Oster (2019) also suggests as a benchmark that $\bar{R}^2$ could be taken as three times the R^2 in the controlled regression. We also report robustness to $\bar{R}^2_{\mathrm{II}} = 0.387$, three times the value of $R^2 = 0.129$ that we find in the controlled regression. Thus, our baseline value of $\hat{\bar{R}}^2 = 0.459$ is more demanding than that suggested by Oster (2019). We finally construct the bias-corrected coefficients assuming $\lambda = 1$, or equal selection on unobservables and observables, for both values of $\bar{R}^2$.

Results We report the results of this exercise in Table A.6. Under our baseline value of $\hat{\bar{R}}^2 = 0.459$, we find that the degree of selection required to induced a zero coefficient is $\hat{\lambda}^* = 1.69$. This is well above the value of $\hat{\lambda}^* = 1$ that Oster (2019) suggests is likely to be conservative. Under the "three times R^2" benchmark, we obtain that $\hat{\lambda}^* = 2.15$. In both cases, we are robust to there being more selection on unobservables than on observables. According to Oster (2019), approximately 50% of the published top-journal articles in their sample are not robust to this extent of selection.

A.5.2 Alternative Empirical Strategy: CEO Change Event Studies

To further isolate variation in the narratives held by firms that is unrelated to fundamentals, we study the effects on hiring of changes in narratives induced by plausibly exogenous managerial turnover.

Data To obtain plausibly exogenous variation in narratives held at the firm level, we will examine the year-to-year change in firm-level narratives stemming from plausibly exogenous CEO changes. To do this, we use the dataset of categorized CEO exits compiled by Gentry, Harrison, Quigley, and Boivie (2021). These data comprise 9,390 CEO turnover events categorized by the reason for the CEO exit. The categorization was performed using primary sources (*e.g.,* press releases, newspaper articles, and regulatory filings) by undergraduate students in a computer lab, supervised by graduate students, with the final dataset checked by both a data outsourcing company and an additional student. We restrict attention to CEO exists caused by death, illness, personal issues, and voluntary retirements. Importantly, we exclude all CEO exits caused by inadequate job performance, quits, and forced retirement.

The Effect of Optimism on Hiring We first revisit our empirical strategy for measuring the effect of optimism on firms' hiring, using the CEO change event studies. For all firms i and years t such that i's CEO leaves because of death, illness, personal issues or voluntary retirements, we estimate the regression equation

$$\Delta \log L_{it} = \delta^{CEO} \text{opt}_{it} + \psi \, \text{opt}_{i,t-1} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it} \tag{404}$$

This differs from our baseline Equation 12 by including parametric controls for lagged values of the narrative loadings, but removing a persistent firm fixed effect.[8] If the studied CEO changes are truly exogenous, as we have suggested, then the narrative loadings of the new CEO are, conditional on the narrative loadings of the previous CEO, solely due to the differences in worldview across these two senior executives. Of course, CEO exits may be disruptive and reduce firm activity. Any time- and

[8] With a firm fixed effect, the regression coefficients of interest would be identified only from firms with multiple plausibly exogenous CEO exits.

industry-varying effects of CEO exits via disruption are controlled for by the intercept of the regression $\chi_{j(i),t}$, since the equation is estimated only on the exit events. Moreover, any within-industry, time-varying, and idiosyncratic disruption is captured through our maintained productivity control. Under this interpretation, the coefficient of interest δ^{CEO} isolates the effect of optimism on hiring purely via the channel of changing managements' narratives.

We present our results in Table A.22. We obtain estimates of δ^{CEO} that are quantitatively similar to our estimates of δ^{OP} in Table 1.1 (columns 1, 2, and 3). In column 4, we estimate a regression equation on the full sample that measures the direct effect of CEO changes and its interaction with the new management's optimism. Specifically, we estimate

$$\begin{aligned}\Delta \log L_{it} = \delta^{\text{NoChange}}\text{opt}_{it} + \delta^{\text{Change}}(\text{opt}_{it} \times \text{ChangeCEO}_{it}) + \alpha^{\text{Change}}\text{ChangeCEO}_{it} \\ + \psi\, \text{opt}_{i,t-1} + \tau' X_{it} + \chi_{j(i),t} + \varepsilon_{it}\end{aligned} \tag{405}$$

where ChangeCEO_{it} is an indicator for our plausibly exogenous CEO change events. We find that CEO changes in isolation reduce hiring ($\alpha^{\text{Change}} < 0$) but also that the effect of optimism is magnified when it accompanies a CEO change ($\delta^{\text{Change}} > 0$). This is further inconsistent with a story under which omitted fundamentals lead us to overestimate the effect of optimism on hiring.

Contagiousness from CEO Change Spillovers We next leverage changes in within-sector and peer-set optimism induced by plausibly exogenous CEO changes as instruments for the level of optimism within these groups. Concretely, we construct an instrument equal to the contribution toward optimism from firms whose CEOs changed for a plausibly exogenous reason, or

$$\overline{\text{opt}}^{\text{ceo}}_{j(i),t-1} = \frac{1}{|M_{j(i),t}|} \sum_{k \in M^c_{j(i),t}} \text{opt}_{k,t-1} \tag{406}$$

where $M_{j(i),t}$ is the set of firms in industry $j(i)$ at time t, and $M^c_{j(i),t}$ is the subset that had plausibly exogenous CEO changes. We construct the peer-set instrument $\overline{\text{opt}}^{\text{ceo}}_{p(i),t-1}$ analogously. We use $(\overline{\text{opt}}^{\text{ceo}}_{j(i),t-1}, \overline{\text{opt}}^{\text{ceo}}_{p(i),t-1})$ as instruments for $(\overline{\text{opt}}_{j(i),t-1}, \overline{\text{opt}}_{p(i),t-1})$ in the estimation of Equation 19. We present the corresponding estimates in Table A.23. We find similar point estimates under IV and OLS, although the IV estimates are significantly noisier.

A.5.3 Narrative Optimism, Beliefs, and Hiring

In this appendix, we study whether narrative optimism, measured using text-analysis methods, matters for firm decisions conditional on firm-manager beliefs, measured from recorded managerial guidance. We find that narrative optimism and measured expectations each have predictive power conditional on the other for explaining hiring and capital investment. These results suggest that textual optimism captures aspects of managers' latent beliefs not captured in traditional measurement of expectations (here, in guidance data).

Data We collect data from IBES (the International Brokers' Estimate System) on quantitative forecasts by company managers for three statistics: sales, capital expenditures (CAPX), and earnings per share (EPS). As described in Section 1.3.1, we restrict to the first recorded forecast per fiscal year of that year's variable. When managers' guidance is reported as a range, we code a point-estimate forecast as the range's midpoint. For each variable $Z \in \{\text{Sales}, \text{Capx}, \text{Eps}\}$, we calculated the manager's predicted growth for fiscal year t as

$$\text{ForecastGrowthZ}_{it} = \log \text{GuidanceForX}_{it} - \log Z_{i,t-1} \tag{407}$$

For example, $\text{GuidanceForSales}_{it}$ is the manager's earliest recorded guidance within fiscal year t for fiscal-year t sales, and $\text{Sales}_{i,t-1}$ are recorded sales from fiscal year $t-1$. Textual narrative optimism opt_{it} is measured as in our main analysis.

Empirical Strategy We re-create our main regression model, predicting hiring by opt_{it} conditional on firm fixed effects and industry-by-time fixed effects. We now include, as control variables, each of the $\text{ForecastGrowthZ}_{it}$ variables:

$$\Delta \log L_{it} = \delta^{OP}\text{opt}_{it} + \delta^{Z}\,\text{ForecastGrowthZ}_{it} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{408}$$

The coefficient δ^{OP} measures the difference in hiring between textually optimistic and non-optimistic firms, holding fixed forecasted growth about variable Z (and the fixed effects). The coefficient δ^{Z} measures the marginal effect of forecasted growth, in variable Z, on hiring, holding fixed whether the firm is optimistic or pessimistic (and the fixed effects). We also estimate a variance with net investment, or $\Delta \log K_{it}$, on the left-hand side.

Results Table A.24 shows the results when hiring is the outcome. We find that forecasted sales, CAPX, and earnings growth have positive effects of hiring, the first two of which are statistically significant (columns 2-4). Nonetheless, conditional on these variables, optimism has a positive effect on hiring of comparable magnitude to the baseline (column 1). The effect is statistically significant conditional on forecasted sales and CAPX growth (respectively, $t = 1.80$ and $t = 4.54$). Both optimism and forecasted EPS growth are insignificant predictors of hiring on the small ($N = 1290$) sample for which we can obtain EPS growth forecasts.

To compare the magnitudes of effects, we can calculate standardized coefficients. These have units of the effect of a one-standard-deviation change in the regressor on standard deviations of the outcome. For column 2, the standardized coefficient on textual optimism is 0.057 (SE: 0.0030) and the coefficient on predicted sales growth is 0.213 (SE: 0.0329). In this sense, predicted sales growth, for the subset of firms for which it is available, explains *larger* variations in hiring than textual optimism; but nonetheless, textual optimism has a statistically and economically significant effect.

Table A.25 shows analogous results when net capital investment is the outcome. As with hiring, we verify that predicted sales and CAPX growth have statistically significant, positive effects on capital investment, and that optimism has a positive effect conditional on these variables. Effects on the sub-sample with earnings guidance are noisy, for both the effects of optimism and the effects of forecasts.

Discussion We interpret our results in a model in which textual optimism, opt_{it}, is one measurement of a non-fundamental shifter in firm managers' beliefs. We validate this interpretation in the paper by showing that opt_{it}: (i) predicts hiring, as reviewed above in this note; (ii) does not predict future positive firm performance; and (iii) does correlate with optimistic manager forecasts, when measured in a variety of ways. We interpret managerial forecasts, about a variety of firm-specific variables, as alternative possible measurements of beliefs and their non-fundamental component. We are agnostic, more or less, about which of these measures explains more variation in firm actions or does so more precisely.

More broadly, while forecasts are quantitative, and provide *hard information* about managers' beliefs, they also capture at best only one or two moments of a probability distribution. By contrast, our measures of text all us to capture information about managers' beliefs that they do not express numerically, *i.e.,* we capture *soft information* about managers' beliefs (Liberti and Petersen, 2019). We find that this soft aspect of managers' beliefs is important for explaining their decisions conditional on hard information. This is consistent with extensive economic and psychological evidence that humans do not naturally think probabilistically (see *e.g.,* Tversky and Kahneman, 1973). Language may reflect nuances not present in the forecasts. These nuances are actually what we want to map to economic models, where we (economists) introduce statistical beliefs to model sentiment. This is the sense in which language might measure aspects of beliefs that are not captured in "measured beliefs." In this way, our results relate to a literature focusing on the decision-relevance of measured beliefs (Gennaioli, Ma, and Shleifer, 2016). They are moreover consistent with the literature focusing on the decision-relevance of textually measured firm-level variables including reported risks (Hassan, Hollander, Van Lent, and Tahoun, 2019) and reported uncertainty (Handley and Li, 2020).

A.5.4 State-Dependent Effects of Sentiment

Our main empirical framework assumes that the effect of narrative sentiment on hiring does not depend on previous sentiment. As one concrete example, this rules out the possibility that switching from relative optimism to relative pessimism has a larger effect than remaining equally pessimistic for two consecutive periods. To test for such state-dependent effects, we estimate augmented regression equations of the form:

$$\begin{aligned}\Delta \log L_{it} = {} & \delta_0 \text{sentiment}_{it} + \delta_1 \text{sentiment}_{i,t-1} + \\ & \delta_2 (\text{sentiment}_{it} \times \text{sentiment}_{i,t-1}) + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it}\end{aligned} \tag{409}$$

where sentiment_{it} is our continuous measure of firm sentiment in language, $(\gamma_i, \chi_{j(i),t})$ are fixed effects at the firm and industry-by-time levels, and X_{it} is a vector of controls. This model allows for the marginal effect of this fiscal year's sentiment to depend on the level of the previous fiscal year's sentiment. In particular, if $\delta_2 > 0$, and the marginal effect of sentiment is positive, then this marginal effect is higher for a previously positive firm; if $\delta_2 < 0$, and the marginal effect of sentiment is positive, then this marginal is lower for a previously positive firm.

Table A.26 shows our results, for different choices of controls. We find significant evidence of positive marginal effects for sentiment_{it} and $\delta_2 < 0$, or larger marginal effects when lagged sentiment is low. This asymmetry is quantitatively small, however, in the following sense. The standard deviation of $\text{sentiment}_{i,t-1}$ is 1.14. Using the estimates of column 1, a one-standard-deviation increase in sentiment, starting from $\text{sentiment}_{i,t-1} = 0$, decreases the marginal effect of sentiment_{it} from 0.022 to 0.016.

A.5.5 Measuring Contagiousness via Granular Instrumental Variables

As an alternative strategy to estimate contagiousness, we apply the methods of Gabaix and Koijen (2020) to construct "granular variables" that aggregate idiosyncratic variation in large firms' narrative loadings. We find evidence that the idiosyncratic optimistic updating of large firms induces optimistic updating, a form of contagiousness.

Constructing the Granular Measures We construct our granular instruments via the following algorithm. We first estimate a firm-level updating regression that controls non-parametrically for aggregate trends and parametrically for firm-level conditions. Specifically, we estimate

$$\text{opt}_{it} = \tau' X_{it} + \chi_{j(i),t} + \gamma_i + u_{it} \tag{410}$$

where $\chi_{j(i),t}$ is an industry-by-time fixed effect (sweeping out industry-specific aggregate shocks), γ_i is a firm fixed effect (sweeping out compositional effects), and X_{it} is the largest vector of controls used in the analysis of Section 1.4.1, consisting of: lagged log employment, current and lagged log TFP, log stock returns, the log book to market ratio, and leverage. We construct the empirical residuals $\hat{u}_{it}$. To construct the aggregate granular variable, $\overline{\text{opt}}_t^{g,sw}$, we take a sales-weighted average of these residuals:

$$\overline{\text{opt}}_t^{g,sw} = \sum_i \frac{\text{sales}_{it}}{\sum_i \text{sales}_{it}} \hat{u}_{it} \tag{411}$$

To construct an industry-level granular variable, $\overline{\text{opt}}_{j(i),t}^{g,sw}$, we take the leave-one-out sales-weighted average of the $\hat{u}_{it}$:

$$\overline{\text{opt}}_t^{g,sw} = \sum_{i':j(i)=j(i'),i'\neq i} \frac{\text{sales}_{i't}}{\sum_i \text{sales}_{i't}} \hat{u}_{i't} \tag{412}$$

We also construct agggregate and industry (leave-one-out) averages of opt_{it} for comparison. We denote these variables as $\overline{\text{opt}}_t^{sw}$ and $\overline{\text{opt}}_{j(i),t}^{sw}$, respectively.

Empirical Strategy At the aggregate level, we first consider a variant of our main model Equation 18, but with one of the sales-weighted variables $Z_t \in \{\overline{\text{opt}}_t^{sw}, \overline{\text{opt}}_t^{g,sw}\}$:

$$\text{opt}_{it} = u \ \text{opt}_{i,t-1} + s \ Z_{t-1} + r \ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it} \tag{413}$$

The coefficient s measures contagiousness with respect to the sales-weighted measures of optimism. We estimate Equation 413 by OLS, and also estimate a version in which the granular variable $\overline{\text{opt}}_t^{g,sw}$ is an instrumental variable for the raw sales-weighted average $\overline{\text{opt}}_t^{sw}$.

Similarly, at the industry level, we estimate the model

$$\text{opt}_{it} = u_{\text{ind}}\ \text{opt}_{i,t-1} + s_{\text{ind}}\ Z_{j(i),t-1} + r_{\text{ind}}\ \Delta \log Y_{j(i),t-1} + \gamma_i + \chi_t + \varepsilon_{it} \tag{414}$$

for $Z_{j(i),t} \in \{\overline{\text{opt}}^{sw}_{j(i),t}, \overline{\text{opt}}^{g,sw}_{j(i),t}\}$. As above, we estimate this first via OLS for each outcome variable, and then via IV where the granular variable $\overline{\text{opt}}^{g,sw}_{j(i),t}$ is an instrument for the raw sales-weighted average $\overline{\text{opt}}^{sw}_{j(i),t}$.

Results We present our results in Table A.27. First, studying aggregate contagiousness, we find strong evidence that $s > 0$ when measured with the raw sales-weighted average or its granular component (columns 1 and 2). We moreover find significant evidence of $s > 0$ in the IV estimation (column 3). Our IV point estimate of $\hat{s} = 0.308$ greatly exceeds the OLS estimate of $\hat{s} = 0.0847$.

At the industry level, we find strong evidence of contagiousness via the sales-weighted measure (column 4). We find imprecise estimates, centered around 0, for contagiousness measured with the granular variable (column 5) or via the granular IV (column 6). However, the granular IV estimate is noisily estimated and is not significantly different from the point estimate of column 4.

A.6 Additional Details on Model Quantification

A.6.1 Solution of Model With Persistent Fundamentals

We first provide the exact solution of the model when fundamentals follow an AR(1) process. We build on the analysis of Appendix A.2.5, which allows for (among other features) persistent fundamentals.

Law of Motion for Output Log aggregate productivity follows the process

$$\log \theta_t = (1-\rho)\mu + \rho \log \theta_{t-1} + \sigma \zeta_t \tag{415}$$

with $\zeta_t \sim N(0,1)$ IID. We continue to assume, as in our main analysis, that there are two narratives associated with high and low values of μ, $\mu_O > \mu_P$, while the true value is $\mu = 0$. Proposition 32 establishes that equilibrium can be written as (f does not depend on θ_{t-1} here as all agents believe persistence is ρ)

$$\log Y_t = a_0 + a_1 \log \theta_t + a_2 \log \theta_{t-1} + f(Q_t) \tag{416}$$

where we normalize $a_0 = 0$. We define the *fundamental component of output* as $\log Y_t - f(Q_t)$:

$$\log Y_t^f = a_1 \log \theta_t + a_2 \log \theta_{t-1} \tag{417}$$

Subtracting $\rho \log Y_{t-1}^f$ from both sides, the above becomes an ARMA$(1,1)$ process:

$$\log Y_t^f - \rho \log Y_{t-1}^f = a_1 \sigma \zeta_t + a_2 \sigma \nu_{t-1} \tag{418}$$

It remains to solve for the coefficients (a_1, a_2). In particular, Equations 334 and 335 give the fixed-point equations which these coefficients must solve. We can simplify these fixed point equations considerably in the case with optimism and pessimism about means and compute $\delta_{t,k}$ for $k \in \{O, P\}$:

$$\begin{aligned}\delta_{t,k} = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \Bigg[& \log\left(\frac{1-\frac{1}{\epsilon}}{\frac{1+\psi}{\alpha}}\right) + \frac{1+\psi}{\alpha}\left[\log \gamma_i + \kappa \log \theta_t + (1-\kappa)((1-\rho)\mu_k + \rho \log \theta_{t-1})\right] \\ & - \frac{1}{2}\left(\frac{1+\psi}{\alpha}\right)^2 \left(\sigma^2_{\theta|s} + \sigma^2_{\tilde{\theta}}\right) + \frac{1}{2} a_1^2 \left(\frac{1}{\epsilon} - \gamma\right)^2 \sigma^2_{\theta|s} \\ & + \left(\frac{1}{\epsilon} - \gamma\right)\left[a_0 + a_1\left(\kappa \log \theta_t + (1-\kappa)((1-\rho)\mu_k + \rho \log \theta_{t-1})\right) + a_2 \log \theta_{t-1} + f(Q_t)\right] \Bigg]\end{aligned} \tag{419}$$

Here, we have used the fact that posterior variances and perceived persistence are the same for the two narratives, and the fact that $\mu(e_k, \theta_{t-1}) = (1-\rho)\mu_k + \rho \log \theta_{t-1}$. Therefore,

$$\alpha \delta^{OP} := \delta_{t,O} - \delta_{t,P} = \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left(\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \right) (1-\kappa)(1-\rho)(\mu_O - \mu_P) \tag{420}$$

is the (time-invariant) average difference in actions between optimists and pessimists, as we identify in the data.

Next, taking $\delta_t(e_1) = \delta_{t,P}$ and Q_t as the fraction of optimists, we write Equation 335 as

$$\log Y_t = \delta_{t,P} + \frac{1}{2} \frac{\epsilon - 1}{\epsilon} \hat{\sigma}^2 + \frac{\epsilon}{\epsilon - 1} \log \left(Q_t \exp \left\{ \frac{\epsilon - 1}{\epsilon} \alpha \delta^{OP} \right\} \right) \tag{421}$$

Substituting in the expression for $\delta_{t,P}$, we can write the above up to a constant C that does not depend on $(\log \theta_t, \log \theta_{t-1}, Q_t)$ as

$$\begin{aligned} \log Y_t = C &+ \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha} \kappa + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \kappa \right] \log \theta_t \\ &+ \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha} (1-\kappa)\rho + \left(\frac{1}{\epsilon} - \gamma \right) (a_1(1-\kappa)\rho + a_2) \right] \log \theta_{t-1} \\ &+ \frac{\epsilon}{\epsilon - 1} \log \left(Q_t \exp \left\{ \frac{\epsilon - 1}{\epsilon} \alpha \delta^{OP} \right\} \right) \end{aligned} \tag{422}$$

To obtain the coefficients in our desired representation, first note that we can write

$$f(Q_t) = \frac{\epsilon}{\epsilon - 1} \log \left(Q_t \exp \left\{ \frac{\epsilon - 1}{\epsilon} \alpha \delta^{OP} \right\} \right) - \frac{\epsilon}{\epsilon - 1} \log \left(\frac{1}{2} \exp \left\{ \frac{\epsilon - 1}{\epsilon} \alpha \delta^{OP} \right\} \right) \tag{423}$$

This is the same form as our main analysis, with a normalization such that $f(1/2) = 0$. Next, from matching coefficients, and noting the definition of $\omega = (1/\epsilon - \gamma)/((1 + \psi - \alpha)/\alpha + 1/\epsilon)$,

$$a_1 = \frac{1}{1 - \kappa\omega} \frac{\frac{1+\psi}{\alpha}\kappa}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \tag{424}$$

Finally, from matching coefficients for a_2,

$$a_2 = \frac{1}{1-\omega} \frac{1}{\frac{1+\psi-\alpha}{\alpha} + \frac{1}{\epsilon}} \left[\frac{1+\psi}{\alpha} + \left(\frac{1}{\epsilon} - \gamma \right) a_1 \right] (1-\kappa)\rho \tag{425}$$

Updating Rule We use the Linear-Associative-Contagious Updating rule introduced as the Main Case (Equation 29), with a normalization. More specifically, we assume transition probabilities

$$
\begin{aligned}
P_O^H(\log Y, Q, \varepsilon) &= [u + r\log Y + sQ + C_P + \varepsilon]_0^1 \\
P_P^H(\log Y, Q, \varepsilon) &= [-u + r\log Y + sQ + C_P + \varepsilon]_0^1
\end{aligned}
\tag{426}
$$

We choose C_P such that an economy with neutral fundamentals ($\log\theta_t = \log\theta_{t-1} = 0$), equal optimists and pessimists ($Q = 1/2$), and no narrative shocks ($\varepsilon = 0$) continues to have equal optimists and pessimists. Specifically, this implies $C_P = \frac{1-s}{2}$.

A.6.2 Calibration Methodology

To calibrate the model, we proceed in four steps.

1. *Setting macro parameters.* We first set $(\epsilon, \gamma, \psi, \alpha)$. In Section 1.7.1 and Table 1.6, we describe our baseline method based on matching estimates of the deep parameters from the literature. We also consider two other strategies as robustness checks. First, to target estimated fiscal multipliers in the literature, we use the same external calibration of α (returns to scale) and ϵ (elasticity of substitution), and set (γ, ψ) to match the desired multiplier. Since the exact choice of these parameters is arbitrary subject to obtain the correct multiplier, we normalize $\gamma = 0$ and vary only ψ. Second, we match an estimate of the multiplier implied by our own data and an exact formula for the omitted variable bias incurred in estimating the effect of optimism on hiring without controlling for general-equilibrium effects via a time fixed effect. We outline that strategy for estimating the multiplier in Section A.6.3 below, and we map this to deep parameters exactly as described in our method for matching the literature's estimated multiplier.

2. *Calibrating the effect of optimism on output.* We observe that, conditional on $(\epsilon, \gamma, \psi, \alpha)$ and an estimate of δ^{OP}, we have identified $f(Q_t)$ as defined in Equation 423. We take our estimate of δ^{OP} from column 1 in Table 1.1. This regression identifies δ^{OP} for the reasons described in Corollary 1.

3. *Calibrating the statistical properties of fundamentals (κ, ρ, σ).*

   (a) *Computing fundamental output.* We construct a cyclical component of output, $\log\hat{Y}_t$, as band-pass filtered US real GDP (Baxter and King, 1999).[9]

[9] Specifically, we filter to post-war quarterly US real GDP data (Q1 1947 to Q1 2022). We use a lead-lag length of 12 quarters, a low period of 6 quarters, and a high period of 32 quarters. We then average these data to the annual level.

We apply our estimated function f to our measured time series of optimism (see Figure 1-1) to get an estimated optimism component of output. we then calculate

$$\log \hat{Y}_t^f = \log \hat{Y}_t - \hat{f}(\hat{Q}_t) \tag{427}$$

(b) *Estimating the ARMA representation.* Using our 24 annual observations of $\log \hat{Y}_t^f$, we estimate a Gaussian-errors ARMA(1,1) model via maximum likelihood. Our point estimates are

$$\log \hat{Y}_t^f - 0.086 \log \hat{Y}_t^f = .0078(\zeta_t + .32\, \nu_{t-1}) \tag{428}$$

This implies $\rho = 0.086$, $a_1\sigma = .0078$, and $a_2\sigma = .32$. ρ is therefore identified immediately.

(c) *Calibrating (κ, σ).* We search non-linearly for values of (κ, σ) that satisfy $a_1\sigma = 0.0078$ and $a_2\sigma = 0.32$. There is a unique such pair, reported in Table 1.6, which also is therefore the maximum likelihood estimate of (κ, σ).

4. *Calibrating the updating rule $(u, r, s, \sigma_\varepsilon^2)$.* The coefficients of the LAC updating model are estimated in column 1 of Table 1.4. Conditional on the previous calibration, we set σ_ε^2 so that within model Q_t has the same standard deviation as the aggregate optimism time series, which is 0.0533.

To calibrate the variant model with multi-dimensional narratives, we follow steps 1-3 exactly as above combined with a different procedure for step 4 described below:

4′. *Calibrating the multi-dimensional updating rule $((\beta^k, u^k, r^k, s^k)_{k=1}^K, \sigma_\varepsilon^2)$.* To map theory to data, we first transform each continuous narrative loading variable $\hat{\lambda}_{k,it}$ (indexed by narrative identifier k, firm i, and time period t) into a binary indicator for being above the sample median,

$$\hat{\lambda}_{k,it}^b = \mathbb{I}\left[\hat{\lambda}_{k,it} \geq \text{med}\left(\hat{\lambda}_{k,it}\right)\right] \in \{0, 1\}$$

We study the 2 decision-relevant Shiller (2020) narratives and 11 decision-relevant topic narratives indicated in Table 1.3. We let $k \in \{1, \ldots, 13\}$ index these narratives below.

(a) *Calibrating the constellation weights β.* To estimate the constellation weights, we first regress optimism at the firm level on the binary indicators for each of the selected narratives, conditional on firm and industry-by-time fixed effects:

$$\text{opt}_{it} = \sum_{k=1}^{13} \tau^k \hat{\lambda}^b_{k,it} + \gamma_i + \chi_{j(i),t} + \nu_{it} \tag{429}$$

For all k such that $\tau^k < 0$, we re-normalize the narrative to $1 - \hat{\lambda}^b_{k,it}$ (*i.e.*, an indicator for being below-median), such that all narratives positively contribute toward the probability of being optimistic. We finally construct an estimator of $\hat{\beta}^k$ which re-normalizes the regression coefficients above to sum to one,

$$\hat{\beta}^k = \frac{|\hat{\tau}^k|}{\sum_{j=1}^{13} |\hat{\tau}^j|} \tag{430}$$

(b) *Calibrating the updating parameters* (u^k, r^k, s^k) *and* σ^2_ε. To estimate the updating rule paramaters, we take the u^k, r^k, and s^k parameters from Figure A-9, flipping the sign of associativeness if we flipped τ^k in step 4'(a). Conditional on the previous calibration, we set σ^2_ε so that within model Q_t has the same standard deviation as the aggregate optimism time series, which is 0.0533.

A.6.3 Estimating a Demand Multiplier in Our Empirical Setting

Here, we describe a method for estimating a demand multiplier in our data on optimism and firm hiring. This circumvents the step of external calibration for the multiplier, but relies on correct specification of the time-series correlates of aggregate optimism. Reassuringly, this method yields a general-equilibrium demand multiplier that is comparable to our baseline calibration and our literature-derived calibration.

Mapping the Model to Data Extending Corollary 1 with the calculations of Appendix A.2.5 and Appendix A.6.1, we first observe that firms' hiring can be written in equilibrium as

$$\Delta \log L_{it} = \tilde{c}_{0,i} + \tilde{c}_{10} \log \theta_t + \tilde{c}_{11} \log \theta_{t-1} + \tilde{c}_2 f(Q_t) + \tilde{c}_3 \log \theta_{it} + \tilde{c}_4 \log L_{i,t-1} + \delta^{OP} \lambda_{it} + \zeta_{it} \tag{431}$$

where ζ_{it} is an IID normal random variable with zero mean and λ_{it} is the indicator for having adopted the optimistic narrative.

In the data, our estimating equation without control variables had the following form

$$\Delta \log L_{it} = \gamma_i + \chi_{j(i),t} + \delta^{OP}\text{opt}_{it} + z_{it} \tag{432}$$

This maps to the structural model with $\gamma_i = \tilde{c}_{0,i}$, $\chi_{j(i),t} = \tilde{c}_{10}\log\theta_t + \tilde{c}_{11}\log\theta_{t-1} + \tilde{c}_2 f(Q_t)$, and $z_{it} = \zeta_{it} + \tilde{c}_3 \log\theta_{it} + \tilde{c}_4 \log L_{i,t-1}$. Under the hypothesis that $\mathbb{E}[z_{it}\text{opt}_{it}] = 0$, then the OLS regression of $\Delta\log L_{it}$ on opt_{it}, conditional on the indicated fixed effects, identifies δ^{OP}.

We consider now an alternative regression equation which is a variant of the above specification without the time fixed effect and with parametric controls for aggregate TFP:

$$\Delta \log L_{it} = \gamma_i + \delta^{OP}\text{opt}_{it} + \tilde{c}_{10}\log\theta_t + \tilde{c}_{11}\log\theta_{t-1} + \tilde{z}_{it} \tag{433}$$

Observe that the new residual, relative to the old residual, is contaminated by the equilibrium effect of optimism. That is, $\tilde{z}_{it} = z_{it} + \tilde{c}_2 f(Q_t)$. To refine this further, we apply the linear approximation $f(Q_t) \approx \frac{\alpha\delta^{OP}}{1-\omega}Q_t$ and the observation that $\tilde{c}_2 = \omega$, so we can write $\tilde{z}_{it} = z_{it} + \frac{\alpha\omega}{1-\omega}\delta^{OP}Q_t$.

We now derive a formula for omitted variables bias in the estimate of δ^{OP} from an OLS estimation of Equation 433. Let X denote a finite-dimensional matrix of data on opt_{it}, firm-level indicators (*i.e.,* the regressors corresponding to the firm fixed effects), and current and lagged aggregate TFP. Similarly, let Y be a finite-dimensional matrix of data on $\Delta\log L_{it}$. The OLS regression coefficient in this finite sample is $\hat{\delta} = ((X'X)^{-1}X'Y)_1$. Using the standard formula for omitted variables bias:

$$\begin{aligned}\mathbb{E}[\hat{\delta}|X] &= \delta^{OP} + \left((X'X)^{-1}\mathbb{E}[X'Q|X]\frac{\alpha\omega}{1-\omega}\delta^{OP}\right)_1 \\ &= \delta^{OP}\left(1 + \frac{\alpha\omega}{1-\omega}\left((X'X)^{-1}\mathbb{E}[X'Q|X]\right)_1\right)\end{aligned} \tag{434}$$

where Q is the vector of observations of Q_t. We can then observe that:

$$(X'X)^{-1}\mathbb{E}[X'Q|X] = \mathbb{E}\left[(X'X)^{-1}X'Q|X\right] \tag{435}$$

Which is the (expected) OLS estimate of β in the following regression:

$$Q_t = \gamma_i + \beta^Q_O\text{opt}_{it} + \beta^Q_\theta\log\theta_t + \beta^Q_{\theta_{-1}}\log\theta_{t-1} + \varepsilon_t \tag{436}$$

But we observe that, averaging both sides, that $\gamma_i = \beta^Q_\theta = \beta^Q_{\theta_{-1}} = 0$ and $\beta^Q_O = 1$.

Thus, $((X'X)^{-1}\mathbb{E}[X'Q|X])_1 = 1$. We therefore obtain that:

$$\mathbb{E}[\hat{\delta} \mid X] = \delta^{OP}\left(1 + \frac{\alpha\omega}{1-\omega}\right) \tag{437}$$

Hence, given a population estimate of the *biased* OLS estimate and an external calibration of α, we can pin down the complementarity ω and the multiplier $\frac{1}{1-\omega}$. Naturally this strategy relies on correctly measuring aggregate TFP as measurement error in that variable would contaminate this estimation. Moreover, it requires us to assume that all variation in aggregate output that is not due to TFP is due to optimism or forces entirely orthogonal to optimism; in view of our running assumption that the spread of optimism is associative, these other forces therefore also have to be completely transitory, lest they be incorporated into current optimism via associative updating in a previous period. These assumptions are strong and are why we do not adopt this strategy for our main quantitative analysis. Nevertheless, we will find similar results, as we now describe.

Empirical Application and Results To operationalize this in practice, we compare estimates of Equation 432 and 433. For the latter, we proxy TFP using the cyclical component of both capacity adjusted and capacity un-adjusted TFP using the data of Fernald (2014).[10] We moreover maintain the assumption of $\alpha = 1$, or constant returns to scale, to map our estimates back to implied multipliers.

Our results are reported in Table A.28, along with the associated values of complementarity ω and the multiplier $\frac{1}{1-\omega}$. Using capacity-adjusted and unadjusted TFP, we respectively obtain estimates of 1.46 and 1.37 for the multiplier. These are lower than our baseline estimate, but comparable to our estimates based on structural modeling in the literature. Both estimates are below our baseline calibration of 1.96 but above our multiplier-literature calibration of 1.33. In Table A.19, we report our quantitative results under the assumed multiplier of 1.46. We find that, as expected, these estimates imply an role for optimism that is an intermediate between the baseline and multiplier-literature calibrations.

[10]Mirroring our filtering of US real GDP, we apply the Baxter and King (1999) band-pass filter to post-war quarterly data using a lead-lag length of 12 quarters, a low period of 6 quarters, and a high period of 32 quarters. We then average these data to the annual level.

A.7 Our Analysis and Shiller's *Narrative Economics*

Shiller (2017, 2020) introduces the notion of narrative economics and identifies "Seven Propositions of Narrative Economics" as a basis for the theoretical and empirical investigation of narratives. In this section, we discuss our work, how our modeling of narratives relates to Shiller's ideas, and the relationship of our modeling, measurement, and results with these propositions. In the process, we highlight how these propositions have informed our analysis, discuss how our analysis contributes new insights, and propose avenues for future work to more fully understand narratives and the macroeconomy.

A.7.1 The modeling of Narratives

We first describe how our modeling and measurement of narratives are designed to capture the salient features of narratives that Shiller (2020) introduces in the preface:

> In using the term narrative economics, I focus on two elements: (1) the word-of-mouth contagion of ideas in the form of stories and (2) the efforts that people make to generate new contagious stories or to make stories more contagious. First and foremost, I want to examine how narrative contagion affects economic events. The word narrative is often synonymous with story. But my use of the term reflects a particular modern meaning given in the Oxford English Dictionary: "a story or representation used to give an explanatory or justificatory account of a society, period, etc." Expanding on this definition, I would add that stories are not limited to simple chronologies of human events. A story may also be a song, joke, theory, explanation, or plan that has emotional resonance and that can easily be conveyed in casual conversation.

To map this verbal definition to our framework (see Section 1.2 for the formal details and notation), consider the following simple verbal rationale for our modeling approach. There is a latent space of economic fundamentals (demand for goods) and endogenous outcomes (aggregate output). An agent has some beliefs about economic fundamentals and corresponding endogenous outcomes (π). They are told the following simple story by another agent about the economy: "I didn't hire (x) because aggregate output (Y) will be low because demand (θ) is weak." This might cause the agent to believe this story that demand is weak and adopt this narrative (placing weight λ on the implied distribution of fundamentals). Moreover, if many of their friends tell them the story, they might be more positively inclined to believe it. Of course, the agent doesn't listen to the story blindly: they can see if demand was previously low (and might even have information about demand from their personal economic activities $\mathcal{I}$) and might regard such a story is fanciful if their own experience contradicts this claim. At the end of this process of contemplation, they update their own weight on the narrative (via P) and arrive at their new belief (π').

Of course, the actual realization of output depends on the circulation of narratives in the population (Q). If an agent believes the "demand is weak" narrative, then they will curtail their hiring. Knowing this, other agents—even if they do not believe that latent demand is weak—will believe that others will curtail their hiring, so that realized demand will be weak. Then, knowing this, all agents further cut hiring. This paradox of thrift induces a hierarchy of higher-order expectations regarding realized demand induced by the distribution of narratives. This converges to a fixed point ($Y^*(Q) : \Theta \to \mathcal{Y}$) describing the mapping of demand into aggregate output under the prevailing circulation of narratives.

Thus, while the primitive narrative began as a story about the strength of demand, in equilibrium it takes on a meaning as not only describing exogenous economic outcomes, but also endogenous economic outcomes. Concretely, given an equilibrium mapping, the narrative endogenously induces the joint belief $N^* \in \Delta(\Theta \times \mathcal{Y})$ given by $N^*(\theta, Y) = N(\theta) \times \mathbb{I}[Y = Y^*(\theta)]$. Importantly, the distribution of narratives Q and endogenous outcomes Y then shape the distribution of narratives tomorrow Q'. The resulting joint dynamics of narratives and endogenous outcomes are the subject of the theoretical and quantitative analysis of this paper.

To ensure that we measure narratives, trace their impact on decisions and study their spread, we operationalize this empirically by measuring narratives in agents' use of language by employing natural language processing methods that we have described in Section 1.3.1. This allows us to use our framework to test if these text-based proxies for narratives shape actions and spread across agents. As described in Section 1.4, we find strong evidence of these premises.

We are, however, essentially silent about the more fundamental determinants of how something comes to be a narrative or what makes a narrative contagious. As a result, we do not speak to the issue of narrative generation suggested by Shiller. We do have one empirical result that hints that firms use narratives to persuade financial analysts. In our IBES strategy, we found that optimistic firms significantly overestimate their sales relative to pessimistic firms. However, we found much weaker evidence that analysts believe that firms are performing this overestimation. As a result, we take this as tentative evidence that firms manage to persuade analysts of their predictions. We view further exploration of this issue as an interesting angle for future work.

A.7.2 Our Work and Shiller's Seven Propositions

Proposition 1: Epidemics Can Be Fast or Slow, Big or Small The model developed in this paper allows for various speeds of narrative dynamics as well as their size and economic impact. Shiller, drawing on the epidemiology literature, identifies two parameters as particularly important in determining these features: the contagion rate and the recovery rate. Viewed through this lens, our structural model from Section 1.5 postulates a recovery rate of $1-P_O(\log Y_t, Q_t)$ and a contagion rate of $P_P(\log Y_t, Q_t)$. Thus, the fundamental parameters determining stubbornness u, associativeness r and contagiousness s are key determinants of the speed and size of narrative epidemics within our model.

Yet further, by moving beyond a purely epidemiological model and studying the two-way feedback between narratives and endogenous outcomes, we endogenize these rates as equilibrium outcomes by characterizing the equilibrium map $Q_t, \theta_t \mapsto Y_t$. Thus, the parameters of P_O and P_P as well as those determining the information and strategic interaction in the economy affect the contagion and recovery rates in ways that we have characterized. Most interestingly, beyond affecting the quantitative features of narrative dynamics (such as speed and size), accounting for the dynamic complementarity of narratives affects their qualitative features. Concretely, these same parameters delineate whether the economy is stable, has a unique steady state, features hysteresis, or hump-shaped and discontinuous impulse responses.

Moreover, we have used our measurement and empirical strategies to place empirical discipline on these parameters. By so doing, we have been able to provide ballpark figures for the likely quantitative importance of the narratives we have uncovered in our data. Future work may lever alternative data sources and identification strategies to study different narratives or more precisely estimate the parameters that we have studied.

Proposition 2: Important Economic Narratives May Comprise a Very Small Percentage of Popular Talk Our empirical analysis found that very little of the total variation in narrative discussion is at the aggregate level (See Appendix Table A.5). For example, only 1% of the variation in optimism is captured in the aggregate time series. Indeed, even for the 75% percentile of our estimated topic narratives, the fraction of variance explained by the time series is less than 10%. Thus, movements in the intensity with which narratives are discussed appear to be largely a cross-sectional phenomenon. As we have shown, this does not at all mean that aggregate movements are unimportant: measured movements in aggregate optimism can account for between 1/6 and 1/3 of GDP movements over significant economic events. Thus, just as idiosyncratic income risk is much larger than aggregate GDP risk, idiosyncratic narrative variation is much larger than aggregate narrative variation.

This echoes the observational account of Shiller that important economic narratives may not feature prominently in popular talk and underlines the conclusion that even if movements in aggregate narratives are not a large fraction of what agents think or discuss, they can nevertheless be critical for understanding economic fluctuations.

Proposition 3: Narrative Constellations Have More Impact Than Any One Narrative A central concept in Shiller's analysis is that of the *narrative constellation*: a grouping of narratives around some basic idea that reinforces contagion. This is a concept about which we are theoretically silent. However, our empirical analysis is designed to allow for the possibility of narrative constellations. Take our analysis of optimism, for example. Our measure does not necessarily capture one coherent economic narrative regarding the strength of the outlook of the economy. What it instead captures is the total sentiment expressed by firms, averaging across the various underlying narratives that they may be adopting at any one instant. We investigate formally the extent to which our data support this by using our more granular narratives as instruments for optimism (see Appendix Table A.14, columns 4) and find similar results to our baseline analysis. Moreover, in our quantitative analysis, we studied how many co-evolving latent narratives can manifest as aggregate, sentiment-driven fluctuations in the economy.

Moreover, our analysis of Shiller's narratives allows us to pick up narrative constellations to the precise extent that Shiller discusses the words comprising the underlying narratives in these constellations in his own analysis. Finally, our topic analysis allows us to pick up narrative constellations to the extent that narratives are used jointly in individual documents. Thus,, we do account for the existence of contellations in our measurement and empirical exercises. We view further analysis of this hypothesis as an interesting avenue for future work.

Proposition 4: The Economic Impact of Narratives May Change Through Time Shiller suggests that the impact of economic narratives has the potential to change through time. First, we evaluate this hypothesis in the context of our study in Section 1.4.1 of how measured optimism affects hiring. Specifically, we consider our baseline regression model

$$\Delta \log L_{it} = \delta^{OP} \mathrm{opt}_{it} + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{438}$$

Our baseline estimate, in column 1 of Table 1.1, is $\hat{\delta}^{OP} = 0.0355$. We now consider a variant in which the coefficient δ^{OP} varies for each year $1996 \leq \tau \leq 2019$ in our sample:[11]

$$\Delta \log L_{it} = \sum_{\tau=1996}^{2019} \delta_{\tau}^{OP} (\mathrm{opt}_{it} \times \mathbb{I}[t = \tau]) + \gamma_i + \chi_{j(i),t} + \varepsilon_{it} \tag{439}$$

We show the coefficient series of δ_{τ}^{OP} graphically in Figure A-16. We observe no strong pattern of a trend or business cycle in the coefficient estimates. We interpret this as evidence that the main narrative studied in our empirical analysis has relatively stable effects on decisions over time.

Second, we evaluate whether the contagiousness of optimism has changed through time. One plausible hypothesis is that the rise of the internet and other information technology may have contributed to a more connected knowledge-economy in which there is a faster diffusion of narratives. To do this, we now estimate a variant of our basic regression equation for estimating industry-level contagiousness in which we allow the coefficients, u, r, and s to vary for each year $1996 \leq \tau \leq 2019$ in our sample:

$$\mathrm{opt}_{it} = \sum_{\tau=1996}^{2019} \left(u_{\tau}\ \mathrm{opt}_{i,t-1} + s_{\tau}\ \overline{\mathrm{opt}}_{j(i),t-1} + r_{\tau}\ \Delta \log Y_{j(i),t-1}\right) \mathbb{I}[t = \tau] + \gamma_i + \chi_t + \varepsilon_{it} \tag{440}$$

We can identify time-specific effects only in the industry-specific setting—the corresponding aggregate level regression with time-specific effects is not point identified. We plot the estimated coefficient series of u_{τ}, r_{τ}, and s_{τ} in Figure A-17. We find that stubbornness and contagiousness increase over time, while associativeness is close to flat. This is despite the fact that, as we just showed, the effects of narrative optimism on hiring are stable through time.

Viewed through the lens of our model, these findings—increased contagiousness and stubbornness combined with stable effects on actions—suggests that narrative optimism's potential to "go viral" and induce dynamic hysteresis in the US economy is increasing over time (see, *e.g.*, Section 1.7.3). Exploring the macroeconomic impli-

[11]The number of firms with data reported for 2019 is very small, so our sample essentially ends in 2018.

cations of this finding may be an interesting avenue for future work. Moreover, this macroeconomic finding underscores the importance of directly testing the hypothesis that the internet and information technology are culpable for the increase in contagiousness and stubbornness; insofar as these changes to narrative propagation could have large effects on macroeconomic dynamics.

Proposition 5: Truth Is Not Enough to Stop False Narratives Shiller emphasizes that narrative epidemics can take place even when patently divorced from fundamentals. Our theoretical analysis shares this feature. Namely, when the contagiousness of a narrative is high, this can swamp any adverse effects on contagion stemming from outcomes that do not fit the narrative. This is made especially clear by our Proposition 2, in which multiple steady states of narrative penetration can coexist even when one narrative (or even both narratives) are false.

Proposition 6: Contagion of Economic Narratives Builds on Opportunities for Repetition Increased exposure to a narrative is likely to cause agents to pick it up. We find evidence that agents are both more likely to retain a narrative they currently have and that exposure to others holding the same narrative increases the chance that an agent both picks up and retains a narrative. These findings are consistent with Shiller's hypothesis that oft-repeated narratives are more likely to result in epidemics. However, we do not explore the idea that repeated exposure through time is likely to increase the persistence or contagiousness of a narrative. We view this as an interesting avenue for future work.

Proposition 7: Narratives Thrive on Attachment: Human Interest, Identity, and Patriotism We do not investigate the idea that more interesting narratives are more likely to be contagious in this paper. Studying this idea requires a deeper model for the origins of the stubbornness, associativeness and contagiousness of narratives, which we do not attempt to provide. We merely measure these parameters and take them as given. Of course, this renders our analysis vulnerable to a form of the Lucas critique: if a policymaker attempted to use our estimates as a guide for how they could affect the economy via manipulating narratives, these coefficients could change if they fail to mimic the human interest, identity, or patriotism that drove attachment to the narrative. While this issue is unimportant for our positive analysis, an understanding of the deeper origins of narrative success is an interesting avenue for future work—especially if a policymaker were to seek to guide narratives to achieve certain economic outcomes.

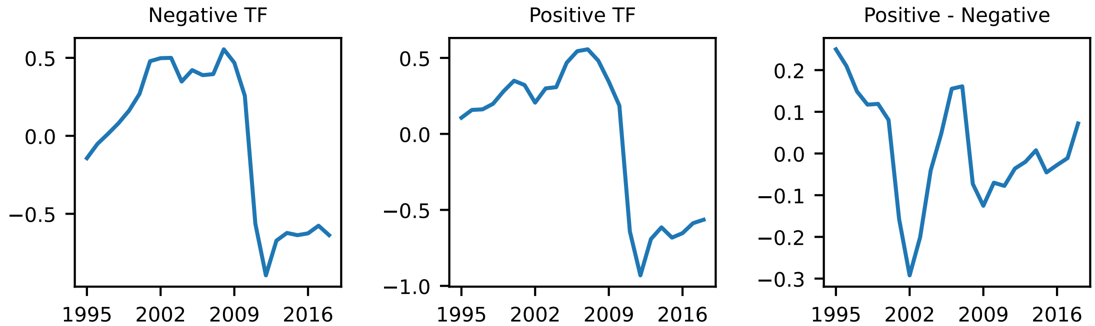
A.7.3 The Perennial Economic Narratives: Our Empirical Findings

Shiller (2020) identifies nine perennial economic narratives based on historical analysis. These narratives correspond to:

1. Panic versus Confidence
2. Frugality versus Conspicuous Consumption
3. The Gold Standard versus Bimetallism
4. Labor-Saving Machines Replace Many Jobs
5. Automation and Artificial Intelligence Replace Almost All Jobs
6. Real Estate Booms and Busts
7. Stock Market Bubbles
8. Boycotts, Profiteers, and Evil Businesses
9. The Wage-Price Spiral and Evil Labor Unions

We have measured the presence of these narratives in firm language and studied which of these matters for firm decision-making. Under our baseline LASSO specification, we found that two of these narratives are relevant for firm hiring: Labor-Saving Machines Replace Many Jobs and Stock Market Bubbles. Discussion of both is positively associated with firm hiring (see column 1 of Appendix Table A.14). Moreover, we find evidence of contagiousness and stubbornness in firms holding onto these narratives (see Appendix Figure A-9).

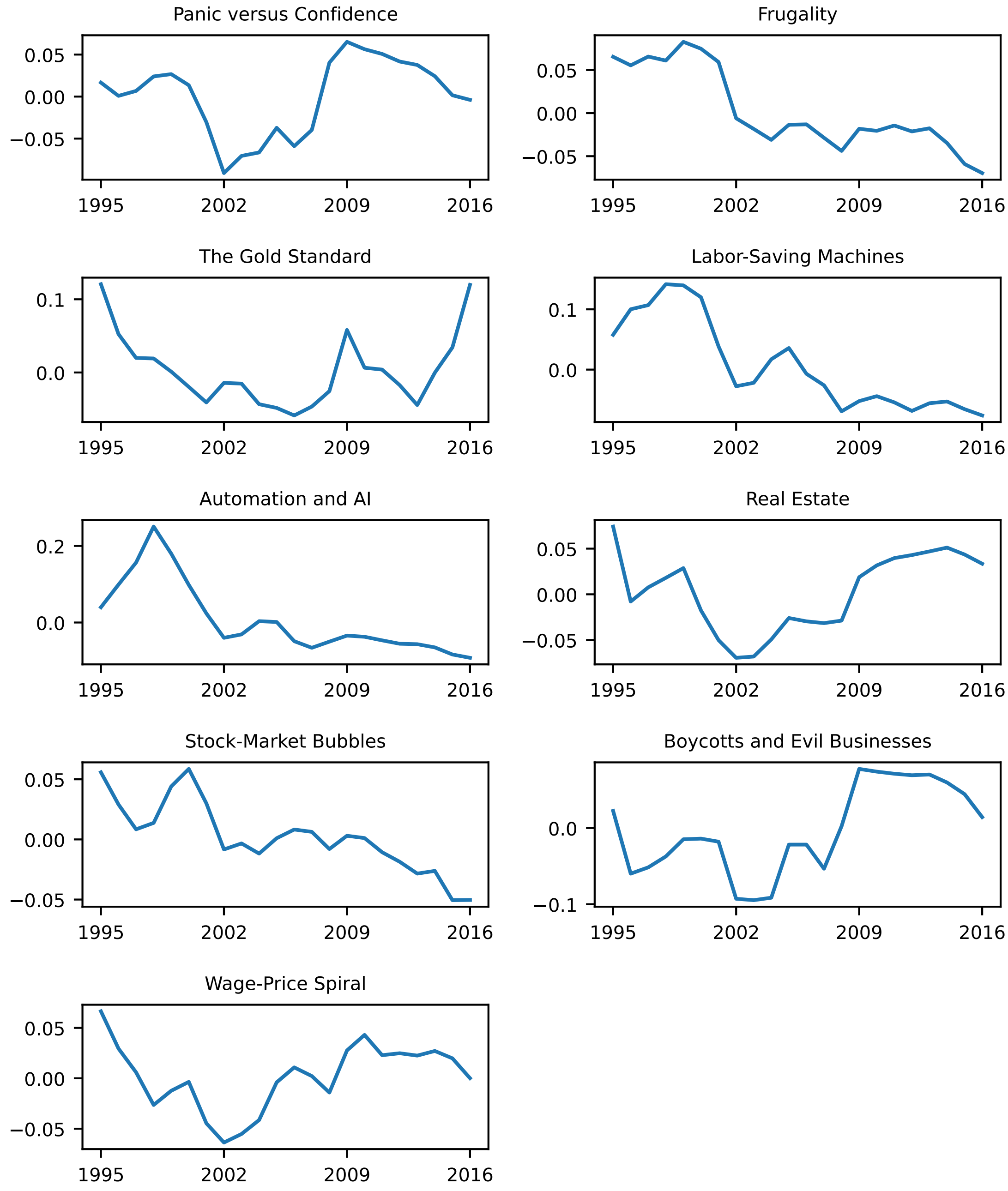
Figure A-1: Time-Series for Positive, Negative, and Their Difference



Notes: Negative and Positive term frequency (first two panels) are cross-sectional averages of z-score transformed variables. The third panel is the cross-sectional average of the difference between the two, or sentiment$_{it}$.

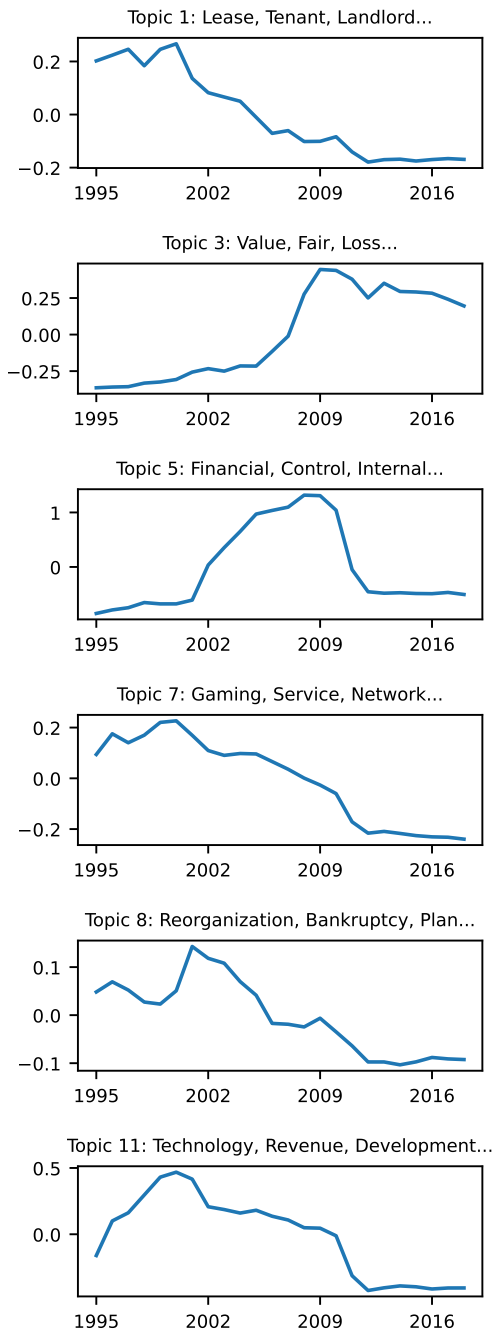
A.8 Additional Figures and Tables

Figure A-2: Time-Series for Shiller's Perennial Economic Narratives



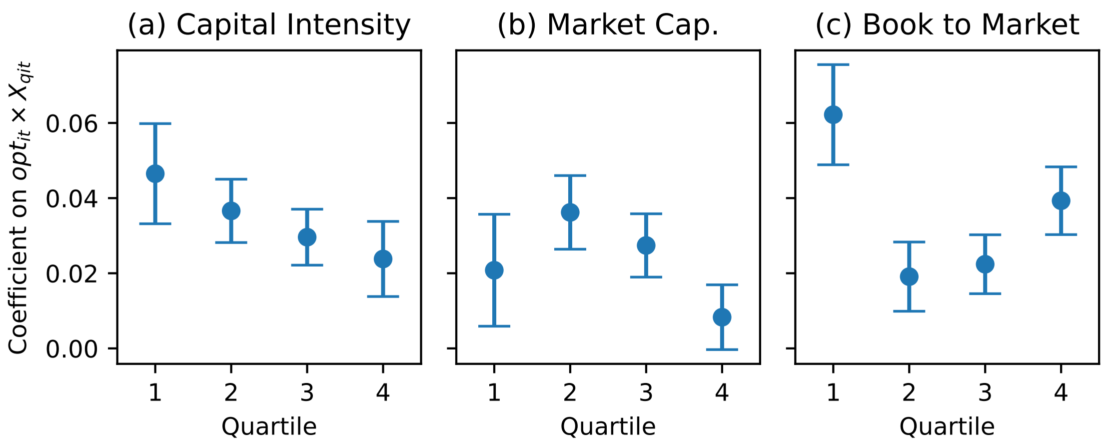
Notes: Each panel plots the time-series average of the narrative variable defined for the corresponding chapter of Shiller (2020)'s *Narrative Economics*. The units are cross-sectional averages of z-score transformed variables.

Figure A-3: Time-Series for the Selected LDA Topics



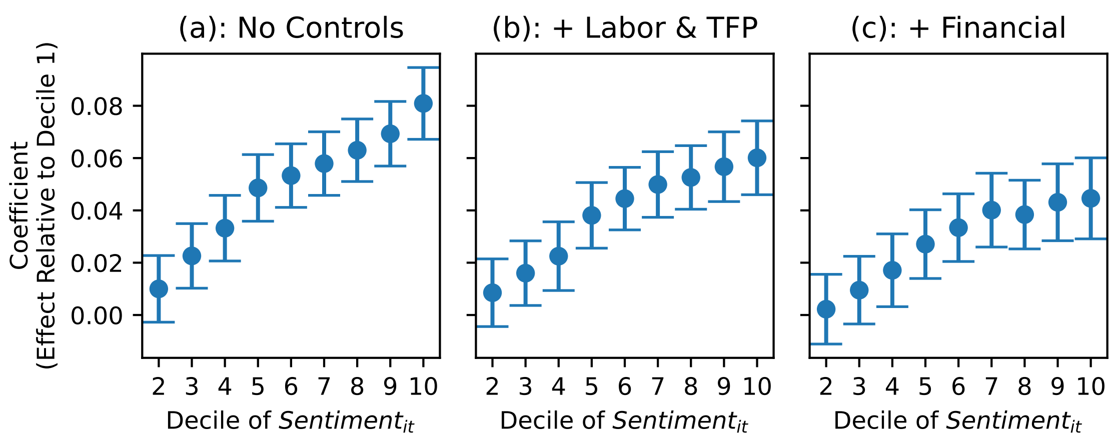
Notes: Each panel plots the time-series average of scores for the corresponding topics, identified by their three most common bigrams. The 11 topics are selected by the LASSO estimation described in Section 1.4.1, and estimates of which are reported in Table A.14. The units in each panel are cross-sectional averages of z-score transformed variables.

Figure A-4: Heterogeneous Effects of Optimism on Hiring



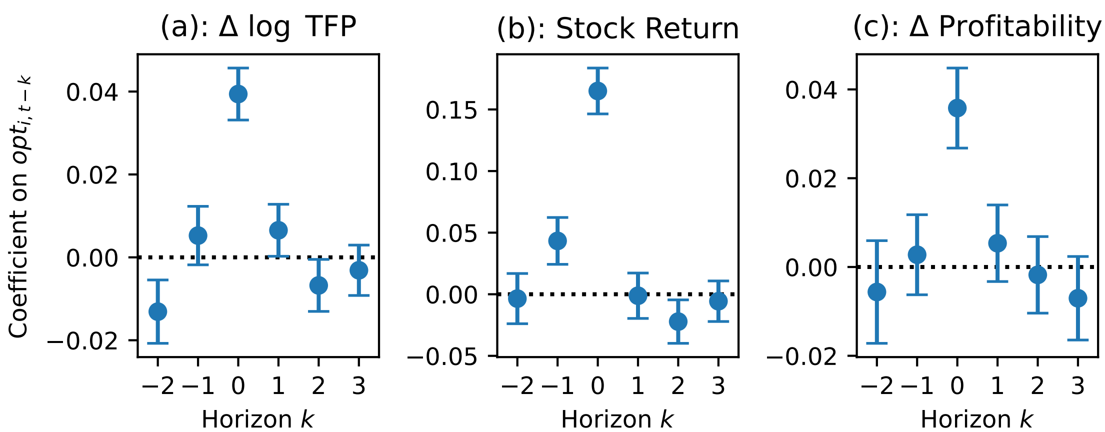
Notes: In each panel, we show estimates from the regression $\Delta \log L_{it} = \sum_{q=1}^{r} \beta_q \cdot (\text{opt}_{it} \times X_{qit}) + \gamma_i + \chi_{j(i),t} + \epsilon_{it}$, where X_{qit} indicates quartile q of the studied variable: one minus the variable cost share of sales, market capitalization, or book-to-market ratio. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are double-clustered by firm ID and industry-year.

Figure A-5: Net Sentiment and Hiring



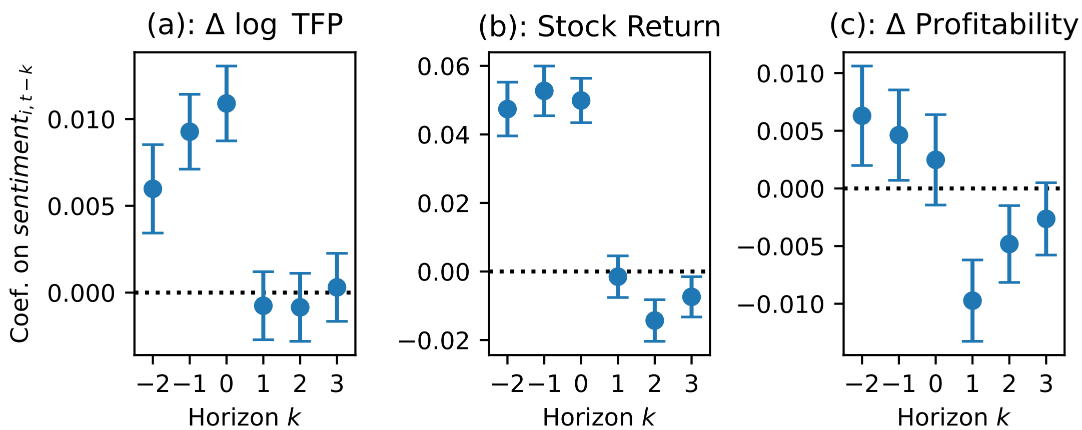
Notes: In each panel, we show estimates from the regression $\Delta \log L_{it} = \sum_{q=1}^{10} \beta_q \cdot (\text{sentiment}_{iqt}) + \tau' X_{it} + \gamma_i + \chi_{j(i),t} + \epsilon_{it}$, where sentiment_{iqt} indicates decile q of the continuous sentiment variable. Panel (a) estimates this equation without controls (like column 1 of Table 1.1); panel (b) adds controls for lagged labor and current and lagged log TFP (like column 2 of Table 1.1); and panel (c) adds controls for the log book to market ratio, log stock return, and leverage (like column 3 of Table 1.1). The excluded category in each regression is the first decile of sentiment_{it}. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are double-clustered by firm ID and industry-year.

Figure A-6: Dynamic Relationship between Optimism and Firm Fundamentals, Conference-Call Measurement



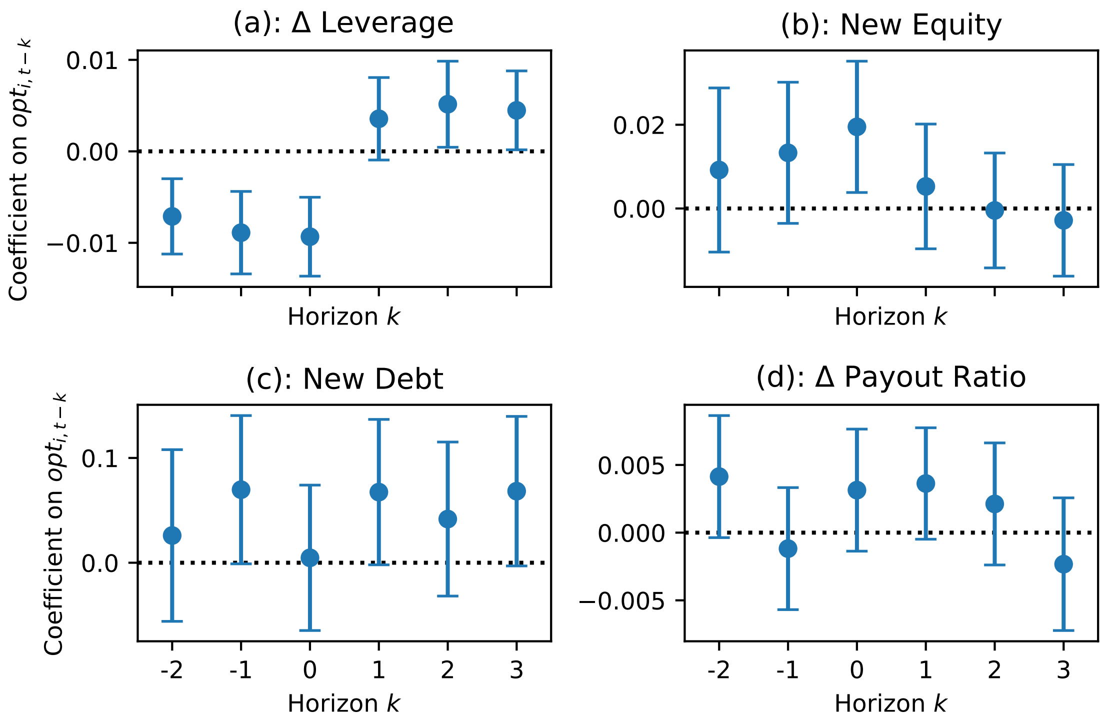
Notes: The regression model is Equation 14 (as in Figure 1-2), but measuring optimism from sales and earnings conference calls. Each coefficient is estimated from a separate projection regression. The outcomes are (a) the log change in TFP, calculated as described in Appendix A.4.2, (b) the log stock return inclusive of dividends, and (c) changes in profitability, defined as earnings before interest and taxes (EBIT) as a fraction of the previous fiscal year's variable costs. In all specifications, we trim the 1% and 99% tails of the outcome variable. Each coefficient is estimated from a separate projection regression. Error bars are 95% confidence intervals.

Figure A-7: Dynamic Relationship between Optimism and Firm Fundamentals, Continuous Sentiment Measurement



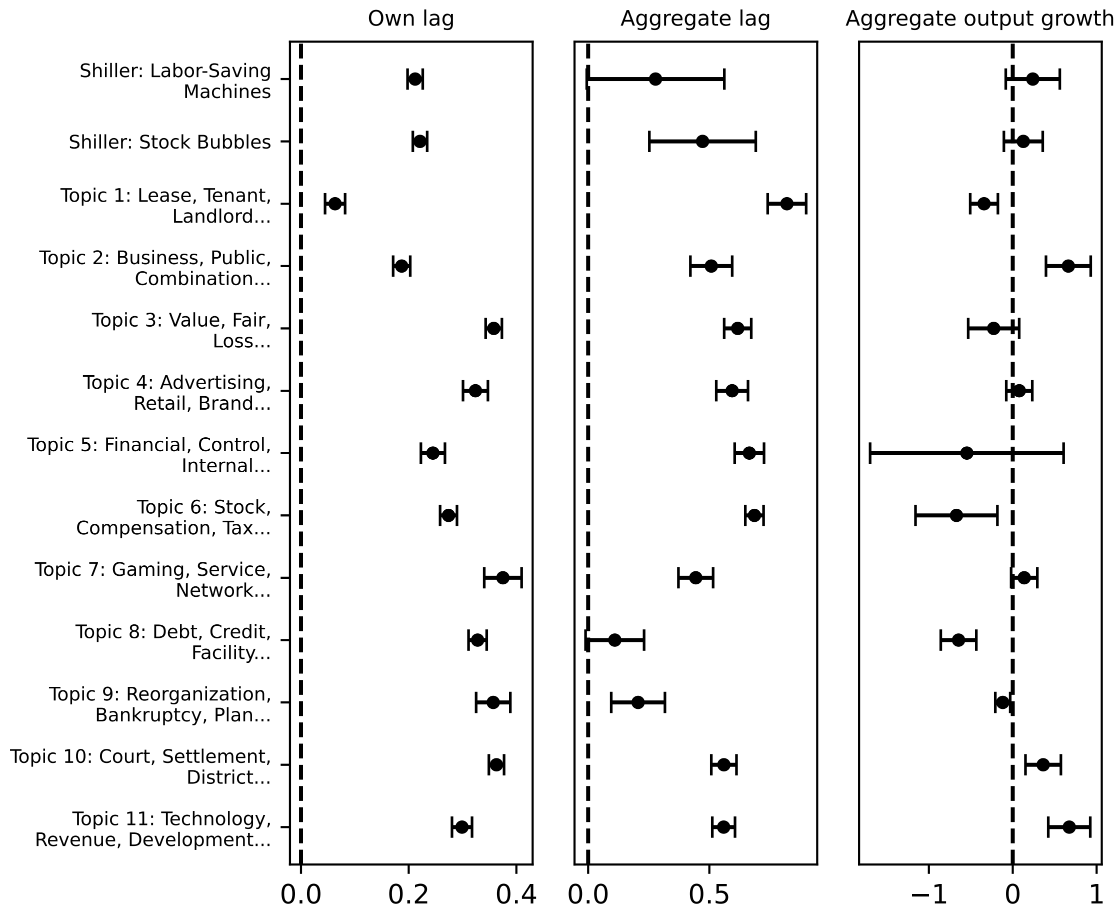
Notes: The regression model is Equation 14 (as in Figure 1-2), but using the continuous variable sentiment$_{it}$. Each coefficient is estimated from a separate projection regression. The outcomes are (a) the log change in TFP, calculated as described in Appendix A.4.2, (b) the log stock return inclusive of dividends, and (c) changes in profitability, defined as earnings before interest and taxes (EBIT) as a fraction of the previous fiscal year's variable costs. In all specifications, we trim the 1% and 99% tails of the outcome variable. In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are two-way clustered by firm ID and industry-year.

Figure A-8: Dynamic Relationship Between Optimism and Financial Variables



Notes: The regression model is Equation 14 (as in Figure 1-2), but with financial fundamentals as outcomes. Each coefficient is estimated from a separate projection regression. The outcome variables are: (a) the fiscal-year-to-fiscal-year difference in leverage, which is total debt (short-term debt plus long-term debt); (b) sale of common and preferred stock minus buybacks, normalized by the total equity outstanding in the previous fiscal year; (c) short-term debt plus long-term debt issuance, normalized by the total debt in the previous fiscal year; and (d) total dividends divided by earnings before interest and taxes (EBIT). In all specifications, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals. Standard errors are two-way clustered by firm ID and industry-year.

Figure A-9: The Contagiousness and Associativeness of Other Identified Narratives



Notes: For each narrative, we estimate the analog of Equation 18. We first transform each continuous narrative loading variable $\hat{\lambda}_{k,it}$ (indexed by narrative identifier k, firm i, and time period t) into a binary indicator for being above the sample median,

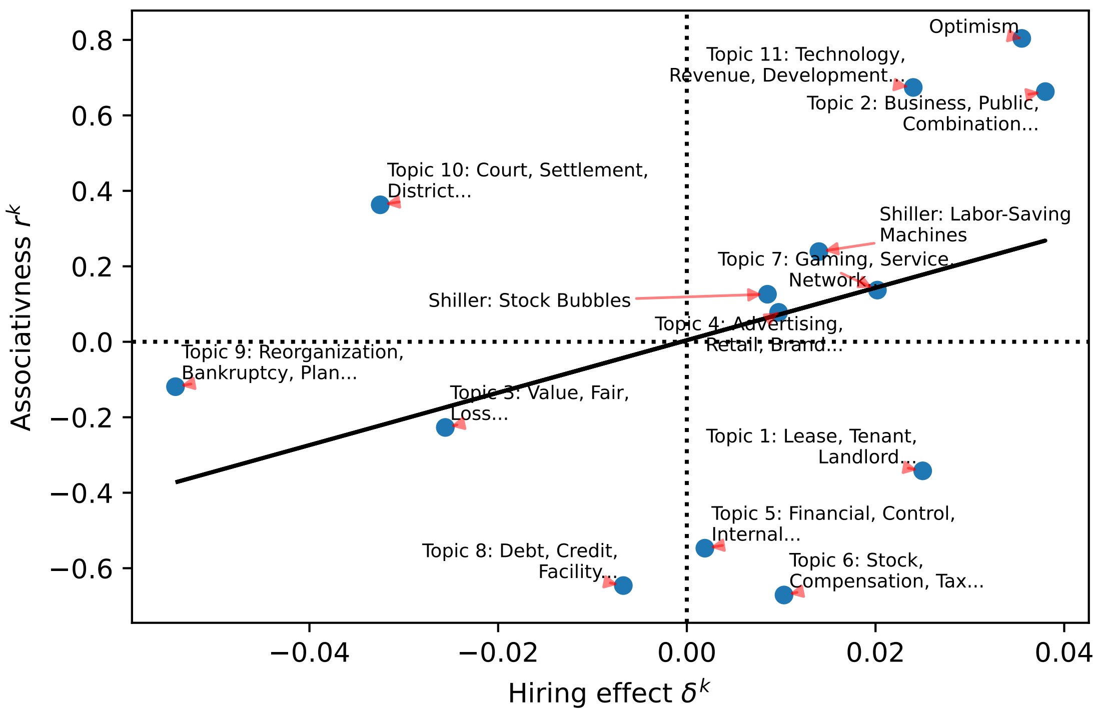
$$\hat{\lambda}^b_{k,it} = \mathbb{I}\left[\hat{\lambda}_{k,it} \geq \text{med}\left(\hat{\lambda}_{k,it}\right)\right] \in \{0,1\}$$

We then estimate

$$\hat{\lambda}^b_{k,it} = u_k\ \hat{\lambda}^b_{k,i,t-1} + s_k\ \overline{\hat{\lambda}^b}_{k,it} + r_k\ \Delta \log Y_{t-1} + \gamma_i + \varepsilon_{it}$$

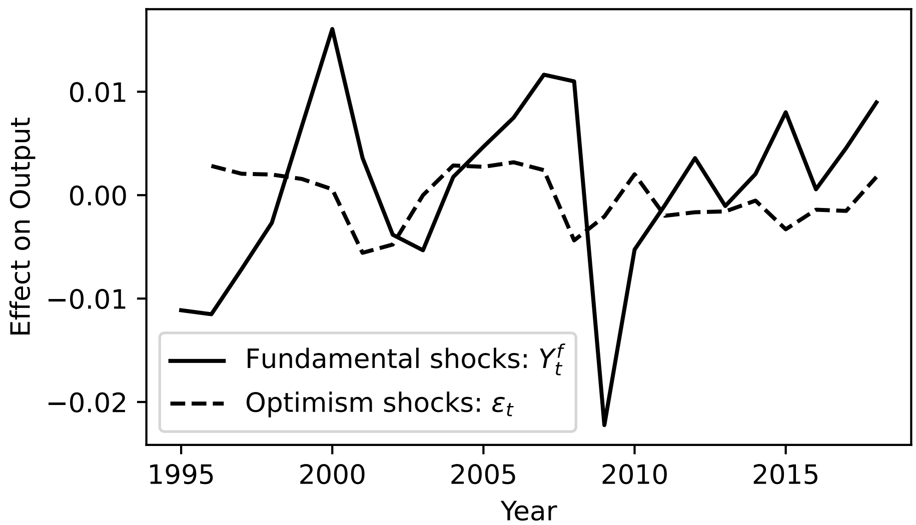
where $\overline{\hat{\lambda}^b}_{k,it}$ denotes an aggregate average of the binary variable. The three panels respectively show our estimates, for each narrative k, of (u_k, s_k, r_k). The error bars are 95% confidence intervals based on double-clustered (firm and industry-year) standard errors.

Figure A-10: Comparing the Hiring Effects and Associativeness of Narratives



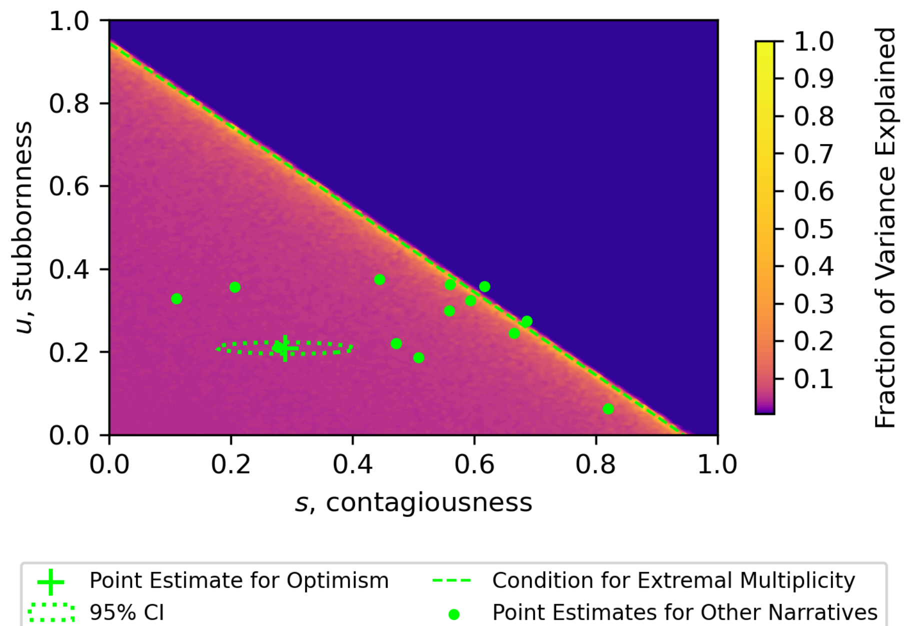
Notes: For each narrative, we estimate the analog of Equation 12 with $\hat{\lambda}^b_{k,it}$, a binary indicator for the narrative weight k being above the sample median (for firm i at time t), in place of opt_{it} (see the notes for Figure A-9 for details about constructing these variables). The x-axis of this figure shows the estimated hiring coefficients δ^k and the y-axis shows the estimated associativeness coefficients r^k (see Figure A-9). Each point is labeled by its running short name, and the solid line is a best-fit trend based on these fourteen points.

Figure A-11: Fundamental and Optimism Shocks That Explain US GDP



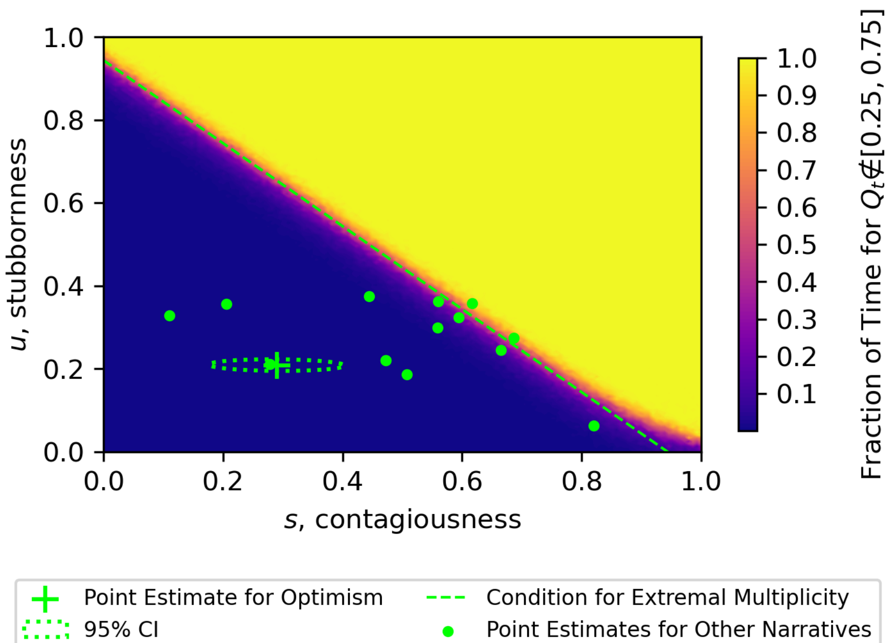
Notes: This figure shows the shocks that rationalize movements in optimism and detrended real GDP in recent US history, as analyzed in Section 1.7.2. The solid line is the exogenous process for fundamental output and the dashed line is the sequence of shocks in narrative evolution. The dashed line is rescaled by $\delta^{OP}(1-\omega)^{-1}$ to be, up to linear approximation of f, in units of output.

Figure A-12: Variance Decomposition for Different Values of Stubbornness and Contagiousness, No Optimism Shocks



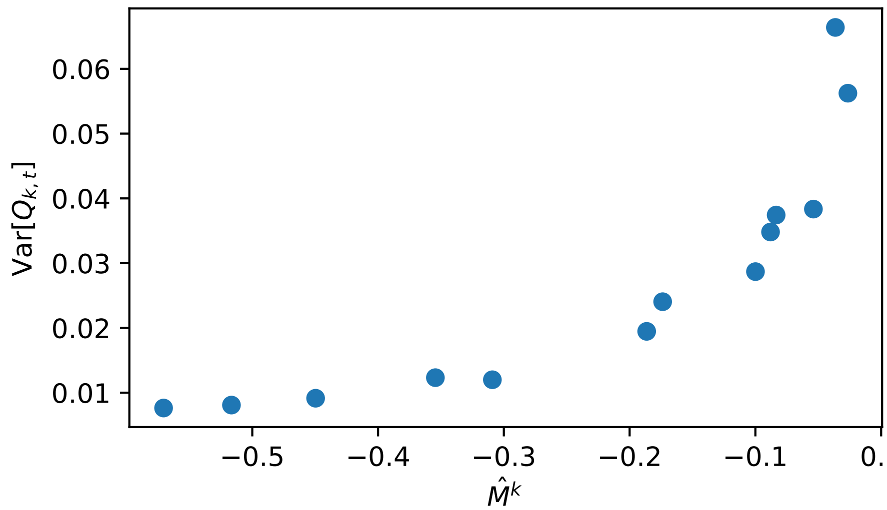
Notes: This Figure replicates Figure A-12, with a different color bar scale, in the variant model with no exogenous shocks to optimism. Calculations vary u and s, holding fixed all other parameters at their calibrated values. The shading corresponds to the fraction of variance explained by optimism, or Share of Variance Explained$_0$ defined in Equation 56. The plus is our calibrated value of (u, s), corresponding to a variance share 4.7%, and the dotted line is the boundary of a 95% confidence set. The dots are calibrated values for other narratives from Figure A-9. The dashed line is the condition of extremal multiplicity from Corollary 3 and Equation 44.

Figure A-13: Tendency Toward Extremal Optimism



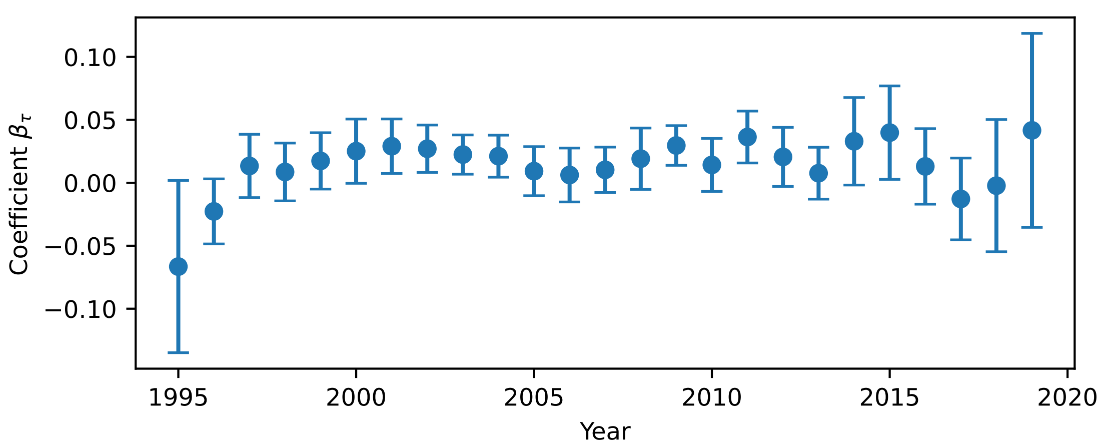
Notes: This Figure plots, in color, the fraction of time that optimism Q_t lies outside of the range $[0.25, 0.75]$ and therefore concentrates at extreme values. Calculations vary u and s, holding fixed all other parameters at their calibrated values. The plus is our calibrated value of (u, s), corresponding to an extremal share of 0%, and the dotted line is the boundary of a 95% confidence set. The dots are calibrated values for other narratives from Figure A-9. The dashed line is the condition of extremal multiplicity from Corollary 3 and Equation 44.

Figure A-14: Variance Contributions Toward Emergent Optimism



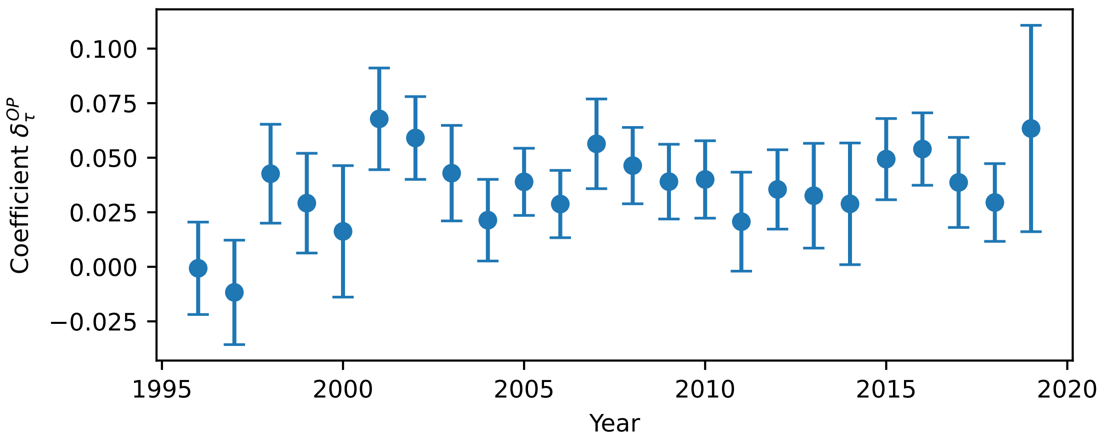
Notes: This Figure plots each granular narrative's contribution toward the variance of emergent optimism in the constellations model (Section 1.7.4) against the condition number for extremal multiplicity (Equation 57). Each dot corresponds to one of the thirteen granular narratives.

Figure A-15: Time-Varying Relationship Between Optimism and TFP



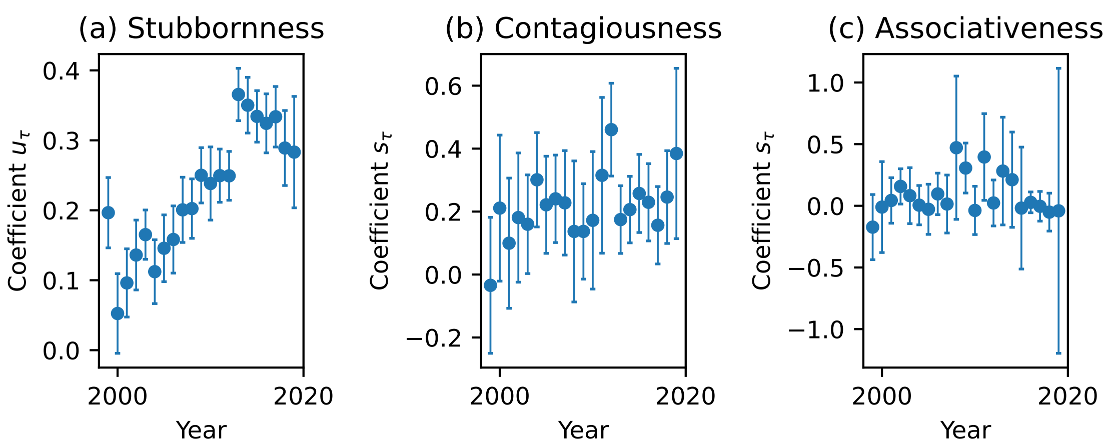
Notes: Each dot is a coefficient β_τ estimated from Equation 348, corresponding to a year-specific effect of binary optimism (opt_{it}) on log TFP ($\log \hat{\theta}_{it}$). The outcome variable is firm-level log TFP, $\log \theta_{it}$, and the regressors are indicators for binary optimism interacted with year dummies. In the regression, we trim the 1% and 99% tails of the outcome variable. Error bars are 95% confidence intervals, based on standard errors clustered by firm and industry-time.

Figure A-16: Time-Varying Effects of Narrative Optimism on Hiring



Notes: Each dot is a coefficient δ_τ^{OP} estimated from Equation 439, capturing a year-specific effect of binary optimism (opt_{it}) on hiring ($\Delta \log L_{it}$). Error bars are 95% confidence intervals, based on standard errors clustered by firm and industry-year.

Figure A-17: Time-Varying Stubbornness, Contagiousness, and Associativeness of Narrative Optimism



Notes: Each dot is a coefficient estimated from Equation 440, and error bars are 95% confidence intervals, based on standard errors clustered by firm and industry-year. Panel (a) plots the stubbornness coefficients on $\text{opt}_{i,t-1}$, u_τ; panel (b) plots the contagiousness coefficients on $\overline{\text{opt}}_{j(i),t-1}$, s_τ; and panel (c) plots the associativeness coefficients on $\Delta \log Y_{j(i),t-1}$, r_τ.

Table A.1: The Twenty Most Common Positive and Negative Words

Positive	Negative
well	loss
good	decline
benefit	disclose
high	subject
gain	terminate
advance	omit
achieve	defer
improve	claim
improvement	concern
opportunity	default
satisfy	limitation
lead	delay
enhance	deficiency
enable	fail
able	losses
best	damage
gains	weakness
improvements	adversely
opportunities	against
resolve	impairment

Notes: The twenty most common lemmatized words among the 230 positive words and 1354 negative words. They are listed in the order of their document frequency. The words are taken from the Loughran and McDonald (2011a) dictionary, as described in Section 1.3.2.

Table A.2: The Twenty Most Common Words for Each Shiller Chapter

Panic	Frugality	Gold Standard	Labor-Saving Machines	Automation and AI	Real Estate	Stock Market	Boycotts	Wage-Price Spiral
bank	help	standard	replac	replac	price	chapter	price	countri
consum	hous	book	produc	appear	appear	peopl	profit	labor
appear	buy	money	technolog	show	real	specul	good	union
show	home	run	appear	question	find	drop	consum	ask
forecast	famili	paper	book	suggest	hous	play	start	wage
economi	lost	peopl	power	labor	estat	depress	fall	inflat
suggest	display	metal	save	ask	buy	warn	buy	strong
run	job	depress	problem	run	home	peak	wage	world
concept	peopl	eastern	labor	worker	citi	great	inflat	mile
peopl	explain	almost	innov	vacat	land	today	world	peopl
grew	phrase	depositor	run	autom	movement	get	cut	happen
around	depress	young	wage	human	world	decad	shop	depress
weather	postpon	today	worker	univers	tend	reaction	peopl	war
figur	car	want	electr	world	peopl	newspap	explain	tri
confid	justifi	went	mechan	machin	never	news	campaign	wrote
wall	cultur	decad	human	job	search	storm	play	peak
happen	fashion	idea	world	peopl	specul	saw	depress	great
depress	unemploy	man	machin	answer	explain	memori	behavior	recess
tri	great	newspap	job	around	popul	interview	postpon	went
unemploy	fault	popular	invent	figur	phrase	watch	war	get

Notes: The twenty most common lemmatized words among the 100 words that typify each Shiller (2020) narrative. Our selection procedure is described in Section 1.3.2.

Table A.3: The Ten Most Common Words for Each Selected Topic

Topic 1		Topic 2		Topic 3		Topic 4		Topic 5		Topic 6	
Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight
lease	0.047	business	0.052	value	0.088	advertising	0.029	financial	0.051	stock	0.049
tenant	0.042	public	0.025	fair	0.082	retail	0.028	control	0.02	compensation	0.039
landlord	0.03	combination	0.024	loss	0.024	brand	0.018	internal	0.019	tax	0.039
lessee	0.017	merger	0.023	investment	0.024	credit	0.018	material	0.013	share	0.028
rent	0.016	class	0.015	asset	0.022	consumer	0.017	affect	0.012	income	0.023
lessor	0.014	offer	0.014	debt	0.02	distribution	0.016	officer	0.011	average	0.019
property	0.012	share	0.013	gain	0.019	card	0.015	base	0.01	expense	0.018
term	0.011	account	0.011	credit	0.019	marketing	0.015	information	0.01	asset	0.016
day	0.009	ordinary	0.01	level	0.017	food	0.013	make	0.01	outstanding	0.016
provide	0.008	private	0.01	financial	0.016	store	0.013	business	0.01	weight	0.015

Topic 7		Topic 8		Topic 9		Topic 10		Topic 11	
Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight	Lem. Word	Weight
gaming	0.035	debt	0.039	reorganization	0.048	court	0.038	technology	0.018
service	0.029	credit	0.039	bankruptcy	0.047	settlement	0.027	revenue	0.017
network	0.022	facility	0.037	plan	0.044	district	0.021	development	0.015
wireless	0.021	senior	0.028	predecessor	0.036	certain	0.019	business	0.013
local	0.019	interest	0.026	successor	0.027	litigation	0.016	customer	0.012
cable	0.015	agreement	0.021	chapter	0.021	action	0.016	stock	0.012
provide	0.014	cash	0.019	asset	0.019	complaint	0.012	product	0.012
equipment	0.013	rate	0.016	court	0.018	damage	0.011	support	0.009
access	0.013	term	0.016	cash	0.016	approximately	0.011	market	0.009
video	0.012	certain	0.014	certain	0.014	case	0.01	service	0.008

Notes: The ten most common words (lemmatized bigrams) in each topic estimated by LDA and selected by our LASSO procedure as relevant for hiring (see Section 1.4.1). Weights correspond to relative importance for scoring the document. The LDA model and our estimation procedure are described in Section 1.3.2.

Table A.4: Persistence and Cyclicality of Narratives

	Correlation with			
Narrative N_t	N_{t-1}	u_{t-1}	u_t	u_{t+1}
Optimism	0.754	-0.283	-0.368	-0.287
Topic Narratives (25th)	0.810	-0.430	-0.307	-0.210
Topic Narratives (median)	0.935	0.003	-0.143	-0.092
Topic Narratives (75th)	0.965	0.339	0.252	0.077
Shiller Narratives (25th)	0.792	-0.379	-0.379	-0.367
Shiller Narratives (median)	0.805	0.043	0.088	-0.034
Shiller Narratives (75th)	0.884	0.541	0.422	0.246

Notes: Calculated with annual data from 1995 to 2019 for optimism and the topics, and 1995 to 2017 for the Shiller Narratives. u_t is the US unemployment rate. The quantiles for Shiller Narratives and Topic Narratives are the quantiles of the distribution of the variable in that column within each set of narratives.

Table A.5: Variance Decomposition of Narratives

	Fraction Variance Explained By Means				
Narrative N_{it}	t	Ind.	Ind. x t	Firm	All
Net Sentiment	0.014	0.053	0.082	0.511	0.530
Optimism	0.011	0.041	0.067	0.427	0.444
Shiller Narratives (25th)	0.002	0.050	0.062	0.758	0.761
Shiller Narratives (median)	0.002	0.071	0.087	0.763	0.770
Shiller Narratives (75th)	0.003	0.095	0.109	0.793	0.794
Topic Narratives (25th)	0.010	0.003	0.049	0.252	0.306
Topic Narratives (median)	0.035	0.014	0.099	0.420	0.575
Topic Narratives (75th)	0.087	0.071	0.237	0.645	0.735

Notes: Each cell is $1 - \mathrm{Var}[N_{it}^{\perp}]/\mathrm{Var}[N_{it}]$, where N_{it} is the narrative intensity and $N_{it}^{\perp}$ is the same after projecting out means at the indicated level. The last column ("All") partials out industry-by-time means and firm means.

Table A.6: Robustness to Assumptions About Unobserved Selection When Estimating the Effect of Narrative Optimism on Hiring

Panel A: Regression Estimates

	(1)	(2)
	Outcome is $\Delta L_{it}^{\perp}$	
$\text{opt}_{it}^{\perp}$	0.0373	0.0305
Controls		✓
N	39,298	39,298
R^2	0.005	0.129

Panel B: Oster (2019) Statistics

	(1)	(2)
	$\bar{R}^2$ is	
	$\hat{\bar{R}}^2 = 0.459$	$\bar{R}^2_{\Pi} = 0.387$
λ^* ($\delta^{OP} = 0$)	1.691	2.151
δ^*_{OP} ($\lambda = 1$)	0.0126	0.0165

Notes: This table summarizes the coefficient stability test described in Appendix A.5.1. Panel A shows estimates of Equation 403, with and without controls for current and lagged log TFP and lagged log labor. The estimate in column 1 differs from that in column 1 of Table 1.1 due to restricting to a common sample in columns 1 and 2. The R^2 values are for the model after partialing out fixed effects, and hence correspond with unreported "within-R^2" values in Table 1.1. Panel B prints the two statistics of Oster (2019). In column 1, we set $\bar{R}^2$ equal to our estimated value of 0.459, calculated as described in the text from an "over-controlled" regression of current hiring on lagged controls and future hiring and productivity. In column 2, we use $\bar{R}^2$ given by three times the R^2 in the controlled hiring regression. The first row (λ^* ($\delta^{OP} = 0$) reports the degree of proportional selection that would generate a null coefficient. The second row ($\delta^*_{OP}(\lambda = 1)$) is the bias corrected effect assuming that unobservable controls have the same proportional effect as observable controls.

Table A.7: Narrative Optimism Predicts Hiring, With More Adjustment-Cost Controls

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0305	0.0257	0.0235	0.0184
	(0.0030)	(0.0034)	(0.0037)	(0.0039)
Firm FE	✓	✓	✓	✓
Industry-by-time FE	✓	✓	✓	✓
$\log L_{i,t-1}$	✓	✓	✓	✓
$(\log \hat{\theta}_{it}, \log \hat{\theta}_{i,t-1})$	✓	✓	✓	✓
$(\log L_{i,t-2}, \log \hat{\theta}_{i,t-2})$		✓	✓	✓
$(\log L_{i,t-3}, \log \hat{\theta}_{i,t-3})$			✓	✓
Log Book to Market				✓
Stock Return				✓
Leverage				✓
N	39,298	31,236	25,156	21,913
R^2	0.401	0.395	0.396	0.415

Notes: The regression model is Equation 12. Column 1 replicates column 2 of Table 1.1. Columns 2 and 3 add more lags of firm-level log employment and firm-level log TFP, and column 4 introduces the baseline financial controls (*i.e.,* those in column 3 of Table 1.1). In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.8: Narrative Optimism Predicts Hiring, Alternative Standard Errors

	(1)	(2)	(3)	(4)	(5)
	Outcome is				
	$\Delta \log L_{it}$				$\Delta \log L_{i,t+1}$
opt_{it}	0.0355	0.0305	0.0250	0.0322	0.0216
	(0.0030)	(0.0030)	(0.0032)	(0.0028)	(0.0037)
	[0.0031]	[0.0026]	[0.0031]	[0.0040]	[0.0034]
	{0.0035}	{0.0026}	{0.0025}	{0.0043}	{0.0036}
Firm FE	✓	✓	✓		✓
Industry-by-time FE	✓	✓	✓	✓	✓
Lag labor		✓	✓	✓	✓
Current and lag TFP		✓	✓	✓	✓
Log Book to Market			✓		
Stock Return			✓		
Leverage			✓		
N	71,161	39,298	33,589	40,580	38,402
R^2	0.259	0.401	0.419	0.142	0.398

Notes: This Table replicates the analysis of Table 1.1 with alternative standard error constructions. Standard errors in parentheses are two-way clustered by firm ID and industry-year; those in square brackets are two-way clustered by firm ID and year; and those in braces are two-way clustered by industry and year. For columns 1-4, the regression model is Equation 12 and the outcome is the log change in firms' employment from year $t-1$ to t. The main regressor is a binary indicator for the optimistic narrative, defined in Section 1.3.2. In all specifications, we trim the 1% and 99% tails of the outcome variable. In column 5, the regression model is Equation 13, the outcome is the log change in firms' employment from year t to $t+1$, and control variables are dated $t+1$.

Table A.9: Narrative Optimism Predicts Hiring, Instrumenting With Lag

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0925	0.106	0.102	0.0470
	(0.0130)	(0.0160)	(0.0168)	(0.0053)
Firm FE	✓	✓	✓	
Industry-by-time FE	✓	✓	✓	✓
Lag labor		✓	✓	✓
Current and lag TFP		✓	✓	✓
Log Book to Market			✓	
Stock Return			✓	
Leverage			✓	
N	63,302	35,768	31,071	36,953
First-stage F	773	478	366	3,597

Notes: All columns come from a two-stage-least-squares (2SLS) estimate of Equation 12, using $\text{opt}_{i,t-1}$ as an instrument for opt_{it}. Specifically, the structural equation is

$$\Delta \log L_{it} = \delta^{OP} \cdot \text{opt}_{it} + \gamma_i + \chi_{j(i),t} + \tau' X_{it} + \varepsilon_{it}$$

the endogenous variable is opt_{it} and the excluded instrument is $\text{opt}_{i,t-1}$. In the last row, we report the first-stage F statistic associated with this equation. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.10: Narrative Optimism Predicts Hiring, Conference-Call Measurement

	(1)	(2)	(3)	(4)	(5)
	Outcome is				
	$\Delta \log L_{it}$				$\Delta \log L_{i,t+1}$
optCC_{it}	0.0277	0.0173	0.0121	0.0237	0.0123
	(0.0038)	(0.0040)	(0.0038)	(0.0038)	(0.0044)
Industry-by-time FE	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓		✓
Lag labor		✓	✓	✓	✓
Current and lag TFP		✓	✓	✓	✓
Log Book to Market			✓		
Stock Return			✓		
Leverage			✓		
N	19,625	11,565	10,851	11,919	11,416
R^2	0.300	0.461	0.467	0.172	0.429

Notes: The regression models are identical to those reported in Table 1.1, but using the measurement of optimism from sales and earnings conference calls. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year. In column 5, control variables are dated $t + 1$.

Table A.11: The Effect of Narrative Optimism on All Inputs

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome is					
	$\Delta \log L_{it}$		$\Delta \log M_{it}$		$\Delta \log K_{it}$	
opt_{it}	0.0355	0.0305	0.0397	0.0193	0.0370	0.0273
	(0.0030)	(0.0030)	(0.0034)	(0.0033)	(0.0034)	(0.0036)
Industry-by-time FE	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓
Lag input		✓		✓		✓
Current and lag TFP		✓		✓		✓
N	71,161	39,298	66,574	39,366	68,864	36,005
R^2	0.259	0.401	0.298	0.418	0.276	0.383

Notes: $\Delta \log M_t$ is the log difference of all variable cost expenditures ("materials"), the sum of cost of goods sold (COGS) and sales, general, and administrative expenses (SGA). $\Delta \log K_t$ is the value of the capital stock is the log difference level of net plant, property, and equipment (PPE) between balance-sheet years $t - 1$ and t. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.12: The Effect of Narrative Optimism on Stock Prices, High-Frequency Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome is stock return on					
	Filing Day		Prior Four Days		Next Four Days	
opt_{it}	0.000145	-0.000142	0.00106	0.000963	0.00173	0.00249
	(0.0007)	(0.0007)	(0.0011)	(0.0014)	(0.0012)	(0.0016)
Firm FE	✓	✓	✓	✓	✓	✓
Industry-by-FY FE	✓	✓	✓	✓	✓	✓
Industry-FF3 inter.		✓		✓		✓
N	39,457	39,457	39,396	17,710	39,346	19,708
R^2	0.189	0.246	0.190	0.345	0.206	0.317

Notes: The regression equation for columns (1), (3), and (5) is $R_{i,w(t)} = \beta \text{opt}_{it} + \gamma_i + \chi_{j(i),y(i,t)} + \varepsilon_{it}$ where i indexes firms, t is the 10K filing day, $w(t)$ is a window around the day (the same day, the prior four days, or the next four days), and $y(i,t)$ is the fiscal year associated with the specific 10-K. In columns (2), (4), and (6), we add interactions of industry codes with the filing day's (i) the market minus risk-free rate, (ii) high-minus-low return, and (iii) small-minus-big return. Standard errors are two-way clustered by firm ID and industry-year.

Table A.13: Textual Optimism and Optimistic Forecasts, Alternative Measurement

	(1)	(2)	(3)	(4)
	$\text{GuidanceOptExPost}_{i,t+1}$		$\text{GuidanceOptExPostC}_{i,t+1}$	
opt_{it}	0.0354		-0.000169	
	(0.0184)		(0.0049)	
sentiment_{it}		0.0152		-0.000219
		(0.0095)		(0.0025)
N	3,817	3,780	3,754	3,719
R^2	0.173	0.174	0.139	0.141

Notes: The regression model is Equation 15. The outcome in columns 1 and 2 is a binary indicator for *ex post* optimism in guidance, and the outcome in columns 3 and 4 is the difference between the log guidance value and the log realized sales. opt_{it} is the binary measure of narrative optimism, and sentiment_{it} is the underlying continuous measure from which opt_{it} is constructed. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.14: The Effects of All Selected Narratives on Hiring

	(1)	(2)	(3)	(4)
		Outcome is $\Delta \log L_{it}$		
	OLS	OLS	OLS	IV
Shiller: Labor-Saving Machines	0.0106			
	(0.0028)			
Shiller: Stock Bubbles	0.00968			
	(0.0031)			
Topic 1: Lease, Tenant, Landlord...		0.0109		
		(0.0017)		
Topic 2: Business, Public, Combination...		0.0266		
		(0.0045)		
Topic 3: Value, Fair, Loss...		-0.00383		
		(0.0016)		
Topic 4: Advertising, Retail, Brand...		0.00864		
		(0.0024)		
Topic 5: Financial, Control, Internal...		-0.000655		
		(0.0025)		
Topic 6: Stock, Compensation, Tax...		0.0135		
		(0.0019)		
Topic 7: Gaming, Service, Network...		0.0146		
		(0.0040)		
Topic 8: Debt, Credit, Facility...		-0.00584		
		(0.0022)		
Topic 8: Reorganization, Bankruptcy, Plan...		-0.00842		
		(0.0018)		
Topic 10: Court, Settlement, District...		-0.00749		
		(0.0019)		
Topic 11: Technology, Revenue, Development...		0.0259		
		(0.0040)		
opt_{it}			0.0305	0.0597
			(0.0030)	(0.0099)
Industry-by-time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Lag labor	✓	✓	✓	✓
Current and lag TFP	✓	✓	✓	✓
N	37,462	39,298	39,298	34,106
R^2	0.413	0.405	0.401	0.130
First-stage F	—	—	—	189.0

Notes: The first-stage equation for column 4 is described in Equation 17. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.15: Narrative Optimism is Contagious and Associative, Alternative Standard Errors

	(1)	(2)	(3)
	Outcome is opt_{it}		
Own lag, $\text{opt}_{i,t-1}$	0.209	0.214	0.135
	(0.0071)	(0.0080)	(0.0166)
	[0.0214]	[0.0220]	[0.0281]
	{0.0218}	{0.0221}	{0.0273}
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290		
	(0.0578)		
	[0.180]		
	{0.179}		
Real GDP growth, $\Delta \log Y_{t-1}$	0.804		
	(0.2204)		
	[0.635]		
	{0.627}		
Industry lag, $\overline{\text{opt}}_{j(i),t-1}$		0.276	0.207
		(0.0396)	(0.0733)
		[0.0434]	[0.0563]
		{0.0496}	{0.0656}
Industry output growth, $\Delta \log Y_{j(i),t-1}$		0.0560	0.0549
		(0.0309)	(0.0632)
		[0.0328]	[0.0668]
		{0.0428}	{0.0772}
Peer lag, $\overline{\text{opt}}_{p(i),t-1}$			0.0356
			(0.0225)
			[0.0259]
			{0.0329}
Firm FE	✓	✓	✓
Time FE		✓	✓
N	64,948	52,258	8,514
R^2	0.481	0.501	0.501

Notes: This Table replicates the analysis of Table 1.4 with alternative standard error constructions. Standard errors in parentheses are two-way clustered by firm ID and industry-year; those in square brackets are two-way clustered by firm ID and year; and those in braces are two-way clustered by industry and year. Aggregate, industry, and peer average optimism are averages of the narrative optimism variable over the respective sets of firms. Industry output growth is the log difference in sectoral value-added calculated from BEA data, linked to two-digit NAICS industries.

Table A.16: Narrative Sentiment is Contagious and Associative

	(1)	(2)	(3)
	Outcome is sentiment$_{it}$		
Own lag, sentiment$_{i,t-1}$	0.259	0.279	0.226
	(0.0091)	(0.0106)	(0.0166)
Aggregate lag, $\overline{\text{sentiment}}_{t-1}$	0.253		
	(0.0519)		
Real GDP growth, $\Delta \log Y_{t-1}$	2.632		
	(0.5305)		
Industry lag, $\overline{\text{sentiment}}_{j(i),t-1}$		0.175	0.108
		(0.0360)	(0.0763)
Industry output growth, $\Delta \log Y_{j(i),t-1}$		0.108	0.142
		(0.0522)	(0.1312)
Peer lag, $\overline{\text{sentiment}}_{p(i),t-1}$			0.0234
			(0.0188)
Firm FE	✓	✓	✓
Time FE		✓	✓
N	63,881	51,555	8,338
R^2	0.568	0.599	0.602

Notes: The regression model is a variant of Equation 18 for column 1, and a variant of Equation 19 for columns 2 and 3, with the continuous variable sentiment$_{it}$ (and averages thereof) substituted for binary optimism. Aggregate, industry, and peer average sentiment are averages of the narrative sentiment variable over the respective sets of firms. Industry output growth is the log difference in sectoral value-added calculated from BEA data, linked to two-digit NAICS industries. In all specifications, we trim the 1% and 99% tails of sentiment$_{it}$. Standard errors are two-way clustered by firm ID and industry-year.

Table A.17: Narrative Sentiment is Contagious and Associative, Controlling for Past and Future Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Outcome is sentiment_{it}			
Aggregate lag, $\overline{\text{sentiment}}_{t-1}$	0.253	0.385	0.410	0.340			
	(0.0519)	(0.0651)	(0.1103)	(0.1785)			
Ind. lag, $\overline{\text{sentiment}}_{j(i),t-1}$					0.175	0.151	0.213
					(0.0360)	(0.0409)	(0.0654)
Time FE					✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓
Own lag, $\text{opt}_{i,t-1}$	✓	✓	✓	✓	✓	✓	✓
$(\Delta \log Y_{t+k})_{k=-2}^{2}$		✓	✓	✓			
$(\Delta \log Y_{j(i),t+k})_{k=-2}^{2}$			✓	✓		✓	✓
$(\Delta \log \hat{\theta}_{i,t+k})_{k=-2}^{2}$				✓			✓
N	63,881	48,889	37,643	13,112	51,555	37,643	13,112
R^2	0.568	0.578	0.599	0.640	0.599	0.601	0.642

Notes: The regression model is a variant of Equation 20 for column 1-4, and an analogous variant of industry-level specification for columns 5-7 (*i.e.*, Equation 19 with past and future controls), with the continuous variable sentiment_{it} (and averages thereof) substituted for binary optimism. Columns 1 and 5 are "baseline estimates" corresponding, respectively, with columns 1 and 3 of Table A.16. The added control variables are two leads, two lags, and the contemporaneous value of: real GDP growth (columns 2-4), industry-level output growth (columns 3-4 and 6-7), and firm-level TFP growth (columns 4 and 7). In all specifications, we trim the 1% and 99% tails of sentiment_{it}. Standard errors are two-way clustered by firm ID and industry-year.

Table A.18: Narrative Optimism and Contagious and Associative, Instrumented With Other Narratives

	(1)	(2)	(3)	(4)
		Outcome is opt_{it}		
	OLS	IV	OLS	IV
Own lag, $\text{opt}_{i,t-1}$	0.209	0.207	0.214	0.200
	(0.0071)	(0.0072)	(0.0080)	(0.0084)
Aggregate lag, $\overline{\text{opt}}_{t-1}$	0.290	0.393		
	(0.0578)	(0.0597)		
Real GDP growth, $\Delta \log Y_{t-1}$	0.804	0.672		
	(0.2204)	(0.2153)		
Ind. lag, $\overline{\text{opt}}_{j(i),t-1}$			0.276	0.437
			(0.0396)	(0.0748)
Ind. output growth, $\Delta \log Y_{j(i),t-1}$			0.0560	0.0390
			(0.0309)	(0.0342)
Firm FE	✓	✓	✓	✓
Time FE		✓	✓	
N	64,948	64,569	52,258	47,536
R^2	0.481	0.050	0.501	0.047
First-stage F	—	795.3	—	19.8

Notes: In column 2, the endogenous variable is $\overline{\text{opt}}_{t-1}$ and the instruments are $(\overline{\text{Shiller}}^k_{t-1})^{K^*_S}_{k=1}$ and $(\overline{\text{topic}}^k_{t-1})^{K^*_T}_{k=1}$ where the sums are over the LASSO-selected narratives (see Table 1.3). In column 4, the endogenous variable is $\overline{\text{opt}}_{j(i),t-1}$ and the instruments are $(\overline{\text{Shiller}}^k_{j(i),t-1})^{K^*_S}_{k=1}$ and $(\overline{\text{topic}}^k_{j(i),t-1})^{K^*_T}_{k=1}$. Standard errors are two-way clustered by firm ID and industry-year.

Table A.19: Sensitivity Analysis for the Quantitative Analysis

	Parameters						Results			
	α	γ	ψ	ϵ	ω	$\frac{1}{1-\omega}$	$\hat{c}_Q(0)$	$\hat{c}_Q(1)$	2000-02	2007-09
Baseline	1.0	0.0	0.4	2.6	0.490	1.962	0.192	0.335	0.316	0.181
High ψ	1.0	0.0	2.5	2.6	0.133	1.154	0.175	0.359	0.186	0.106
High γ	1.0	1.0	0.4	2.6	-0.784	0.560	0.041	0.184	0.090	0.052
Empirical Multiplier	1.0	0.0	1.15	2.6	0.250	1.333	0.167	0.329	0.215	0.123
Calibrated Multiplier	1.0	0.0	0.845	2.6	0.313	1.455	0.168	0.324	0.235	0.134
High ϵ	1.0	0.0	0.21	5.0	0.490	1.962	0.109	0.240	0.317	0.181
Decreasing RtS	0.75	0.0	0.05	2.6	0.490	1.962	0.125	0.238	0.237	0.135

Notes: This table summarizes the quantitative results under alternative calibrations of the macroeconomic parameters, which we report along side their implied complementarity ω and demand multiplier $\frac{1}{1-\omega}$. We report four statistics as the "results" in the last four columns. The first two are the fraction of output variance explained statically, $\hat{c}_Q(0)$, and at a one-year horizon, $\hat{c}_Q(1)$, by optimism. The second two are the fraction of output losses in the 2000-02 downturn and 2007-09 downturn explained by fluctuations in narrative optimism. Baseline corresponds to our main calibration. High ψ increases the inverse Frisch elasticity to 2.5, or decreases the Frisch elasticity to 0.4. High γ increases the curvature of consumption utility (indexing income effects in labor supply) from 0.0 to 1.0. Empirical Multiplier adjusts ψ to match an output multiplier in line with estimates from Flynn, Patterson, and Sturm (2021). Calibrated multiplier adjusts ψ to match our own calculation of the multiplier in our setting in Appendix A.6.3. High ϵ increases the elasticity of substitution from 2.6 to 5.0, with ψ adjusting to hold fixed the multiplier. Decreasing RtS reduces the returns-to-scale parameter α from 1.0 to 0.75, with ψ adjusting to hold fixed the multiplier.

Table A.20: An Empirical Test for Cycles and Chaos

	(1) Outcome is opt_{it}
α: Constant	-0.051
	(0.244)
α_1: $\text{opt}_{i,t-1}$	0.655
	(0.062)
β_1: $\text{opt}_{i,t-1} \cdot \overline{\text{opt}}_{i,t-1}$	0.052
	(1.021)
β_2: $(1 - \text{opt}_{i,t-1}) \cdot \overline{\text{opt}}_{i,t-1}$	0.952
	(1.006)
τ: $(\overline{\text{opt}}_{i,t-1})^2$	-0.062
	(1.034)
η: Logistic parameter	1.443
	(0.698)
Firm FE	✓
N	67,648
R^2	0.480

Notes: The regression model is Equation 365. η is a function of the regression coefficients defined in Equation 366, and interpretable in the model of cycles and chaos in Appendix A.2.7. Standard errors are two-way clustered by firm ID and industry-year. The standard error for η is calculated using the delta method.

Table A.21: Data Definitions in Compustat

	Quantity	Expenditure
Production, x_{it}	—	`sale`
Employment, L_{it}	`emp`	`emp` × industry wage
Materials, M_{it}	—	`cogs` + `xsga` − `dp`
Capital, K_{it}	`ppegt` plus net investment	—

Table A.22: The Effect of Optimism on Hiring, CEO Change Strategy

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0253	0.0404	0.0362	0.0253
	(0.0131)	(0.0131)	(0.0132)	(0.0029)
$\text{opt}_{it} \times \text{ChangeCEO}_{it}$				0.0220
				(0.0099)
ChangeCEO_{it}				-0.0232
				(0.0088)
Industry-by-time FE	✓	✓	✓	✓
Lag optimism	✓	✓	✓	✓
Lag labor		✓	✓	✓
Current and lag TFP		✓	✓	✓
Log Book to Market			✓	
Stock Return			✓	
Leverage			✓	
N	1,725	982	905	36,953
R^2	0.243	0.375	0.375	0.134

Notes: The regression model is Equation 404 for columns 1-3, and Equation 405 for column 4. The outcome is the log change in firms' employment. opt_{it} is a binary indicator for the optimistic narrative, defined in Section 1.3.2. ChangeCEO_{it} is a binary indicator for whether firm i changed CEO in fiscal year t due to death, illness, personal issues or voluntary retirement. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.23: The Contagiousness of Optimism, CEO Change Strategy

	(1)	(2)	(3)	(4)
	Outcome is opt_{it}			
	OLS	IV	OLS	IV
Industry lag, $\overline{\text{opt}}_{j(i),t-1}$	0.275	0.260	0.195	0.272
	(0.0407)	(0.2035)	(0.0760)	(0.5270)
Peer lag, $\overline{\text{opt}}_{p(i),t-1}$			0.0437	0.129
			(0.0236)	(0.1677)
Firm FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Industry output growth, $\Delta \log Y_{j(i),t-1}$	✓	✓	✓	✓
N	50,604	50,604	7,873	7,873
R^2	0.503	0.051	0.508	0.020
First-stage F	—	29.7	—	36.8

Notes: The IV strategies instrument the industry and/or peer lag with the CEO-change variation in those averages. Standard errors are two-way clustered by firm ID and industry-year.

Table A.24: Narrative Optimism Predicts Hiring, Conditional on Measured Beliefs

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log L_{it}$			
opt_{it}	0.0355	0.0232	0.0311	0.0203
	(0.0030)	(0.0129)	(0.0068)	(0.0164)
$\text{ForecastGrowthSales}_{it}$		0.157		
		(0.0329)		
$\text{ForecastGrowthCapx}_{it}$			0.0564	
			(0.0062)	
$\text{ForecastGrowthEps}_{it}$				0.000961
				(0.0104)
Ind.-by-time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
N	71,161	2,908	7,312	1,290
R^2	0.259	0.506	0.425	0.638

Notes: opt_{it} is textual optimism from the 10-K for fiscal year t. $\text{ForecastGrowthZ}_{it}$ is defined in the text as the log difference between manager guidance about statistic Z, for fiscal year t, with last fiscal year's realized value. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.25: Narrative Optimism Predicts Investment, Conditional on Measured Beliefs

	(1)	(2)	(3)	(4)
	Outcome is $\Delta \log K_{it}$			
opt_{it}	0.0370	0.0238	0.0251	0.00503
	(0.0034)	(0.0177)	(0.0072)	(0.0193)
$\text{ForecastGrowthSales}_{it}$		0.172		
		(0.0423)		
$\text{ForecastGrowthCapx}_{it}$			0.0943	
			(0.0079)	
$\text{ForecastGrowthEps}_{it}$				-0.0147
				(0.0102)
Ind.-by-time FE	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
N	68,864	2,748	7,048	1,245
R^2	0.276	0.496	0.472	0.661

Notes: This table is identical to Table A.24, but has net capital investment ΔK_{it} as the outcome. opt_{it} is textual optimism from the 10-K for fiscal year t. $\text{ForecastGrowthZ}_{it}$ is defined in the text as the log difference between manager guidance about statistic Z, for fiscal year t, with last fiscal year's realized value. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.26: State-Dependent Effects of Sentiment on Hiring

	(1)	(2)	(3)
	Outcome is $\Delta \log L_{it}$		
sentiment$_{it}$	0.0218	0.0172	0.0130
	(0.0017)	(0.0018)	(0.0020)
sentiment$_{i,t-1}$	0.00605	0.00877	0.00830
	(0.0015)	(0.0016)	(0.0016)
sentiment$_{it}$ $\times$ sentiment$_{i,t-1}$	-0.00497	-0.00501	-0.00404
	(0.0008)	(0.0008)	(0.0008)
N	63,302	35,768	31,071
R^2	0.257	0.394	0.416
Ind.-by-time FE	✓	✓	✓
Firm FE	✓	✓	✓
Lag labor		✓	✓
Current and lag TFP		✓	✓
Log Book to Market			✓
Stock Return			✓
Leverage			✓

Notes: This table reports estimates from Equation 409 with our baseline sets of controls. In all specifications, we trim the 1% and 99% tails of the outcome variable. Standard errors are two-way clustered by firm ID and industry-year.

Table A.27: Optimism is Contagious and Associative, Granular IV Strategy

	(1)	(2)	(3)	(4)	(5)	(6)
			Outcome is opt_{it}			
	OLS	OLS	IV	OLS	OLS	IV
Own lag, $\text{opt}_{i,t-1}$	0.212	0.213	0.210	0.219	0.220	0.219
	(0.0071)	(0.0071)	(0.0073)	(0.0080)	(0.0081)	(0.0081)
Agg. sales-wt. lag, $\overline{\text{opt}}^{sw}_{t-1}$	0.0847		0.308			
	(0.0421)		(0.1044)			
Real GDP growth, $\Delta \log Y_{t-1}$	1.058	1.104	0.768			
	(0.2205)	(0.2110)	(0.2607)			
Agg. sales-wt. granular lag, $\overline{\text{opt}}^{g,sw}_{t-1}$		0.150				
		(0.0506)				
Ind. sales-wt. lag, $\overline{\text{opt}}^{sw}_{j(i),t-1}$				0.0728		0.0195
				(0.0209)		(0.0459)
Ind. output growth, $\Delta \log Y_{j(i),t-1}$				0.0851	0.0903	0.0886
				(0.0325)	(0.0336)	(0.0333)
Ind. sales-wt. granular lag, $\overline{\text{opt}}^{g,sw}_{j(i),t-1}$					0.00913	
					(0.0216)	
Firm FE	✓	✓	✓	✓	✓	✓
Time FE				✓	✓	✓
N	64,948	64,948	64,948	52,258	50,842	50,842
R^2	0.481	0.481	0.049	0.500	0.503	0.051
First-stage F	—	—	99.1	—	—	262.3

Notes: This table estimates Equations 18 and 19, respectively modeling the spread of optimism at the aggregate and industry level, using granular identification of spillovers (contagiousness). $\overline{\text{opt}}^{sw}_{t-1}$ and $\overline{\text{opt}}^{sw}_{j(i),t-1}$ are sales-weighted averages of aggregate and industry optimism, respectively. $\overline{\text{opt}}^{g,sw}_{t-1}$ and $\overline{\text{opt}}^{g,sw}_{j(i),t-1}$ are (lagged) sales-weighted averages of the *non-fundamentally-predictable* components of firm-level optimism in the aggregate and in the industry, respectively, as explained in Appendix A.5.5. In columns 3 and 6, we use the granular variables as instruments for the raw sales-weighted averages. Standard errors are two-way clustered by firm ID and industry-year.

Table A.28: Multiplier Calibrations via Under-Controlled Regressions of Hiring on Optimism

	(1)	(2)	(3)
	Outcome is ΔL_{it}		
opt_{it}	0.0355	0.0516	0.0486
	(0.0030)	(0.0034)	(0.0033)
Complementarity ω	—	0.313	0.270
Multiplier $\frac{1}{1-\omega}$	—	1.455	1.370
Industry-by-time FE	✓		
Firm FE	✓	✓	✓
Current and lagged adjusted TFP		✓	
Current and lagged unadjusted TFP			✓
N	71,161	65,508	65,508
R^2	0.259	0.207	0.216

Notes: The regression models are introduced in Appendix A.6.3. The first column replicates Column 1 of Table 1.1. The second two columns remove the industry-by-time FE and control for the contemporaneous and lagged value of seasonally adjusted log TFP, respectively with and without capacity utilization adjustment, as reported by the updated data series of Fernald (2014). The sample size is lower in columns 2 and 3 due to the band-pass filtering being impossible for the last part of the sample. The remaining rows give the implied complementarity ω and demand multiplier $\frac{1}{1-\omega}$, by comparing the coefficients with that of column 1 and applying the formula in Equation 437. Standard errors are double-clustered by industry-year and firm ID.

Appendix B

Appendix to Attention Cycles

B.1 Omitted Proofs

B.1.1 Proof of Proposition 6

Proof. Consider a firm of type λ_i, with a payoff $u : \mathcal{X} \times \mathcal{Z} \to \mathbb{R}$ and prior density $\pi \in \Delta(\mathcal{Z})$. The firm's stochastic choice problem can be written as

$$\max_{p \in \mathcal{P}} \int_{\mathcal{X}} \int_{\mathcal{Z}} u(x,z) p(x|z) \, \mathrm{d}x \, \pi(z) \, \mathrm{d}z - \lambda_i \int_{\mathcal{X}} \int_{\mathcal{Z}} p(x|z) \log p(x|z) \, \mathrm{d}x \, \pi(z) \, \mathrm{d}z \tag{441}$$

We can formulate this problem as constrained optimization for choosing $p(x|z)$ point-wise, with constraints embodying non-negativity and the restriction that conditional distributions integrate to one. We can then write a Lagrangian for this problem, giving these constraints multipliers $\kappa(x,z)$ and $\gamma(z)$, respectively:

$$\begin{aligned} \mathcal{L}(\{p(x|z), \kappa(x,z)\}, \{\gamma(z)\}) = & \int_{\mathcal{Z}} \int_{\mathcal{X}} u(x,z) p(x|z) \, \mathrm{d}x \, \pi(z) \, \mathrm{d}z \\ & - \lambda_i \int_{\mathcal{Z}} \int_{\mathcal{X}} p(x|z) \log p(x|z) \, \mathrm{d}x \, \pi(z) \, \mathrm{d}z \\ & + \int_{\mathcal{Z}} \int_{\mathcal{X}} \kappa(x,z) p(x|z) \, \mathrm{d}x \, \pi(z) \, \mathrm{d}z \\ & + \int_{\mathcal{Z}} \gamma(z) \left(\int_{\mathcal{X}} p(x|z) \, \mathrm{d}x - 1 \right) \pi(z) \, \mathrm{d}z \end{aligned} \tag{442}$$

The Lagrangian is concave in the collection $\{p(x \mid z)\}$, since the expected utility term and the two constraint terms are linear in these variables, and the control-cost term is convex in these variables. Taking the first-order condition of the Lagrangian with respect to $p(x|z)$ yields the necessary first-order condition

$$u(x,z) - \lambda_i(\log p(x|z) + 1) + \kappa(x,z) + \gamma(z) = 0 \tag{443}$$

Re-arranging this expression and applying the normalization that the density integrates to one, we get the solution

$$p(x|z) = \frac{\exp\left(\lambda_i^{-1} u(x,z)\right)}{\int_{\mathcal{X}} \exp\left(\lambda_i^{-1} u(x',z)\right) \mathrm{d}x'} \tag{444}$$

This solution is invariant to the prior distribution $\pi(z)$, and hence can be indexed solely by the *ex post* realized state z.

To solve our firm's problem, we replace u in the above with $\tilde{\Pi}$ and z with z_i. Performing this substitution, and ignoring the normalizing constant, we get

$$p(x|z_i) \propto \exp\left(-\frac{(x - x^*(z_i))^2}{2\lambda_i |\Pi_{zz}(z_i)|^{-1}}\right) \tag{445}$$

Taking $\mathcal{X} = \mathbb{R}$, it is then immediate that $p(x|z_i)$ is a Gaussian random variable with mean $x^*(z_i)$ and variance $\lambda_i |\Pi_{xx}(z_i)|^{-1}$. Observing that $|\Pi_{xx}(z_i)| = |\pi_{xx}(z_i)| M(X)$, we can re-write the variance as $\lambda_i(|\pi_{xx}(z_i)| M(X))^{-1}$. Finally, we observe that the stochasticity in each firm's action conditional on z_i is independent from z_i and/or any other firm's action. Thus:

$$x_i = x^*(z_i) + \sqrt{\frac{\lambda_i}{|\pi_{xx}(z_i)| M(X)}} \cdot v_i, \qquad v_i \sim \text{Normal}(0,1) \tag{446}$$

□

B.1.2 Proof of Proposition 7

Proof. We first re-define the state variable as the $\epsilon - 1$ quasi-arithmetic mean of θ_i

$$\theta := \left(\mathbb{E}_{\theta_i}[\theta_i^{\epsilon-1} \mid \theta]\right)^{\frac{1}{\epsilon-1}} \tag{447}$$

The expectation is taken over the distribution $G(\theta)$, which is by assumption increasing in θ via first-order stochastic dominance. Because this redefinition of the state preserves the strict ordering of realizations of θ, and lies within the domain Θ, it is without loss of generality.

To prove existence, we first study the problem of a single firm i who is best replying to the conjecture that the law of motion of the aggregate is $X : \Theta \to \mathbb{R}$. In particular, they believe that output is given by $X(\theta)$ in each state θ.

As established by Proposition 6, the firm's best-response is invariant to the firm's prior state $z_{i,t-1}$ and described by the following random variable conditional on each realization of z_{it}:

$$x_{it} = x^*(z_{it}) + \sqrt{\frac{\lambda_i}{|\Pi_{xx}(z_{it})|}} \cdot v_{it}, \qquad v_{it} \sim \text{Normal}(0,1) \tag{448}$$

As derived in Appendix B.1.7, the mean and variance scalings are the following, after substituting in the equilibrium conjecture $X_t = X(\theta_t)$:

$$\begin{aligned} x^*(z_{it}) &= v_x(\epsilon, \chi, \bar{w}, \bar{X}) \cdot X(\theta_t)^{1-\chi\epsilon}\theta_{it}^{\epsilon} \\ |\Pi_{xx}(z_{it})| &= v_\Pi(\epsilon, \chi, \bar{w}, \bar{X}) \cdot X(\theta_t)^{-1-\gamma+\chi(1+\epsilon)}\theta_{it}^{-1-\epsilon} \end{aligned} \tag{449}$$

for constants $v_x, v_\Pi > 0$ given by:

$$\begin{aligned} v_x &:= \epsilon^{-\epsilon}\left(\epsilon-1\right)^{\epsilon}\bar{w}^{-\epsilon}\bar{X}^{\chi\epsilon} \\ v_\Pi &:= (\epsilon-1)^{-\epsilon}\epsilon^{\epsilon-1}\bar{w}^{1+\epsilon}\bar{X}^{-\chi(1+\epsilon)} \end{aligned} \tag{450}$$

Conditional on the realization of any state θ, aggregate output must solve the fixed point equation defined by combining the aggregate-good production function (72) and the firms' best responses. Applying the law of iterated expectations, we can re-write the aggregate-good production function as

$$X = X^* - \frac{1}{2\epsilon(X^*)^{-\frac{1}{\epsilon}}}\mathbb{E}_{\theta_i}\left[(x^*(z_i))^{-1-\frac{1}{\epsilon}}\mathbb{E}_{\lambda_i,v_i}\left[(x_i - x^*(z_i))^2 \mid \theta_i, \theta\right] \mid \theta\right] \tag{451}$$

where

$$X^* = \left(\mathbb{E}_{\theta_i}[x^*(z_i)^{1-\frac{1}{\epsilon}} \mid \theta]\right)^{\frac{\epsilon}{\epsilon-1}} \tag{452}$$

We now specialize the expressions above using the structure of the best response in Equations 448, 449, and 450. We first compute X^* as

$$X^* = v_x X^{1-\chi\epsilon}\theta^{\epsilon} \tag{453}$$

where θ is the transformation defined in Equation 447. We next calculate the the second, "variance" term. We start with the "misoptimization variance"

$$\mathbb{E}_{\lambda_i, v_i}\left[(x_i - x^*(z_i))^2\right] = \frac{\lambda}{v_\Pi} X^{1+\gamma-\chi(1+\epsilon)} \theta_i^{1+\epsilon} \tag{454}$$

and then calculate the full term

$$(X^*)^{\frac{1}{\epsilon}} \mathbb{E}_{\theta_i}\left[(x^*(z_i))^{-1-\frac{1}{\epsilon}} \mathbb{E}_{\lambda_i, v_i}\left[(x_i - x^*(z_i))^2\right]\right] = \frac{\lambda}{v_\Pi v_x} X^{\gamma-\chi} \theta \tag{455}$$

Substituting in Equations 453 and 455, we derive that equilibrium output solves:

$$X(\theta) = v_x X(\theta)^{1-\chi\epsilon} \theta^\epsilon - \frac{\lambda}{2\epsilon v_x v_\Pi} X(\theta)^{\gamma-\chi} \theta \tag{456}$$

There is always a trivial equilibrium $X = 0$ arising from our approximations. Toward proving existence and uniqueness of a non-trivial equilibrium, define:

$$g(X, \theta) = a_0 X^{1-\chi\epsilon} \theta^\epsilon - a_1 X^{\gamma-\chi} \theta \tag{457}$$

where $a_0 = v_x > 0$ and $a_1 = \frac{\lambda}{2\epsilon v_x v_\Pi} > 0$. We now compute this function's derivatives in X:

$$\begin{aligned} g_X(X, \theta) &= a_0(1-\chi\epsilon) X^{-\chi\epsilon} \theta^\epsilon - a_1(\gamma - \chi) X^{\gamma-\chi-1} \theta \\ g_{XX}(X, \theta) &= -a_0(1-\chi\epsilon)\chi\epsilon X^{-\chi\epsilon-1} \theta^\epsilon - a_1(\gamma-\chi)(\gamma-\chi-1) X^{\gamma-\chi-2} \theta \end{aligned} \tag{458}$$

If $1 - \chi\epsilon > 0$ and $\gamma > 1 + \chi$, then

$$\lim_{X \to 0} g_X(X, \theta) = +\infty \qquad \lim_{X \to \infty} g_X(X, \theta) = -\infty \tag{459}$$

Moreover, if $\gamma > \chi + 1$ we have that $g_{XX}(X, \theta) < 0$ on $(0, \infty)$. Thus, when $\gamma > \chi + 1$ and $\chi\epsilon < 1$, $g(X, \theta)$ crosses X from above and there exists a unique, positive fixed point for each θ. Iterated for all states $\theta \in \Theta$, this reasoning shows the existence of a unique, positive equilibrium mapping $X : \Theta \to \mathbb{R}_+$.

We now show monotonicity of the fixed point. To this end, we implicitly differentiate the fixed point condition:

$$\begin{aligned} \frac{\mathrm{d}X(\theta)}{\mathrm{d}\theta} &= \left[a_0(1-\chi\epsilon) X(\theta)^{-\chi\epsilon} \theta^\epsilon - a_1(\gamma-\chi) X(\theta)^{\gamma-\chi-1} \theta\right] \frac{\mathrm{d}X(\theta)}{\mathrm{d}\theta} \\ &\quad + \left[a_0 \epsilon X(\theta)^{1-\chi\epsilon} \theta^{\epsilon-1} - a_1 X(\theta)^{\gamma-\chi}\right] \end{aligned} \tag{460}$$

Yielding:

$$\frac{\mathrm{d}X(\theta)}{\mathrm{d}\theta} = \frac{a_0\epsilon X(\theta)^{1-\chi\epsilon}\theta^{\epsilon-1} - a_1 X(\theta)^{\gamma-\chi}}{1 - [a_0(1-\chi\epsilon)X(\theta)^{-\chi\epsilon}\theta^{\epsilon} - a_1(\gamma-\chi)X(\theta)^{\gamma-\chi-1}\theta]} \tag{461}$$

Multiplying both sides by a factor of $\frac{\theta}{X}$:

$$\frac{\mathrm{d}\log X(\theta)}{\mathrm{d}\log\theta} = \frac{a_0\epsilon X(\theta)^{1-\chi\epsilon}\theta^{\epsilon} - a_1 X(\theta)^{\gamma-\chi}\theta}{X(\theta) - [a_0(1-\chi\epsilon)X(\theta)^{1-\chi\epsilon}\theta^{\epsilon} - a_1(\gamma-\chi)X(\theta)^{\gamma-\chi}\theta]} \tag{462}$$

We first show that the numerator is positive

$$\begin{aligned} a_0\epsilon X(\theta)^{1-\chi\epsilon}\theta^{\epsilon} - a_1 X(\theta)^{\gamma-\chi}\theta &= a_0(\epsilon-1)X(\theta)^{1-\chi\epsilon}\theta^{\epsilon} + X(\theta) \\ &> 0 \end{aligned} \tag{463}$$

The first equality substitutes in the original fixed-point equation. The second follows from observing that $\epsilon > 1$, $a_0 > 0$, $X(\theta) > 0$, and $\theta > 0$. To show $\frac{\mathrm{d}\log X(\theta)}{\mathrm{d}\log\theta} > 0$ it now suffices to show that the denominator is positive, which follows from $\chi\epsilon < 1$ and $\gamma > \chi + 1 > \chi$:

$$\begin{aligned} X(\theta) &= a_0 X(\theta)^{1-\chi\epsilon}\theta^{\epsilon} - a_1 X(\theta)^{\gamma-\chi}\theta \\ &\geq a_0 \underbrace{(1-\chi\epsilon)}_{\chi\epsilon<1} X(\theta)^{1-\chi\epsilon}\theta^{\epsilon} - a_1 \underbrace{(\gamma-\chi)}_{\gamma-\chi>1} X(\theta)^{\gamma-\chi}\theta \end{aligned} \tag{464}$$

This shows that $\frac{\mathrm{d}\log X(\theta)}{\mathrm{d}\log\theta} > 0$ and implies that $X(\theta)$ is an increasing function. □

B.1.3 Proof of Proposition 8

Proof. Building on the discussion in the main text, we first provide a formal definition of how to rank average attention and misoptimization:

Definition 17 (Aggregate Attention and Misoptimization). *Fix equilibrium laws of motion $\{X(\theta), w(\theta)\}$. Firms' aggregate attention in state θ, $a(\theta)$, is their average realized cognitive cost*

$$a(\theta) := \mathbb{E}_{\theta_i,\lambda_i,v_i}\left[\lambda_i p^*(x|z_i(\theta);\lambda_i)\log p^*(x|z_i(\theta);\lambda_i) \mid \theta\right] \tag{465}$$

where $z_i(\theta) = (\theta_i, X(\theta), w(\theta))$ and $p^*(\cdot \mid z_i;\lambda_i)$ is the uniquely optimal state-contingent plan of a type-λ_i firm contingent on realized state z_i. Firms aggregate misoptimization *is their average mean-squared-error around the ex post optimal action, or*

$$m(\theta) := \mathbb{E}_{\theta_i,\lambda_i,v_i}\left[(x_i - x^*(z_i(\theta)))^2 \mid \theta\right] \tag{466}$$

We now prove the result, starting with the monotonicity of misoptimization. Using the result of Proposition 6, and the substitution of $(x^*(z_i), |\Pi_{xx}(z_i)|)$ as in the proof of Proposition 7, we show that the average "misoptimization variance" of actions conditional on (z_i, λ_i) is

$$m(z_i,\lambda_i,\theta) := \mathbb{E}_{v_i}\left[(x_i - x^*(z_i(\theta)))^2 \mid z_i,\lambda_i,\theta\right] = \frac{\lambda_i X(\theta)^{1+\gamma-\chi(1+\epsilon)}\theta_i^{1+\epsilon}}{v_\Pi(\epsilon,\chi,\bar{w},\bar{X})} \tag{467}$$

where $v_\Pi > 0$ is defined in Equation 450. Using the law of iterated expectations, we can write $m(\theta) = \mathbb{E}_{\theta_i,\lambda_i}[m(z_i,\lambda_i,\theta)]$. Assessing this outer expectation, we derive

$$m(\theta) = \frac{\lambda}{v_\Pi(\epsilon,\chi,\bar{w},\bar{X})} \cdot X(\theta)^{1+\gamma-\chi(1+\epsilon)} \cdot \mathbb{E}_{\theta_i}[\theta_i^{1+\epsilon} \mid \theta] \tag{468}$$

We observe that $\mathbb{E}_{\theta_i}[\theta_i^{1+\epsilon} \mid \theta]$ increases in θ, as $y \mapsto y^{1+\epsilon}$ is an increasing function and $\theta' > \theta \implies G(\theta') \succsim_{FOSD} G(\theta)$. We next observe that $X(\theta)^{1+\gamma-\chi(1+\epsilon)}$ increases in X if $\gamma > \chi(1+\epsilon) - 1$. This condition is guaranteed by $\gamma > \chi + 1$ and $\chi\epsilon < 1$. Moreover, by Proposition 7, the stated conditions ensure that $X(\theta)$ is an increasing function. This proves that $m(\theta)$ increases in θ, or misoptimization is higher in the higher-productivity, higher-output state.

We next consider the monotonicity of attention. The entropy of a Gaussian random variable with variance σ^2 is proportional, up to scaling and constants, to $\log(\sigma^2)$. We therefore derive, up to scaling and constants,

$$a(\theta) = (-1 - \gamma + \chi(1+\epsilon))\log X(\theta) - (1+\epsilon)\mathbb{E}_{\theta_i}[\log\theta_i \mid \theta] \tag{469}$$

This is monotone decreasing in X if $\gamma > \chi(1+\epsilon) - 1$, as desired. It is monotone decreasing in θ if $\mathbb{E}_{\theta_i}[\log\theta_i \mid \theta]$ increases θ. This is true because $y \mapsto \log y$ is an increasing function and $\theta' > \theta \to G(\theta') \succsim_{FOSD} G(\theta)$. Thus $a(\theta)$ is a decreasing function of θ, and attention is lower in the higher-productivity, higher-output states. □

B.1.4 Proof of Proposition 9

Proof. We first derive output in the fully attentive $\lambda = 0$ limit, which we define by some mapping $X_0 : \Theta \to \mathbb{R}$. Recall the fixed-point equation for output from

Proposition 7:

$$X(\theta) = a_0 X(\theta)^{1-\chi\epsilon}\theta^{\epsilon} - a_1 X(\theta)^{\gamma-\chi}\theta \tag{470}$$

When $\lambda = 0$, we have that $a_1 = 0$. Thus,

$$X_0(\theta) = a_0 X_0(\theta)^{1-\chi\epsilon}\theta^{\epsilon} \tag{471}$$

Or simply:

$$X_0(\theta) = a_0^{\frac{1}{\chi\epsilon}}\theta^{\frac{1}{\chi}} \tag{472}$$

We now define the proportional wedge between equilibrium output and output without the attention friction as:

$$W(\theta;\lambda) := \frac{X(\theta)}{X_0(\theta)} = \frac{X(\theta)}{a_0^{\frac{1}{\chi\epsilon}}\theta^{\frac{1}{\chi}}} \tag{473}$$

Via this definition, we re-write output in the form claimed in the Proposition.

$$\log X(\log\theta) = X_0 + \chi^{-1}\log\theta + \log W(\log\theta) \tag{474}$$

where $X_0 := \frac{1}{\chi\epsilon}\log a_0$.

We next prove that the wedge is positive. To prove this and other properties, we write a fixed-point equation for $W(\theta)$. Combining the definition of the wedge with Equation 470, we obtain

$$W(\theta) = W(\theta)^{1-\chi\epsilon} - a_1(\lambda) a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}} W(\theta)^{\gamma-\chi}\theta^{\frac{\gamma-1}{\chi}} \tag{475}$$

Based on identical arguments to those in the proof of Proposition 7, the wedge is positive and unique under the exact same conditions that $X(\theta)$ is positive and unique: $\chi\epsilon < 1$ and $\gamma > \chi + 1$. Moreover, $W(\theta)$ crosses the 45 degree line from above. To show that $W(\theta) \leq 1$, it then suffices to show that the right-hand-side of the fixed point equation is less than unity when evaluated at $W(\theta) = 1$. As $a_1, a_0 > 0$, this is immediate. Thus $\log W(\theta) \leq 0$, as claimed. Moreover, given that $\frac{\partial a_1}{\partial\lambda} > 1$, it is immediate to show $\frac{\partial W}{\partial\lambda} < 0$.

Toward the final claim, we show that $\log W(\theta)$ is monotone decreasing in θ. First, we implicitly differentiate the fixed point condition:

$$\begin{aligned}\frac{\mathrm{d}W}{\mathrm{d}\theta} &= \left[(1-\chi\epsilon)W(\theta)^{-\chi\epsilon} - a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}(\gamma-\chi)W(\theta)^{\gamma-\chi-1}\theta^{\frac{\gamma-1}{\chi}}\right]\frac{\mathrm{d}W}{\mathrm{d}\theta} \\ &\quad - a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}\frac{\gamma-1}{\chi}W(\theta)^{\gamma-\chi}\theta^{\frac{\gamma-\chi-1}{\chi}}\end{aligned} \tag{476}$$

or:

$$\frac{\mathrm{d}W}{\mathrm{d}\theta} = \frac{-a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}\frac{\gamma-1}{\chi}W(\theta)^{\gamma-\chi}\theta^{\frac{\gamma-\chi-1}{\chi}}}{1-\left[(1-\chi\epsilon)W(\theta)^{-\chi\epsilon} - a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}(\gamma-\chi)W(\theta)^{\gamma-\chi-1}\theta^{\frac{\gamma-1}{\chi}}\right]} \tag{477}$$

which we can rewrite as, after multiplying by $\frac{\theta}{W}$, as

$$\frac{\mathrm{d}\log W}{\mathrm{d}\log\theta} = \frac{-a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}\frac{\gamma-1}{\chi}W(\theta)^{\gamma-\chi}\theta^{\frac{\gamma-1}{\chi}}}{W(\theta)-\left[(1-\chi\epsilon)W(\theta)^{1-\chi\epsilon} - a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}(\gamma-\chi)W(\theta)^{\gamma-\chi}\theta^{\frac{\gamma-1}{\chi}}\right]} \tag{478}$$

By positivity of a_1 and a_0 and the assumption that $\gamma > \chi + 1$, the numerator of this expression is negative. To show that the wedge is monotone decreasing, we need to show that the denominator is positive. To this end, we see that:

$$\begin{aligned}W(\theta) &= W(\theta)^{1-\chi\epsilon} - a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}W(\theta)^{\gamma-\chi}\theta^{\frac{\gamma-1}{\chi}} \\ &\geq \underbrace{(1-\chi\epsilon)}_{\chi\epsilon<1}W(\theta)^{1-\chi\epsilon} - a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}}\underbrace{(\gamma-\chi)}_{\gamma>\chi+1}W(\theta)^{\gamma-\chi}\theta^{\frac{\gamma-1}{\chi}}\end{aligned} \tag{479}$$

This completes the proof. □

B.1.5 Proof of Corollary 5

Proof. The labor demand of any given firm i is given by $L_i = \frac{x_i}{\theta_i}$. Total labor demand in the economy is then given by:

$$L = \int_{[0,1]} L_i \,\mathrm{d}i \tag{480}$$

Using Proposition 6, the definition of $x^*(z_i)$ in the proof of Proposition 7, and the equilibrium law of motion $X(\theta)$, we write the production of each firm as

$$x_i = x^*(z_i) + \tilde{v}_i = v_x X(\theta)^{1-\chi\epsilon}\theta_i^{\epsilon} + \tilde{v}_i \tag{481}$$

where $\tilde{v}_i$ is the misoptimization scaled by its endogenous standard deviation. Plugging this into the expression for L, we derive

$$L = v_x X(\theta)^{1-\chi\epsilon} \int_0^1 \theta_i^{\epsilon-1} \, \mathrm{d}i \tag{482}$$

Simplifying and applying a law of large numbers, we write this as

$$L_t = L(\theta) = v_x X(\theta)^{1-\chi\epsilon} \theta^{\epsilon-1} \tag{483}$$

where we define, as in the main text, $\theta := \left(\mathbb{E}_{\theta_i}[\theta_i^{\epsilon-1} \mid \theta]\right)^{\frac{1}{\epsilon-1}}$.

Combining the definition $\log A(\theta) = \log X(\theta) - \log L(\theta)$ with Equation 483, we derive

$$\log A(\theta) = -\log v_x + \chi\epsilon \log X(\theta) - (\epsilon - 1)\log\theta \tag{484}$$

Using our representation of aggregate output from Proposition 9, we obtain:

$$\log A(\theta) = (\chi\epsilon X_0 - \log v_x) + \log\theta + \chi\epsilon \log W(\theta) \tag{485}$$

where $W(\cdot)$ inherits all of the properties proved in Proposition 9. We finally observe that $X_0 = \frac{\log v_x}{\chi\epsilon}$, as defined in the proof of Proposition 9, so $\chi\epsilon X_0 - \log v_x = 0$. This completes the proof. □

B.1.6 Proof of Corollary 6

Proof. From Proposition 9, the expression for output is

$$\log X(\log\theta) = X_0 + \chi^{-1}\log\theta + \log W(\log\theta) \tag{486}$$

Moreover, if $\lambda = 0$, then $\log W(\log\theta) = 0$.

First, consider the response of output to a small shock ν_t starting from θ_{t-1}. We have immediately that:

$$\left.\frac{\partial \log X(\log\theta)}{\partial \log\theta}\right|_{\theta=\theta_{t-1}} = \chi^{-1} + \left.\frac{\partial \log W(\log\theta)}{\partial\log\theta}\right|_{\theta=\theta_{t-1}} \tag{487}$$

When $\lambda = 0$, then $\log W \equiv 0$ according to Proposition 9. Thus $\frac{\partial \log X(\log\theta)}{\partial\log\theta} \equiv \chi^{-1}$ in all states, if inattention is removed from the model. If instead $\lambda > 0$ then $\frac{\partial \log W(\log\theta)}{\partial\log\theta}|_\theta$

is a non-linear function of θ which satisfies

$$\frac{\mathrm{d}\log W}{\mathrm{d}\log\theta} = \frac{-a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}} \frac{\gamma-1}{\chi} W(\theta)^{\gamma-\chi} \theta^{\frac{\gamma-1}{\chi}}}{W(\theta) - \left[(1-\chi\epsilon)W(\theta)^{1-\chi\epsilon} - a_1 a_0^{\frac{\gamma-\chi-1}{\chi\epsilon}} (\gamma-\chi) W(\theta)^{\gamma-\chi} \theta^{\frac{\gamma-1}{\chi}}\right]} \tag{488}$$

Toward showing endogenous stochastic volatility, we approximate $\log X$ to the first order in the shock ν_t:

$$\log X(\nu_t;\theta_{t-1}) \approx \log X(\theta_{t-1}) + \left.\frac{\partial \log X(\log\theta)}{\partial\log\theta}\right|_{\theta=\theta_{t-1}} \nu_t \tag{489}$$

By the same logic above, this is state independent if $\lambda = 0$ and $\log W \equiv 0$. Taking the variance of this expression conditional on θ_{t-1} yields:

$$\mathrm{Var}[\log X_t \mid \theta_{t-1}] = \left(\chi^{-1} + \left.\frac{\partial \log W(\log\theta)}{\partial\log\theta}\right|_{\theta=\theta_{t-1}}\right)^2 \sigma_\theta^2 \tag{490}$$

Taking a second-order approximation in ν_t and differentiating yields the asymmetric shock propagation equation:

$$\left.\frac{\partial \log X(\log\theta)}{\partial\log\theta}\right|_{\theta=\theta_{t-1}} = \chi^{-1} + \frac{\partial \log W(\log\theta)}{\partial\log\theta}\Big|_{\theta=\theta_{t-1}} + \left(\left.\frac{\partial^2 \log W(\log\theta)}{\partial\log\theta^2}\right|_{\theta=\theta_{t-1}}\right)\nu_t \tag{491}$$

Again, this is state independent if $\lambda = 0$ and $\log W \equiv 0$. □

B.1.7 Additional Calculations

Quadratic Approximation of Risk-Adjusted Profits

Using the expressions for dollar profits and marginal utility in Equation 67, we can write firms' risk-adjusted profits as the following:

$$\Pi(x, z_i) := X^{-\gamma}\left(x^{1-\frac{1}{\epsilon}} X^{\frac{1}{\epsilon}} - x\frac{w}{\theta_i}\right) \tag{492}$$

where, as throughout, we define the decision state vector $z_i = (\theta_i, X, w)$. The optimal action in the absence of stochastic choice solves the first-order condition

$$\left(1 - \frac{1}{\epsilon}\right) x^{*^{-\frac{1}{\epsilon}}} X^{\frac{1}{\epsilon}} = \frac{w}{\theta_i} \tag{493}$$

which can be re-arranged to define

$$x^*(z_i) = \left(1 - \frac{1}{\epsilon}\right)^{\epsilon} X \left(\frac{w}{\theta_i}\right)^{-\epsilon} \tag{494}$$

We now approximate the firm's profit function to second order in x around $x^*(z_i)$:

$$\begin{aligned}\Pi(x, z_i) &= \Pi(x^*(z_i), z_i) + \Pi_x(z_i)(x - x^*(z_i)) + \frac{1}{2}\Pi_{xx}(z_i)(x - x^*(z_i))^2 + O^3(x) \\ &=: \tilde{\Pi}(x, z_i) + O^3(x)\end{aligned} \tag{495}$$

where $\Pi_x(z_i) := \Pi_x(x, z_i)|_{x=x^*(z_i)}$ and $\Pi_{xx}(z_i) := \Pi_{xx}(x, z_i)|_{x=x^*(z_i)}$. By the envelope theorem, $\Pi_x(z_i) = 0$. Thus, our approximation reduces to the quadratic utility function in the Linear-Quadratic equilibrium:

$$\tilde{\Pi}(x, z_i) = \Pi(x^*(z_i), z_i) + \frac{1}{2}\Pi_{xx}(z_i)(x - x^*(z_i))^2 \tag{496}$$

It remains to characterize the intercept and curvature. We first derive the intercept:

$$\begin{aligned}\Pi(x^*(z_i), z_i) &= X^{-\gamma} \left(X \left(\frac{w}{\theta_i}\right)^{1-\epsilon}\right) \left(\left[1 - \frac{1}{\epsilon}\right]^{\epsilon\left(1-\frac{1}{\epsilon}\right)} - \left[1 - \frac{1}{\epsilon}\right]^{\epsilon}\right) \\ &= X^{-\gamma} \left(X \left(\frac{w}{\theta_i}\right)^{1-\epsilon}\right) \epsilon^{-\epsilon} (\epsilon - 1)^{\epsilon - 1}\end{aligned} \tag{497}$$

We now characterize the curvature, which is the product of marginal utility with the curvature of the dollar-profit function:

$$\Pi_{xx}(z_i) = X^{-\gamma} \cdot \pi_{xx}(z_i) \tag{498}$$

We calculate, using the form of the profit function from Equation 67, the dollar profit function's second derivative:[1]

$$\begin{aligned}\pi_{xx}(x^*(z_i), X) &= -\frac{1}{\epsilon}\left(1-\frac{1}{\epsilon}\right)(x^*(z_i))^{-1-\frac{1}{\epsilon}} X^{\frac{1}{\epsilon}} \\ &= -\frac{1}{\epsilon}\left(1-\frac{1}{\epsilon}\right)\left(1-\frac{1}{\epsilon}\right)^{-\left(1+\frac{1}{\epsilon}\right)\epsilon} X^{-1-\frac{1}{\epsilon}} X^{\frac{1}{\epsilon}} \left(\frac{w}{\theta_i}\right)^{\epsilon\left(1+\frac{1}{\epsilon}\right)} \\ &= -\epsilon^{\epsilon-1}(\epsilon-1)^{-\epsilon} X^{-1}\left(\frac{w}{\theta_i}\right)^{1+\epsilon}\end{aligned} \tag{499}$$

We substitute in the wage rule, Equation 63, to derive

$$\pi_{xx}(z_i) = -v_\pi(\epsilon, \chi, \bar{X}, \bar{w}) \cdot \theta_i^{-1-\epsilon} X^{\chi(1+\epsilon)-1} \tag{500}$$

as in Equation 73, where the constant is

$$v_\pi(\epsilon, \chi, \bar{X}, \bar{w}) := (\epsilon-1)^{-\epsilon}\epsilon^{\epsilon-1}\bar{w}^{1+\epsilon}\bar{X}^{-\chi(1+\epsilon)} > 0 \tag{501}$$

Quadratic Approximation of Final-Goods Technology

We now consider the second-order approximation of the aggregator, which is reprinted below

$$X(\{x_i\}_{i\in[0,1]}) = \left(\int_0^1 x_i^{\frac{\epsilon-1}{\epsilon}}\, \mathrm{d}i\right)^{\frac{\epsilon}{\epsilon-1}} \tag{502}$$

Technically speaking, we take a quadratic approximation of a discretized version of this aggregator, and then take consider the limit of this approximation. First, we suppose that there are $K \times K' \times K''$ discrete firms. Define the firm-level state for any firm $kk'k''$ as $\omega_{kk'k''} = (\theta_k, \lambda_{k'}, v_{k''})$ with corresponding production level $x(\omega_{kk'k''})$. Define the CES aggregator in this economy as:

$$X_{KK'K''}(\{x_{kk'k''}\}) = \left(\frac{1}{K}\sum_{k=1}^{K}\frac{1}{K'}\sum_{k'=1}^{K'}\frac{1}{K''}\sum_{k''=1}^{K''} x(\omega_{kk'k''})^{1-\frac{1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}} \tag{503}$$

[1] Because marginal costs are constant, this curvature arises purely from the curvature of the revenue function.

Second, we take a quadratic approximation of this function around the firm-level optimal production points $x_{kk'k''} = x^*(\theta_k)$:

$$X_{KK'K''} = X_K^* + \frac{1}{K}\sum_{k=1}^{K}\frac{1}{K'}\sum_{k'=1}^{K'}\frac{1}{K''}\sum_{k''=1}^{K''} D_k(x(\omega_{kk'k''}) - x^*(\theta_k))$$
$$+\frac{1}{K}\sum_{k=1}^{K}\frac{1}{K'}\sum_{k'=1}^{K'}\frac{1}{K''}\sum_{k''=1}^{K''}\frac{1}{K}\sum_{\tilde{k}=1}^{K}\frac{1}{K'}\sum_{\tilde{k}'=1}^{K'}\frac{1}{K''}\sum_{\tilde{k}''=1}^{K''}\frac{1}{2}D^2_{k\tilde{k}}(x(\omega_{kk'k''}) - x^*(\theta_k))(x(\omega_{\tilde{k}\tilde{k}'\tilde{k}''}) - x^*(\theta_{\tilde{k}})) \tag{504}$$

where:

$$X_K^* = \left(\frac{1}{K}\sum_{k=1}^{K} x^*(\theta_k)^{1-\frac{1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}} \tag{505}$$

and:

$$D_k = (X_K^*)^{\frac{1}{\epsilon}} x^*(\theta_k)^{-\frac{1}{\epsilon}} \tag{506}$$

and:

$$D^2_{k\tilde{k}} = \begin{cases} -KK'K''\frac{1}{\epsilon}x^*(\theta_k)^{-\frac{1}{\epsilon}-1}(X_K^*)^{\frac{1}{\epsilon}} + \frac{1}{\epsilon}\frac{\partial X_K^*}{\partial x_{kk'k''}}(X_K^*)^{\frac{1}{\epsilon}-1}x^*(\theta_k)^{-\frac{1}{\epsilon}} & \text{if } kk'k'' = \tilde{k}\tilde{k}'\tilde{k}'' \\ \frac{1}{\epsilon}(X_K^*)^{\frac{1}{\epsilon}-1}(x^*(\theta_k))^{-\frac{1}{\epsilon}}(x^*(\theta_{\tilde{k}}))^{-\frac{1}{\epsilon}} & \text{if } kk'k'' \neq \tilde{k}\tilde{k}'\tilde{k}''. \end{cases} \tag{507}$$

We now take limits of this approximation in the following order. We first send $K'' \to \infty$. Observe that, for fixed k, k', we have that each k'' firm by Proposition 6 has action distributed as $N(x^*(\theta_k), \sigma^2_{kk'})$. Thus, as $K'' \to \infty$, by the law of large numbers:

$$\frac{1}{K''}\sum_{k''=1}^{K''} D_k(x(\omega_{kk'k''}) - x^*(\theta_k)) \to^{a.s} 0 \tag{508}$$

Thus, the second term in the quadratic expansion above is zero almost surely in the large firm limit.

We can perform the same exercise for the third term in the quadratic expansion, which we can write as:

$$Q = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{K'}\sum_{k'=1}^{K'}\frac{1}{K}\sum_{\tilde{k}=1}^{K}\frac{1}{K'}\sum_{\tilde{k}'=1}^{K'}\left(\frac{1}{K''^2}\sum_{k''=1}^{K''}\sum_{\tilde{k}''=1}^{K''}\frac{1}{2}D^2_{k\tilde{k}}(x(\omega_{kk'k''}) - x^*(\theta_k))(x(\omega_{\tilde{k}\tilde{k}'\tilde{k}''}) - x^*(\theta_{\tilde{k}}))\right)$$
$$= \frac{1}{K}\sum_{k=1}^{K}\frac{1}{K'}\sum_{k'=1}^{K'}\frac{1}{K}\sum_{\tilde{k}=1}^{K}\frac{1}{K'}\sum_{\tilde{k}'=1}^{K'}\frac{1}{2}D^2_{k\tilde{k}}\left(\frac{1}{K''^2}\sum_{k''=1}^{K''}\sum_{\tilde{k}''=1}^{K''}(x(\omega_{kk'k''}) - x^*(\theta_k))(x(\omega_{\tilde{k}\tilde{k}'\tilde{k}''}) - x^*(\theta_{\tilde{k}}))\right) \tag{509}$$

Fix $k = \tilde{k}$, $k' = \tilde{k}'$ and consider the summation in brackets. This has two terms. First, for $\tilde{k}'' = k''$, the summand is simply $(x(\omega_{kk'k''}) - x^*(\theta_k))^2$. Second, $\tilde{k}'' \neq k''$, the summand is the product of two independent normal random variables with common distribution distribution $N(0, \sigma^2_{kk'})$. Thus, in the $K'' \to \infty$ limit we have that:

$$\frac{1}{K''^2} \sum_{k''=1}^{K''} \sum_{\tilde{k}'' \neq k''}^{K''} (x(\omega_{kk'k''}) - x^*(\theta_k))(x(\omega_{kk'\tilde{k}''}) - x^*(\theta_k)) \to^{a.s} 0 \tag{510}$$

and:

$$\frac{1}{K''} \sum_{k''=1}^{K''} (x(\omega_{kk'k''}) - x^*(\theta_k))(x(\omega_{\tilde{k}\tilde{k}'k''}) - x^*(\theta_{\tilde{k}})) \to^{a.s} \sigma^2_{kk'} \tag{511}$$

Thus, using the observation that $\lim_{K'' \to \infty} \frac{\partial X^*_K}{\partial x_{kk'k''}} = 0$, we substitute in D^2_{kk} to obtain:

$$Q = -\frac{1}{2\epsilon} \frac{1}{K} \sum_{k=1}^{K} \frac{1}{K'} \sum_{k'=1}^{K'} \frac{\sigma^2_{kk'}}{x^*(\theta_k)^{1+\frac{1}{\epsilon}} (X^*_K)^{-\frac{1}{\epsilon}}} \tag{512}$$

We now observe that $\sigma^2_{kk'} = \frac{\lambda_{k'}}{\lambda} \sigma^2_k$. Thus, taking the $K' \to \infty$ limit we have that:

$$Q \to^{a.s.} -\frac{1}{2\epsilon} \frac{1}{K} \sum_{k=1}^{K} \frac{\lambda \sigma^2_k}{x^*(\theta_k)^{1+\frac{1}{\epsilon}} (X^*_K)^{-\frac{1}{\epsilon}}} \tag{513}$$

Now taking the limit as $K \to \infty$, we can express this as:

$$Q \to^{a.s} -\frac{1}{2\epsilon} \mathbb{E}\left[\frac{\lambda \sigma^2_k}{x^*(\theta_k)^{1+\frac{1}{\epsilon}} (X^*_K)^{-\frac{1}{\epsilon}}}\right] \tag{514}$$

Moreover, in the same limit, by applying the law of large numbers and the continuous mapping theorem we have that:

$$X^*_K \to^{a.s.} \left(\mathbb{E}[x^*(\theta_k)^{1-\frac{1}{\epsilon}}]\right)^{\frac{\epsilon}{\epsilon-1}} \tag{515}$$

Combining all of the above, we have shown that, in the limit, almost surely:

$$X \approx \left(\mathbb{E}[(x^*(\theta_k))^{1-\frac{1}{\epsilon}}]\right)^{\frac{\epsilon}{\epsilon-1}} - \frac{1}{2\epsilon} \mathbb{E}\left[\frac{\lambda \sigma^2_k}{x^*(\theta_k)^{1+\frac{1}{\epsilon}} (X^*_K)^{-\frac{1}{\epsilon}}}\right] \tag{516}$$

Which we denote by the (somewhat imprecise, but standard) integral form over agents:

$$X = \left(\int_0^1 x^*(z_i)^{1-\frac{1}{\epsilon}}\,\mathrm{d}i\right)^{\frac{\epsilon}{\epsilon-1}} - \frac{1}{2\epsilon}\int_0^1 \frac{(x_i - x^*(z_i))^2}{(X^*)^{-\frac{1}{\epsilon}}(x^*(z_i))^{1+\frac{1}{\epsilon}}}\,\mathrm{d}i \tag{517}$$

Mapping Misoptimization Dispersion to the Model

Here, we explicitly calculate the within-model analogue to Misoptimization Dispersion. We show that monotone misoptimization dispersion implies our within-model measure of misoptimization is monotone and therefore confirms the misoptimization cycles prediction.

Recall from Definition 17 the definition of aggregate misoptimization,

$$m(\theta) := \mathbb{E}_{\theta_i,\lambda_i,v_i}\left[(x_i - x^*(z_i(\theta)))^2 \mid \theta\right] \tag{518}$$

which, as derived in the proof of Proposition 8, had expression

$$m(\theta) = \frac{\lambda}{v_\Pi(\epsilon,\chi,\bar{w},\bar{X})}\cdot X(\theta)^{1+\gamma-\chi(1+\epsilon)}\cdot \mathbb{E}_{\theta_i}[\theta_i^{1+\epsilon} \mid \theta] \tag{519}$$

Misoptimization Dispersion is the optimal-sales-weighted population average of the *normalized* mean-squared error of actions. Let us define this model object as

$$\tilde{m}(\theta) = \mathbb{E}_{\theta_i,\lambda_i,v_i}\left[\hat{s}^*(\theta_i)\left(\frac{(x_i - x^*(z_i(\theta)))}{x^*(z_i)}\right)^2 \mid \theta\right] \tag{520}$$

where $s^*(\theta_i)$ are sales weights evaluated at the optimal production levels. We can use the model's structure to simplify these weights:

$$\hat{s}^*(\theta_i) := \frac{q^*(z_i)x^*(z_i)}{\mathbb{E}_{\theta_i}[q^*(z_i)x^*(z_i)]} = \frac{X^{\frac{1}{\epsilon}}(v_x X^{1-\chi\epsilon}\theta_i^{\epsilon})^{1-\frac{1}{\epsilon}}}{\mathbb{E}_{\theta_i}[X^{\frac{1}{\epsilon}}(v_x X^{1-\chi\epsilon}\theta_i^{\epsilon})^{1-\frac{1}{\epsilon}}]} = \frac{\theta_i^{\epsilon-1}}{\theta^{\epsilon-1}} \tag{521}$$

where, as throughout, $\theta = \left(\mathbb{E}_{\theta_i}[\theta_i^{\epsilon-1}]\right)^{\frac{1}{\epsilon-1}}$. We can therefore write the expected variance of normalized misoptimizations, conditioning on a specific firm, as

$$\tilde{m}(z_i,\lambda_i,\theta) := \mathbb{E}_{v_i}\left[s^*(\theta_i)\left(\frac{x_i - x^*(z_i(\theta))}{x^*(z_i)}\right)^2 \mid z_i,\lambda_i,\theta\right] = \frac{\lambda_i X(\theta)^{\gamma+\chi(\epsilon-1)-1}\theta^{1-\epsilon}}{v_\Pi v_x^2} \tag{522}$$

where $v_\Pi, v_X > 0$ are defined in Equation 450. It is trivial to integrate over (θ_i,λ_i) to derive

$$\tilde{m}(\theta) = \frac{\lambda X(\theta)^{\gamma+\chi(\epsilon-1)-1}\theta^{1-\epsilon}}{v_\Pi v_x^2} \tag{523}$$

We can relate this to $m(\theta)$ by writing

$$\frac{m(\theta)}{\tilde{m}(\theta)} = v_x^2 \theta^{\epsilon-1} \mathbb{E}_{\theta_i}[\theta_i^{1+\epsilon} \mid \theta] X^{2(1-\chi\epsilon)} \tag{524}$$

See that, given $\epsilon > 1$ and $\chi\epsilon < 1$, this is an increasing function of both θ and X. Therefore, if $\tilde{m}(\theta)$ is monotone increasing in θ in an equilibrium with monotone $X(\theta)$, then $m(\theta)$ is also monotone increasing in θ.

B.2 Extended Model

In this appendix, we formally develop the extension of the baseline model from Section 2.3 to include multiple inputs and market clearing wages. In the process, we will provide more direct model micro-foundations for the wage rule and the stock return regression analysis in Section 2.5.3.

B.2.1 Set-up

Time is discrete, and indexed by $t \in \mathbb{N}$. There are three kinds of firms: perfectly competitive materials firms who use labor to produce materials; intermediate goods producers who differ in their productivity and who use labor and materials to produce a monopolistic variety indexed by $i \in [0, 1]$; and final goods firms who produce consumption goods as a constant elasticity of substitution aggregate of intermediate goods. There are two types of households: capitalists who own the firms in the economy, do not work and have constant relative risk aversion (CRRA) preferences over consumption; workers who supply labor, are hand-to-mouth (consuming all of their labor income in each period), and have Greenwood, Hercowitz, and Huffman (1988) (GHH) preferences over consumption and labor. Finally, as in our baseline model, the stochastic choice friction is embedded in the production of intermediate goods: intermediate goods producers perfectly cost-minimize but find it hard to produce the optimal amount.

Firms

Materials are produced by perfectly competitive firms with linear production technology in labor so that aggregate production of materials M_t is given by:

$$M_t = \theta_t^M L_t^M \tag{525}$$

where θ_t^M is the productivity of the materials sector and L_t^M is its labor input.

Intermediate goods producers of variety i are the monopoly producers of that variety. They have firm-specific productivity θ_{it} and use materials m_{it} and labor L_{it} to produce output x_{it} with Cobb-Douglas production technology:

$$x_{it} = \theta_{it} L_{it}^{\alpha} m_{it}^{1-\alpha} \tag{526}$$

where $\alpha \in (0, 1)$. To the extent that other intermediate goods (e.g. capital) exist and are combined in a CRS Cobb-Douglas production function with labor, this is fully general.

The stochastic process of productivity is exactly as described in Section 2.3.2. There is an aggregate productivity state $\theta_t \in \Theta$, which follows a first-order Markov process with transition density given by $h(\theta_t \mid \theta_{t-1})$. The cross-sectional productivity distribution is given in state Θ by the mapping $G : \Theta \to \Delta(\Theta)$, where we denote the productivity distribution in any state θ_t by $G_t = G(\theta_t)$ with corresponding density g_t. We assume that the total order on θ_t ranks distributions G_t by first-order stochastic dominance, or $\theta \geq \theta'$ implies $G(\theta) \succsim_{FOSD} G(\theta')$. Finally, materials productivity θ_t^M is determined as an increasing function of the overall productivity state θ_t.

Intermediate goods producers perfectly cost-minimize facing wages w_t and intermediate goods prices p_t^M. That is, for given production level x_{it}, they always choose the cost-minimizing input bundle. We define the firm-level *decision state* $z_{it} = (\theta_{it}, X_t, w_t, p_t^M) \in \mathcal{Z}$ as the concatenation of all decision-relevant variables that the firm takes as given; unlike in the baseline model, this definition includes the materials price. All firms believe that the vector z_{it} follows a first-order Markov process with transition densities described by $f : \mathcal{Z} \to \Delta(\mathcal{Z})$, with $f(z_{it}|z_{i,t-1})$ being the density of z_{it} conditional on last period's state being $z_{i,t-1}$. At time t, each firm i knows the sequence of previous $\{z_{is}\}_{s<t}$ but not the contemporaneous value z_{it}.

Given this firms have risk-adjusted profits given by $\Pi(x_{it}, z_{it})$. They then choose stochastic choice rules to maximize expected profits net of control costs, as captured by the following program which is identical to Equation 68 in the main model, with a different definition of the decision state and profits function:

$$\max_{p \in \mathcal{P}} \int_{\mathcal{Z}} \int_{\mathcal{X}} \Pi(x, z_{it})\, p(x \mid z_{it})\, \mathrm{d}x \ f(z_{it} \mid z_{i,t-1})\, \mathrm{d}z - c\,(p, \lambda_i, z_{i,t-1}, f) \tag{527}$$

Intermediate goods firms generate profits in units of consumption goods given by π_{it}. The firms store these consumption goods and pay them out as dividends d_{it} to their owners in the following period:

$$d_{it+1} = \pi_{it} \tag{528}$$

A unit supply of stock in the firm, which confers the right to the dividend stream, is available at price P_{it}.

The output of intermediate goods firms is combined to produce consumption goods with a CES production technology. Thus, if the intermediate producers produce $\{x_{it}\}_{i\in[0,1]}$, then the aggregate supply of consumption goods is:

$$X_t = X(\{x_{it}\}_{i\in[0,1]}) = \left(\int_{[0,1]} x_{it}^{\frac{\epsilon-1}{\epsilon}}\right)^{\frac{\epsilon}{\epsilon-1}} \tag{529}$$

Households

There are two types of households: capitalists and workers. Capitalists own all firms in the economy and workers are hand-to-mouth. Capitalists have preferences over streams of consumption $\{C_{t+j}\}_{j\in\mathbb{N}}$ given by:

$$\mathcal{U}^C(\{C_{t+j}\}_{j\in\mathbb{N}}) = \mathbb{E}_t \sum_{j=0}^{\infty} \beta_C^j \frac{C_{t+j}^{1-\gamma}}{1-\gamma} \tag{530}$$

where $\beta_C \in [0,1), \gamma \geq 0$. The dynamic budget constraint of capitalists is given by:

$$C_t + A_{t+1} + \int_{[0,1]} P_{it} S_{it+1}\,\mathrm{d}i \leq \int_{[0,1]} d_{it} S_{it}\,\mathrm{d}i + (1+r_t)A_t + \int_{[0,1]} P_{it} S_{it}\,\mathrm{d}i \tag{531}$$

where S_{it} is their stock-holding in intermediate firm i at time t and A_t is their bond-holding at time t.

Workers have preferences over streams of consumption and labor $\{C^W_{t+j}, L_{t+j}\}_{j\in\mathbb{N}}$ given by:

$$\mathcal{U}^W(\{C^W_{t+j}, L_{t+j}\}_{j\in\mathbb{N}}) = \mathbb{E}_t \sum_{j=0}^{\infty} \beta_W^j U\left(C_t^W - \frac{L_{t+j}^{1+\psi}}{1+\psi}\right) \tag{532}$$

where $U' > 0, U'' < 0, \psi > 0, \beta^W \in [0,1)$. Workers are hand-to-mouth and they supply labor L_t at wage w_t, meaning that they consume:

$$C_t^W = w_t L_t \tag{533}$$

Equilibrium

An equilibrium is simply a set of all endogenous variables:

$$\{L_t^M, M_t, p_t^M, p_t^*, w_t, \{x_{it}, L_{it}, m_{it}, \pi_{it}, d_{it}, P_{it}, S_{it}\}_{i\in[0,1]}, X_t, L_t, C_t, A_t\} \tag{534}$$

such that all agents optimize as described above and markets clear given the exogenous process $\{\theta_t, \theta_t^M, \{\theta_{it}\}_{i\in[0,1]}\}_{t\in\mathbb{N}}$.

We will primarily be interested, as in the main text, in linear-quadratic equilibria where Π is approximated around its optimal level and the CES aggregator is approximated as described in Section 2.3.4.

B.2.2 Characterizing Equilibrium

We now reduce the description of equilibrium to a scalar fixed-point equation that can equivalently be formulated in terms of total production or capitalist consumption. This simplifies the analysis of the model and allows us to establish some equilibrium properties.

Production by Intermediate Goods Firms

Owing to CES aggregation, intermediate goods firms face the following iso-elastic demand curve:

$$q_{it} = X_t^{\frac{1}{\epsilon}} x_{it}^{-\frac{1}{\epsilon}} \tag{535}$$

They moreover perfectly cost-minimize. As a result, given production level x_{it} their unit input choices solve the following program:

$$\min_{L_{it}, m_{it}} w_t L_{it} + p_t^M m_{it} \quad \text{s.t.} \quad x_{it} = \theta_{it} L_{it}^{\alpha} m_{it}^{1-\alpha} \tag{536}$$

Taking the ratio of the two FOCs and rearranging:

$$m_{it} = \frac{1-\alpha}{\alpha} \frac{w_t}{p_t^M} L_{it} \tag{537}$$

Thus, given x_{it}, the optimal labor and materials choices are given by:

$$\begin{aligned} L_{it} &= \frac{1}{\theta_{it}} \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} \left(\frac{p_t^M}{w_t}\right)^{1-\alpha} x_{it} \\ m_{it} &= \frac{1}{\theta_{it}} \left(\frac{\alpha}{1-\alpha}\right)^{-\alpha} \left(\frac{p_t^M}{w_t}\right)^{-\alpha} x_{it} \end{aligned} \tag{538}$$

It follows that the cost of producing x_{it} is given by:

$$w_t L_{it} + p_t^M m_{it} = \frac{q_t}{\theta_{it}} x_{it} \tag{539}$$

where we define the unit marginal cost up to constant $c_\alpha > 0$:

$$q_t := \left[\left(\frac{\alpha}{1-\alpha}\right)^{-\alpha}\frac{1}{1-\alpha}\right] w_t^\alpha (p_t^M)^{1-\alpha} = c_\alpha w_t^\alpha (p_t^M)^{1-\alpha} \tag{540}$$

We now turn to solving the firm's stochastic choice problem. From the above, firm dollar profits are given by:

$$\pi_{it} = X_t^{\frac{1}{\epsilon}} x_{it}^{1-\frac{1}{\epsilon}} - \frac{q_t}{\theta_{it}} x_{it} \tag{541}$$

Recall that this is paid out as a dividend at period $t+1$, $d_{it+1} = \pi_{it}$. Note moreover in equilibrium by market clearing that $A_t = 0$ and $S_{it} = 1$ for all $i \in [0,1]$ and $t \in \mathbb{N}$. Thus, $C_{t+1} = \int_{[0,1]} d_{it+1}\, \mathrm{d}i = \int_{[0,1]} \pi_{it}\, \mathrm{d}i$. The firm's risk-adjusted profit is then given by:

$$\Pi_{it} = C_{t+1}^{-\gamma} \pi_{it} \tag{542}$$

where the firm takes C_{t+1} as given.

As in the main text, we define the optimal production level $x^*(\Lambda, \theta)$ which solves:

$$x^*(\Lambda, \theta_{it}) := \arg\max_{x \in \mathcal{X}} \ \Pi(x; \Lambda, \theta_{it}) \tag{543}$$

and $\bar{\Pi}(\Lambda, \theta_{it})$ as the maximized objective. Now let $\Pi_{xx}(\Lambda, \theta_{it})$ denote the second derivative of the profits function in x, evaluated at x^*:

$$\Pi_{xx}(\Lambda, \theta_{it}) := \left.\frac{\partial^2 \Pi}{\partial x^2}\right|_{x^*(\Lambda,\theta_{it});\Lambda,\theta} \tag{544}$$

The approximate objective of the intermediate goods firm is:

$$\tilde{\Pi}(x; \Lambda, \theta_{it}) := \bar{\Pi}(\Lambda, \theta_{it}) + \frac{1}{2}\Pi_{xx}(\Lambda, \theta_{it})(x - x^*(\Lambda, \theta_{it}))^2 \tag{545}$$

Under this approximate objective, it follows by a slight algebraic variation of the arguments in Proposition 6 that optimal choices follow:

$$x_{it} \sim N\left(x_{it}^*, \frac{\lambda}{|\Pi_{xx,it}|}\right) \tag{546}$$

where:

$$\begin{aligned} x_{it}^* &= \left(1 - \frac{1}{\epsilon}\right)^{\epsilon} X_t \theta_{it}^{\epsilon} q_t^{-\epsilon} \\ |\Pi_{xx,it}| &= (\epsilon - 1)^{-\epsilon} \epsilon^{\epsilon-1} C_{t+1}^{-\gamma} X_t^{-1} \theta_{it}^{-1-\epsilon} q_t^{1+\epsilon} \end{aligned} \tag{547}$$

These expressions mirror those in the main model, with $C_{t+1}^{-\gamma}$ replacing $X_t^{-\gamma}$ as the marginal utility and q_t replacing w_t as the marginal cost.

Finding Materials Prices and Wages

Materials producers maximize profits:

$$p_t^M \theta_t^M L_t^M - w_t L_t^M \tag{548}$$

Thus, in equilibrium, it follows that:

$$p_t^M = \frac{1}{\theta_t^M} w_t \tag{549}$$

The workers' labor supply condition is given by their Euler equation:

$$w_t = L_t^{\psi} \tag{550}$$

Moreover, we know that aggregate labor is equal to the sum of labor used to produce intermediates and materials:

$$L_t = \int_{[0,1]} L_{it} \mathrm{d}i + L_t^M = \int_{[0,1]} L_{it} \mathrm{d}i + \frac{1}{\theta_t^M} \int_{[0,1]} m_{it} \mathrm{d}i \tag{551}$$

where the second equality follows by market clearing for intermediates as $\int_{[0,1]} m_{it} \mathrm{d}i = M_t = \theta_t^M L_t^M$. We next substitute our expression of materials demand as a function of labor demand for intermediate goods firms to simplify the labor supply condition further:

$$\begin{aligned} L_t &= \int_{[0,1]} L_{it} di + \frac{1}{\theta_t^M} \int_{[0,1]} m_{it} di \\ &= \int_{[0,1]} L_{it} \,\mathrm{d}i + \frac{1}{\theta_t^M} \int_{[0,1]} \frac{1-\alpha}{\alpha} \frac{w_t}{p_t^M} L_{it} \,\mathrm{d}i \\ &= \left(1 + \frac{1}{\theta_t^M} \frac{1-\alpha}{\alpha} \frac{w_t}{p_t^M}\right) \int_{[0,1]} L_{it} \,\mathrm{d}i \end{aligned} \tag{552}$$

We finally substitute in the fact that the material input is priced at marginal cost to simplify further

$$
\begin{aligned}
L_t &= \left(1 + \frac{1-\alpha}{\alpha}\right) \int_{[0,1]} L_{it} \mathrm{d}i \\
&= \frac{1}{\alpha} \int_{[0,1]} L_{it} \mathrm{d}i
\end{aligned}
\tag{553}
$$

We now write this in terms of prices and output choices by substituting in, from the intermediate goods firm's cost-minimization, $L_{it} = \frac{1}{\theta_{it}} \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} \left(\frac{p_t^M}{w_t}\right)^{1-\alpha} x_{it}$. Thus:

$$
L_t = \frac{1}{\alpha} \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} \left(\frac{p_t^M}{w_t}\right)^{1-\alpha} \int_{[0,1]} \frac{x_{it}}{\theta_{it}} \mathrm{d}i \tag{554}
$$

We can then use our earlier characterization of the solution to the intermediate goods producers' stochastic choice problem to compute:

$$
\begin{aligned}
\int_{[0,1]} \frac{x_{it}}{\theta_{it}} \mathrm{d}i &= \mathbb{E}\left[\frac{x_{it}}{\theta_{it}}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{x_{it}}{\theta_{it}} | \theta_{it}\right]\right] \\
&= \mathbb{E}\left[\frac{1}{\theta_{it}} x_{it}^*\right] \\
&= \mathbb{E}\left[\frac{1}{\theta_{it}} \left(1 - \frac{1}{\epsilon}\right)^{\epsilon} X_t q_t^{-\epsilon} \theta_{it}^{\epsilon}\right] \\
&= \left(1 - \frac{1}{\epsilon}\right)^{\epsilon} X_t q_t^{-\epsilon} \theta^{\epsilon-1}
\end{aligned}
\tag{555}
$$

where we use the definition $\theta = \left(\mathbb{E}_{\theta_i}[\theta_i^{\epsilon-1} \mid \theta]\right)^{\frac{1}{\epsilon-1}}$. By combining the previous two equations, we derive that total labor demand is given by:

$$
L_t = \frac{1}{\alpha} \left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha} \left(\frac{p_t^M}{w_t}\right)^{1-\alpha} \left(1 - \frac{1}{\epsilon}\right)^{\epsilon} X_t q_t^{-\epsilon} \theta^{\epsilon-1} \tag{556}
$$

Substituting this into the workers' intratemporal Euler equation, and using Equation 540 to write the marginal cost in terms of materials prices and wages, we obtain:

$$\begin{aligned} w_t^{\frac{1}{\psi}} &= \frac{1}{\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\theta_t^M}\right)^{1-\alpha}\left(1-\frac{1}{\epsilon}\right)^{\epsilon}\theta_t^{\epsilon-1}c_\alpha^{-\epsilon}X_t\left(w_t\left(\frac{p_t^M}{w_t}\right)^{1-\alpha}\right)^{-\epsilon} \\ &= \frac{1}{\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha}(\theta_t^M)^{(\epsilon-1)(1-\alpha)}\left(1-\frac{1}{\epsilon}\right)^{\epsilon}\theta^{\epsilon-1}c_\alpha^{-\epsilon}X_t w_t^{-\epsilon} \end{aligned} \tag{557}$$

Moreover, we can write the marginal cost for the firm as

$$q_t = c_\alpha w_t^\alpha (p_t^M)^{1-\alpha} = \bar{q}_t X_t^\chi \tag{558}$$

where we define coefficient $\chi = \frac{\psi}{1+\epsilon\psi}$ and intercept

$$\bar{q}_t = \left[\left(\frac{\alpha}{1-\alpha}\right)^{-\alpha}\frac{1}{1-\alpha}\right]^{1-\alpha}\left(\frac{1}{\alpha}\left(\frac{\alpha}{1-\alpha}\right)^{1-\alpha}(\theta_t^M)^{(\epsilon-1)(1-\alpha)}\left(1-\frac{1}{\epsilon}\right)^{\epsilon}\theta^{\epsilon-1}c_\alpha^{-\epsilon}\right)^{\chi}(\theta_t^M)^{-1+\alpha} \tag{559}$$

Marginal costs, holding fixed productivity, increase in output due to upward-sloping labor supply or convex disutilty of effort. The intercept of this "cost rule" varies as a function of productivity in the intermediate-goods and materials sectors. Observe that Equation 558 is the "fully Neoclassical" analogue to our wage rule Equation 63; indeed, when $\alpha = 1$ or there is no materials factor, it reduces to a wage rule

$$w_t = \bar{w}_t X_t^\alpha \tag{560}$$

where $\bar{w}_t = \bar{q}_t|_{\alpha=1}$. This verifies our claim in the main text that the wage rule can be (essentially) micro-founded in the simple model. Indeed, in a model with materials or $\alpha < 1$, we obtain exactly the wage rule studied in the main text if $\theta_t^M = \theta_t^\beta$ where $\beta = \frac{\chi(\epsilon-1)}{(1-\chi(\epsilon-1))(1-\alpha)} > 0$, thereby canceling out the direct effect of productivity on the intercept of the wage rule.

Finding Equilibrium Output

We have characterized all endogenous objects in period t in terms of output X_t and capitalists' consumption C_{t+1}. It remains only to characterize these variables.

To this end, we approximate as we have throughout:

$$X = \left(\int_0^1 x^*(z_i)^{1-\frac{1}{\epsilon}}\,\mathrm{d}i\right)^{\frac{\epsilon}{\epsilon-1}} - \frac{1}{2\epsilon}\int_0^1 \frac{(x_i - x^*(z_i))^2}{(X^*)^{-\frac{1}{\epsilon}}(x^*(z_i))^{1+\frac{1}{\epsilon}}}\,\mathrm{d}i \tag{561}$$

where the mean and variance are taken over the realizations of θ_{it}, conditional on the aggregate state θ_t. Substituting in Equation 547, this provides one equation in terms of (X_t, C_{t+1}). Consider first the computation of X_t^*. Observe that we can write:

$$x_{it}^* = \delta_t X_t^{1-\chi\epsilon} \theta_{it}^{\epsilon} \tag{562}$$

where:

$$\delta_t = \left(1 - \frac{1}{\epsilon}\right)^{\epsilon} \left(c_\alpha \bar{w}_t \left(\frac{1}{\theta_t^M}\right)^{1-\alpha}\right)^{-\epsilon} \tag{563}$$

Substituting this into the expression for X_t^*, we obtain:

$$X_t^* = \left(\delta_t^{\frac{\epsilon-1}{\epsilon}} X_t^{\frac{\epsilon-1}{\epsilon}(1-\chi\epsilon)} \theta_t^{\epsilon-1}\right)^{\frac{\epsilon}{\epsilon-1}} = \delta_t X_t^{1-\chi\epsilon} \theta_t^{\epsilon} \tag{564}$$

Now consider the computation of the dispersion term. See that we can write:

$$\begin{aligned} \mathbb{E}\left[\frac{(x_{it} - x^*(z_i))^2}{(X^*)^{-\frac{1}{\epsilon}}(x^*(z_i))^{1+\frac{1}{\epsilon}}}\right] &= (X_t^*)^{\frac{1}{\epsilon}} \mathbb{E}\left[x^*(z_{it})^{-1-\frac{1}{\epsilon}} \mathbb{E}[(x_{it} - x^*(z_{it}))^2 | z_{it}]\right] \\ &= (X_t^*)^{\frac{1}{\epsilon}} \mathbb{E}\left[x^*(z_{it})^{-1-\frac{1}{\epsilon}} \frac{\lambda}{|\Pi_{xx,it}|}\right] \end{aligned} \tag{565}$$

To simplify this, observe that we can write:

$$\frac{1}{|\Pi_{xx,it}|} = \zeta_t^{-1} C_{t+1}^{\gamma} X_t^{1-\chi(\epsilon+1)} \theta_{it}^{1+\epsilon} \tag{566}$$

where:

$$\zeta_t = (\epsilon - 1)^{-\epsilon} \epsilon^{\epsilon-1} \left(c_\alpha \bar{w}_t \left(\frac{1}{\theta_t^M}\right)^{1-\alpha}\right)^{1+\epsilon} \tag{567}$$

So we may express:

$$\begin{aligned} \mathbb{E}\left[\frac{(x_{it} - x^*(z_i))^2}{(X^*)^{-\frac{1}{\epsilon}}(x^*(z_i))^{1+\frac{1}{\epsilon}}}\right] &= (X_t^*)^{\frac{1}{\epsilon}} \mathbb{E}\left[\delta_t^{-1-\frac{1}{\epsilon}} X_t^{-(1+\frac{1}{\epsilon})(1-\chi\epsilon)} \theta_{it}^{-1-\epsilon} \lambda \zeta_t^{-1} C_{t+1}^{\gamma} X_t^{1-\chi(\epsilon+1)} \theta_{it}^{1+\epsilon}\right] \\ &= \lambda \zeta_t^{-1} \delta_t^{-1} C_{t+1}^{\gamma} X_t^{-\chi} \end{aligned} \tag{568}$$

Putting all of the above together we have that:

$$X_t = \delta_t X_t^{1-\chi\epsilon} \theta_t^{\epsilon} - \frac{\lambda}{2\epsilon} \zeta_t^{-1} \delta_t^{-1} C_{t+1}^{\gamma} X_t^{-\chi} \theta_t \tag{569}$$

The final equation we require comes from equating capitalists' consumption with the previous period's dividends, which is implied by market clearing in the securities market and the fact that workers are hand-to-mouth. Thus:

$$C_{t+1} = \int_{[0,1]} \pi_{it} \,\mathrm{d}i \tag{570}$$

Using our running approximation:

$$\pi_{it} = \pi_{it}(x_{it}^*) + \frac{1}{2}\pi_{xx,it}(x_{it} - x_{it}^*)^2 \tag{571}$$

we obtain:

$$\begin{aligned}
C_{t+1} &= \int_{[0,1]} \pi_{it} di \\
&= \mathbb{E}\left[\mathbb{E}\left[\pi_{it}(x_{it}^*) + \frac{1}{2}\pi_{xx,it}(x_{it}^*)(x_{it} - x_{it}^*)^2 \mid x_{it}^*\right] \mid \delta_t, \theta_t\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\pi_{it}(x_{it}^*) + \frac{1}{2}\pi_{xx,it}(x_{it}^*)\frac{\lambda}{|\Pi_{xx,it}|} \mid x_{it}^*\right] \mid \delta_t, \theta_t\right] \\
&= \mathbb{E}\left[\pi_{it}(x_{it}^*) \mid \delta_t, \theta_t\right] - \frac{\lambda}{2}C_{t+1}^{\gamma} \\
&= (\epsilon - 1)^{-1}\delta_t \theta_t^{\epsilon-1} X_t^{1-\chi(\epsilon-1)} - \frac{\lambda}{2}C_{t+1}^{\gamma}
\end{aligned} \tag{572}$$

We can therefore solve for X_t as a function of C_{t+1}:

$$X_t = \left(\frac{C_{t+1} + \frac{\lambda}{2}C_{t+1}^{\gamma}}{(\epsilon-1)^{-1}\delta_t\theta_t^{\epsilon-1}}\right)^{\frac{1}{1-\chi(\epsilon-1)}} \tag{573}$$

Plugging this into the scalar fixed point equation for output then boils down the equilibrium of the model to a scalar fixed-point equation for the consumption of capitalists:

$$\begin{aligned}
\left(\frac{C_{t+1} + \frac{\lambda}{2}C_{t+1}^{\gamma}}{(\epsilon-1)^{-1}\delta_t\theta_t^{\epsilon-1}}\right)^{\frac{1}{1-\chi(\epsilon-1)}} &= \delta_t \left(\frac{C_{t+1} + \frac{\lambda}{2}C_{t+1}^{\gamma}}{(\epsilon-1)^{-1}\delta_t\theta_t^{\epsilon-1}}\right)^{\frac{1-\chi\epsilon}{1-\chi(\epsilon-1)}} \theta_t^{\epsilon} \\
&\quad - \frac{\lambda}{2\epsilon}\zeta_t^{-1}\delta_t^{-1}C_{t+1}^{\gamma}\left(\frac{C_{t+1} + \frac{\lambda}{2}C_{t+1}^{\gamma}}{(\epsilon-1)^{-1}\delta_t\theta_t^{\epsilon-1}}\right)^{\frac{-\chi}{1-\chi(\epsilon-1)}} \theta_t
\end{aligned} \tag{574}$$

The above can be summarized in the following result:

Proposition 36. *Equilibria of the model are characterized by the solutions to Equa-*

tion 574.

B.2.3 Existence, Uniqueness, and Monotonicity of Equilibrium

To establish existence of equilibrium, all we require is that the above equation has a solution. As there is always a trivial equilibrium with $C_{t+1} = 0$, we will focus on when there exists an equilibrium with positive output, when it is unique, and when it is monotone. In this more general setting, we show that so long as cognitive frictions are not too large, these properties apply.

Proposition 37. *Suppose $\chi(\epsilon - 1) < 1$. There exists $\bar{\lambda} > 0$ such that there exists a unique equilibrium with positive output whenever $\lambda < \bar{\lambda}$. Moreover, equilibrium output is monotone increasing in aggregate productivity θ.*

Proof. Following Equation 574, define:

$$\begin{aligned} g_\lambda(C) = \left(\frac{C + \frac{\lambda}{2}C^\gamma}{(\epsilon-1)^{-1}\delta\theta^{\epsilon-1}}\right)^{\frac{1}{1-\chi(\epsilon-1)}} - \Bigg[\delta\left(\frac{C + \frac{\lambda}{2}C^\gamma}{(\epsilon-1)^{-1}\delta\theta^{\epsilon-1}}\right)^{\frac{1-\chi\epsilon}{1-\chi(\epsilon-1)}}\theta^\epsilon \\ - \frac{\lambda}{2\epsilon}\zeta^{-1}\delta^{-1}C^\gamma\left(\frac{C + \frac{\lambda}{2}C^\gamma}{(\epsilon-1)^{-1}\delta\theta^{\epsilon-1}}\right)^{\frac{-\chi}{1-\chi(\epsilon-1)}}\theta\Bigg] \end{aligned} \tag{575}$$

Observe that g_λ is continuous in λ in the sup-norm. Thus, if we can show that there is a unique value of $C \in \mathbb{R}_{++}$ such that $g_0(C) = 0$ and $g_0'(C) \neq 0$, then there exists λ such that for all $\lambda < \bar{\lambda}$ there will be a unique $C' \in \mathbb{R}_{++}$ such that $g_\lambda(C') = 0$.

To prove the result, it remains to show that there is a unique value of $C \in \mathbb{R}_{++}$ such that $g_0(C) = 0$ and $g_0'(C) \neq 0$ when $\chi(\epsilon-1) < 1$. To this end, define $\tilde{C} = \frac{C}{(\epsilon-1)^{-1}\delta\theta^{\epsilon-1}}$ and see that:

$$g_0(\tilde{C}) = \tilde{C}^{\frac{1}{1-\chi(\epsilon-1)}} - b_{\tilde{C}}\tilde{C}^{\frac{1-\chi\epsilon}{1-\chi(\epsilon-1)}} \tag{576}$$

where $b_{\tilde{C}} = \delta\theta^\epsilon$. We can then compute:

$$\begin{aligned} g_0'(\tilde{C}) &= \frac{1}{1-\chi(\epsilon-1)}\tilde{C}^{\frac{\chi(\epsilon-1)}{1-\chi(\epsilon-1)}} - b_{\tilde{C}}\frac{1-\chi\epsilon}{1-\chi(\epsilon-1)}\tilde{C}^{\frac{-\chi}{1-\chi(\epsilon-1)}} \\ g_0''(\tilde{C}) &= \frac{\chi(\epsilon-1)}{(1-\chi(\epsilon-1))^2}\tilde{C}^{\frac{\chi(\epsilon-1)}{1-\chi(\epsilon-1)}-1} + b_{\tilde{C}}\frac{\chi(1-\chi\epsilon)}{(1-\chi(\epsilon-1))^2}\tilde{C}^{\frac{-\chi}{1-\chi(\epsilon-1)}-1} \end{aligned} \tag{577}$$

From which we observe the following when $\chi(\epsilon-1) < 1$:

$$\lim_{\tilde{C}\to 0} g_0'(\tilde{C}) = -\infty \qquad \lim_{\tilde{C}\to\infty} g_0'(\tilde{C}) = \infty \qquad g_0''(\tilde{C}) > 0 \quad \text{for all } \tilde{C} \in \mathbb{R}_{++} \tag{578}$$

We now establish monotonicity. If we can show that the unique value of $C \in \mathbb{R}_{++}$ such that $g_0(C) = 0$ and $g_0'(C) \neq 0$ is monotone increasing in θ, then there exists $\bar{\lambda}$ such that for all $\lambda < \bar{\lambda}$ the same will be true of the unique $C' \in \mathbb{R}_{++}$ such that $g_\lambda(C') = 0$.

To this end, see that the solution when $\lambda = 0$ is given by:

$$\ln \tilde{C} = \frac{(1 - \chi(\epsilon - 1))}{\chi\epsilon} \ln b_{\tilde{C}} \tag{579}$$

We also know that $\ln \tilde{C} = \ln C + \ln(\epsilon - 1) - \ln \delta - (\epsilon - 1) \ln \theta$. Thus, we have that:

$$\ln C = \frac{(1 - \chi(\epsilon - 1))}{\chi\epsilon} (\ln \delta + \epsilon \ln \theta) - \ln(\epsilon - 1) + \ln \delta + (\epsilon - 1) \ln \theta \tag{580}$$

As $\epsilon > 1$ and $1 > \chi(\epsilon - 1)$ by hypothesis, and δ is increasing in θ, the result follows. $\square$

B.2.4 Attention and Misoptimization Cycles in the Extended Model

Having shown that equilibrium output is monotone and increasing in the extended model, we now provide conditions under which the analogue of Proposition 8 that establishes monotonicity of attention and mistakes holds in this setting:

Proposition 38. *Assume $\chi\epsilon < 1$ and $1 > \chi(\epsilon + 1)$. There exists $\bar{\lambda}$ such that when $\lambda < \bar{\lambda}$, intermediate goods firms pay more attention and misoptimize less in lower-productivity, lower-output states.*

Proof. Recall that:

$$m(z_{it}) = \frac{\lambda_i}{|\Pi_{xx,it}|} = \lambda_i \zeta_t^{-1} C_{t+1}^{\gamma} X_t^{1-\chi(\epsilon+1)} \theta_{it}^{1+\epsilon} \tag{581}$$

Thus, the average extent of misoptimization in aggregate state θ is:

$$m(\theta) = \lambda \zeta(\theta)^{-1} C(\theta)^{\gamma} X(\theta)^{1-\chi(\epsilon+1)} \mathbb{E}[\theta_{it}^{1+\epsilon} \mid \theta] \tag{582}$$

See that $1 > \chi(\epsilon + 1)$ implies $1 > \chi(\epsilon - 1)$. As we assumed $\chi\epsilon < 1$, Proposition 37 implies that C and X are both increasing in θ. By the assumed FOSD ordering on θ, we have that $\mathbb{E}[\theta_{it}^{1+\epsilon} \mid \theta]$ is monotone increasing in θ. We moreover have that $\xi \propto \delta^{-\frac{1+\epsilon}{\epsilon}}$. Thus, as δ is increasing in θ, we have that ξ^{-1} is increasing in θ. This establishes that $m(\theta)$ is increasing in θ, and therefore that intermediate goods firms misoptimize less in lower productivity and lower output states. By the same arguments as in Proposition 8, it is immediate that the opposite pattern holds for attention. □

B.2.5 Macroeconomic Dynamics in the Extended Model

We can moreover derive an analogous representation of the impact of inattention on macroeconomic dynamics through an attention wedge that depresses output relative to the fully-attentive benchmark. Formally:

Proposition 39. *Output can be written in the following way:*

$$\log X(\log\theta,\lambda) = \frac{1}{\chi}\log\tilde{\theta} + \log W(\log\theta,\lambda) \tag{583}$$

where $\tilde{\theta} = \theta\delta^{\frac{1}{\epsilon}}$ *and* $\log W(\log\theta, 0) = 0$ *for all* $\theta \in \Theta$.

Proof. The representation follows immediately by combining Equations 579 and 573. That the wedge is 0 when $\lambda = 0$ follows immediately from the same equations. □

This formula differs from Proposition 9 only in so far as θ is replaced by $\tilde{\theta}$ which captures the effect of the inclusion of other factors of production and the endogenous labor supply of agents. Note that this result does not establish any properties of the wedge in this case, as the fixed point equation is challenging to manipulate. The nature of the wedge is then a quantitative question. Similarly to the main text, a concave attention wedge implies higher shock responsiveness in low states, greater responsiveness to negative than positive shocks, and volatility of output that is greater in low states.

B.2.6 Micro-foundation and Interpretation of the Stock Return Regressions

In Section 2.5.3, we showed that mistakes of the same size by firms lead to more adverse impacts on stock returns when the aggregate stock market return is low. We interpreted this as direct evidence in favor of our mechanism that risk-pricing is a key determinant of attention cycles. The simple model of Section 2.3 is too stylized to formally map to this regression. However, in the extended model developed in this section, we can derive exactly the regression we run from the theory and show how the estimated regression coefficients map to the risk-pricing channel in the theory.

First, from the Euler equation of capitalists, the equilibrium price of firm i at time t, P_{it} is given by:

$$u'(C_t)P_{it} = \mathbb{E}_t[\beta u'(C_{t+1})(P_{it+1} + d_{it+1})] \tag{584}$$

where $d_{it+1} = \pi_{it}$. Thus we may write:

$$u'(\pi_{t-1})P_{it} = \beta u'(\pi_t)\pi_{it} + \beta u'(\pi_t)\mathbb{E}_t[P_{it+1}] \tag{585}$$

where $\pi_t = \int_{[0,1]} \pi_{it}\,\mathrm{d}i$. It follows that:

$$P_{it} = \beta\frac{u'(\pi_t)}{u'(\pi_{t-1})}\pi_{it} + \beta\frac{u'(\pi_t)}{u'(\pi_{t-1})}\mathbb{E}_t[P_{it+1}] \tag{586}$$

A production mistake $m_{it} \equiv x_{it} - x_{it}^*$ leads to profits (under our running quadratic approximation) of:

$$\pi_{it} = \pi_{it}(x_{it}^*) + \pi_{xx,it}m_{it}^2 \tag{587}$$

Thus, the firm's stock price follows:

$$P_{it} = \beta\frac{u'(\pi_t)}{u'(\pi_{t-1})}\left(\pi_{it}(x_{it}^*) + \pi_{xx,it}m_{it}^2\right) + \beta\frac{u'(\pi_t)}{u'(\pi_{t-1})}\mathbb{E}_t[P_{it+1}] \tag{588}$$

Thus:

$$\frac{\partial P_{it}}{\partial m_{it}^2} = \beta\frac{u'(\pi_t)}{u'(\pi_{t-1})}\pi_{xx,it} \tag{589}$$

and:

$$\frac{\partial P_{it}}{\partial \pi_{it}} = \beta\frac{u'(\pi_t)}{u'(\pi_{t-1})} \tag{590}$$

To simplify this further, observe by the Euler equation for trading an equally weighted portfolio of all intermediate goods firms must satisfy, where P_t is the price of this portfolio (the stock market):

$$u'(C_t)P_t = \mathbb{E}_t[\beta u'(C_{t+1})(P_{t+1} + \pi_t)] \tag{591}$$

Or:

$$\beta\frac{u'(\pi_t)}{u'(\pi_{t-1})} = \frac{P_t}{\mathbb{E}_t[P_{t+1}] + D_{t+1}} = \frac{1}{R_t} \tag{592}$$

Which is the inverse aggregate return on equity between period t and period $t+1$, R_t. We therefore have that:

Proposition 40. *The equilibrium effect of mistakes on stock returns is given by:*

$$\frac{\partial P_{it}}{\partial m_{it}^2} = -\frac{1}{R_t}|\pi_{xx,it}| \tag{593}$$

If a mistake is measured in terms of its impact in profit units, then one obtains the *simpler:*

$$\frac{\partial P_{it}}{\partial \pi_{it}} = -\frac{1}{R_t} \tag{594}$$

Proof. Given in the text above. □

Of course, it is trivial to reformulate the above comparative statics in terms of firm level returns as P_{it-1} is invariant to innovations in m_{it}.

When equity returns are high, mistakes should (all else equal) have a lower price impact. Mapping this slightly more formally to our exact regression analysis: when we instrument for profits with mistakes, we should obtain a negative and significant coefficient on the interaction between profits and the aggregate stock market return. This is exactly what we find. The OLS regressions of returns on mistakes retain a similar structure but are intermediated by the curvature of dollar profits across firms. These regression models therefore provide a less sharp test of the risk-pricing channel, although empirically they produce entirely consistent results.

B.3 Measuring Productivity and Misoptimization

This appendix describes in greater detail our data construction and empirical methodologies for our firm-level analysis of production and misoptimization. It serves in particular as a companion for Section 2.5.1 of the paper.

B.3.1 Sample Selection and Data Construction

Our main dataset is Compustat Annual Fundamentals, which compiles detailed information from public firms' financial statements. Table B.10 summarizes the main sample variables that we use and their definitions. Production, in value terms, is defined as reported sales. Employment in Compustat is reported as the number of employees. To calculate a wage bill, we multiply this by the average industry wage calculated from the Census Bureau's County Business Patterns dataset in the same year, as the sector's total national wage bill divided by the number of employees. From 1998 onward, we use the 2- or 3-digit NAICS classification that is consistent with our main analysis. Prior to 1997, and the introduction of NAICS codes in the CBP data, we use 2-digit SIC industries. For materials expenditure, we measure the sum of reported variable costs (`cogs`) and sales and administrative expense (`xsga`) net of depreciation (`dp`) and the aforementioned wage bill. To measure the capital stock, we use a perpetual inventory method as in Ottonello and Winberry (2020) starting with the first reported observation of gross value of plant, property, and equipment and adding net investment or the differences in net value of plant, property, and equipment.[2]

We restrict the sample to firms based in the United States, reporting statistics in US Dollars, and present in the "Industrial" dataset. Within this sample, we apply the following additional filters:

1. Sales, material expenditures, and capital stock are strictly positive;
2. Employees exceed 10;
3. 2-digit NAICS is not 52 (Finance and Insurance) or 22 (Utilities);
4. Acquisitions as a proportion of assets (`aqc` over `at`) does not exceed 0.05.

The first two ensure that all companies meaningfully report all variables of interest for our production function estimation; the second applies a stricter cut-off to eliminate firms that are very small, and lead to outlier estimates of productivity and choices. The third filter eliminates firms in two industries that, respectively, may have highly non-standard production technology and non-standard market structure. The fourth is a simple screening device for large acquisitions which may spuriously show up as large innovations in firm choices and/or productivity. We finally restrict attention to firms operating on a fiscal calendar that ends in December, for more straightforward calculations of aggregate time trends.

We categorize the data into 44 sectors. These are defined at the 2-digit NAICS level, but for the Manufacturing (31-33) and Information (51) sectors, which we classify at the 3-digit level to achieve better balance of sector size. Table B.13 lists the sectors along with summary statistics for their relative size, in terms of sales and employment, in cross-sections corresponding to 1990 and 2010, in the full (not selected) Compustat sample. Overall, the full dataset covers between 15-20% of US employment and 60-80% of US output, modulo the clarification that not all Compustat sales necessarily occur in the United States.

B.3.2 Production Function and Productivity Estimation

Our primary method for estimating production functions, and thereby recovering total factor productivity, is a cost share approach. In brief, we use cost shares for materials and labor to back out production elasticities, and treat the elasticity of capital as the

[2]Note that, because of our later usage of fixed effects and lack of direct calculations using capital "expenditures" evaluated at an imputed rental rate, it is inessential to deflate the value of the capital stock.

implied "residual" given an assumed mark-up $\mu > 1$ (in our baseline, $\mu = 4/3$) and constant returns to scale. We validate, in subsection B.3.3 of this Appendix and in particular Lemma 11, that this method is consistent in sample up to an essentially negligible correction term, due to the underlying logic that input choices are "right on average" even in the presence of mistakes. The exact procedure is the following:

1. For all firms in industry j, calculate the estimated materials and labor shares:

$$\begin{aligned}\text{Share}_{M,j'} &= \frac{\sum_{i:j(i)=j'} \sum_t \text{MaterialExpenditure}_{it}}{\sum_{i:j(i)=j'} \sum_t \text{Sales}_{it}} \\ \text{Share}_{L,j'} &= \frac{\sum_{i:j(i)=j'} \sum_t \text{WageBill}_{it}}{\sum_{i:j(i)=j'} \sum_t \text{Sales}_{it}}\end{aligned} \tag{595}$$

2. If $\text{Share}_{M,j'} + \text{Share}_{L,j'} \leq \mu^{-1}$, then set

$$\begin{aligned}\alpha_{M,j'} &= \mu \cdot \text{Share}_{M,j'} \\ \alpha_{L,j'} &= \mu \cdot \text{Share}_{L,j'} \\ \alpha_{K,j'} &= 1 - \alpha_{M,j'} - \alpha_{L,j'}\end{aligned} \tag{596}$$

3. Otherwise, adjust shares to match the assumed returns to scale, or set

$$\begin{aligned}\alpha_{M,j'} &= \frac{\text{Share}_{M,j'}}{\text{Share}_{M,j'} + \text{Share}_{L,j'}} \\ \alpha_{L,j'} &= \frac{\text{Share}_{L,j'}}{\text{Share}_{M,j'} + \text{Share}_{L,j'}} \\ \alpha_{K,j'} &= 0\end{aligned} \tag{597}$$

A more flexible method for production function estimation might allow for the returns to scale and/or markups to vary across sectors. We opt not to make such a method a baseline because uniform returns to scale and markups are consistent with our subsequent empirical calibration of the model. In robustness checks, we have experimented in particular with extracting production function parameters under different assumed markups and found overall stable results for the empirical behavior of production misoptimizations.

To translate our production function estimates into productivity, we first calculate a "Sales Solow Residual" $\tilde{\theta}_{it}$ of the following form:

$$\log \tilde{\theta}_{it} = \log \text{Sales}_{it} - \frac{1}{\mu}\left(\alpha_{M,j(i)} \cdot \log \text{MatExp}_{it} - \alpha_{L,j(i)} \cdot \log \text{Empl}_{it} - \alpha_{K,j(i)} \cdot \log \text{CapStock}_{it}\right) \tag{598}$$

where we intentionally use the "literal" accounting notation to highlight the fact that not all variables are specified in economically meaningful *quantity* units. To alleviate this issue in a conservative way, we define our final estimate of TFP as the residual of the previous from industry-by-time fixed effects. This procedure, under our presumed model of industry-level variation in factor prices, identifies (log) TFP rescaled by μ^{-1}. The rescaling is immaterial for our analysis. The exact mapping from theory to data is shown in Section B.3.3.

As a robustness check, we also calculate TFP using the method of Olley and Pakes (1996), applied separately to estimate the production function of each industry.[3] The methodology of Olley and Pakes (1996) aims, in particular, to correct the bias in standard least-squares estimates that under-states the output elasticity to capital, since that input is likely pre-determined and therefore mechanically less responsive to productivity. Since the underlying timing assumptions of Olley and Pakes (1996) are not directly compatible with our specifically assumed structure, we do not prefer it as a main estimate. But we are re-assured by these estimates' "upstream" and "downstream" similarity to our baseline estimates. To the first point, Table B.14 shows the results from regressing the two TFP measures on one another in a common sample, including various levels of fixed effects. In each case, the slope is close to one and the within-R^2, or goodness of fit net of fixed effects, exceeds 0.6. To the second point, the relevant columns of Tables B.3, B.5, and B.8 demonstrate how our main aggregate and firm-level results replicate under the alternative measurement scheme, with similar quantitative and qualitative take-aways.

B.3.3 Theory to Data: Micro-foundations

In this subsection, we outline more exactly the mapping from our model to our production function estimation via cost shares and our log-linear estimating equations.

Firms are monopolists within their unique products. We assume that this demand curve has a constant elasticity of substitution form, so prices q_{it} lie on the demand curve $\log q_{it} = \gamma_i - \frac{1}{\epsilon}(\log x_{it} - \log X_t)$ for some inverse elasticity $\epsilon > 1$ and aggregate

[3] In particular, we use the implementation by Yasar, Raciborski, and Poi (2008) of the `opreg` package in Stata. We use log investment as the proxy variable and year dummies as additional controls. We throw out estimates that imply individual elasticities that are negative or greater than 1, but do not otherwise enforce any returns to scale normalization.

output X_t.[4] Finally, firms face sector-specific input prices $(q_{j(i),L,t}, q_{j(i),M,t}, q_{j(i),K,t})$ for the three inputs, respectively.

We model firm (mis-) optimization in the following way that is uniform across inputs. Conditional on any chosen level of production, firms cost minimize over their input bundle conditional on observed input prices. Let $q_{i,T,t}$ denote the associated "Total" input cost per unit of produced output, and $x^*(q_T, X)$ denote the unconditionally profit-maximizing level of production. Firms choose a production level which differs from this level by a *misoptimization* m_{it}:

$$\log x_{it} = \log x^*(q_{i,T,t}, X_t) + m_{it} \tag{599}$$

And the dynamics of the misoptimization, as given in the main text's Equation 81 and re-printed here, are described by an AR(1) process:

$$m_{it} = \rho m_{i,t-1} + \left(\sqrt{1-\rho^2}\right) u_{it} \tag{600}$$

in which innovations u_{it} are mean zero with variance σ_{it}^2.

First, we characterize the firm's optimal production level x^* as the total input price (which itself will depend on productivity) and aggregate demand:

Lemma 9 (Optimal Output Choice). *The firm's optimal output choice is*

$$\log x^*(\theta, q_T, X) = \epsilon \log\left(1 - \frac{1}{\epsilon}\right) + \epsilon\gamma_i + \log X - \epsilon \log q_T \tag{601}$$

Proof. Immediate from the first-order conditions of the program

$$x^*(\theta, q_T, X) = \arg\max_x \left\{ x \left(e^{\gamma_i} x^{-\frac{1}{\epsilon}} X^{\frac{1}{\epsilon}} - q_T \right) \right\} \tag{602}$$

□

We next characterize the optimal choice of each input, based on cost-minimization and the expression for chosen output as a function of optimal output and the misoptimization (Equation 599):

Lemma 10 (Input Choice). *For any input $Z \in \{L, M, K\}$,*

$$\log Z_{it} = \eta_i + \chi_{j(i),Z,t} + (\epsilon - 1)\log\theta_{it} + m_{it} \tag{603}$$

[4] It is straightforward, and consistent with our modeling approach, also to allow substitution within more narrowly defined industries.

where

$$\eta_i = \epsilon\gamma_i + \epsilon \log\left(1 - \frac{1}{\epsilon}\right) \tag{604}$$

and

$$\chi_{j(i),Z,t} = \log \alpha_{Z.j(i)} - \log q_{j(i),Z,t} + \log X_t + (1-\epsilon)\sum_Z \alpha_{Z,j(i)}(\log q_{j(i),Z,t} - \log \alpha_{Z,j(i)}) \tag{605}$$

Proof. In the cost minimization step, for any planned output choice Q, the firm solves

$$\min_{L_{it},M_{it},K_{it}} \sum_{z\in\{L,M,K\}} q_{j(i),Z,t}Z_{it} \qquad \text{s.t. } \theta_{it}L_{it}^{\alpha_{L,j(i)}}M_{it}^{\alpha_{M,j(i)}}K_{it}^{\alpha_{K,j(i)}} \geq Q \tag{606}$$

Standard first-order methods yield the solution, for each input,

$$\log Z_{it} = \log q_{i,T,t} + \log Q + \log \alpha_{Z,j(i)} - \log q_{j(i),Z,t} \tag{607}$$

where the price index $q_{i,T,t}$, which is also the Lagrange multiplier on the constraint, is

$$\log q_{i,T,t} = \sum_Z \alpha_{Z,j(i)}(\log q_{j(i),Z,t} - \log \alpha_{Z,j(i)}) - \log \theta_{it} \tag{608}$$

The desired expression comes from substituting in Equations 599 and 602 into the above, for Q. □

This calculation validates our log-linear regression model Equation 80. It also has the same loading on the misoptimization m_{it} and log productivity $\log\theta_{it}$ for all inputs Z. Thus, all separate inputs, as well as total production in physical units, inherit the "optimal choice plus error" structure.

We finally describe and validate our method for recovering production function parameters from a cost shares approach. The following result shows how cost shares are recovered at the firm level, if all data were observed without noise:

Lemma 11 (Production Function Estimation)**.** *For any input $Z \in \{L, M, K\}$, and firm i,*

$$\alpha_{j(i),Z} = \left(1 - \frac{1}{\epsilon}\right)^{-1} \frac{q_{j(i),Z,t}Z_{it}}{q_{it}x_{it}} \exp\left(-\frac{m_{it}}{\epsilon}\right) \tag{609}$$

Proof. This can be calculated directly using the results of Lemmas 9 and 10. We work backwards starting from the result and substitute input demand from Equation

607 and calculate.

$$
\begin{aligned}
\alpha_{j(i),Z} &= \left(1-\frac{1}{\epsilon}\right)^{-1} \frac{q_{j(i),Z,t} Z_{it}}{q_{it} x_{it}} \exp\left(-\frac{m_{it}}{\epsilon}\right) \\
&= \left(1-\frac{1}{\epsilon}\right)^{-1} \frac{\exp\left(\log q_{i,T,t} + \log x_{it} + \log \alpha_{j(i),Z,t}\right)}{\exp\left(\log q_{it} + \log x_{it}\right)} \exp\left(-\frac{m_{it}}{\epsilon}\right) \\
&= \left(1-\frac{1}{\epsilon}\right)^{-1} \exp\left(\log q_{i,T,t} - \log q_{it} + \log \alpha_{j(i),Z,t} - \frac{m_{it}}{\epsilon}\right)
\end{aligned} \tag{610}
$$

where the third line cancels out x_{it} in the fraction. We take the demand curve $\log q_{it} = \gamma_i - \frac{1}{\epsilon}(\log x_{it} - \log X_t)$ and observe that, using the expression for the unconstrained optimal output in Equation 601 and the fact that $\log x_{it} = \log x_{it}^* + m_{it}$:

$$
\begin{aligned}
\log q_{it} &= \gamma_i + \frac{1}{\epsilon}\log X_t - \frac{1}{\epsilon}\left(\epsilon \log\left(1-\frac{1}{\epsilon}\right) + \epsilon\gamma_i + \log X_t - \epsilon \log q_{i,T,t} + \log m_{it}\right) \\
&= -\log\left(1-\frac{1}{\epsilon}\right) + \log q_{i,T,t} - \frac{m_{it}}{\epsilon}
\end{aligned} \tag{611}
$$

We then substitute the above back into Equation 610 to get

$$
\begin{aligned}
\alpha_{j(i),Z} &= \left(1-\frac{1}{\epsilon}\right)^{-1} \exp\left(\log q_{i,T,t} - \left(-\log\left(1-\frac{1}{\epsilon}\right) + \log q_{i,T,t} - \frac{m_{it}}{\epsilon}\right) + \log \alpha_{j(i),Z} - \frac{m_{it}}{\epsilon}\right) \\
&= \left(1-\frac{1}{\epsilon}\right)^{-1} \exp\left((\log q_{i,T,t} - \log q_{i,T,t}) + \log\left(1-\frac{1}{\epsilon}\right) + \log \alpha_{j(i),Z} + \left(\frac{m_{it}}{\epsilon} - \frac{m_{it}}{\epsilon}\right)\right) \\
&= \alpha_{j(i),Z}
\end{aligned} \tag{612}
$$

as desired. □

In words, this result says that the ratio of expenditures on input Z to total sales, multiplied by the markup and a correction factor related to the mistake, equals the production elasticity. In principle, we could simultaneously estimate the production function and the statistical properties of mistakes to correct for the fact that the term $\exp\left(-\frac{m_{it}}{\epsilon}\right)$ is not zero on average. In practice, our mistakes are zero mean by construction and have a variance of about 0.08 in sample. Using a log-linear calculation, and our standard value of $\epsilon = 4$, this implies an average correction factor of $\exp\left(\frac{1}{2 \cdot 4^2} \cdot 0.08\right) = 1.0025$ which is essentially negligible.

We finally show, in the theory, how the calculation of Equation 598, net of fixed effects, recovers a re-scaling of TFP. In this subsection's language, that calculation is

$$\log \tilde{\theta}_{it} = \log x_{it} + \log q_{it} - \frac{1}{\mu}\left(\sum_Z \alpha_{Z,j(i)}(\log q_{j(i),Z,t} + \log Z_{it})\right) \tag{613}$$

Substituting in the demand curve $\log q_{it} = \gamma_i - \frac{1}{\epsilon}(\log x_{it} - \log X_t)$

$$\log \tilde{\theta}_{it} = \left(1 - \frac{1}{\epsilon}\right)\log x_{it} + \gamma_i + \frac{1}{\epsilon}\log X_t - \frac{1}{\mu}\left(\sum_Z \alpha_{Z,j(i)}(\log q_{j(i),Z,t} + \log Z_{it})\right) \tag{614}$$

We next observe that $\mu = 1 - \frac{1}{\epsilon}$ and that $\log \theta_{it} = \log x_{it} - \left(\sum_Z \alpha_{Z,j(i)} \log Z_{it}\right)$. Thus,

$$\log \tilde{\theta}_{it} = \frac{1}{\mu}\log \theta_{it} + \gamma_i + \frac{1}{\epsilon}\log X_t - \frac{1}{\mu}\left(\sum_Z \alpha_{Z,j(i)} \log q_{j(i),Z,t}\right) \tag{615}$$

Grouping terms into fixed effects, this is

$$\log \tilde{\theta}_{it} = \frac{1}{\mu}\log \theta_{it} + \tau_{j(i),t} + \gamma_i \tag{616}$$

where $\tau_{j(i),t} = \frac{1}{\epsilon}\log X_t - \frac{1}{\mu}\left(\sum_Z \alpha_{Z,j(i)} \log q_{j(i),Z,t}\right)$ is the industry-by-time fixed effect capturing aggregate demand and factor prices. and γ_i is the firm fixed effect from the demand curve. Thus, net of fixed effects, we recover $\frac{1}{\mu}\log \theta_{it}$.

B.4 Alternative Specifications of Stochastic Choice

In this section, we extend our basic class of cost function to allow for persistent mistakes as in the empirical analysis. In particular, we micro-found the AR(1) structure of mistakes that we uncovered in the data but abstracted from in the simple model. Further, we show how the core logic of attention cycles carries over to settings with alternative foundations for stochastic choice in terms of information acquisition of two forms: Gaussian signal extraction, and optimal signal processing with mutual information costs. Concretely, we will provide simple sufficient conditions under which an increase in the stakes of making mistakes—the curvature of intermediate goods firms' profits—leads to increased attention and smaller mistakes. In our baseline model, this corresponds to firms making smaller mistakes when market risk pricing is more severe and leads to an attention cycle.

B.4.1 Persistent Mistakes

In our empirical work we showed that firms' mistakes are persistent. The basic model we have developed, however, places no restrictions on the auto-correlation of mistakes across time within a firm. In this section, we introduce a more general class of cost functional that allows us to place restrictions on the within-firm correlation of mistakes across time and, in particular, to derive the AR(1) formulation of mistakes that we use in the empirical analysis.

We have so far considered likelihood-separable cost functions $c : \mathcal{P} \times \mathcal{Z} \to \mathbb{R}$ of the form:

$$c(p; \theta_{t-1}) = \int_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta)) \, \mathrm{d}x \, f(z|z_{t-1}) \, \mathrm{d}z \tag{617}$$

for some convex ϕ that we take to be $\phi(y) = y \log y$. To allow for persistent mistakes we now allow the cost functional to depend on the previous period's mistake v_{t-1} and today's optimal action $c : \mathcal{P} \times \mathcal{Z} \times \mathbb{R} \to \mathbb{R}$ of the form:

$$c(p; z_{t-1}, v_{t-1}) = \int_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta); v_{t-1}, x^*, x) \, \mathrm{d}x \, f(z|z_{t-1}) \, \mathrm{d}z \tag{618}$$

for ϕ convex in its first argument. In this formulation, the full non-parametric distribution of mistakes now depends on the previous period's mistake and today's optimal action.

To derive the Gaussian AR(1) formulation of mistakes, we now suppose that:

$$\phi(y; m, x^*, x) = \lambda y \log y + \omega y((x - x^*) - m)^2 \tag{619}$$

Concretely, this leads to the following cost functional:

$$\begin{aligned} c(p; z_{t-1}, v_{t-1}) = \int_{\Theta} \Bigg[& \lambda \int_{\mathcal{X}} p(x|\theta) \log p(x|\theta) \, \mathrm{d}x \\ & + \omega \int_{\mathcal{X}} ((x - x^*(\theta)) - v_{t-1})^2 p(x|\theta) \, \mathrm{d}x \Bigg] f(z|z_{t-1}) \, \mathrm{d}z \end{aligned} \tag{620}$$

which penalizes sharply peaked distributions and those where average mistakes differ greatly from the previous period's mistake. If we moreover suppose that firm risk-adjusted profits are of their quadratic form:

$$\tilde{\Pi}(x, z) := \bar{\Pi}(z) + \frac{1}{2} \Pi_{xx}(z)(x - x^*(z))^2 \tag{621}$$

and we suppose that firms solve the problem:

$$\max_{p \in \mathcal{P}} \int_{\Theta} \int_{\mathcal{X}} \Pi(x, z)\, p(x \mid \theta)\, \mathrm{d}x\, f(z|z_{t-1})\, \mathrm{d}z - c\,(p; z_{t-1}, v_{t-1}) \tag{622}$$

Solving this problem yields the AR(1) structure for mistakes, with Gaussian innovations.

Proposition 41. *The optimal stochastic choice pattern is given by:*

$$x = x^*(z) + v \tag{623}$$

where:

$$v = \rho v_{t-1} + u \tag{624}$$

and

$$\rho = \rho(z) = \frac{\omega}{\frac{1}{2}|\Pi_{xx}(z)| + \omega} \tag{625}$$

and

$$u \sim N\left(0, \frac{2\lambda}{\frac{1}{2}|\Pi_{xx}(z)| + \omega}\right) \tag{626}$$

Proof. We will denote $x^*(z)$ by γ and $\frac{1}{2}|\Pi_{xx}(z)|$ by β to simplify notation. We observe that the FOC characterizing optimal stochastic choice is given by:

$$-\beta(x-\gamma)^2 - \lambda\left[1 + \log p(x|z)\right] - \omega(x - \gamma - m)^2 + \mu(z) + \kappa(x, z) = 0 \tag{627}$$

where $\mu(z)$ is the Lagrange multiplier on the constraint that $p(x|z)$ integrates to unity and $\kappa(x, z)$ is the Lagrange multiplier on the non-negativity constraint that $p(x|z) \geq 0$. We can then observe that this has solution:

$$p(x|z) = \frac{\exp\left(-\tilde{\beta}(x - \tilde{z})^2\right)}{\int_{\mathcal{X}} \exp\left(-\tilde{\beta}(x' - \tilde{z})^2\right) \mathrm{d}x'} \tag{628}$$

where $\tilde{\beta} = \frac{\beta+\omega}{\lambda}$ and $\tilde{\gamma} = \gamma + \frac{\omega}{\beta+\omega} v_{t-1}$. It follows that:

$$x|z \sim N\left(\gamma + \frac{\omega}{\beta + \omega} v_{t-1}, \frac{2\lambda}{\beta + \omega}\right) \tag{629}$$

Putting this in more explicit terms, and substituting for γ and β, we obtain the desired representation. □

B.4.2 Transformed Gaussian Signal Extraction

In this section, we analyze attention cycles in a setting with Gaussian signal extraction and show how the basic logic of our main model carries over to this setting. For notational simplicity, we describe this alternative model under the assumption that there is a common, scalar state variable θ, which represents each firm's productivity.

Set-up

When the state of the world is θ, the previous state is θ_{-1}, agents have priors $\pi_{\theta_{-1}} \in \Delta(\Theta)$, and the equilibrium level of output is $X(\theta, \theta_{-1})$, intermediates goods firms have payoffs given by:

$$\tilde{\Pi}(x, X(\theta, \theta_{-1}), \theta) = \alpha(X(\theta, \theta_{-1}), \theta) - \beta(X(\theta, \theta_{-1}), \theta)(x - \gamma(X(\theta, \theta_{-1}), \theta))^2 \quad (630)$$

where we will write $\beta(\theta, \theta_{-1}) = \beta(X(\theta, \theta_{-1}), \theta)$ and similarly for α and γ, and as we micro-found via a second-order approximation of their true profit functions around the unconditionally optimal level of production in the main text.

Suppose moreover that agents receive a private Gaussian signal regarding their stakes-adjusted optimal action given by:

$$s_i = \frac{\beta(\theta, \theta_{-1})}{\int_\Theta \beta(\theta, \theta_{-1}) \mathrm{d}\pi(\theta|\theta_{-1})} \gamma(\theta, \theta_{-1}) + \frac{1}{\tau(\theta_{-1})} \varepsilon_i \quad (631)$$

where ε_i is a $N(0, 1)$ variable that is independent across agents and time periods; $\tau(\theta_{-1})$ is the (soon-to-be endogenized) square-root precision; and the agents' prior $\pi_{\theta_{-1}}$ is such that:

$$\frac{\beta(\theta, \theta_{-1})}{\int_\Theta \beta(\theta, \theta_{-1}) \mathrm{d}\pi(\theta|\theta_{-1})} \gamma(\theta, \theta_{-1}) \sim N\left(\mu(\theta_{-1}), \sigma^2(\theta_{-1})\right) \quad (632)$$

This model incorporates the tractability of linear signal extraction into our non-quadratic tracking problem.

Conditional on such a signal s, the best reply of any firm is equal to the conditional expectation of the stakes-adjusted optimal action:

$$\begin{aligned} x(s) &= \mathbb{E}_{\pi_{\theta_{-1}}}\left[\frac{\beta(\theta, \theta_{-1})}{\int_\Theta \beta(\theta, \theta_{-1}) \mathrm{d}\pi(\theta|\theta_{-1})} \gamma(\theta, \theta_{-1}) | s\right] \\ &= \lambda(\theta_{-1}) s + (1 - \lambda(\theta_{-1}))\mu(\theta_{-1}) \end{aligned} \quad (633)$$

where $\lambda(\theta_{-1}) = \frac{\tau^2(\theta_{-1})}{\tau^2(\theta_{-1})+\frac{1}{\sigma^2(\theta_{-1})}}$ is the appropriate signal-to-noise ratio. Thus, the cross-sectional distribution of actions is given by:

$$x|\theta,\theta_{-1} \sim N\Bigg(\lambda(\theta_{-1})\frac{\beta(\theta,\theta_{-1})}{\int_\Theta \beta(\theta,\theta_{-1})\mathrm{d}\pi(\theta|\theta_{-1})}\gamma(\theta,\theta_{-1}) \\ + (1-\lambda(\theta_{-1}))\mathbb{E}_{\pi_{\theta_{-1}}}\left[\frac{\beta(\theta,\theta_{-1})}{\int_\Theta \beta(\theta,\theta_{-1})\mathrm{d}\pi(\theta|\theta_{-1})}\gamma(\theta,\theta_{-1})\right], \frac{\lambda^2(\theta_{-1})}{\tau^2(\theta_{-1})}\Bigg) \tag{634}$$

Say we empirically estimate an equation of the form:

$$x_{it} = \gamma_i + \chi_{j(i),t} + f_t(\theta_{it},\theta_{i,t-1}) + \varepsilon_{it} \tag{635}$$

which differs from our baseline specification in controlling flexibly for observed and lagged productivity.[5] The fitted values span $\mathbb{E}[x \mid \theta, \theta_{-1}]$ and capture state-dependent anchoring toward the prior mean. The residual ε_{it} captures the noise in the firm's action coming from the noise in the signal. The fact that the average action is no longer the unconditionally optimal action is an important departure from our baseline models in Section 2.3 and Appendix B.2. In the signal extraction model, the behavior of the stochastic residual captures some, but not all, of the effects of the "cognitive friction," since it does not directly speak to anchoring.

Interpreting Monotone Misoptimization

We now discuss the interpretation of our empirical exercise of studying stochastic volatility in ε_{it}. The variance of the residual is given in this model by:

$$\mathbb{V}_t[\varepsilon_{it}] = \frac{\lambda^2(\theta_{t-1})}{\tau^2(\theta_{t-1})} = \frac{\tau^2(\theta_{t-1})}{\left(\tau^2(\theta_{t-1}) + \frac{1}{\sigma^2(\theta_{t-1})}\right)^2} \tag{636}$$

[5]In our main analysis, we consider some specifications with time-varying responsiveness to the fundamental shock. We have also considered specifications which depend more flexibly on lagged TFP and found broadly similar results to our baseline, but do not print these in the paper for brevity.

Our empirical findings are consistent with $\theta_{t-1} \mapsto \mathbb{V}_t[\varepsilon_{it}]$ being an increasing function.[6] This holds in this model exactly when:

$$\frac{\partial \tau^2(\theta_{-1})}{\partial \theta_{-1}} \left(\frac{1}{2\lambda(\theta_{-1})} - 1 \right) > \frac{\partial \frac{1}{\sigma^2(\theta_{-1})}}{\partial \theta_{-1}} \tag{637}$$

for all $\theta_{-1} \in \Theta$. Our own regression analysis in Section 2.5.5, as well as comprehensive studies of manufacturing establishments by Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018) and Kehrig (2015), suggests that there is less fundamental dispersion in higher aggregate productivity states of the world. As a result, the RHS of this expression condition must be positive. Thus, there are two potential conditions under which this variant of our model is consistent with pro-cyclical misoptimization and counter-cyclical fundamentals dispersion:

1. Firms acquire sufficiently less precise signals in higher states $\frac{\partial \tau^2(\theta_{-1})}{\partial \theta_{-1}} < 0$ and the signal to noise ratio is always such that $\lambda(\theta_{-1}) > \frac{1}{2}$
2. Firms acquire sufficiently more precise signals in lower states $\frac{\partial \tau^2(\theta_{-1})}{\partial \theta_{-1}} > 0$ and the signal to noise ratio is always $\lambda(\theta_{-1}) < \frac{1}{2}$

In the former case, "high attention" measured by high signal precision correlates with low residual variance. In the latter case, "low attention" measured by low precision correlates with low residual variance. An important difference between the signal-extraction model from our baseline, then, is that additional information is required to separately identify patterns in attention and residual variance. In both models, "misoptimization" in payoff terms and attention are perfectly correlated by construction. But residual variance is monotone in misoptimization in our baseline model, but not in the signal extraction model due to the role of anchoring.

We interpret our finding that firms discuss macroeconomic developments more during recessions as qualitatively inconsistent with model case 2, and therefore an identifying piece of evidence for case 1. Elsewhere in the literature, Coibion, Gorodnichenko, and Kumar (2018a) find that firms have higher demand for information when presented with bad macroeconomic news. Chiang (2021) also finds evidence of higher attention (interpreted as precision of signals) in downturns. Thus, our preferred interpretation of the model is one in which residual variance inherits the monotonicity of signal precision and attention.

[6] Our specific empirical specification measured the correlation between contemporaneous output and contemporaneous dispersion of ϵ_{it}. If output is monotone in the state of nature and, along with the state of nature, very persistent, the translation to $\partial \mathbb{V}_t[\varepsilon_{it}]/\partial \theta_{t-1} \geq 0$ is immediate.

Monotone Endogenous Precision

We now extend the model to include endogenous choice of signal precision and derive conditions under which firms obtain less precise signals in high productivity states in the model. To this end, suppose that after θ_{-1} is realized, but before θ is realized, that the agent can pay a cost $\tilde{\phi}(\tau^2)$ to achieve signal precision of τ^2, and where $\tilde{\phi}', \tilde{\phi}'' > 0$. Concretely, the optimal $\tau^2(\theta_{-1})$ solves:

$$\max_{\tau^2(\theta_{-1})\in\mathbb{R}_+} \mathbb{E}_{\pi_{\theta_{-1}}}\left[-\beta(\theta,\theta_{-1})\left(x^*(s,\tau^2(\theta_{-1})) - \gamma(\theta,\theta_{-1})\right)^2\right] - \tilde{\phi}(\tau^2(\theta_{-1})) \tag{638}$$

We moreover parameterize the scaling of quadratic losses by writing $\beta(\theta,\theta_{-1}) = \kappa\hat{\beta}(\theta,\theta_{-1})$ for all (θ,θ_{-1}) and some $\kappa \geq 1$. Our first goal will be to derive conditions under which the optimally chosen τ^2 in Program 638 is monotone increasing in κ. This demonstrates the natural incentives for firms to choose more precise information when the utility cost of a fixed posterior variance about the stakes-adjusted optimal action is higher.

Toward this end, we first simplify the agent's objective function. Using the distribution of optimal actions condition on τ^2 from in Equation 634, we write

$$\begin{aligned}
&\mathbb{E}_{\pi_{\theta_{-1}}}\left[-\beta(\theta,\theta_{-1})\left(x^*(s,\tau^2(\theta_{-1})) - \gamma(\theta,\theta_{-1})\right)^2\right] \\
&= \mathbb{E}_{\pi_{\theta_{-1}}}\left[\mathbb{E}\left[-\beta(\theta,\theta_{-1})\left(x^*(s,\tau^2(\theta_{-1})) - \gamma(\theta,\theta_{-1})\right)^2|\theta\right]\right] \\
&= \mathbb{E}_{\pi_{\theta_{-1}}}\left[\mathbb{E}\left[-\beta(\theta,\theta_{-1})\left(x^*(s,\tau^2(\theta_{-1})) - \bar{x}(\theta,\theta_{-1}) + \bar{x}(\theta,\theta_{-1}) - \gamma(\theta,\theta_{-1})\right)^2|\theta\right]\right] \\
&= \mathbb{E}_{\pi_{\theta_{-1}}}\left[-\beta(\theta,\theta_{-1})\mathbb{E}\left[\left(x^*(s,\tau^2(\theta_{-1})) - \bar{x}(\theta,\theta_{-1})\right)^2 + \left(\bar{x}(\theta,\theta_{-1}) - \gamma(\theta,\theta_{-1})\right)^2|\theta\right]\right] \\
&= \mathbb{E}_{\pi_{\theta_{-1}}}\left[-\beta(\theta,\theta_{-1})\left(\frac{\lambda^2(\theta_{-1})}{\tau^2(\theta_{-1})} + \left[\lambda(\theta_{-1})\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1}) + (1-\lambda(\theta_{-1}))\mu(\theta_{-1}) - \gamma(\theta,\theta_{-1})\right]^2\right)\right]
\end{aligned} \tag{639}$$

where $\bar{x}(\theta,\theta_{-1})$ is the mean of the distribution in Equation 634 and $\bar{\beta}(\theta_{-1}) = \int_\Theta \beta(\theta,\theta_{-1})\mathrm{d}\pi(\theta|\theta_{-1})$. Observe moreover that we simplify the second term as the following:

$$\begin{aligned}
&\mathbb{E}_{\pi_{\theta_{-1}}}\left[-\beta(\theta,\theta_{-1})\left(\lambda(\theta_{-1})\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1}) + (1-\lambda(\theta_{-1}))\mu(\theta_{-1}) - \gamma(\theta,\theta_{-1})\right)^2\right] \\
&= \mathbb{E}_{\pi_{\theta_{-1}}}\left[-\beta(\theta,\theta_{-1})\left((1-\lambda(\theta_{-1}))\left(\mu(\theta_{-1}) - \frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1})\right) + \left(\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})} - 1\right)\gamma(\theta,\theta_{-1})\right)^2\right]
\end{aligned} \tag{640}$$

A necessary condition for an interior and optimal $\tau^2(\theta_{-1})$ is then the following first-order condition:

$$\begin{aligned}
\tilde{\phi}'(\tau^2(\theta_{-1})) &= -\bar{\beta}(\theta_{-1})\frac{\partial}{\partial\tau^2(\theta_{-1})}\left[\frac{\lambda^2(\theta_{-1})}{\tau^2(\theta_{-1})}\right] \\
&+ 2(1-\lambda(\theta_{-1}))\frac{\partial\lambda(\theta_{-1})}{\partial\tau^2(\theta_{-1})}\mathbb{E}_{\pi_{\theta_{-1}}}\left[\beta(\theta,\theta_{-1})\left(\mu(\theta_{-1})-\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1})\right)^2\right] \\
&+ 2\frac{\partial\lambda(\theta_{-1})}{\partial\tau^2(\theta_{-1})}\mathbb{E}_{\pi_{\theta_{-1}}}\left[\beta(\theta,\theta_{-1})\left(\mu(\theta_{-1})-\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1})\right)\left(\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}-1\right)\gamma(\theta,\theta_{-1})\right]
\end{aligned} \tag{641}$$

which reduces to:

$$\begin{aligned}
\tilde{\phi}'(\tau^2(\theta_{-1})) &= -\bar{\beta}(\theta_{-1})\frac{\frac{1}{\sigma^2(\theta_{-1})}-\tau^2}{\left(\tau^2+\frac{1}{\sigma^2(\theta_{-1})}\right)^3} \\
&+ 2\frac{\frac{1}{\sigma^2(\theta_{-1})}}{\tau^2+\frac{1}{\sigma^2(\theta_{-1})}}\frac{\frac{1}{\sigma^2(\theta_{-1})}}{\left(\tau^2+\frac{1}{\sigma^2(\theta_{-1})}\right)^2}\mathbb{E}_{\pi_{\theta_{-1}}}\left[\beta(\theta,\theta_{-1})\left(\mu(\theta_{-1})-\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1})\right)^2\right] \\
&+ 2\frac{\frac{1}{\sigma^2(\theta_{-1})}}{\left(\tau^2+\frac{1}{\sigma^2(\theta_{-1})}\right)^2}\mathbb{E}_{\pi_{\theta_{-1}}}\left[\beta(\theta,\theta_{-1})\left(\mu(\theta_{-1})-\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1})\right)\left(\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}-1\right)\gamma(\theta,\theta_{-1})\right]
\end{aligned} \tag{642}$$

We can now ask how $\tau^2(\theta_{-1})$ moves with κ. In particular, see that we can write:

$$\tilde{\phi}'(\tau^2(\theta_{-1})) = \xi(\tau^2(\theta_{-1}))\kappa \tag{643}$$

where:

$$\begin{aligned}
\xi(\tau^2(\theta_{-1})) &= \frac{1}{\left(\tau^2+\frac{1}{\sigma^2(\theta_{-1})}\right)^3}\Bigg[-\bar{\beta}(\theta_{-1})\left(\frac{1}{\sigma^2(\theta_{-1})}-\tau^2\right) \\
&\qquad + 2\left(\frac{1}{\sigma^2(\theta_{-1})}\right)^2\xi_1(\theta_{-1}) + 2\frac{1}{\sigma^2(\theta_{-1})}\left(\tau^2(\theta_{-1})+\frac{1}{\sigma^2(\theta_{-1})}\right)\xi_2(\theta_{-1})\Bigg] \\
\xi_1(\theta_{-1}) &= \mathbb{E}_{\pi_{\theta_{-1}}}\left[\beta(\theta,\theta_{-1})\left(\mu(\theta_{-1})-\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1})\right)^2\right] \\
\xi_2(\theta_{-1}) &= \mathbb{E}_{\pi_{\theta_{-1}}}\left[\beta(\theta,\theta_{-1})\left(\mu(\theta_{-1})-\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}\gamma(\theta,\theta_{-1})\right)\left(\frac{\beta(\theta,\theta_{-1})}{\bar{\beta}(\theta_{-1})}-1\right)\gamma(\theta,\theta_{-1})\right]
\end{aligned} \tag{644}$$

Applying the implicit function theorem we then have that:

$$\frac{\partial \tau^2(\theta_{-1})}{\partial \kappa} = \frac{\xi(\tau^2(\theta_{-1}))}{\tilde{\phi}''(\tau^2(\theta_{-1})) - \xi'(\tau^2(\theta_{-1}))} = \frac{\tilde{\phi}'(\tau^2(\theta_{-1}))}{\tilde{\phi}''(\tau^2(\theta_{-1})) - \xi'(\tau^2(\theta_{-1}))} \tag{645}$$

where the denominator is positive as the marginal cost of precision is always positive. Thus, we have that:

$$\frac{\partial \tau^2(\theta_{-1})}{\partial \kappa} > 0 \iff \tilde{\phi}''(\tau^2(\theta_{-1})) > \xi'(\tau^2(\theta_{-1})) \tag{646}$$

A sufficient condition for $\frac{\partial \tau^2(\theta_{-1})}{\partial \kappa} > 0$, therefore, by convexity of the costs of precision is that $\xi'(\tau^2(\theta_{-1})) < 0$ for all $\tau^2(\theta_{-1})$. In words, if the benefit of precision is a concave function, then optimally set precision is increasing in κ.

Having shown the desired general comparative static, we now return to the context of our macroeconomic model. Recall that the curvature of firms profits is given by:

$$\beta(\theta, \theta_{-1}) = v_\Pi X(\theta, \theta_{-1})^{-1-\gamma+\chi(1+\epsilon)} \theta^{-1-\epsilon} \tag{647}$$

Thus:

$$\bar{\beta}(\theta_{-1}) = \mathbb{E}_{\pi_{\theta_{-1}}}[v_\Pi X(\theta, \theta_{-1})^{-1-\gamma+\chi(1+\epsilon)} \theta^{-1-\epsilon}] \tag{648}$$

Thus, whenever aggregate output is monotonically increasing in both θ and θ_{-1} and the prior $\pi_{\theta_{-1}}$ is monotone increasing in the FOSD order and $\gamma > \chi(1+\epsilon) - 1$, we have that $\bar{\beta}(\theta_{-1})$ is monotone decreasing in θ_{-1}. It then follows that $\tau^2(\theta_{-1})$ is monotone decreasing in θ_{-1} in equilibrium whenever $\xi' < 0$. Thus, the core logic of our baseline model translates exactly over to this setting with Gaussian signal extraction.

B.4.3 Rational Inattention

We now extend our results to the case of mutual information cost. As in the previous subsection, for notational simplicity, we describe this alternative model under the assumption that there is a uniform, scalar state variable θ, which represents each firm's productivity.

We first introduce the class of posterior-separable cost functionals. Denti (2018) provides this formulation as a representation theorem in stochastic choice space of the usual posterior-based definition of Caplin and Dean (2013):

Definition 18 (Posterior-Separable Cost Functionals)**.** *A cost functional c has a posterior-separable representation if and only if there exists a convex and continuous*

ϕ such that:

$$c(p) = \int_{\mathcal{X}} \hat{\phi}(\{p(x|\theta)\}_{\theta \in \Theta}) \, \mathrm{d}x \tag{649}$$

where:

$$\hat{\phi}(\{p(x|\theta)\}_{\theta\in\Theta}) = p(x)\phi\left(\left\{\frac{p(x|\theta)\pi(\theta)}{p(x)}\right\}_{\theta\in\Theta}\right) \tag{650}$$

whenever $p(x) > 0$ and $\hat{\phi} = 0$ otherwise.

Intuitively, such a cost functional considers the cost to the agent of arriving at any given posterior and adds that up over the distribution of posteriors that are realized. Important cost functionals such as the mutual information cost functional considered in the literature on rational inattention are members of this class. Indeed, mutual information is the special case of the above where ϕ returns the entropy of the distribution that is its argument.

The mathematical structure of posterior-separable cost functionals does not admit the same prior-independence property as likelihood-separable cost functionals. As a result, we will not be able to carry all of our results over to this setting. Nevertheless, as we will argue, the key qualitative forces apply.

In the setting with likelihood separable choice in the single-agent context, we showed that greater curvature of payoffs leads to more precise actions (Proposition 6). With posterior-separable choice, the above result does not hold in general. This is because the prior also influences the states in which the agent would like to learn precisely. In particular, even if a state features high curvature, if it is unlikely to arise, the agent may not care to acquire precise information in that state. A particular case where this complication can be bypassed is when costs are given by mutual information and all actions are exchangeable in the prior in the sense that all actions are *ex ante* equally attractive (Matějka and McKay, 2015). This is a natural case to consider and yields a particularly revealing structure to the optimal policy: the agent's actions in state θ are given by a normal distribution centered on the objective optimum and with variance inversely proportional to the curvature of their objective in that state – a normal mixture model.

Proposition 42. *Suppose that $u(x, \theta) = \alpha(\theta) - \beta(\theta)(x - \gamma(\theta))^2$ and costs are posterior separable with entropy kernel $\lambda\phi(\cdot)$ for some $\lambda > 0$. If all actions are exchangeable in the prior, then in the limit of the support of the action set to infinity, $\hat{x} \to \infty$ for*

$\overline{x} = -\underline{x} = \hat{x}$, *the optimal stochastic choice rule is given by:*[7]

$$p(x|\theta) = \frac{1}{\sqrt{\frac{\pi\lambda}{\beta(\theta)}}} \exp\left\{ -\frac{1}{2} \left(\frac{x - \gamma(\theta)}{\sqrt{\frac{\lambda}{2\beta(\theta)}}} \right)^2 \right\} \tag{652}$$

Which is to say that the agent's actions follow a normal mixture model with conditional action density given by:

$$x|\theta \sim N\left(\gamma(\theta), \frac{\lambda}{2\beta(\theta)}\right) \tag{653}$$

Proof. We first show that mutual information can be written in the claimed stochastic choice form. These arguments follow closely Matějka and McKay (2015) and Denti (2018). The agent can design an arbitrary signal space $\mathcal{S}$ and choose a joint distribution between signals and states $g \in \Delta(\mathcal{S} \times \Theta)$. As in Sims (2003), the mutual information is the reduction in entropy from having access to this signal relative to the prior:

$$I(g) = \int_{\mathcal{S}} \int_{\Theta} g(s,\theta) \log\left(\frac{g(s,\theta)}{\pi(\theta) \int_{\Theta} g(s,\tilde{\theta})\,\mathrm{d}\tilde{\theta}} \right) \mathrm{d}\theta\,\mathrm{d}s \tag{654}$$

We now argue that it is without loss to consider a choice over stochastic choice rules $p : \Theta \to \Delta(\mathcal{X})$. Suppose x is an optimal action conditional on receiving any $s \in S_x$. Suppose that there exist $S_x^1, S_x^2 \in S_x$ of positive measure such that $g(\theta|s_1) \neq g(\theta|s_2)$ for all $s_1 \in S_x^1, s_2 \in S_x^2$. Now generate a new signal structure g' such $\tilde{s} \in S_x^1 \cup S_x^2$ is sent whenever any $s \in S_x^1 \cup S_x^2$ was sent under g. Clearly, x is optimal conditional on receiving $\tilde{s}$. Thus, expected payoffs under g' are the same as those under g. Moreover, g' is simply a garbling of g in the sense of Blackwell. Thus $C(g') < C(g)$ for any convex cost functional C. As I is convex, this is a contradiction. Thus, there must be at most one posterior (realized with positive density) associated with each action. As $g(s,\theta) = g(s|\theta)\pi(\theta)$, the choice of $g(s,\theta) \in \Delta(\mathcal{S} \times \Theta)$ is a choice over $g(\cdot|\cdot) : \Theta \to \Delta(S)$. Moreover, there is a unique posterior $\mu(\theta|s)$ associated with each (non-dominated) action which is determined exactly by $g(\cdot|\cdot)$. Hence, the agent directly chooses a mapping $p(\cdot|\cdot) : \Theta \to \Delta(\mathcal{X})$. The agent's problem can then be

[7] Formally, all actions are exchangeable in the prior if:

$$\int_{\Theta} \frac{\exp\{\beta(\theta)\lambda^{-1}(x - \gamma(\theta))^2\}}{\int_{\mathcal{X}} \exp\{\beta(\theta)\lambda^{-1}(\tilde{x} - \gamma(\theta))^2\} d\tilde{x}} \pi(\theta) d\theta = 1 \quad \forall x \in \mathcal{X} \tag{651}$$

directly re-written in the claimed stochastic choice form for some c_I:

$$\max_{P \in \mathcal{P}} \int_\Theta \int_\mathcal{X} u(x,\theta)\,\mathrm{d}P(x|\theta)\,\mathrm{d}\pi(\theta) - c_I(P) \tag{655}$$

Moreover, separating terms, one achieves the following representation of c_I:

$$c_I(p) = \int_\Theta \int_\mathcal{X} p(x|\theta)\log p(x|\theta)\,\mathrm{d}x\,\mathrm{d}\pi(\theta) - \int_\mathcal{X} p(x)\log p(x)\,\mathrm{d}x \tag{656}$$

where:

$$p(x) = \int_\Theta p(x|\theta)\,\mathrm{d}\pi(\theta) \tag{657}$$

The stochastic choice problem can now be expressed by the Lagrangian: ($\kappa(x,\theta)$ are the non-negativity constraints and $\tilde{\gamma}(\theta)$ are the constraints that all action distributions integrate to unity)

$$\begin{aligned} \mathcal{L}(\{p(x|\theta),\kappa(x,\theta)\}_{x\in\mathcal{X},\theta\in\Theta},\{\gamma(\tilde{\theta})\}_{\theta\in\Theta}) &= \int_\Theta \int_\mathcal{X} u(x,\theta)p(x|\theta)\,\mathrm{d}x\,\mathrm{d}\pi(\theta) \\ - \lambda\left(-\int_\mathcal{X} p(x)\log p(x)\,\mathrm{d}x + \int_\Theta\int_\mathcal{X} p(x|\theta)\log p(x|\theta)\,\mathrm{d}x\,\mathrm{d}\pi(\theta)\right) \\ + \kappa(x,\theta)p(x|\theta) + \tilde{\gamma}(\theta)\left(\int_\mathcal{X} p(x|\theta)\,\mathrm{d}x - 1\right) \end{aligned} \tag{658}$$

Any time that $p(x|\theta) > 0$, taking the FOC pointwise with respect to $p(x|\theta)$ and rearranging we have that:

$$p(x|\theta) = \frac{p(x)\exp\{u(x,\theta)\}}{\int_\mathcal{X} p(\tilde{x})\exp\{u(\tilde{x},\theta)\}\,\mathrm{d}\tilde{x}} \tag{659}$$

Moreover, we can plug the above back into the general problem and take the FOC. Re-arranging we have that for all x such that $p(x) > 0$:

$$\int_\Theta \frac{\exp\{u(x,\theta)\}}{\int_\mathcal{X} p(\tilde{x})\exp\{u(\tilde{x},\theta)\}\,\mathrm{d}\tilde{x}}\,\mathrm{d}\pi(\theta) = 1 \tag{660}$$

Up to now we have applied standard techniques from Matějka and McKay (2015). We now use our utility function and exchangeability assumption to derive our result. In particular, we take the utility function as:

$$u(x,\theta) = \alpha(\theta) - \beta(\theta)(x-\gamma(\theta))^2 \tag{661}$$

And assume exchangeability in the prior such that all actions are *ex-ante* equally attractive in the limit:

$$\int_{\Theta} \frac{\exp\{-\beta(\theta)\lambda^{-1}(x-\gamma(\theta))^2\}}{\int_{\mathcal{X}} \exp\{-\beta(\theta)\lambda^{-1}(\tilde{x}-\gamma(\theta))^2\}\,\mathrm{d}\tilde{x}} \pi(\theta)\,\mathrm{d}\theta = 1 \quad \forall x \in \mathcal{X} \tag{662}$$

Under this condition, in the limit of the support to infinity, the unconditional action distribution converges to the improper uniform distribution $p(x) = p(x')$ for all $x \in \mathcal{X}$. The conditional action distribution then becomes:

$$p(x|\theta) = \frac{\exp\{-\beta(\theta)\lambda^{-1}(x-\gamma(\theta))^2\}}{\int_{\mathcal{X}} \exp\{-\beta(\theta)\lambda^{-1}(\tilde{x}-\gamma(\theta))^2\}\,\mathrm{d}\tilde{x}} \tag{663}$$

The denominator of this expression can be computed:

$$\begin{aligned}
\int_{\mathcal{X}} \exp\{-\beta(\theta)\lambda^{-1}(x-\gamma(\theta))^2\}\,\mathrm{d}x &= \int_{\mathcal{X}} \frac{\sqrt{2\pi\frac{\lambda}{2\beta(\theta)}}}{\sqrt{2\pi\frac{\lambda}{2\beta(\theta)}}} \exp\left\{-\frac{1}{2}\left(\frac{x-\gamma(\theta)}{\sqrt{\frac{\lambda}{2\beta(\theta)}}}\right)^2\right\} \mathrm{d}x \\
&= \sqrt{2\pi\frac{\lambda}{2\beta(\theta)}} \int_{\mathcal{X}} \frac{1}{\sqrt{2\pi\frac{\lambda}{2\beta(\theta)}}} \exp\left\{-\frac{1}{2}\left(\frac{x-\gamma(\theta)}{\sqrt{\frac{\lambda}{2\beta(\theta)}}}\right)^2\right\} \\
&= \sqrt{2\pi\frac{\lambda}{2\beta(\theta)}}
\end{aligned} \tag{664}$$

It follows that:

$$p(x|\theta) = \frac{1}{\sqrt{\frac{\pi\lambda}{\beta(\theta)}}} \exp\left\{-\frac{1}{2}\left(\frac{x-\gamma(\theta)}{\sqrt{\frac{\lambda}{2\beta(\theta)}}}\right)^2\right\} \tag{665}$$

Which is to say that $X|\theta$ is a Gaussian random variable with mean $\gamma(\theta)$ and variance $\frac{\lambda}{2\beta(\theta)}$. □

This result extends the known results on Gaussian optimality of stochastic choice with mutual information (Sims, 2003) to a domain with a stochastic weight on the deviation from optimality. For our purposes, the novel and interesting feature is that the variance of the action distribution in any given state is inversely-proportional to curvature. It follows that if all actions are exchangeable in the prior when:

$$\begin{aligned}
\gamma(\theta) &= x^*(X(\theta), \theta) \\
\beta(\theta) &= \frac{1}{2}|\Pi_{xx}(X(\theta), \theta)|
\end{aligned} \tag{666}$$

where $X(\theta)$ is the unique equilibrium level of aggregate production, then the model with mutual information is exactly equivalent to the model with entropic likelihood-separable cost that we studied. All results from Section 2.4 then carry directly.

B.5 Additional Numerical Results

In this Appendix, we discuss robustness of our numerical findings as well as how the macroeconomic implications of attention cycles change under counterfactual scenarios.

B.5.1 Sensitivity of Main Results

Parameter Choice

To probe robustness to our choice of elasticity of substitution ϵ and the wage rule slope χ, we re-calibrate the model for alternative choices. For brevity, we summarize these experiments by considering "high" and "low" deviations for each parameter, holding fixed the others at baseline values, and present the proportional difference from the baseline in three summary statistics introduced in Section 2.6.3:

1. The relative output effect of negative and positive shocks, normalized in $\log \theta$ units such that the latter increases output by 3%;

2. The relative output effect of a "double dip" versus positive shock, holding fixed the size of the shock to productivity as above;

3. The ratio of output-growth volatility from the 10th to the 90th percentile of the output distribution.

We present our results in Figure B-13. Lowering the elasticity of substitution or increasing the implied average markups can have ambiguous effects because it simultaneously increases the bite of a fixed level of misoptimization on misallocation, productivity, and output, while decreasing the bite of the profit-curvature channel toward cyclical attention. We find numerically that increasing markups or decreasing ϵ, toward the level implied by De Loecker, Eeckhout, and Unger (2020), significantly increases the extent of our predicted asymmetries (≈ 1.75x), while decreasing markups or increasing ϵ, toward the level implied by Edmond, Midrigan, and Xu (2018), modestly increases the extent of our predicted asymmetries.

Increasing the slope of the wage rule dampens our predictions, due to its dampening the economy's Keynesian-cross feedback. Decreasing the slope increases the bite of our predictions substantially by amplifying the same general-equilibrium effects.

Classical Labor Markets

For tractability in the theoretical section we equipped our model with a reduced-form *wage rule* rather than a micro-founded labor supply curve. As an alternative, we use the preferences of Greenwood, Hercowitz, and Huffman (1988) which replace Equation 61 with the following:

$$\mathcal{U}(\{C_{t+j}, L_{t+j}\}_{j\in\mathbb{N}}) = \mathbb{E}_t \sum_{j=0}^{\infty} \beta^j \frac{\left(C_{t+j} - \frac{L_{t+j}{}^{1+\phi}}{1+\phi}\right)^{1-\gamma}}{1-\gamma} \tag{667}$$

and remove the wage rule, Equation 63. These preferences generate a labor supply curve $w_t = L_t^{\phi}$ which closely resembles our reduced-form wage rule, but also takes seriously the implications for risk-pricing by making marginal utility a function of hours worked. We choose a parameterization of $\phi = \frac{\chi}{1-\chi\epsilon} = 0.153$, where χ and ϵ take our benchmark values indicated in Table 2.4. As indicated in the richer model of Appendix B.2, this calibration replicates an elasticity of $\chi = 0.095$ between real wages and real output.

Figure B-14 is this model's analogue to Figure 2-4, showing output, the attention wedge, and labor productivity. We find comparable behavior of the attention wedge and losses from misoptimization and inattention. Figure B-15 is this model's analogue to Figure 2-5, showing state-dependent shock response and stochastic volatility. Our results are quantitatively similar to our baseline calibration. These results demonstrate that classical labor markets do not undermine our main results in calibrations that are consistent with our calibrated wage rigidity.

B.5.2 Attention Cycles Under Counterfactual Scenarios

Because the main amplification mechanism in our model, the reallocation of attention, is endogenous to economic conditions, we can use our framework to study how attention cycles and all associated macro phenomena would behave under counterfactual conditions. In this subsection, we explore the interaction of our findings with recent trends in product mark-ups; the decreasing pro-cyclicality of real wages; and external uncertainty shocks.

The Rise of Markups, via Lower Substitutability

An extensive recent literature documents a secular increase in markups charged by US public firms over the last half century (see, e.g., De Loecker, Eeckhout, and Unger, 2020; Edmond, Midrigan, and Xu, 2018; Demirer, 2020). In our empirical calibration, we targeted "modern" average mark-ups as informed by this literature. Our framework would interpret any trends in aggregate mark-ups as arising from changes in the elasticity of substitution ϵ between products, which would need to have been higher in the previous, low-markup era than it is today. In our model, a lower elasticity of substitution or higher markup increases the output cost of a fixed amount of misoptimization dispersion, as it intuitively makes each individual product more "essential" to the consumed good; and it has *a priori* ambiguous effects on the extent of equilibrium attention cycles.

In our model, we run the following simple experiment. First, we adjust ϵ upward to simulate a 15 percentage point decrease in the aggregate markup, to match the estimate of Demirer (2020) for markups since the 1960s; and second, we adjust ϵ downward to match a 15 percentage point increase. We find that lower markups correspond to more severe effects of attention cycles on business cycles, as summarized by the asymmetry and state-dependence of dynamics (second and third panel of Figure B-16). These results provide only speculative clues about the future, owing to both uncertainty about the true "cause" of rising markups and the plausibility of the trend continuing. But, regardless of their precision, they highlight a potentially important pathway linking market structure, firm decisionmaking, and the aggregate cost of misallocation.

More Rigid Real Wages

The inverse relationship between wage inflation and real conditions has proved elusive in modern data, particularly since the financial crisis (see, e.g., Galí and Gambetti, 2019). Our baseline estimate of slope 0.095 between (detrended) real output and real wages since 1987 reflects this reality. In our model, more rigid real wages corresponded to a steeper Keynesian cross, and a steeper incentive toward high attention in low states of the world. For this reason, we may expect that the growing disconnect between factor prices and real conditions contributes toward the severity of our estimated macro effects.

In parallel to the previous experiment, we simulate both a "calibrated past" and "extrapolated future." For the former, we plug in the estimate of Galí and Gambetti (2019) that the wage Phillips curve has flattened by a factor of 1.9 over the last half century; for the latter, we extrapolate the same multiplicative trend as additional flattening.[8] We find, as shown in panels four and five of Figure B-16, stronger effects of attention cycles in the regime with more rigid wages. This underscores the

[8] In particular, we use the ratio of the 1964-2007 estimate and 2007-2017 estimates in Table 3A of Galí and Gambetti (2019).

complementarity between attention cycles and the steepness of the Keynesian cross, and suggests a novel pathway by which factor price rigidity can influence patterns of macroeconomic volatility.

Elevated Uncertainty

Spikes in uncertainty around exceptional economic and political events have large documented effects on financial markets and firm decisionmaking (Bloom, 2009). Moreover, large, disorienting shocks are often either a natural consequence of poor economic performance (e.g., policy surprises during the 2007-2009 financial crisis) or their root cause (e.g., the Covid-19 pandemic). For this reason it is natural to study how changes in the "level of uncertainty," formalized in our model as variation in the attention cost λ, might interact with our main business cycle predictions. Proposition 9 showed that increases in uncertainty depress output in our model. These shocks also, according to the results of Proposition 6, increase the *sensitivity* of dispersion to macroeconomic conditions and hence, based on extrapolation of this partial-equilibrium logic, may amplify the extent of misoptimization cycles. For this reason, we might predict that elevated uncertainty is also complementary to the asymmetry and state-dependence generated at the macro level.

We explore this relationship by solving for the model equilibrium under scenarios with depressed and elevated attention costs, and numerically verify the predicted complementarity (panels six and seven of Figure B-16). Thus our theory predicts that business cycles caused and/or amplified by background uncertainty-inducing events may induce sharper fluctuations in aggregate volatility due to endogenous reallocation of attention.

B.6 Additional Empirical Results

B.6.1 Macro Attention With Conference Calls

In this Appendix, we describe an alternative construction of our textual Macro Attention measure using sales and earnings conference calls as an alternative source of information about firms' relative attention to different risks and events. This analysis mirrors and supplements our main analysis of forms 10-K and 10-Q.

Data and Measurement We obtain data from the Fair Disclosure (FD) Wire service, which records transcripts of sales and earnings conference calls for public companies around the world. We obtain an initial sample of 294,900 calls which cover 2003 to 2014. We next subset to documents that have reported firm names and stock tickers, which are automatically associated with documents by Lexis Nexis. When matches are probabilistic, we use the first (highest probability) match.[9] We finally restrict to firms that are listed on one of three US stock exchanges: the NYSE, the NASDAQ, or the NYSE-MKT (Small Cap). We finally connect tickers to firm identifiers (GVKEY) using the master cross-walk available on Wharton Research Data Services (WRDS). These operations together reduce the sample size to 164,805 calls.

We finally restrict to conference calls that are sales or earnings reports. This further reduces the sample to 158,810 total observations, by removing conference calls related to other activities (e.g., mergers). All in all, this sample is about 3,600 firm observations per quarter, or about 60% of the per-quarter observations we obtained via the SEC filings.

To tabulate histograms of words within documents, we use the CountVectorizer function in the FeatureExtraction module of the standard Python package Scikit Learn. We then replicate the exact methodology of Section 2.2.1, generating a new list of 73 macroeconomic words. Like the main list of words, they are a combination of very interpretable choices ("government," "unemployment," "monetary") and false positives related to structure and pedagogy ("theory," "chapter").

Results Figure B-11 plots the conference-call-derived measure, in log units and with quarterly fixed effects taken out, alongside the US unemployment rate. Conference-call-derived macro attention, like our main measure derived from forms 10-Q/K, is cyclical and persistent. To benchmark these facts in the same way we did in the main text, we first run linear regressions of the form

$$\log \text{MacroAttentionCC}_t = \alpha + \beta_Z \cdot Z_t + \epsilon_t \tag{668}$$

for $Z_t \in \{\text{Unemployment}_t/100, \log \text{SPDetrend}_t\}$. The first two columns of Table B.11 show the coefficients, which are slightly larger in absolute value than their equivalents with our 10K/Q measure (1.529 and -0.104, respectively). The third column gives our estimate of an AR(1) process, which is close to a unit root.

Casual comparison of Figure B-11 and Figure 2-1 in the main text suggests that, while our two measures of attention have similar cyclical patterns, they do not closely track each other at the aggregate level. Conference-call-derived attention is more sharply peaked around the onset of the Great Recession while 10-K/Q-derived attention remains elevated for several subsequent years. The correlation between the two

[9] In the essentially zero-measure cases in which there is a tie, we take the alphabetically first ticker.

measures on a common sample is a (statistically insignificant) 0.091. The relationship is closer, however, at the firm level. Columns 4-6 of Table B.11 show the results of regressing the conference-call-derived measure 10K/Q-derived measure at the firm level, with increasingly more stringent fixed effects. The correlation is consistently positive, though strongest in terms of cross-firm differences as opposed to within-firm differences. Moreover, as indicated in column 1 of Table B.8, our finding linking firm-level misoptimization with firm-level attention is stable based on the conference call measure.

B.6.2 Macro Attention With Word Stemming

Measurement Our main method for constructing Macro Attention treats individual words as the unit of measurement. For this reason, words like "unemployment" and "unemployed" are counted separately despite likely communicating the same meaning in all contexts. This method, while appealingly simple, may systematically under-count words that have a number of different forms or tenses, while allowing the multiple forms of certain ubiquitous words to crowd out other distinct concepts.

As an alternative method, which allays some of these concerns, we re-do our calculation of macroeconomic language using *word stems*. For each word w in the macroeconomics references and/or regulatory filings, we use the Porter Stemmer implemented in Python's `nltk` software to determine a stem $s(w)$. Stemming is an algorithmic and imperfect process. In examples relevant to our context, the Porter Stemmer associates "unemployment" and "unemployed" with the common stem "unemploy." But it also, employing the same logic, associates "nominal" with "nomin," a stem which may match to words less often used to describe aggregate prices (e.g., "nominate").

We adapt our tf-idf calculation to the stem level by calculating, for each stem s that appears in the regulatory filings,

$$\text{tf-idf}(s)_{it} := \text{tf}(s)_{it} \cdot \log\left(\frac{1}{\text{df}(s)}\right) \tag{669}$$

where $\text{tf}(s)$ is the total term frequency of all words mapped to stem s, and $\text{df}(s)$ is the minimum document frequency among words associated with the stem.[10] We calculate the top macro stems using the approach described in the main text (Section 2.2.1); construct the set of macro words $\mathcal{W}_M$ as the set of all words associated with a macro

[10] We use the minimum instead of the overall frequency due to a data limitation of having document frequencies at the word, not stem, level. We expect either method to produce broadly similar results.

stem; and proceed in the standard way to calculate firm-level and aggregate macro attention.

Results Table B.12, in analogy to Table B.11, presents a summary of the cyclical patterns of the stemmed Macro Attention measure as well as its relationship to our main measure. The two measures behave very similarly in the time series and are tightly connected at the firm level. Moreover, when we replicate our main model linking firm-level Macro Attention to firm-level misoptimization as in Table B.8, we estimate a coefficient of -0.020 (SE: 0.004), which is comparable within error bars to our baseline estimate of -0.009 (SE 0.003).

B.6.3 Dispersion in TFPR and Value Marginal Products

In this Appendix, we review theoretically and empirically the behavior of revenue-TFP (TFPR) dispersion in our analysis. This analysis builds a bridge between our findings and those of Kehrig (2015), who shows that cross-firm variance in revenue total-factor-productivity, or the product of prices and physical productivity, is counter-cyclical for the US durable manufacturing sector between 1972 and 2007.

Definitions and Theoretical Context In our model, with a three-input production function, log physical TFP is defined as

$$\log\theta_{it} = \log x_{it} - (\alpha_{j(i),M}\log M_{it} + \alpha_{j(i),K}\log K_{it} + \alpha_{j(i),L}\log L_{it}) \tag{670}$$

where x_{it} is physical sales, in quantity units; (M_{it}, K_{it}, L_{it}) are materials, capital, and labor; and $(\alpha_{j(i),M}, \alpha_{j(i),K}, \alpha_{j(i),L})$ are (industry-specific) weights on these inputs, which sum to one. Prices are defined by the demand curves $\log p_{it} = \frac{1}{\epsilon}(\log X_t - \log x_{it})$ and revenue-based TFP is therefore

$$\log\theta^R_{it} = \log\theta_{it} + p_{it} = \log \text{Sales}_{it} - (\alpha_{j(i),M}\log M_{it} + \alpha_{j(i),K}\log K_{it} + \alpha_{j(i),L}\log L_{it}) \tag{671}$$

where $\text{Sales}_{it} = p_{it}x_{it}$. We can calculate exactly what TFPR is in our empirical model with inattentive firms, introduced in Appendix B.3.3, by combining this definition with the input-choice policy functions derived in Lemma 10:

Lemma 12. *TFPR in our model is given by*

$$\log\theta^R_{it} = \tilde{\gamma}_i + \Xi_{j(i),t} - \frac{1}{\epsilon}m_{it} \tag{672}$$

where $\tilde{\gamma}_i$ is a constant at the firm level and $\Xi_{j(i),t}$ is a constant that varies at the *industry-by-time level.*

Thus, the only sources of within-industry variation in revenue-based TFP in our model are the firm fixed effects and misoptimizations. This is another way of stating our theoretical result that misoptimizations matter for aggregate output and productivity via their effects on "misallocation," as TFPR measures the value marginal product of the (minimal cost) input bundle. A simple corollary, under our assumption of cost minimization, is that the value marginal product $\log\theta^Z_{it} = \log\frac{\text{Sales}_{it}}{Z_{it}}$, for any of the three inputs, can also be written as $\log\theta^Z_{it} = \tilde{\gamma}_{Z,i} + \Xi_{j(i),Z,t} - \frac{1}{\epsilon}m_{it}$ with now input-specific fixed effects (firm-level and industry-by-time level).

In our model, therefore, the presence of TFPR dispersion or value-marginal-product dispersion indicates that there is non-zero misoptimization. The monotonicity of TFPR or value-marginal-product dispersion over the business cycle, once we project out firm and industry-by-time fixed effects, therefore provides an alternative test of the model's prediction of procyclical misoptimization dispersion.

TFPR Dispersion and VMPL Dispersion We calculate log TFPR in our data by first calculating

$$\log\tilde{\theta}^R_{it} = \log\text{Sales}_{it} - \big(\alpha_{M,j(i)}\cdot\log\text{MatExp}_{it} - \alpha_{L,j(i)}\cdot\log\text{Empl}_{it} - \alpha_{K,j(i)}\cdot\log\text{CapStock}_{it}\big) \tag{673}$$

where variable definitions follow the convention of Appendix B.3. We then remove industry-by-time fixed effects and firm fixed effects to remove factor prices, as suggested by Lemma 12, to generate the variable $\log\hat{\theta}^R_{it}$.[11] We also calculate the log value marginal product of labor (VMPL) by first calculating

$$\log\tilde{\theta}^L_{it} = \log\text{Sales}_{it} - \log L_{it} \tag{674}$$

and similarly removing industry-by-time and firm fixed effects to generate the final measure $\log\hat{\theta}^L_{it}$.

Appendix Figure B-12 shows the cyclical behavior of TFPR dispersion, the variance of $\log\hat{\theta}^R_{it}$, and VMPL dispersion, the variance of $\log\hat{\theta}^L_{it}$. In line with our earlier results, VMPL dispersion is markedly pro-cyclical with respect to both the unemployment rate and the return on the S&P500. TFPR dispersion has no significant relationship with either unemployment or the S&P500. Taken together, these exercises suggest that value-marginal-product-based measures of misallocation give results

[11] Compare with Equation 598, which resembles Equation 673 but for deflating the input shares by the mark-up. This is exactly the within-model adjustment for prices p_{it}.

consistent with our main finding of pro-cyclical misoptimization.

B.7 State-Dependent Attention in Survey Data

In this Appendix, we test our interpretation of attention and misoptimization cycles using the dataset of Coibion, Gorodnichenko, and Kumar (2018a) (henceforth, CGK), one of the most comprehensive datasets of firm-level operations and macro backcasts in an advanced economy. These data were assembled from a detailed survey of the general managers of a representative panel of firms in New Zealand from 2013 to 2016. The final subsection of this appendix, B.7.3, contains more details about the survey questions and data.

B.7.1 Reported Attention and the Business Cycle

Although the CGK survey took place during relatively tranquil times for the New Zealand economy, it did ask two *hypothetical* questions directly revealing of the premise for this paper. Each concerned firm's desire to collect information on the macroeconomy conditional on either good (or poor) conditions:

> Suppose that you hear on TV that the economy is doing well [or poorly]. Would it make you more likely to look for more information?

Table B.15 reports the percentage of answers in each of five bins, given the conditions of the economy doing "well" or "poorly." This self-reported demand for information increases in the context of *bad* news about the macroeconomy and, if anything, contracts with *good* news of the macroeconomy. This is consistent with our hypothesis that bad conditions increase the stakes for firms' decisions and hence make keen attention to macroeconomic conditions more important, while good news does not have a symmetric effect.

B.7.2 Reported Profit Function Curvature and Attention

A second test possible in the CGK data relates to our more specific prediction that higher curvature of the firm's objective, as a function of decision variables, should increase attentiveness to decision-relevant variables including macroeconomic aggregates. The CGK survey indirectly elicits information on this shifter via questions about purely hypothetical price changes and revenue increases to an "optimal point." In Section B.7.3 at the end of this appendix, we show exactly how one can use a pair of linked questions about firms' hypothetical optimal reset price, and the hypothetical percentage increase in profits that would be associated with that change, to develop an elicited measure of *firm profit curvature* in non-risk-adjusted units.[12]

As outcomes for macro attention, we can turn to two sources. The first is the absolute-value error in firms' one-year back-casts for three macro variables: inflation, output growth, and unemployment. The second is firm managers' reported (binary) interest in *tracking* one of the aforementioned variables.

For each of the aforementioned firm-level outcomes Y_{it}, we run the following regression on the firm-level profit curvature variable ProfitCurv_{it} and a vector of controls Z_{it}:

$$Y_{it} = \alpha + \beta \cdot \text{ProfitCurv}_{it} + \gamma' Z_{it} + \epsilon_{it} \tag{675}$$

Our prediction that profit curvature drives stakes for attention corresponds to $\beta < 0$ for back-cast errors and $\beta > 0$ for reported attention. We control for five bins in the firms' total reported output and the firms' 3-digit ANZ-SIC code industries. Finally, we cluster all standard errors by 3-digit industry.

Table B.16 shows the results. For inflation we find strong evidence that higher-curvature firms make smaller errors, with some of the effect being absorbed by control variables when added. For GDP growth we find estimates that are much less precise but have the same signs; and for unemployment, results that are further imprecise and have the wrong signs. We take this as further suggestive evidence for the mechanism that our theory proposes: that the differential stakes of making mistakes is a contributing factor to macro attention.

B.7.3 Details of Data Construction

In this final subsection, we describe more precisely how we use the raw survey results of Coibion, Gorodnichenko, and Kumar (2018a) in our empirical analysis. We make use of the full dataset contained in the replication files posted on the article's page hosted by the *American Economic Review* and OpenICPSR (Coibion, Gorodnichenko, and Kumar, 2018b). All direct references to survey questions by wave or number match the "Appendix 5: Selected Survey Questions" in the online appendix available at the same link.

Profit Function Curvature

We draw our measure of profit function curvature from the answers to two survey questions about hypothetical price changes. These are jointly asked as Question 17 of Wave 5, Part B:

[12]Table B.17 shows that this curvature measure is higher for smaller firms with more within-industry competitors, though these patterns are immaterial for our reduced-form verification of our prediction.

> If this firm was free to change its price (i.e. suppose there was no cost to renegotiating contracts with clients, no costs of reprinting catalogues, etc.) right now, by how much would it change its price? Please provide a percentage answer. By how much do you think profits would change as a share of revenues? Please provide a numerical answer in percent.

Denote the answer for prices as Δp_i and the answer for profits as $\Delta\Pi_i$. Under the assumption that the following second-order approximation holds for the deviation of profits from their frictionless optimum (e.g., a version of (70), in percentage units for the outcome and the choice variable), the following relationship holds between the measurable quantities and the profit function curvature ProfitCurv_i:

$$\Delta\Pi_i = \text{ProfitCurv}_i \cdot \Delta p_i^2 \tag{676}$$

We use this expression to calculate an empirical analogue of profit curvature. The top panel of Table B.17 provides summary statistics of measured profit curvatures among the 3,153 firms for which we can measure it. The median reported curvature is 0.12, which means that a one-percentage-point deviation from the optimal price for such a firm corresponds to a 0.12-percentage-point deviation from optimal profits as a fraction of revenue.

The bottom panel of Figure B.17 shows firm and manager-level correlates for our measure in the CGK data. The table reports coefficients of the following regression:

$$\widehat{\text{ProfitCurv}}_i = \beta \cdot \hat{X}_i + e_{it} \tag{677}$$

where the hat denotes that both variables have been normalized to z-score units (i.e., with means subtracted and standard deviation divided out), so the coefficient β is a "normalized" metric of the standard-deviation-to-standard-deviation effect. We find strong evidence that the firms with higher profit function curvature are smaller and have more competitors. There is only weaker evidence that the associated managers are more skilled and/or better rewarded. We interpret this cautiously as evidence that likely confounds via manager skill and firm sophistication (i.e., better managers grow firms larger, and make better forecasts) are going the "wrong direction" to explain our reduced-form correlations between profit curvature and forecasting accuracy.

Outcomes: Back-cast Errors

For back-cast errors, we use the following questions that are split among waves of the survey. In survey wave 1, firms are asked the following question:

> During the last twelve months, by how much do you think prices changed overall in the economy?

Although the wording of the question is not entirely clear about what indicator is being referred to, we follow CGK and interpret this as the annual percent change in CPI, with realized value 1.6%. Firms are asked a similar question in wave 4, but we prefer the wave 1 version because the sample size is slightly larger. Table B.18 recreates Table B.16 from the main text, first for the wave 1 back-cast of inflation (reported for the main text) and next for the wave 4 back-cast of inflation (not reported in the main text, but quantitatively very similar).

For GDP growth, we use the following question from wave 4:

> What do you think the real GDP growth rate has been in New Zealand during the last 12 months? Please provide a precise quantitative answer in percentage terms.

and compare with a realized value of 2.5%. Finally, for unemployment, we use the following question also from wave 4:

> What do you think the unemployment rate currently is in New Zealand? Please provide a precise quantitative answer in percentage terms.

and compare with a realized value of 5.7%. All realized values are taken from the replication files of CGK, to deal with any ambiguity about statistical releases, and ensure comparability with that study.

Outcomes: Tracking Indicators

We finally use, for the lower panel of Table B.16, the following questions from wave 4 about tracking different variables:

> Which macroeconomic variables do you keep track of? Check each variable that you keep track of.
>
> 1. Unemployment rate
> 2. GDP
> 3. Inflation
> 4. None of these is important to my decisions

We code for each variable a binary indicator of whether the firm lists the variable of interest. We combine GDP in this question (by implication, in levels) with quantitative forecasts of GDP Growth in Table B.16.

Table B.1: Misoptimization and Firm Performance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Outcome: R_{it}				Outcome: π_{it}			
$\hat{u}_{it}^2$	-0.236	-0.230	-0.060	-0.051	-0.316	-0.316	-0.106	-0.105
	(0.026)	(0.026)	(0.032)	(0.032)	(0.024)	(0.024)	(0.018)	(0.017)
Sector x Time FE	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE			✓	✓			✓	✓
TFP Control		✓		✓		✓		✓
N	41,578	41,578	41,206	41,206	51,015	51,015	50,966	50,996
R^2	0.238	0.261	0.384	0.403	0.117	0.131	0.663	0.681

Notes: R_{it} is the firm-level log stock return and π_{it} is firm-level profitability. $\hat{u}_{it}^2$ is the squared firm-level misoptimization residual, constructed using the methods described in Section 2.5.1. Standard errors are double-clustered at the year and firm level.

B.8 Supplemental Tables and Figures

Table B.2: Dynamic Effects of Misoptimization

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
					Outcome:				
		$\Delta \log \hat{\theta}_{i,t+k}$			$R_{i,t+k}$			$\pi_{i,t+k}$	
Horizon k	0	1	2	0	1	2	0	1	2
$\hat{u}_{it}^2$	-0.009	0.014	-0.007	-0.236	-0.252	-0.251	-0.316	-0.286	-0.265
	(0.007)	(0.008)	(0.010)	(0.026)	(0.027)	(0.038)	(0.024)	(0.018)	(0.019)
Sector x Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
N	50,455	40,671	32,362	41,578	34,643	28,103	51,015	42,014	33,934
R^2	0.231	0.245	0.263	0.238	0.241	0.248	0.117	0.123	0.126

Notes: Each column is the estimate of a separate projection regression. The outcomes are TFP growth $\Delta \log \hat{\theta}_{it}$ (first three columns), log stock returns R_{it} (second three columns), and profitability π_{it} (last three columns). $\hat{u}_{it}^2$ is the squared firm-level misoptimization residual, constructed using the methods described in Section 2.5.1. Standard errors are double-clustered at the year and firm level.

Table B.3: Cyclicality of Misoptimization, with Alternative Measurement Schemes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Outcome: MisoptimizationDispersion$_t$										
Unemployment$_t$/100	-0.841	-0.439	-0.822	-0.749	-0.501	-0.695	-0.802	-0.813	-0.605	-0.812	-1.034
	(0.341)	(0.196)	(0.336)	(0.292)	(0.159)	(0.280)	(0.337)	(0.330)	(0.293)	(0.335)	(0.549)
Correlation	-0.493	-0.468	-0.485	-0.466	-0.546	-0.505	-0.479	-0.489	-0.494	-0.462	-0.379
Baseline	✓										
Adj. Control		✓									
Leverage Control			✓								
t, t^2 Control				✓							
Manufacturing					✓						
Sector Policy Fn.						✓					
t-varying Policy Fn.							✓				
Quadratic Policy Fn.								✓			
Pre-Period TFP									✓		
t-varying Prod. Fn.										✓	
OP (96) TFP											✓
N	31	31	31	31	31	31	31	31	20	31	31
R^2	0.243	0.219	0.235	0.326	0.298	0.255	0.230	0.239	0.244	0.213	0.144

Notes: The first row reports the coefficient from the regression of MisoptimizationDispersion$_t$ on Unemployment$_t$/100, with standard errors that are HAC-robust with a 3-year Bartlett Kernel. The following row reports the correlation of these variables (in column 4, conditional on projecting out controls). The "Adjustment Cost" and "Leverage" controls are described in the main text. The "Sector Policy Fn." estimates the policy function separately for each sector. The "t-varying Policy Fn." model interacts all coefficients in the policy function with time fixed effects. The "Quadratic Policy Fn." allows for quadratic dependence on TFP. The "Pre-Period TFP" model uses cost shares from before 1997 to construct the production function, and data after 1998 to estimate the policy function and misoptimizations. The "t-varying Prod. Fn." model estimates the Solow residual using industry-by-year-specific cost shares. The "OP (96)" model estimates productivity using the proxy-variable strategy of Olley and Pakes (1996), as detailed in Appendix B.3.2.

Table B.4: The Effects of Misoptimization in Levels

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome:					
	R_{it}		π_{it}		$\hat{m}_{it}^2$	
$\hat{m}_{it}^2$	-0.042	-0.076	-0.021	-0.025		
	(0.028)	(0.033)	(0.010)	(0.011)		
$\hat{m}_{it}^2 \times R_t$		0.177		0.038		
		(0.078)		(0.053)		
$\log \text{MacroAttention}_{it}$					-0.010	-0.020
					(0.006)	(0.007)
Firm FE	✓	✓	✓	✓		✓
Sector x Time FE	✓	✓	✓	✓	✓	✓
N	41,247	41,247	57,646	57,646	34,421	33,841
R^2	0.385	0.385	0.656	0.656	0.053	0.488

Notes: $\hat{m}_{it}$ is the firm-level misoptimization, constructed using the methods described in Section 2.5.1 These specifications replicate results in Tables 2.1, 2.2, 2.3, and B.1 with with $\hat{m}_{it}$ in place of $\hat{u}_{it}$. Standard errors are double-clustered at the year and firm level.

Table B.5: Pricing of Misoptimization, with Alternative Measurement Schemes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
					Outcome: R_{it}					
$\hat{u}_{it}^2$	-0.097	-0.239	-0.101	-0.168	-0.090	-0.099	-0.109	-0.062	-0.057	-0.096
	(0.034)	(0.941)	(0.035)	(0.045)	(0.036)	(0.035)	(0.034)	(0.028)	(0.021)	(0.034)
$\hat{u}_{it}^2 \times R_t$	0.443	0.941	0.415	0.680	0.420	0.330	0.447	0.227	0.231	0.417
	(0.171)	(0.370)	(0.169)	(0.182)	(0.156)	(0.163)	(0.163)	(0.130)	(0.098)	(0.168)
Sector x Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Baseline	✓									
Adj. Control		✓								
Leverage Control			✓							
Manufacturing				✓						
Sector Policy Fn.					✓					
t-varying Policy Fn.						✓				
Quadratic Policy Fn.							✓			
Pre-Period TFP								✓		
t-varying Prod. Fn.									✓	
OP (96) TFP										✓
N	41,206	35,388	41,016	22,902	41,197	41,203	41,203	26,206	40.078	41,166
R^2	0.385	0.387	0.385	0.367	0.385	0.384	0.385	0.429	0.382	0.385

Notes: R_{it} is the firm-level log stock return. $\hat{u}_{it}$ is the firm-level misoptimization residual, constructed using the methods described in Section 2.5.1 and the indicated variants described in the main text and the notes of Table B.3. R_t is the log return of the S&P 500. Standard errors are double-clustered at the year and firm level. The scenarios are described in the main text and the notes of Table B.3.

Table B.6: Markets Punish Misoptimizations Harder in Low States, Additional Controls

	(1)	(2)	(3)	(4)	(5)	(6)
			Outcome: R_{it}			
$\hat{u}_{it}^2 \times R_t$	0.376	0.378	0.345	0.321	0.330	0.489
	(0.123)	(0.109)	(0.118)	(0.173)	(0.094)	(0.296)
Sector x Time FE	✓	✓	✓	✓	✓	✓
$\hat{u}_{it}^2$	✓	✓	✓	✓	✓	✓
TFP and Interaction		✓				
Leverage and Interaction			✓			
Lag Return and Interaction				✓		
Industry FE and Interaction					✓	
Firm FE and Interaction						✓
N	41,578	41,578	41,429	34,805	41,206	41,206
R^2	0.239	0.261	0.246	0.239	0.379	0.498

Notes: R_{it} is the firm-level log stock return. $\hat{u}_{it}$ is the firm-level misoptimization residual, constructed using the methods described in Section 2.5.1. R_t is the log return of the S&P 500. Column 1 reports the baseline estimate, Column 1 of Table 2.1. Columns 2-6 add additional variables and their interaction with $\hat{u}_{it}^2$. These variables are: the level of log firm TFP, $\log \hat{\theta}_{it}$; leverage, Lev_{it}; the previous year's stock return, $R_{i,t-1}$; an industry fixed effect $\chi_{j(i)}$; and a firm fixed effect γ_i. Standard errors are double-clustered at the year and firm level.

Table B.7: Alternative Timing for Relationship Between Attention and Misoptimization

	(1)	(2)	(3)	(4)	(5)	(6)
			Outcome: $\hat{u}_{it}^2$			
$\log \text{MacroAttention}_{i,t-1}$	-0.0031	-0.0010			0.0060	0.0079
	(0.0034)	(0.0045)			(0.0048)	(0.0045)
$\log \text{MacroAttention}_{it}$					-0.0006	-0.0058
					(0.0048)	(0.0042)
$\log \text{MacroAttention}_{i,t+1}$			-0.0064	-0.0052	-0.012	-0.0067
			(0.0030)	(0.0039)	(0.004)	(0.0048)
Joint F-statistic					4.17	2.41
Joint p-value					0.019	0.097
Sector x Time FE	✓	✓	✓	✓	✓	✓
Firm FE		✓		✓		✓
N	25,657	25,122	24,094	23,312	19,330	18,649
R^2	0.054	0.375	0.062	0.382	0.065	0.383

Notes: $\hat{u}_{it}^2$ is the squared firm-level misoptimization residual, constructed using the methods described in Section 2.5.1. $\log \text{MacroAttention}_{it}$ is the measure of firm-level macroeconomic attention, constructed using the methods described in Section 2.2.1. Standard errors are double-clustered at the year and firm level.

Table B.8: Attention and Misoptimization, with Alternative Measurement Schemes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
						Outcome: $\hat{u}_{it}^2$					
$\log \text{MacroAttention}_{it}$	-0.0081	-0.0163	-0.0035	-0.0076	-0.0127	-0.0107	-0.0084	-0.0062	-0.0140	-0.0080	-0.0140
	(0.0028)	(0.0066)	(0.0015)	(0.0028)	(0.0037)	(0.0028)	(0.0028)	(0.0026)	(0.0042)	(0.0028)	(0.0042)
Sector x Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Baseline	✓										
Conf. Call		✓									
Adj. Control			✓								
Leverage Control				✓							
Manufacturing					✓						
Sector Policy Fn.						✓					
t-varying Policy Fn.							✓				
Quadratic Policy Fn.								✓			
Pre-Period TFP									✓		
t-varying Prod. Fn.										✓	
OP (96) TFP											✓
N	28,279	5,997	24,024	28,133	14,891	28,283	28,275	28,275	24,785	28,266	24,785
R^2	0.053	0.072	0.060	0.053	0.041	0.054	0.051	0.056	0.046	0.053	0.046

Notes: $\hat{u}_{it}^2$ is the squared firm-level misoptimization residual, constructed using the methods described in Section 2.5.1 and the indicated variants described in the main text and the notes of Table B.3. $\log \text{MacroAttention}_{it}$ is the measure of firm-level macroeconomic attention, constructed using the methods described in Section 2.2.1 and Appendix B.6.1. Standard errors are double-clustered at the year and firm level. The scenarios are described in the main text and the notes of Table B.3.

Table B.9: Policy Function Estimation

	Baseline	Adj. Cost	Leverage	Quadratic
	Panel A: Persistence of Misoptimization Outcome: $\hat{m}^0_{it}$			
$\hat{m}^0_{i,t-1}$	0.696	0.016	0.696	0.683
	(0.021)	(0.005)	(0.003)	(0.003)
	Panel B: Quasi-Differenced Policy Function Outcome: $\log L_{it} - \hat{\rho} \log L_{i,t-1}$			
$\log \hat{\theta}_{it}$	0.418	0.381	0.419	0.463
	(0.024)	(0.026)	(0.025)	(0.029)
$\log \hat{\theta}_{i,t-1}$	-0.031	-0.090	-0.026	0.006
	(0.018)	(0.015)	(0.019)	(0.020)
$\log \hat{\theta}_{it} \times \text{Lev}_{it}$			-0.008	
			(0.003)	
$\log \hat{\theta}_{i,t-1} \times \text{Lev}_{i,t-1}$			-0.025	
			0.006	
Lev_{it}			-0.020	
			(0.005)	
$\text{Lev}_{i,t-1}$			-0.050	
			(0.011)	
$(\log \hat{\theta}_{it})^2$				0.045
				(0.008)
$(\log \hat{\theta}_{i,t-1})^2$				0.031
				(0.006)
$\log L_{i,t-1}$		0.811		
		(0.012)		
$\log L_{i,t-2}$		-0.041		
		(0.010)		
N	51,891	44,051	51,664	51,891
R^2	0.896	0.990	0.896	0.904

Notes: The four columns correspond to four of our policy-function estimation methods, as described in the main text. The coefficient on $\hat{m}^0_{i,t-1}$ in panel A corresponds in the model to parameter ρ. Standard errors are double clustered by firm and year.

Table B.10: Data Definitions in Compustat

	Quantity	Expenditure
Production, x_{it}	—	`sale`
Employment, L_{it}	`emp`	`emp` $\times$ industry wage
Materials, M_{it}	—	`cogs` + `xsga` − `dp` − wage bill
Capital, K_{it}	`ppegt` plus net investment	—

Table B.11: Time-Series and Cross-Sectional Properties of Conference-Call Attention

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome:					
	$\log \text{MacroAttnCC}_t$			$\log \text{MacroAttnCC}_{it}$		
$\frac{\text{Unemployment}_t}{100}$	2.481					
	(0.596)					
$\log \text{SPDetrend}_t$		-0.270				
		(0.056)				
$\log \text{MacroAttnCC}_{t-1}$			0.949			
			(0.068)			
$\log \text{MacroAttn10K}_{it}$				0.463	0.372	0.121
				(0.034)	(0.036)	(0.028)
Firm FE						✓
Sector x Time FE					✓	✓
N	46	46	45	8,023	7,994	7,670
R^2	0.376	0.593	0.873	0.123	0.308	0.804

Notes: The "CC" MacroAttention is constructed using the methods described in Appendix B.6.1. The "10K" MacroAttention is our baseline measure from Section 2.2. In the first three columns, standard errors are HAC robust with a bandwidth (Bartlett kernel) of four quarters. In the second three columns, standard errors are double-clustered by year and firm ID.

Table B.12: Time-Series and Cross-Sectional Properties of Word-Stemmed Attention

	(1)	(2)	(3)	(4)	(5)	(6)
	Outcome:					
	$\log \text{MacroAttnStem}_t$			$\log \text{MacroAttnStem}_{it}$		
$\frac{\text{Unemployment}_t}{100}$	0.994					
	(0.330)					
$\log \text{SPDetrend}_t$		-0.062				
		(0.031)				
$\log \text{MacroAttnStem}_{t-1}$			0.811			
			(0.057)			
$\log \text{MacroAttn10K}_{it}$				0.553	0.542	0.518
				(0.010)	(0.010)	(0.008)
Firm FE						✓
Sector x Time FE					✓	✓
N	92	92	92	46,612	46,590	45,458
R^2	0.118	0.140	0.675	0.561	0.639	0.867

Notes: The "Stem" MacroAttention is constructed using the methods described in Appendix B.6.2. The "10K" MacroAttention is our baseline measure from Section 2.2. In the first three columns, standard errors are HAC robust with a bandwidth (Bartlett kernel) of four quarters. In the second three columns, standard errors are double-clustered by year and firm ID.

Table B.13: Selected Summary Statistics of Firm Micro-Data

Code	Name	1990 Sales millions	1990 Sales share	1990 Employees thousands	1990 Employees share	2010 Sales millions	2010 Sales share	2010 Employees thousands	2010 Employees share
11	Agriculture, Forestry, Fishing and Hunting	7935.69	0.22	117.03	0.54	16028.05	0.13	157.12	0.49
21	Mining, Quarrying, and Oil and Gas Extraction	143716.64	3.95	557.43	2.59	728283.11	5.88	1020.34	3.18
23	Construction	20221.05	0.56	79.43	0.37	76357.58	0.62	249.05	0.78
42	Wholesale Trade	92141.14	2.53	316.32	1.47	220566.67	1.78	286.52	0.89
44	Retail Trade (I)	90746.30	2.49	697.08	3.24	755083.01	6.10	1802.16	5.61
45	Retail Trade (II)	71844.36	1.98	606.25	2.82	103823.53	0.84	393.61	1.23
48	Transportation and Warehousing (I)	202207.84	5.56	1278.65	5.94	490525.14	3.96	1484.93	4.62
49	Transportation and Warehousing (II)	22033.22	0.61	337.61	1.57	60473.08	0.49	505.72	1.57
53	Real Estate and Rental and Leasing	29186.20	0.80	335.31	1.56	97644.47	0.79	414.75	1.29
54	Professional, Scientific, and Technical Services	39096.89	1.07	367.54	1.71	142888.17	1.15	878.65	2.74
56	Administrative and Support and Waste Management and Remediation Services	30757.47	0.85	1055.62	4.91	97139.76	0.78	1432.94	4.46
61	Educational Services	2830.78	0.08	21.45	0.10	12350.02	0.10	93.87	0.29
62	Health Care and Social Assistance	14676.48	0.40	278.49	1.29	113475.92	0.92	874.35	2.72
71	Arts, Entertainment, and Recreation	4299.00	0.12	77.61	0.36	14592.82	0.12	140.96	0.44
72	Accommodation and Food Services	28474.41	0.78	741.09	3.45	126452.99	1.02	1949.83	6.07
81	Other Services (except Public Administration)	818.90	0.02	13.61	0.06	6028.48	0.05	71.80	0.22
99	Nonclassifiable Establishments	72587.03	2.00	402.04	1.87	337554.38	2.73	957.49	2.98
311	Food Manufacturing	89888.93	2.47	393.90	1.83	280503.94	2.26	773.89	2.41
312	Beverage and Tobacco Product Manufacturing	81514.55	2.24	527.06	2.45	323117.33	2.61	1128.77	3.51
313	Textile Mills	8281.43	0.23	109.85	0.51	1698.84	0.01	14.17	0.04
314	Textile Product Mills	1642.40	0.05	15.52	0.07	6512.22	0.05	31.62	0.10
315	Apparel Manufacturing	18195.52	0.50	179.87	0.84	50193.90	0.41	204.02	0.64
316	Leather and Allied Product Manufacturing	3220.64	0.09	18.56	0.09	8007.68	0.06	26.47	0.08
321	Wood Product Manufacturing	17079.99	0.47	94.34	0.44	32329.60	0.26	79.89	0.25

Code	Name	1990 Sales millions	1990 Sales share	1990 Employees thousands	1990 Employees share	2010 Sales millions	2010 Sales share	2010 Employees thousands	2010 Employees share
322	Paper Manufacturing	80736.02	2.22	485.96	2.26	137924.31	1.11	366.40	1.14
323	Printing and Related Support Activities	6959.89	0.19	64.35	0.30	17960.41	0.14	98.55	0.31
324	Petroleum and Coal Products Manufacturing	707106.33	19.44	1054.61	4.90	2666606.41	21.53	1718.55	5.35
325	Chemical Manufacturing	376182.22	10.34	2146.13	9.98	1234732.55	9.97	2362.52	7.36
326	Plastics and Rubber Products Manufacturing	23886.12	0.66	206.51	0.96	35081.57	0.28	139.19	0.43
327	Nonmetallic Mineral Product Manufacturing	22485.57	0.62	190.56	0.89	72492.53	0.59	242.54	0.76
331	Primary Metal Manufacturing	81200.98	2.23	438.98	2.04	308524.32	2.49	883.98	2.75
332	Fabricated Metal Product Manufacturing	34872.77	0.96	296.50	1.38	51755.70	0.42	172.89	0.54
333	Machinery Manufacturing	126838.80	3.49	823.90	3.83	303757.26	2.45	1060.24	3.30
334	Computer and Electronic Product Manufacturing	169692.71	4.66	1377.51	6.40	575557.43	4.65	1976.28	6.15
335	Electrical Equipment, Appliance, and Component Manufacturing	48369.82	1.33	447.71	2.08	96056.38	0.78	394.72	1.23
336	Transportation Equipment Manufacturing	553821.60	15.22	3359.30	15.62	1210603.25	9.77	3201.16	9.97
337	Furniture and Related Product Manufacturing	2516.33	0.07	23.81	0.11	15259.16	0.12	68.65	0.21
339	Miscellaneous Manufacturing	16722.13	0.46	143.46	0.67	63225.24	0.51	259.74	0.81
511	Publishing Industries (except Internet)	18829.43	0.52	149.15	0.69	42552.96	0.34	170.81	0.53
512	Motion Picture and Sound Recording Industries	13821.56	0.38	53.44	0.25	36968.97	0.30	112.21	0.35
515	Broadcasting (except Internet)	23400.47	0.64	187.74	0.87	189150.10	1.53	435.67	1.36
517	Telecommunications	158686.32	4.36	1045.88	4.86	1107220.16	8.94	3041.71	9.47
518	Data Processing, Hosting, and Related Services	1032.31	0.03	9.77	0.05	74483.99	0.60	273.40	0.85
519	Other Information Services	77053.51	2.12	384.84	1.79	45026.66	0.36	168.24	0.52
Total		3637612	100	21512	100	12386570	100	32120	100
USA		5963000		125840		14990000		153890	
Compustat share of USA			61.00		17.09		82.63		20.87

Table B.14: Comparison of TFP Measures

	(1)	(2)	(3)
	Outcome: Cost-Share TFP		
OP TFP	1.117	1.141	1.025
	(0.009)	(0.008)	(0.006)
Firm FE		✓	✓
Sector x Time FE			✓
N	68,825	68,821	67,395
R^2	0.649	0.721	0.977

Notes: "Cost-Share TFP" is our baseline measure of $\log \hat{\theta}_{it}$ used in the main analysis. "OP TFP" is the alternative measure based on the method of Olley and Pakes (1996). Standard errors are double-clustered by year and firm ID.

Table B.15: Changing Macro Attention in Response to News

Response	Poorly	Well
Much more likely	44.96	9.77
Somewhat more likely	30.91	19.42
No change	12.56	8.67
Somewhat less likely	7.16	53.35
Much less likely	4.40	8.79
Total	100.00	100.00

Notes: Data are from the Coibion, Gorodnichenko, and Kumar (2018a) survey, as described in Appendix B.7.

Table B.16: Profit-Function Curvature and Attention to Macro Variables

Panel 1: Back-cast Error (Absolute Value)

Variable	Inflation		GDP Growth		Unemployment	
ProfitCurv$_{it}$	-1.172	-0.328	-0.072	-0.042	0.075	0.121
	(0.195)	(0.091)	(0.041)	(0.041)	(0.072)	(0.077)
Controls		✓		✓		✓
R^2	0.024	0.457	0.001	0.006	0.001	0.032
N	3,153	3,145	1,257	1,237	1,257	1,256

Panel 2: Keeping Track

Variable	Inflation		GDP Growth		Unemployment	
ProfitCurv$_{it}$	0.170	0.050	0.015	0.019	-0.005	-0.022
	(0.039)	(0.029)	(0.022)	(0.028)	(0.035)	(0.081)
Controls		✓		✓		✓
R^2	0.032	0.332	0.000	0.074	0.000	0.065
N	1,255	1,235	1,255	1,235	1,255	1,235

Notes: Standard errors are clustered by three-digit industry. Data are from the Coibion, Gorodnichenko, and Kumar (2018a) survey, as described in Appendix B.7.

Table B.17: Profit Curvature in the Data

Summary Statistics

Mean	Quantiles 5	25	50	75	95
0.280	0.020	0.05	0.12	0.28	1.00

Correlates

	Variable	Norm. coef.	t-stat	R^2
Firm	Frequency of price review	-0.106	-7.80	0.011
	log output	-0.066	-9.43	0.015
	Firm age	-0.117	-10.17	0.014
	Employment	-0.122	-7.19	0.015
	Labor share	-0.138	-7.98	0.020
	Number of competitors	0.130	6.81	0.017
Manager	log income	0.015	0.55	0.000
	Some or more college	0.043	1.87	0.002
	Tenure at firm	-0.117	-5.73	0.014
	Tenure in industry	-0.058	-2.33	0.003
	Manager age	-0.091	-3.25	0.008

Notes: The top panel gives summary statistics. The bottom panel gives normalized regression coefficients for a number of possible correlates. Standard errors, used to calculate the t-statistics, are clustered by three-digit industry.

Table B.18: Curvature and Inflation BCE in Waves 1 versus 4

	Outcome: absolute Inflation BCE			
Wave	1		4	
ProfitCurv$_{it}$	-1.172	-0.328	-0.884	-0.330
	(0.195)	(0.091)	(0.181)	(0.126)
Controls		✓		✓
R^2	0.024	0.457	0.033	0.268
N	3,153	3,145	1,257	1,256

Notes: Standard errors are clustered by three-digit industry.

Figure B-1: Frequency over Time of Each Word in MacroAttention (Part I)

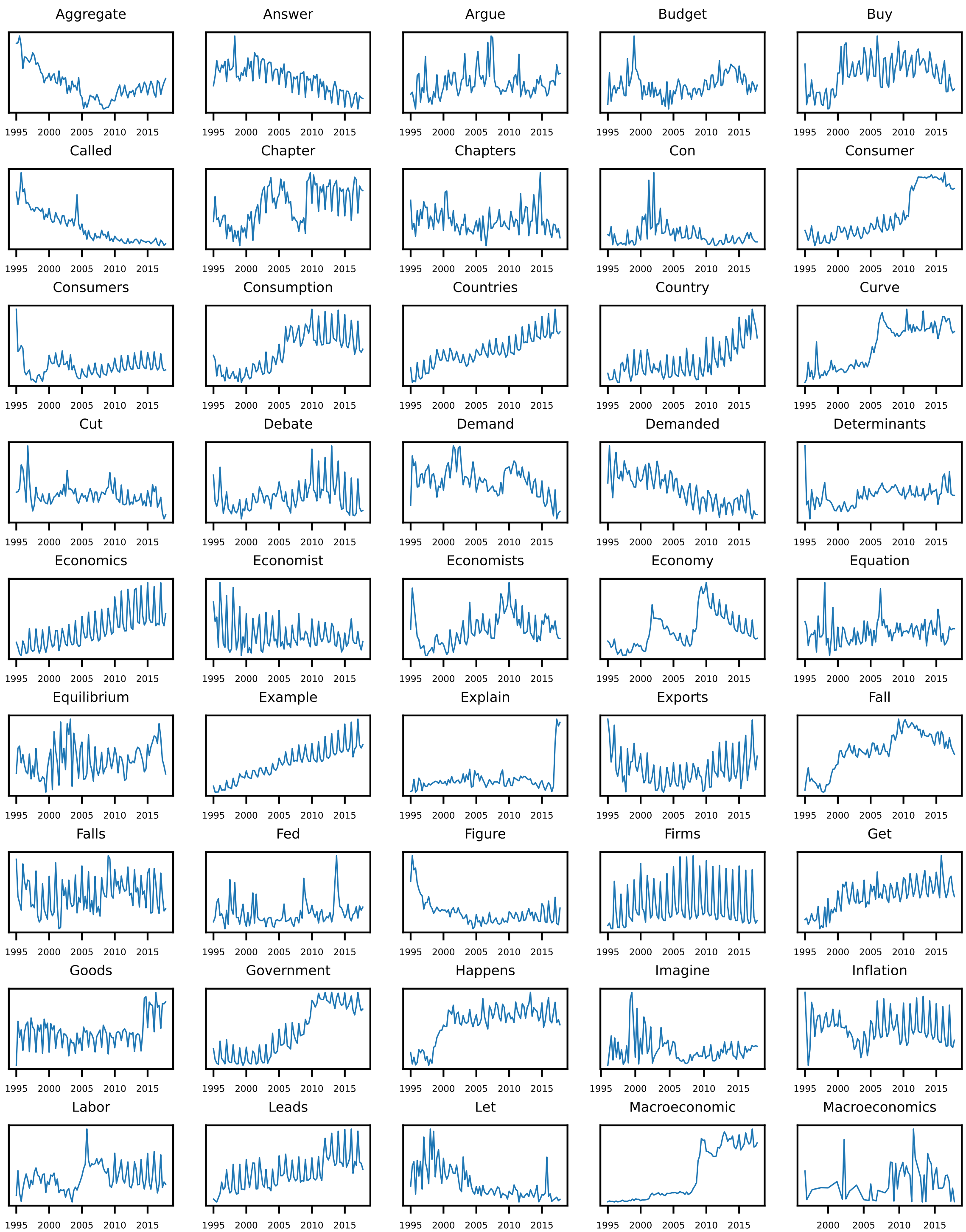


Figure B-1: Frequency over Time of Each Word in MacroAttention (Part II)

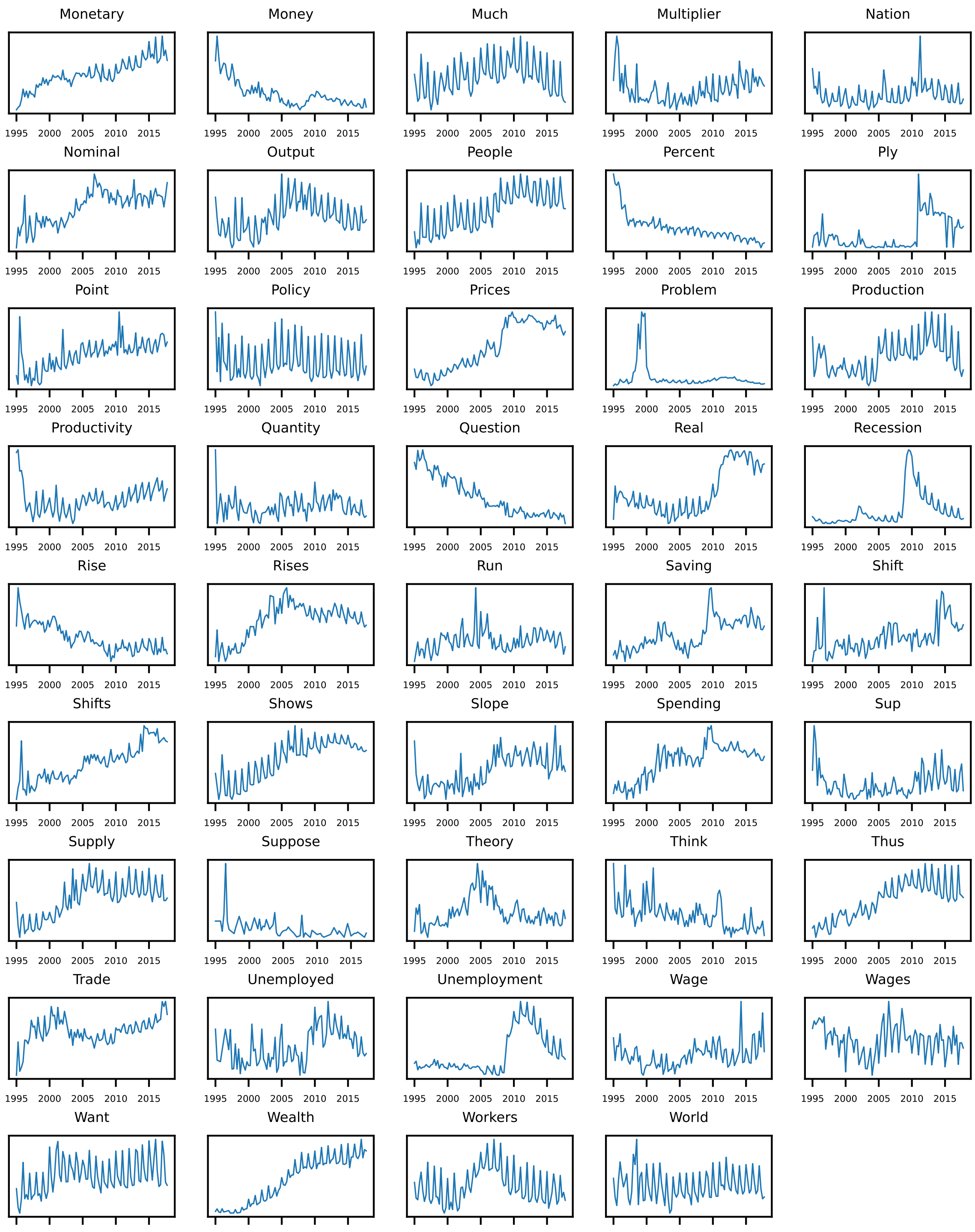
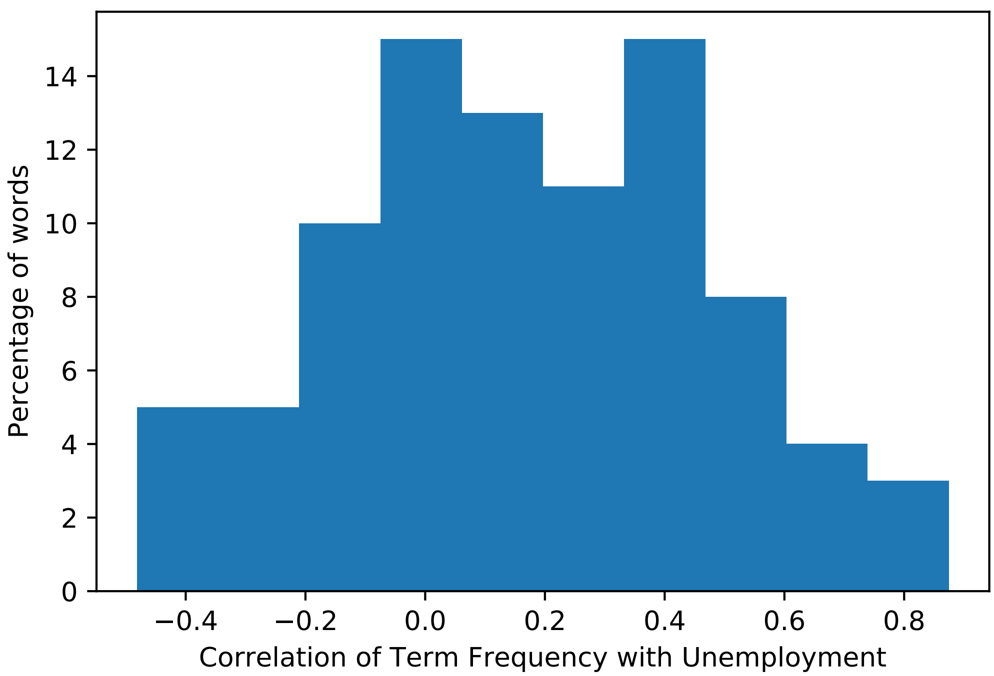
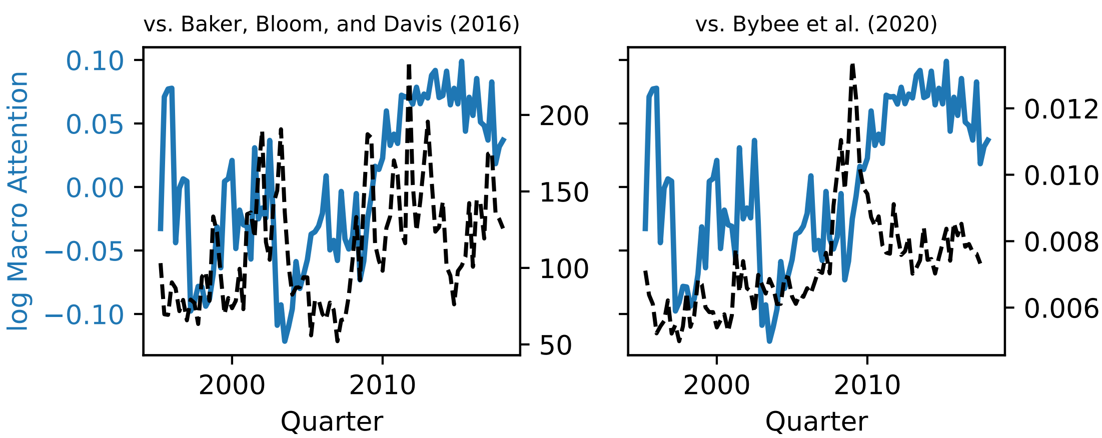


Figure B-2: Correlations with Unemployment by Word



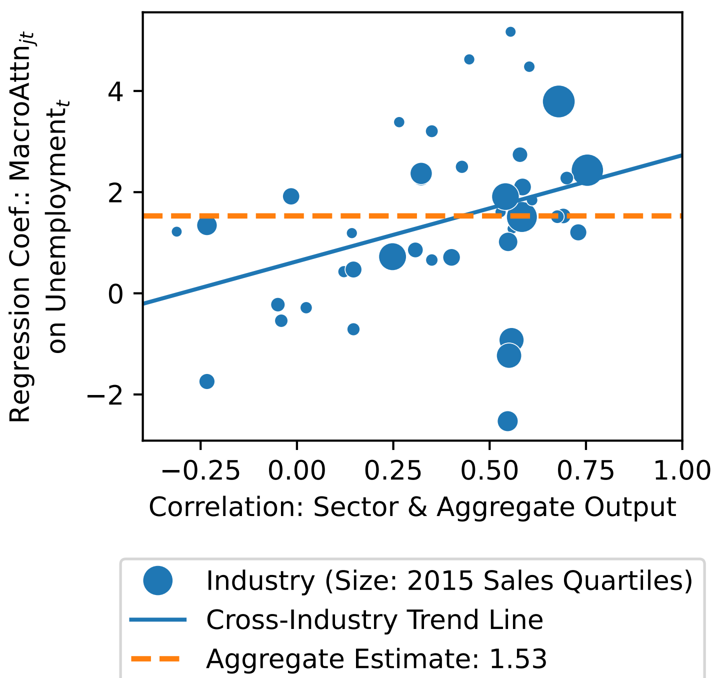
Notes: Correlations are calculated at the quarterly frequency.

Figure B-3: Relationship of Macro Attention to News Indices



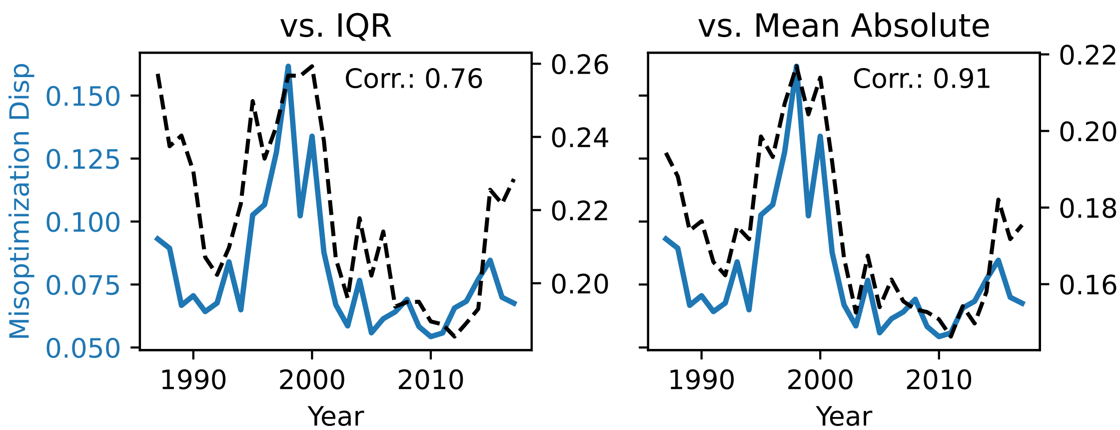
Notes: The blue line, measured by the left axis of each plot, is the log of Macro Attention. The left panel is the news index of Baker, Bloom, and Davis (2016). The right panel is the sum of five macroeconomic topics from the Bybee, Kelly, Manela, and Xiu (2021b) Structure of News dataset: "economic growth," "Federal Reserve," "financial crisis", "recession," and "macroeconomic data."

Figure B-4: Industry-Specific Cyclicality of Macro Attention



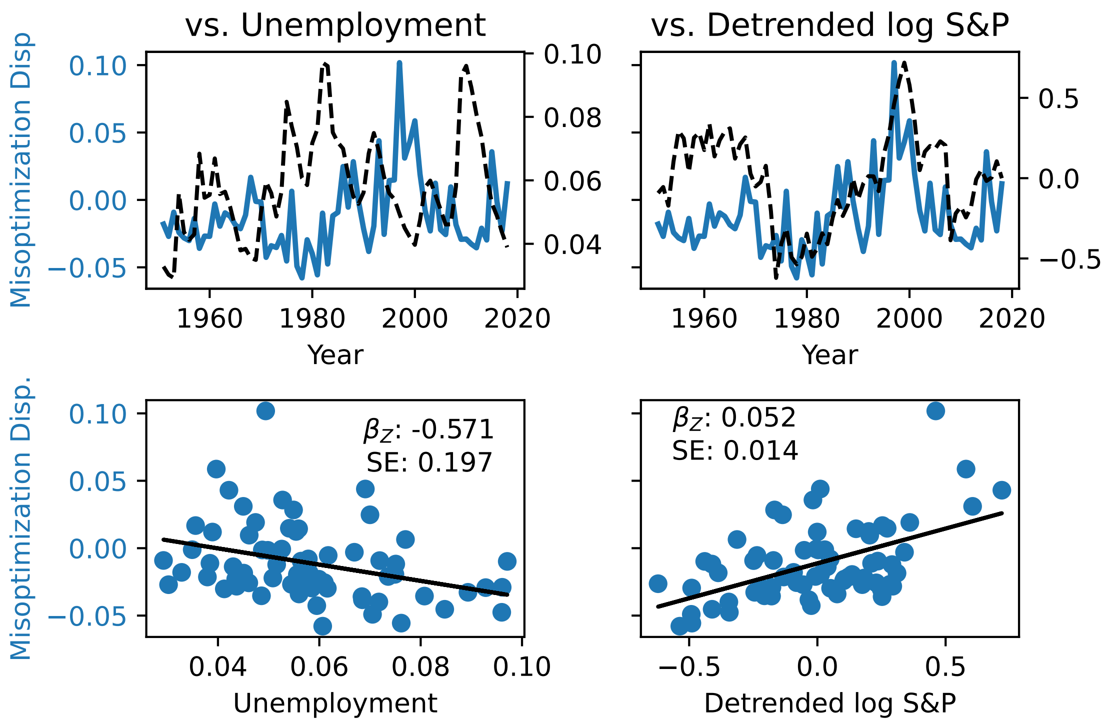
Notes: The horizontal axis is the correlation of sectoral and aggregate nominal GDP, calculated as described in Section 2.2.3 of the main text. The vertical axis is the regression coefficient of log sectoral macro attention, net of quarterly fixed effects, on the US unemployment rate. The dashed orange line is the estimate of the same using aggregate Macro Attention. The dots are sized based on quartiles of total sales in Compustat in 2015. The blue solid line is a cross-industry linear regression line.

Figure B-5: Relationship of Misoptimization Dispersion with Other Statistics



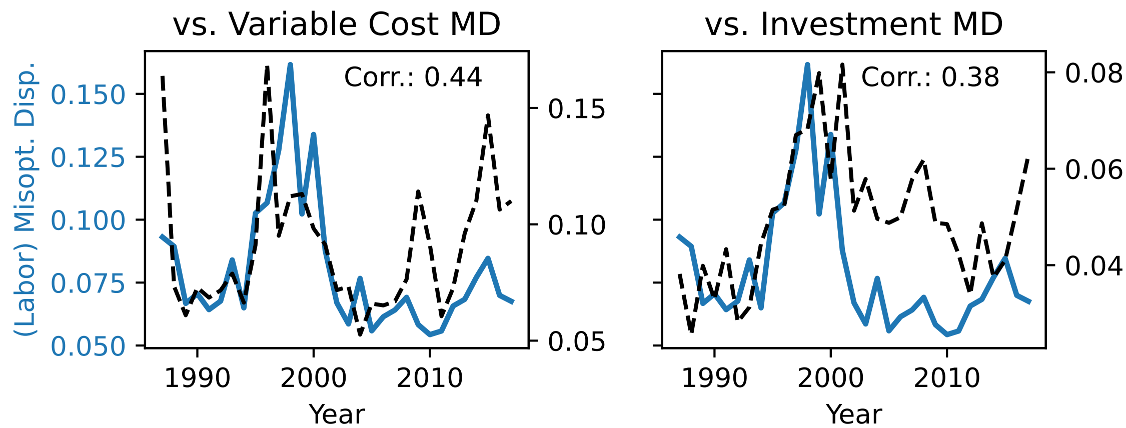
Notes: The blue line, measured by the left axis of each plot, is Misoptimization Dispersion as defined in Section 2.5. The black dashed line on the left is the (optimal-sale-weighted) interquartile range of the distribution of $\hat{u}_{it}$. The black dashed line on the right is the (optimal-sale-weighted) average of $|\hat{u}_{it}|$.

Figure B-6: Misoptimization Dispersion is Pro-Cyclical, Long Sample



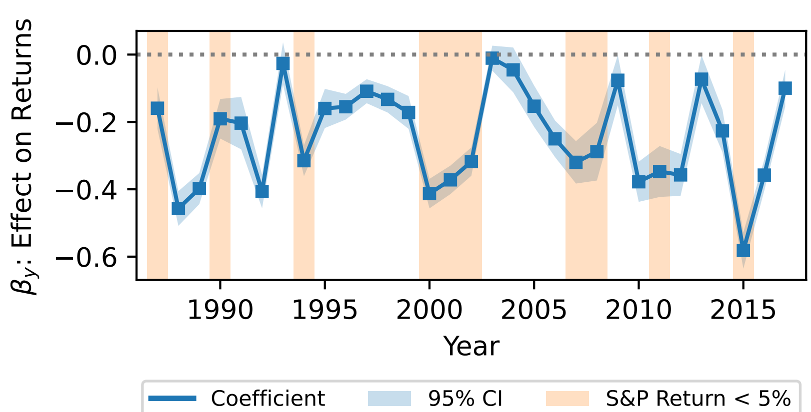
Notes: This Figure replicates Figure 2-3 using our long-sample (1950-2018) calculation of Misoptimization Dispersion, described in Section 2.5.2 ("Robustness to Studied Time Period."). The top two panels plot Misoptimization Dispersion (blue line, left axis) along with, respectively, unemployment and the linearly detrended S&P 500 price (black dashed lines, right axis). Because of the composition adjustment, the metric can be negative. The bottom two panels are scatterplots of Misoptimization Dispersion versus each macroeconomic aggregate. The black solid line is the linear regression fit. The standard errors are HAC robust based on a Bartlett kernel with a three-year bandwidth.

Figure B-7: Relationship of Misoptimization Dispersion Across Inputs



Notes: The blue line, measured by the left axis of each plot, is Misoptimization Dispersion as defined in Section 2.5. The black dashed line on the left is Misoptimization Dispersion for total variable cost expenditures. The black dashed line on the right is the same for investment rates.

Figure B-8: Time-Varying Punishments of Misoptimizations

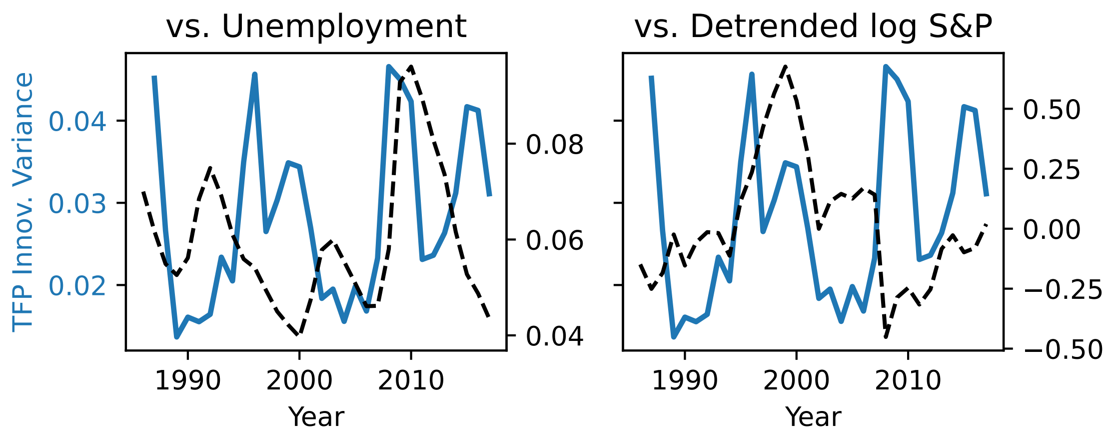


Notes: This Figure shows the estimates β_y from the following model of year-specific effects of misoptimizations on stock returns:

$$R_{it} = \sum_y \beta_y \cdot \hat{u}_{it}^2 \cdot \mathbb{I}[t = y] + \chi_{j(i),t} + \epsilon_{it}$$

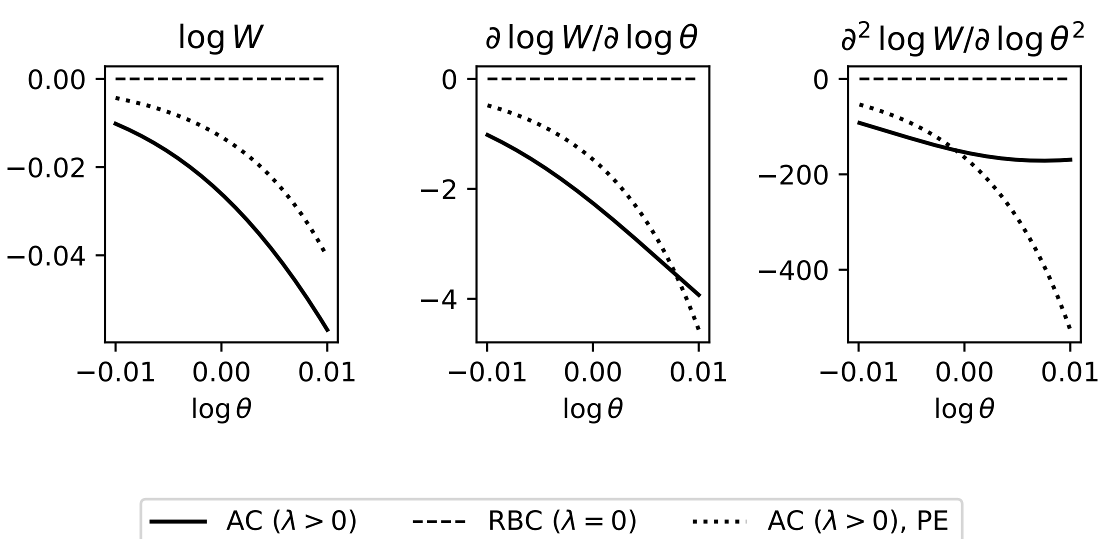
The blue squares are the point estimates $\hat{\beta}_y$ and the shaded bands are 95% confidence intervals (based on standard errors clustered by firm and year). The orange shading indicates years in which the S&P 500 return was less than 5%.

Figure B-9: Relationship of TFP Innovation Variance with Macro Variables



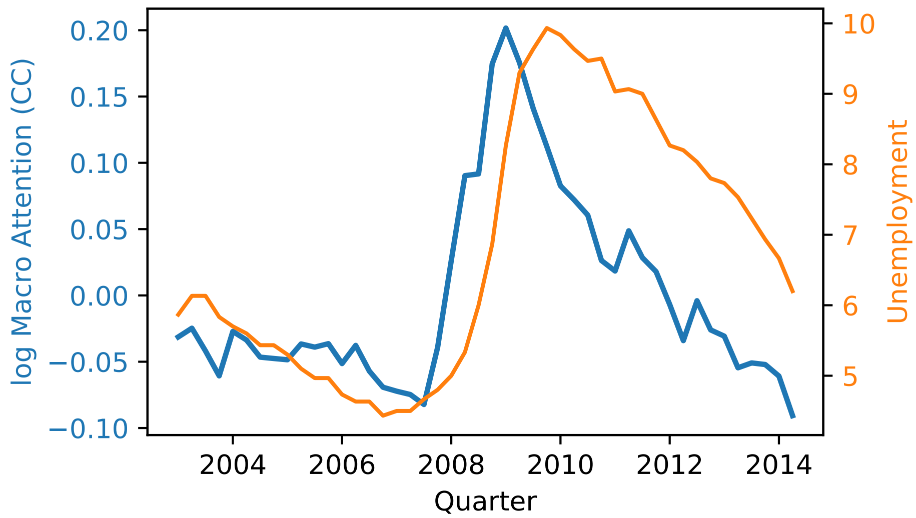
Notes: "TFP Innovations" are the residuals from an AR(1) model for $\log \theta_{it}$ with firm and sector-by-time fixed effects, as described in Section 2.5.5 of the main text. We calculate their average, consistent with our main calculations of Misoptimization Dispersion, using optimal-sales-weights.

Figure B-10: The Attention Wedge and Its Derivatives



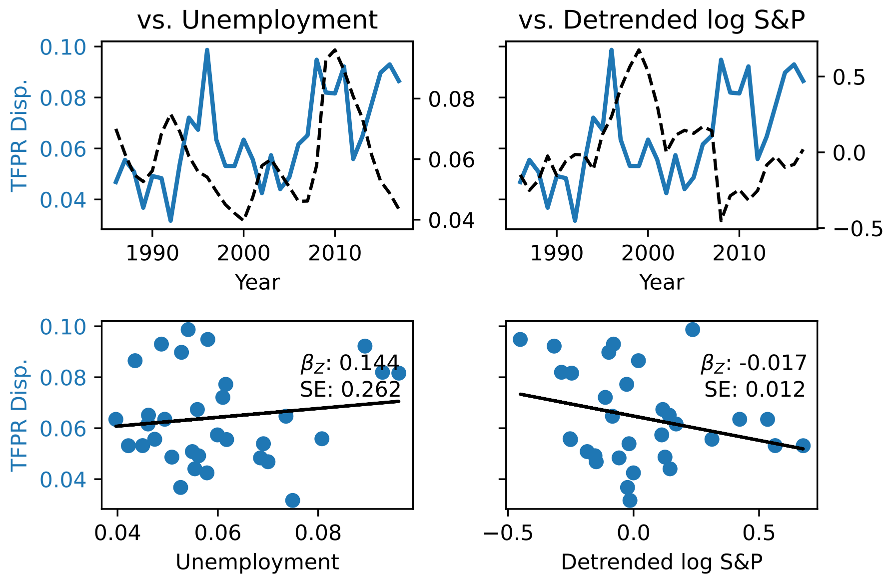
Notes: The panels respectively show the level, first derivative, and second derivative of the log attention wedge in the log state. The "Partial Equilibrium" thought experiment, plotted in each figure as a dotted line, is for firms to best-reply to the output and wages of the counterfactual RBC equilibrium.

Figure B-11: Conference-Call Macro Attention and Unemployment

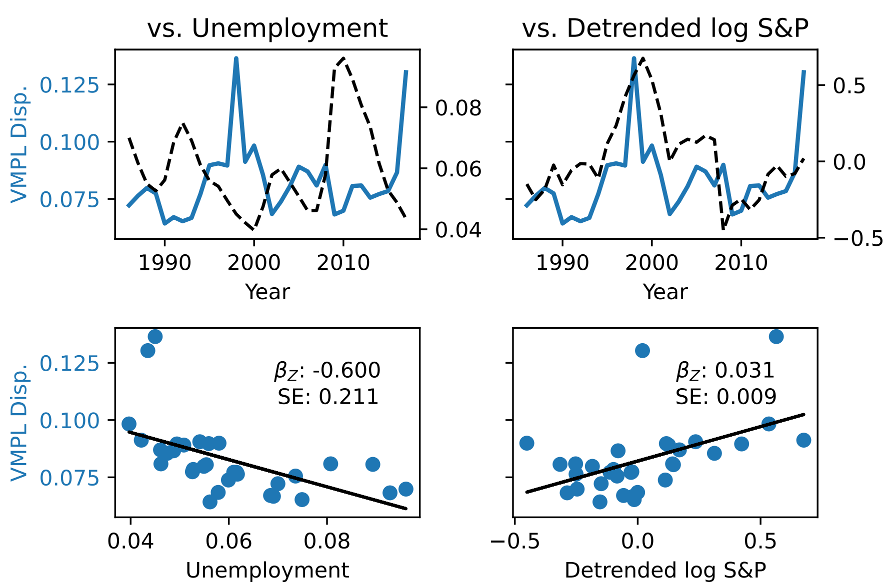


Notes: The left axis and blue line show our estimate of Macro Attention based on conference-call data, in log units net of seasonal trends. The right axis and orange line show the US unemployment rate.

Figure B-12: The Cyclical Behavior of TFPR and VMPL Dispersion



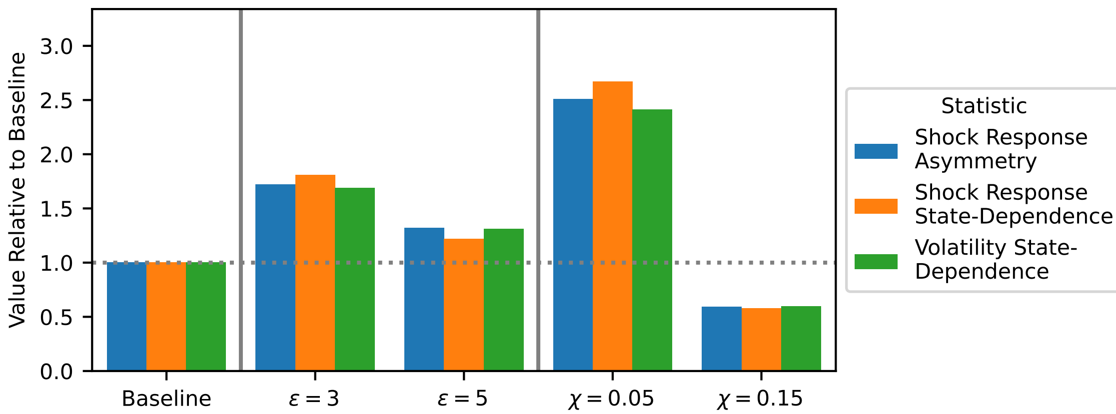
(a) TFPR Dispersion



(b) VMPL Dispersion

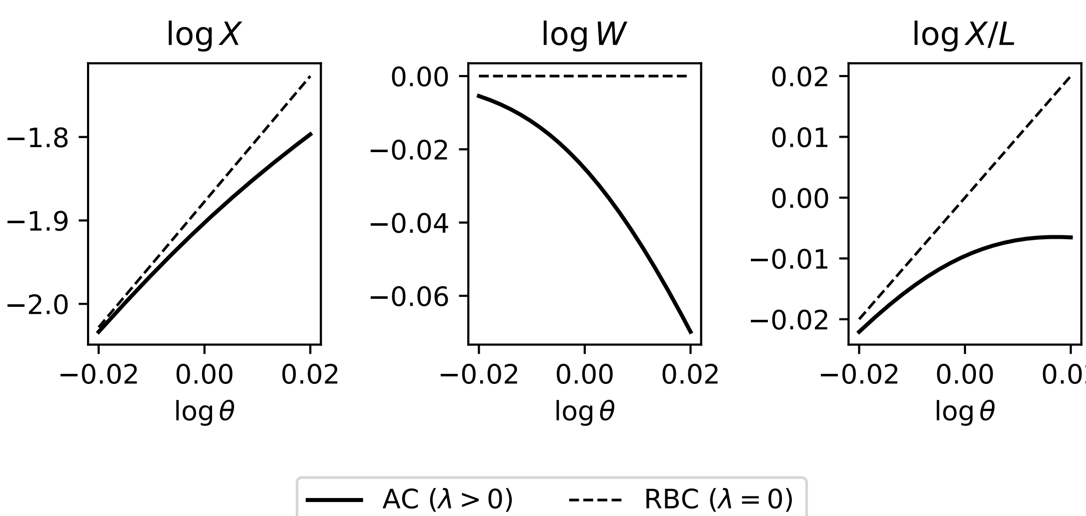
Notes: Each half of this Figure replicates Figure 2-3 using our measures of TFPR Dispersion (top) and VMPL Dispresion (bottom). Variable construction is defined in Appendix B.6.3. Because of the composition adjustment, the variance metrics can be negative. The standard errors for the linear fits are HAC robust baesd on a Bartlett kernel with a three-year bandwidth.

Figure B-13: Robustness of Numerical Results to Parameter Choices



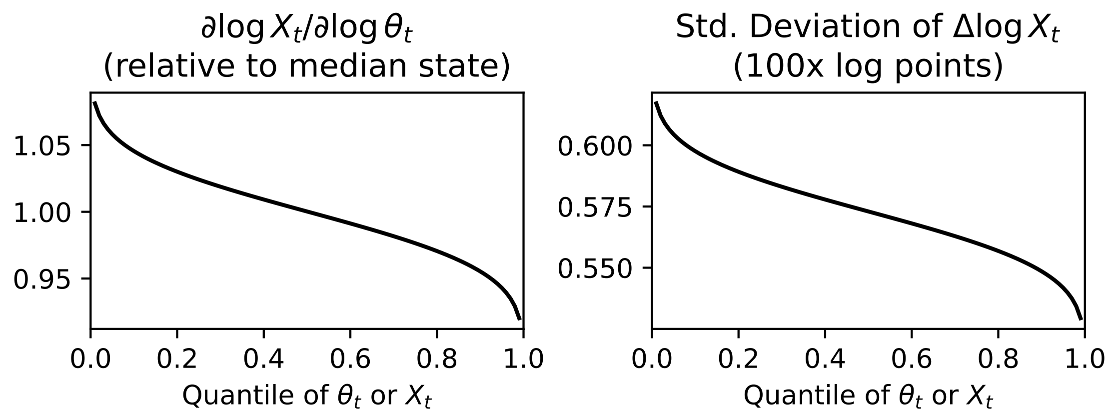
Notes: We re-calibrate the model under each indicated parameter choice. The outcomes and their interpretation are described in Appendix B.5.1.

Figure B-14: Output, Wedge, and Labor Productivity with GHH Preferences



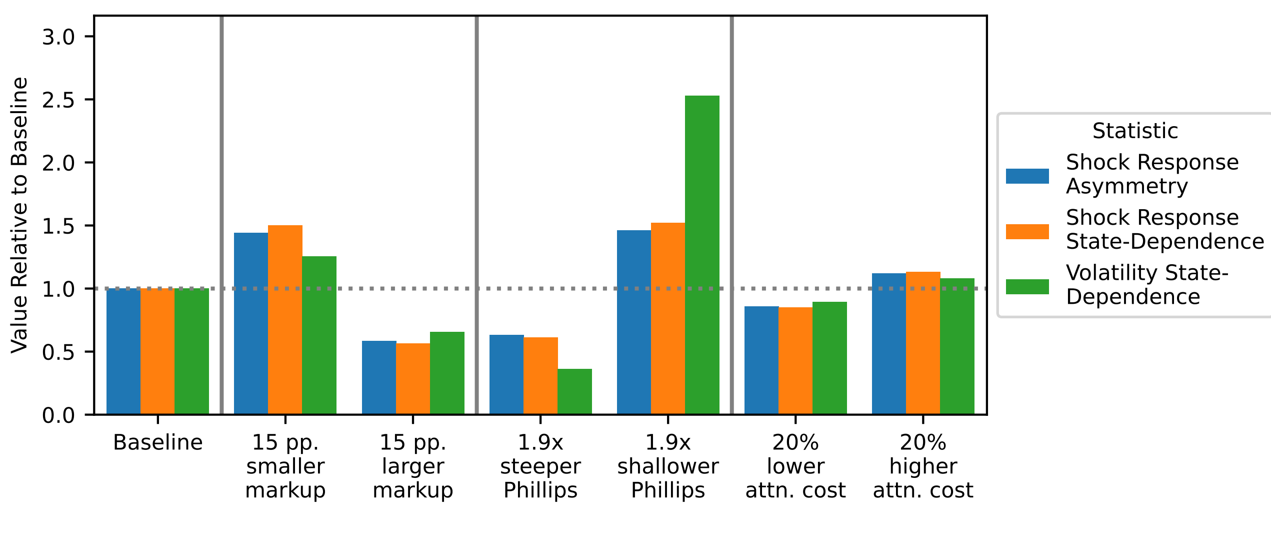
Notes: This recreates Figure 2-4 from the main analysis in the variant model with Greenwood, Hercowitz, and Huffman (1988) preferences, described in Appendix B.5.

Figure B-15: Asymmetric Shock Response and Stochastic Volatility with GHH Preferences



Notes: This recreates Figure 2-5 from the main analysis in the variant model with Greenwood, Hercowitz, and Huffman (1988) preferences, described in Appendix B.5.

Figure B-16: Predictions in Counterfactual Scenarios



Notes: The scenarios are described in the main text. The outcomes are the same as in Figure B-13, and are described in Appendix B.5.1.

Appendix C

Appendix to Strategic Mistakes

C.1 Omitted Proofs

C.1.1 Proof of Lemma 2

Proof. Define $\tilde{u} = \frac{u}{\lambda}$. Inequality 107 can be re-expressed as:

$$\tilde{u}(x', X', \theta) + \tilde{u}(x, X, \theta) \geq \tilde{u}(x', X, \theta) + \tilde{u}(x, X', \theta) \tag{678}$$

which is the statement that $\tilde{u}$ is a supermodular function in (x, X). By Topkis' Characterization Theorem (see *e.g,* Milgrom and Roberts, 1990), when $\tilde{u}$ is twice continuously differentiable in (x, X), this is equivalent to the statement that $\tilde{u}_{xX}(x, X, \theta) \geq 0$. As we have assumed that u and λ are both twice continuously differentiable in (x, X), Inequality 107 is equivalent to:

$$\tilde{u}_{xX}(x, X, \theta) = \frac{u_{xX}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_X(X,\theta)}{\lambda(X,\theta)}}{\lambda(X, \theta)} \geq 0 \tag{679}$$

Inequality 108 can be re-expressed as:

$$\tilde{u}(x', X + \alpha, \theta) + \tilde{u}(x - \alpha, X, \theta) \leq \tilde{u}(x, X + \alpha, \theta) + \tilde{u}(x' - \alpha, X, \theta) \tag{680}$$

Define $f(y, \gamma; \theta, X) = \tilde{u}(y + \gamma, X + \gamma, \theta)$. Set $y' = x' - \alpha$, $y = x - \alpha$, $\gamma' = \alpha$ and $\gamma = 0$ and observe that $y' \geq y$ and $\gamma' \geq \gamma$. Inequality 680 is equivalent to:

$$f(y', \gamma'; \theta, X) + f(y, \gamma; \theta, X) \leq f(y, \gamma'; \theta, X) + f(y', \gamma; \theta, X) \tag{681}$$

Which is equivalent to submodularity of $f(\cdot\,; \theta, X)$ in (y, γ). Again by Topkis' Characterization Theorem, and by twice continuous differentiability of f in (y, γ), this is equivalent to:

$$f_{y\gamma}(y, \gamma; \theta, X) = \tilde{u}_{xx}(y + \gamma, X + \gamma, \theta) + \tilde{u}_{xX}(y + \gamma, X + \gamma, \theta) \leq 0 \tag{682}$$

Moreover, $\tilde{u}_{xx} = \frac{u_{xx}}{\lambda}$. Thus, Inequality 108 is equivalent to:

$$-\frac{u_{xx}(x, X, \theta)}{\lambda(X, \theta)} \geq \tilde{u}_{xX}(x, X, \theta) = \frac{u_{xX}(x, X, \theta) - u_x(x, X, \theta)\frac{\lambda_X(X,\theta)}{\lambda(X,\theta)}}{\lambda(X, \theta)} \tag{683}$$

Combining Inequalities 679 and 683 and multiplying by $\lambda > 0$, we obtain the claimed result. □

C.1.2 Proof of Lemma 4

To establish the result, as the entropy kernel has derivative $\phi'(x) = 1 + \log x$ and $\phi'(x) = x$, it is sufficient to show the following:

Lemma 13. *$\mathcal{F}$, the class of functions satisfying quasi-MLRP, contains* $\{\log(\cdot), Id(\cdot)\}$.

Proof. To see that quasi-MLRP is satisfied for $f(x) = \log x$ (and $1 + \log x$), the required condition (Equation 114) becomes:

$$\left(\frac{g'(x')}{g'(x)} \geq \frac{g(x')}{g(x)} \ \forall x' \geq x \right) \implies g' \succeq_{FOSD} g \tag{684}$$

The left-hand side of this implication is simply the MLRP property. Moreover, MLRP implies FOSD. We now prove that $f(x) = x$ satisfies quasi-MLRP. This requires us to prove that for any two distributions $g', g \in \Delta(\mathcal{X})$:

$$\left(g'(x') - g'(x) \geq g(x') - g(x) \ \forall x' \geq x \right) \implies g' \succeq_{FOSD} g \tag{685}$$

To do this, we first prove a technical lemma, which may be of future use for characterizing other functions that satisfy quasi-MLRP:

Lemma 14. *For any two distributions $g', g \in \Delta(\mathcal{X})$, the following holds:*

$$\begin{gathered}\left(f(g'(x')) - f(g'(x)) \geq f(g(x')) - f(g(x)) \ \forall x' \geq x \right) \implies \\ \left(\frac{\int_x^{\overline{x}} [f(g'(\tilde{x})) - f(g(\tilde{x}))] \, \mathrm{d}\tilde{x}}{\overline{x} - x} \geq \frac{\int_{\underline{x}}^{x} [f(g'(\tilde{x})) - f(g(\tilde{x}))] \, \mathrm{d}\tilde{x}}{x - \underline{x}} \quad \forall x \in \mathcal{X} \right)\end{gathered} \tag{686}$$

Proof. To prove the required implication, we begin with the hypothesis:

$$f(g'(x')) - f(g'(x)) \geq f(g(x')) - f(g(x)) \ \forall x' \geq x \tag{687}$$

Which can be rewritten as:

$$f(g'(x')) + f(g(x)) \geq f(g(x')) + f(g'(x)) \ \forall x' \geq x \tag{688}$$

We now integrate from $\underline{x}$ to x' with respect to x to obtain the inequality:

$$(x' - \underline{x}) f(g'(x')) + \int_{\underline{x}}^{x'} f(g(x)) \,\mathrm{d}x \geq (x' - \underline{x}) f(g(x')) + \int_{\underline{x}}^{x'} f(g'(x)) \,\mathrm{d}x \tag{689}$$

Imposing $x' = x$ we obtain:

$$(x - \underline{x}) \left[f(g'(x)) - f(g(x)) \right] \geq \int_{\underline{x}}^{x} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \,\mathrm{d}\tilde{x} \tag{690}$$

Applying the same procedure but this time integrating from x to $\overline{x}$ with respect to x' and evaluate at $x' = x$ to obtain this inequality:

$$\int_{x}^{\overline{x}} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \,\mathrm{d}\tilde{x} \geq (\overline{x} - x) \left[f(g'(x)) - f(g(x)) \right] \tag{691}$$

Combining our two inequalities we obtain the required one:

$$\frac{\int_{x}^{\overline{x}} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \,\mathrm{d}\tilde{x}}{\overline{x} - x} \geq \frac{\int_{\underline{x}}^{x} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \,\mathrm{d}\tilde{x}}{x - \underline{x}} \quad \forall x \in \mathcal{X} \tag{692}$$

Which completes the proof. □

Thus, if it can be established that:

$$\begin{aligned} &\left(\frac{\int_{x}^{\overline{x}} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \,\mathrm{d}\tilde{x}}{\overline{x} - x} \geq \frac{\int_{\underline{x}}^{x} \left[f(g'(\tilde{x})) - f(g(\tilde{x})) \right] \,\mathrm{d}\tilde{x}}{x - \underline{x}} \quad \forall x \in \mathcal{X} \right) \\ &\implies g' \succeq_{FOSD} g \end{aligned} \tag{693}$$

then we will have established that function f satisfies quasi-MLRP.

We now use this to prove that $f(x) = x$ satisfies quasi-MLRP. Plugging in to the derived integral condition, we obtain:

$$\frac{G(x) - G'(x)}{\overline{x} - x} \geq \frac{G'(x) - G(x)}{x - \underline{x}} \quad \forall x \in \mathcal{X} \tag{694}$$

Re-arranging this:

$$G(x) \geq G'(x) \quad \forall x \in \mathcal{X} \tag{695}$$

which is the definition that $g' \succeq_{FOSD} g$. This completes the proof and establishes that quasi-MLRP is a strict weakening of MLRP. □

C.1.3 Proof of Proposition 10

Proof. This follows immediately from the step proving the monotonicity and discounting conditions in Theorem 1. Note that this invokes only Assumptions 1 and 3. □

C.1.4 Proof of Theorem 1

Proof. We first study the problem of a single agent i who is best replying to the conjecture that the law of motion of the aggregate is $\hat{X} : \Theta \to \mathbb{R}$. See that this agent faces the problem:

$$\mathcal{P}^*(\hat{X}) = \arg\max_{P \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta) \, \mathrm{d}P(x|\theta) \, \pi(\theta) - c(P, \hat{X}) \tag{696}$$

First, let us examine the set of stochastic choice rules:

$$\mathcal{P} = \{P : \Theta \to \Delta(\mathcal{X})\} = \prod_{\theta \in \Theta} \Delta(\mathcal{X}) \tag{697}$$

See that $\Delta(\mathcal{X})$ is compact as $\mathcal{X}$ is compact. It therefore follows by finiteness of Θ that $\mathcal{P}$ is compact.

Define $k : \mathcal{P} \times \mathcal{B} \to \bar{\mathbb{R}}$, where $\mathcal{B} = \{\hat{X} : \Theta \to \mathbb{R}\}$ as:

$$k(P, \hat{X}) = \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta) \, \mathrm{d}P(x|\theta) \, \pi(\theta) - c(P, \hat{X}) \tag{698}$$

As ϕ is strictly convex and u is bounded, it is without loss of optimality to restrict to optimizing over the set of stochastic choice rules with density bounded above by some $M \in \mathbb{R}$, $\mathcal{P}_M$. This is a closed set, which is a subset of a compact set $\mathcal{P}$, and therefore also compact. Moreover, k is continuous in P, by continuity of u and continuity of c over $\mathcal{P}_M$ for any M. Thus, by Weierstrass' theorem, there exists a maximum. Moreover, by strict convexity of $k(\cdot, \hat{X})$, it is unique. It immediately follows that in any equilibrium $P_i^* = P^* = \mathcal{P}^*(\hat{X})$ for all i and thus that there cannot exist asymmetric equilibria.

To show existence of an equilibrium it suffices to show that there exists a $\hat{X}$ such that:

$$\hat{X} = X \circ \mathcal{P}^*(\hat{X}) \tag{699}$$

To this end define the operator $T : \mathcal{B} \to \mathcal{B}$ such that:

$$T(\hat{X}) = X \circ \mathcal{P}^*(\hat{X}) \tag{700}$$

We wish to show that T has a fixed point. We will moreover prove that this fixed point is unique as under the stated assumptions we can prove that T is a contraction map. To this end, we wish to apply Blackwell's sufficient conditions for an operator to be a contraction. More specifically, if T operates on the space of bounded functions and is endowed with the sup norm, then the following are sufficient for T to be a contraction:

1. Monotonicity: $\hat{X}' \geq \hat{X} \implies T(\hat{X}') \geq T(\hat{X})$ for any $\hat{X}', \hat{X} \in \mathcal{B}$
2. Discounting: there exists $\beta \in (0,1)$ such that $T(\hat{X} + \alpha) \leq T(\hat{X}) + \beta\alpha$ for all $\alpha \in \mathbb{R}_+$ and any $\hat{X} \in \mathcal{B}$

Toward proving these properties, we first derive some necessary conditions for optimal stochastic choice. To this end, see that the stochastic choice program under an equilibrium conjecture $\hat{X}$ is given by:

$$\max_{p \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta)\, \mathrm{d}P(x|\theta)\, \pi(\theta) - \sum_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta))\, \mathrm{d}x\, \pi(\theta)\lambda(\hat{X}(\theta), \theta) \tag{701}$$

Take the optimal policy p and now consider a family or perturbations of p around actions $x, x' \in \mathcal{X}$ in state $\theta \in \Theta$ such that $p(x|\theta; \hat{X}), p(x'|\theta; \hat{X}) > 0$ by $\varepsilon > 0$ and $\delta \geq 0$ such that:

$$\begin{aligned} \tilde{p}(\tilde{x}|\theta; \hat{X}) &= p(\tilde{x}|\theta; \hat{X}) + \varepsilon,\ \ \tilde{x} \in [x', x' - \delta] \\ \tilde{p}(\tilde{x}|\theta; \hat{X}) &= p(\tilde{x}|\theta; \hat{X}) - \varepsilon,\ \ \tilde{x} \in [x, x - \delta] \end{aligned} \tag{702}$$

For p that has full support on $[x', x' - \delta]$, $[x, x - \delta]$, we have that $\tilde{p} \in \mathcal{P}$. Moreover, as u is continuous, if δ is sufficiently small, such a full-support perturbation is possible by the property that $p(x|\theta;\hat{X}), p(x'|\theta;\hat{X}) > 0$ and the fact that p is optimal.

Taking the derivative of the value of $\tilde{p}$ in ε and evaluating at $\varepsilon = 0$, we obtain the necessary optimality condition:

$$\begin{aligned}\int_{x'-\delta}^{x'} &\left[u(\tilde{x},\hat{X}(\theta),\theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta;\hat{X}))\pi(\theta)\lambda(\hat{X}(\theta),\theta)\right] \mathrm{d}\tilde{x} \\ &= \int_{x-\delta}^{x} \left[u(\tilde{x},\hat{X}(\theta),\theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta;\hat{X}))\pi(\theta)\lambda(\hat{X}(\theta),\theta)\right] \mathrm{d}\tilde{x}\end{aligned} \tag{703}$$

Normalizing both sides by $\delta > 0$, we obtain:

$$\begin{aligned}&\frac{\int_{x'-\delta}^{x'} \left[u(\tilde{x},\hat{X}(\theta),\theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta;\hat{X}))\pi(\theta)\lambda(\hat{X}(\theta),\theta)\right] \mathrm{d}\tilde{x}}{\delta} \\ &\qquad = \frac{\int_{x-\delta}^{x} \left[u(\tilde{x},\hat{X}(\theta),\theta)\pi(\theta) - \phi'(p(\tilde{x}|\theta;\hat{X}))\pi(\theta)\lambda(\hat{X}(\theta),\theta)\right] \mathrm{d}\tilde{x}}{\delta}\end{aligned} \tag{704}$$

Taking the limit of both sides as $\delta \to 0$, applying L'Hôpital's rule and Leibniz's rule we obtain:

$$u(x',\hat{X}(\theta),\theta) - \lambda(\hat{X}(\theta),\theta)\phi'(p(x'|\theta;\hat{X})) = u(x,\hat{X}(\theta),\theta) - \lambda(\hat{X}(\theta),\theta)\phi'(p(x|\theta;\hat{X})) \tag{705}$$

This condition is necessary for all $x, x' \in \mathcal{X}$ that have a positive density in state θ.

By the previous necessary condition and the supermodularity assumption (Assumption 1) we have that, for all $x' \geq x$ in the support of both stochastic choice rules, all θ, and any conjectures $\hat{X}$ and $\hat{X}'$ such that $\hat{X}' \geq \hat{X}$:

$$\begin{aligned}\phi'(p(x'|\theta;\hat{X}')) - \phi'(p(x|\theta;\hat{X}')) &= \frac{u(x',\hat{X}'(\theta),\theta)}{\lambda(\hat{X}'(\theta),\theta)} - \frac{u(x,\hat{X}'(\theta),\theta)}{\lambda(\hat{X}'(\theta),\theta)} \\ &\geq \frac{u(x',\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} - \frac{u(x,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} \\ &= \phi'(p(x'|\theta;\hat{X})) - \phi'(p(x|\theta;\hat{X}))\end{aligned} \tag{706}$$

We now need to check the cases where the stochastic choice rules do not have full support. Define the support in state θ under law of motion $\hat{X}$ as $\mathcal{X}(\theta,\hat{X}) = \mathrm{cl}_{\mathcal{X}}\{x \in \mathcal{X} : p^*(x|\theta,\mathcal{X}) > 0\}$. Let $\hat{x} \in \mathcal{X}(\theta,\hat{X})$, $\tilde{x} \in \mathcal{X}(\theta,\hat{X}')$ and define $x' = \max\{\hat{x},\tilde{x}\}$, $x = \min\{\hat{x},\tilde{x}\}$. We proceed to show that $\mathcal{X}(\theta,\hat{X})$ is monotone in the strong set order in $\hat{X}$. That is, for $\hat{X}' \geq \hat{X}$, we have that $x' \in \mathcal{X}(\theta,\hat{X}')$ and $x \in \mathcal{X}(\theta,\hat{X})$. By Assumption 1, u is a concave function of x. This implies that $\mathcal{X}(\theta,\hat{X})$ is an interval.

We will denote its lower end-point by $\underline{x}(\theta, \hat{X})$ and its upper end-point by $\overline{x}(\theta, \hat{X})$. We also note that $p(\underline{x}(\theta, \hat{X})|\theta, \hat{X}) = p(\overline{x}(\theta, \hat{X})|\theta, \hat{X}) = 0$. Showing that $x \in \mathcal{X}(\theta, \hat{X})$ is increasing in the strong set order therefore reduces to showing that $\underline{x}(\theta, \hat{X}) \leq \underline{x}(\theta, \hat{X}')$ and $\overline{x}(\theta, \hat{X}) \leq \overline{x}(\theta, \hat{X}')$. Without loss of generality (the other case follows by identical arguments), we will show that $\overline{x}(\theta, \hat{X}) \leq \overline{x}(\theta, \hat{X}')$.

Toward a contradiction, suppose that $\overline{x}(\theta, \hat{X}) > \overline{x}(\theta, \hat{X}')$. There are two cases to consider: the case in which the supports strictly overlap $\underline{x}(\theta, \hat{X}) < \overline{x}(\theta, \hat{X}')$, and the case in which they do not $\underline{x}(\theta, \hat{X}) \geq \overline{x}(\theta, \hat{X}')$. First, consider the case in which $\underline{x}(\theta, \hat{X}) < \overline{x}(\theta, \hat{X}')$. By continuity of $p(\cdot|\theta, \hat{X}')$ and $p(\cdot|\theta, \hat{X})$, the fact that $0 = p(\underline{x}(\theta, \hat{X})|\theta, \hat{X}) < p(\underline{x}(\theta, \hat{X})|\theta, \hat{X}')$, and the fact that $p(x|\theta, \hat{X}) > 0$ for $x \in (\underline{x}(\theta, \hat{X}), \overline{x}(\theta, \hat{X}'))$, there exists an $x \in (\underline{x}(\theta, \hat{X}), \overline{x}(\theta, \hat{X}'))$ such that $p(x|\theta, \hat{X}') > p(x|\theta, \hat{X})$. Fix also any $x' \in (\overline{x}(\theta, \hat{X}'), \overline{x}(\theta, \hat{X}))$. Consider a perturbation, as per Equation 702 that moves density from (a neighborhood of) x to (a neighborhood of) x' in state θ under conjecture $\hat{X}'$. We have that the following holds:

$$\begin{aligned}
\phi'(0) - \phi'(p(x|\theta, \hat{X}')) &< \phi'(p(x'|\theta, \hat{X})) - \phi'(p(x|\theta, \hat{X})) \\
&= \frac{u(x', \theta, \hat{X}(\theta))}{\lambda(\hat{X}(\theta), \theta)} - \frac{u(x, \theta, \hat{X}(\theta))}{\lambda(\hat{X}(\theta), \theta)} \\
&\leq \frac{u(x', \theta, \hat{X}'(\theta))}{\lambda(\hat{X}'(\theta), \theta)} - \frac{u(x, \theta, \hat{X}'(\theta))}{\lambda(\hat{X}'(\theta), \theta)}
\end{aligned} \tag{707}$$

where the first line follows from the strict convexity of ϕ, the fact that $p(x'|\theta, \hat{X}) > 0$, and the fact that $p(x|\theta, \hat{X}') > p(x|\theta, \hat{X})$. The second line follows from the optimality of $p(\cdot|\theta, \hat{X})$. The third line follows by Assumption 1. However, Equation 707 implies that the considered perturbation provides a strict gain relative to $p(\cdot|\theta, \hat{X}')$, which contradicts the optimality of $p(\cdot|\theta, \hat{X}')$.

Consider now the case in which $\underline{x}(\theta, \hat{X}) \geq \overline{x}(\theta, \hat{X}')$. By the fact that $p(\cdot|\theta, \hat{X})$ is strictly preferred to $p(\cdot|\theta, \hat{X}')$ when the aggregate follows $\hat{X}$, we have that:

$$\begin{aligned}
&\int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \frac{u(x, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} p(x|\theta, \hat{X}) \, \mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \frac{u(x, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} p(x|\theta, \hat{X}') \, \mathrm{d}x \\
&\quad > \int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \phi(p(x|\theta, \hat{X})) \, \mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \phi(p(x|\theta, \hat{X}')) \, \mathrm{d}x
\end{aligned} \tag{708}$$

Moreover, as $\underline{x}(\theta, \hat{X}) > \overline{x}(\theta, \hat{X}')$, we have by Assumption 1 that:

$$\int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \frac{u(x, \hat{X}'(\theta), \theta)}{\lambda(\hat{X}'(\theta), \theta)} p(x|\theta, \hat{X})\, \mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \frac{u(x, \hat{X}'(\theta), \theta)}{\lambda(\hat{X}'(\theta), \theta)} p(x|\theta, \hat{X}')\, \mathrm{d}x$$
$$\geq \int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \frac{u(x, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} p(x|\theta, \hat{X})\, \mathrm{d}x - \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \frac{u(x, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} p(x|\theta, \hat{X}')\, \mathrm{d}x \tag{709}$$

Combining these inequalities, we obtain that:

$$\int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} u(x, \hat{X}'(\theta), \theta) p(x|\theta, \hat{X})\, \mathrm{d}x - \lambda(\hat{X}'(\theta), \theta) \int_{\underline{x}(\theta;\hat{X})}^{\overline{x}(\theta;\hat{X})} \phi(p(x|\theta, \hat{X}))\, \mathrm{d}x >$$
$$\int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} u(x, \hat{X}'(\theta), \theta) p(x|\theta, \hat{X}')\, \mathrm{d}x - \lambda(\hat{X}'(\theta), \theta) \int_{\underline{x}(\theta;\hat{X}')}^{\overline{x}(\theta;\hat{X}')} \phi(p(x|\theta, \hat{X}'))\, \mathrm{d}x \tag{710}$$

which implies that $p(\cdot|\theta, \hat{X})$ is strictly better than $p(\cdot|\theta, \hat{X}')$ when the aggregate follows $\hat{X}'$, which contradicts the optimality of $p(\cdot|\theta, \hat{X}')$.

Thus, we have shown that $\mathcal{X}(\theta, \hat{X})$ is monotone in the strong set order in $\hat{X}$ and we have derived (by Equation 705) that:

$$\phi'(p(x'|\theta; \hat{X}')) - \phi'(p(x|\theta; \hat{X}')) \geq \phi'(p(x'|\theta; \hat{X})) - \phi'(p(x|\theta; \hat{X})) \tag{711}$$

for all $x' \geq x$ such that $x', x \in \mathcal{X}(\theta, \hat{X}) \cap \mathcal{X}(\theta, \hat{X}')$. Thus, if ϕ' satisfies quasi-MLRP (Assumption 3), then we have that $p(\theta; \hat{X}') \succeq_{FOSD} p(\theta; \hat{X})$ for all θ. It then follows by the monotonicity property of the aggregator (Assumption 2) that $X(p(\theta; \hat{X}')) \geq X(p(\theta; \hat{X}))$ for all θ and therefore that $T(\hat{X}') \geq T(\hat{X})$, which establishes the required monotonicity property of the equilibrium operator.

We now prove discounting. To this end, we will show that when we take $\hat{X}' = \hat{X} + \alpha$ for $\alpha \in \mathbb{R}_+$ that the resulting stochastic choice is dominated by an α right translation of the original stochastic choice under $\hat{X}$. Under this transformation, observe by the necessary condition for optimality and the sufficient concavity condition on utility (Assumption 1), we can apply the same arguments as above to derive that for all

$x' \geq x$ such that $x', x \in \mathcal{X}(\theta, \hat{X}) \cap \mathcal{X}(\theta, \hat{X}+\alpha)$:

$$
\begin{aligned}
\phi'(p_{-\alpha}(x'|\theta;\hat{X})) - \phi'(p_{-\alpha}(x|\theta,\hat{X})) &= \frac{u(x'-\alpha,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} - \frac{u(x-\alpha,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} \\
&\geq \frac{u(x',\hat{X}(\theta)+\alpha,\theta)}{\lambda(\hat{X}(\theta)+\alpha,\theta)} - \frac{u(x,\hat{X}(\theta)+\alpha,\theta)}{\lambda(\hat{X}(\theta)+\alpha,\theta)} \\
&= \phi'(p(x'|\theta;\hat{X}+\alpha)) - \phi'(p(x|\theta;\hat{X}+\alpha))
\end{aligned}
\tag{712}
$$

We now show that $\underline{x}(\theta,\hat{X})+\alpha \geq \underline{x}(\theta,\hat{X}+\alpha)$ and $\overline{x}(\theta,\hat{X})+\alpha \geq \overline{x}(\theta,\hat{X}+\alpha)$. Without loss of generality (the other case follows by identical arguments), we will show that $\overline{x}(\theta,\hat{X})+\alpha \geq \overline{x}(\theta,\hat{X}+\alpha)$. Toward a contradiction, suppose that $\overline{x}(\theta,\hat{X})+\alpha < \overline{x}(\theta,\hat{X}+\alpha)$. As in the previous argument, there are two cases to consider, the case in which the supports strictly overlap $\overline{x}(\theta,\hat{X})+\alpha > \underline{x}(\theta,\hat{X}+\alpha)$ and the case in which the supports are disjoint $\overline{x}(\theta,\hat{X})+\alpha \leq \underline{x}(\theta,\hat{X}+\alpha)$. In the overlapping support case, fix an $x \in (\underline{x}(\theta,\hat{X}+\alpha), \overline{x}(\theta,\hat{X})+\alpha)$ such that $p_{-\alpha}(x|\theta,\hat{X}) > p(x|\theta,\hat{X}+\alpha) > 0$ (which is possible by the same arguments used in the first part of the proof). Fix next a point $x' \in (\overline{x}(\theta,\hat{X})+\alpha, \overline{x}(\theta,\hat{X}+\alpha))$. And consider a perturbation that moves density from $x-\alpha$ to $x'-\alpha$ in state θ under conjecture $\hat{X}$. The following holds:

$$
\begin{aligned}
\phi'(0) - \phi'(p_{-\alpha}(x|\theta,\hat{X})) &< \phi'(p(x'|\theta,\hat{X}+\alpha)) - \phi'(p(x|\theta,\hat{X}+\alpha)) \\
&= \frac{u(x',\hat{X}(\theta)+\alpha,\theta)}{\lambda(\hat{X}(\theta)+\alpha,\theta)} - \frac{u(x,\hat{X}(\theta)+\alpha,\theta)}{\lambda(\hat{X}(\theta)+\alpha,\theta)} \\
&\leq \frac{u(x'-\alpha,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)} - \frac{u(x-\alpha,\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)}
\end{aligned}
\tag{713}
$$

where the first inequality follows by construction, the second by optimality, and the third by Assumption 1. This contradicts the optimality of $p(\cdot|\theta,\hat{X})$. In the non-overlapping case, we can follow the same steps as the first part of the proof, adapted in the obvious way (using the sufficient concavity inequality in place of the supermodularity inequality).

Thus, by quasi-MLRP of ϕ' (Assumption 3), we have that $p_{-\alpha}(\theta,\hat{X}) \succeq_{FOSD} p(\theta,\hat{X}+\alpha)$ where $p_{-\alpha}$ is the described right translation by α of p. Moreover, by the discounting property of the aggregator (Assumption 2), we then have that:

$$
T(\hat{X}+\alpha) \leq X \circ p_{-\alpha}(\hat{X}) \leq T(\hat{X}) + \beta\alpha \tag{714}
$$

which establishes the discounting property of T. We have now shown that T satisfies

Blackwell's sufficient conditions and is a contraction map. By the Banach fixed point theorem, there then exists a unique equilibrium Ω. $\square$

C.1.5 Proof of Theorem 2

Proof. To show that the unique equilibrium aggregate law of motion of monotone in θ, we use Corollary 1 from Chapter 3 of Stokey, Lucas, and Prescott (1989).

Define the set of monotone increasing and bounded functions $\mathcal{M} = \{\hat{X} \in \mathcal{B} | \hat{X}(\theta') \geq \hat{X}(\theta) \quad \forall \theta, \theta' \in \Theta : \theta' \geq \theta\}$. See that this set is closed. If we can show that $T(\hat{X}) \in \mathcal{M}$ for any $\hat{X} \in \mathcal{M}$, then we know that the unique fixed point of T is in $\mathcal{M}$ and therefore that the unique equilibrium law of motion is in $\mathcal{M}$ according to Corollary 1 of Stokey, Lucas, and Prescott (1989). To this end, we wish to show that:

$$\hat{X}(\theta') \geq \hat{X}(\theta) \quad \forall \theta, \theta' \in \Theta : \theta' \geq \theta \implies T(\hat{X})(\theta') \geq T(\hat{X})(\theta) \quad \forall \theta, \theta' \in \Theta : \theta' \geq \theta \tag{715}$$

This follows immediately from the necessary condition used in the proof of Theorem 1. More precisely, by the necessary optimality condition (Equation 705) from the proof of Theorem 1 and Assumption 4, we have that for all $x' \geq x$ such that $x', x \in \mathcal{X}(\theta) \cap \mathcal{X}(\theta')$

$$\begin{aligned} \phi'(p(x'|\theta', \hat{X})) - \phi'(p(x|\theta', \hat{X})) &\geq \frac{u(x', \hat{X}(\theta'), \theta')}{\lambda(\hat{X}(\theta'), \theta')} - \frac{u(x, \hat{X}(\theta'), \theta')}{\lambda(\hat{X}(\theta'), \theta')} \\ &\geq \frac{u(x', \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} - \frac{u(x, \hat{X}(\theta), \theta)}{\lambda(\hat{X}(\theta), \theta)} \\ &= \phi'(p(x'|\theta, \hat{X})) - \phi'(p(x|\theta, \hat{X})) \end{aligned} \tag{716}$$

In the case where optimal action distributions do not have full support, the same arguments for the monotonicity of $\mathcal{X}(\theta, \hat{X})$ imply monotonicity of $\mathcal{X}(\theta) = \mathcal{X}(\theta, \hat{X}(\theta))$ in the strong set order when $\hat{X}$ is montotone increasing. Thus, by the quasi-MLRP property of ϕ' (Assumption 3) we then have that $p(\theta'; \hat{X}) \succeq_{FOSD} p(\theta; \hat{X})$ and thus by the monotonicity of the aggregator (Assumption 2) that $T(\hat{X})(\theta') \geq T(\hat{X})(\theta)$. $\square$

C.1.6 Proof of Theorem 3

Proof. Recall also by Theorem 1, that the unique symmetric stochastic choice rule consistent with the unique equilibrium $\hat{X}$ solves the following program:

$$p \in \arg\max_{p \in \mathcal{P}} \sum_{\Theta} \int_{\mathcal{X}} u(x, \hat{X}(\theta), \theta) \, \mathrm{d}P(x|\theta) \, \pi(\theta) - \sum_{\Theta} \int_{\mathcal{X}} \phi(p(x|\theta)) \mathrm{d}x \, \pi(\theta) \lambda(\hat{X}(\theta), \theta) \tag{717}$$

where we will suppress the dependence of the optimal policy on $\hat{X}$ as it is unique. Applying the necessary optimal condition from the proof of Theorem 1 (Equation 705), for a given x such that $p(x|\theta) > 0$, we have that:

$$u(\gamma(\hat{X}(\theta),\theta),\hat{X}(\theta),\theta) - u(x,\hat{X}(\theta),\theta) = \lambda(\hat{X}(\theta),\theta)\left(\phi'(p(\gamma(\hat{X}(\theta),\theta)|\theta)) - \phi'(p(x|\theta))\right) \tag{718}$$

Under Assumption 5, we moreover have that

$$u(x,X,\theta) = \alpha(X,\theta) - \beta(X,\theta)\Gamma(|x-\gamma(X,\theta)|) \tag{719}$$

Thus our necessary condition simplifies to:

$$\frac{\beta(\hat{X}(\theta),\theta)}{\lambda(\hat{X}(\theta),\theta)}\Gamma(|x-\gamma(\hat{X}(\theta),\theta)|) = \phi'(p(\gamma(\hat{X}(\theta),\theta)|\theta)) - \phi'(p(x|\theta)) \tag{720}$$

Now consider any θ,θ' such that $\tilde{\beta}(\theta',\hat{X}(\theta')) \geq \tilde{\beta}(\theta,\hat{X}(\theta))$ (where $\tilde{\beta} = \frac{\beta}{\lambda}$). Note that, by Theorem 2, the aggregate $\hat{X}$ is monotone increasing in the state θ. Thus if $\tilde{\beta}(\theta,X)$ is decreasing in both arguments, the stated case corresponds to $\theta' \leq \theta$. If instead $\tilde{\beta}(\theta,X)$ is increasing in both arguments, the stated case corresponds to $\theta' \geq \theta$. Therefore, to verify the desired result, we now prove that the action distribution in state θ' is more precise about $\gamma(\theta',\hat{X}(\theta'))$ than the action distribution in state θ is about $\gamma(\theta,\hat{X}(\theta))$, with respect to ϕ'.

To that end, we take x,x' such that:

$$|x-\gamma(\hat{X}(\theta),\theta)| = |x'-\gamma(\hat{X}(\theta'),\theta')| \tag{721}$$

It follows that:

$$\begin{aligned}\phi'(p(\gamma(\hat{X}(\theta),\theta)|\theta)) - \phi'(p(x|\theta)) &= \tilde{\beta}(\hat{X}(\theta),\theta)\Gamma(|x-\gamma(\hat{X}(\theta),\theta)|)\\ &\geq \tilde{\beta}(\hat{X}(\theta'),\theta')\Gamma(|x'-\gamma(\hat{X}(\theta'),\theta')|)\\ &= \phi'(p(\gamma(\hat{X}(\theta'),\theta')|\theta')) - \phi'(p(x'|\theta))\end{aligned} \tag{722}$$

We now take care of those points that have no density. To this end consider the first-order condition for $p(x|\theta)$:

$$u(x,\hat{X}(\theta),\theta) - \phi'(p(x|\theta)) - \lambda(\theta) - \kappa(x,\theta) = 0 \tag{723}$$

where $\lambda(\theta)$ is the Lagrange multiplier on the constraint that $\int_{\mathcal{X}} p(x|\theta) = 1$ and $\kappa(x,\theta)$

is the Lagrange multiplier on the constraint that $p(x|\theta) \geq 0$. When $p(x|\theta) = 0$, we have that $\kappa(x, \theta) \leq 0$. Given our assumption on utility, this is given by:

$$\kappa(x, \theta) = -\beta(\hat{X}(\theta), \theta)\Gamma(|x - \gamma(\hat{X}(\theta), \theta)|) + \alpha(\hat{X}(\theta), \theta) - \lambda(\theta) \tag{724}$$

which is monotonically decreasing in $|x - \gamma(\hat{X}(\theta), \theta)|$. Thus, if there is an x such that $p(x|\theta) = 0$, then there exists an $\bar{x}(\theta)$ such that $p(x|\theta) = 0$ if and only if $|x - \gamma(\hat{X}(\theta), \theta)| \geq |\bar{x}(\theta) - \gamma(\hat{X}(\theta), \theta)|$. Moreover, by monotonicity of $\tilde{\beta}(\hat{X}(\theta), \theta)$ in θ, we have that $|\bar{x}(\theta) - \gamma(\hat{X}(\theta), \theta)| \leq |\bar{x}(\theta') - \gamma(\hat{X}(\theta'), \theta')|$. Hence, so long as $x \in [\gamma(\hat{X}(\theta), \theta) - \bar{x}(\theta), \gamma(\hat{X}(\theta), \theta) + \bar{x}(\theta)]$, we always have that $x' \in [\gamma(\hat{X}(\theta'), \theta') - \bar{x}(\theta'), \gamma(\hat{X}(\theta'), \theta') + \bar{x}(\theta')]$. Thus, every element of the support of $p(\theta)$ satisfies:

$$\phi'(p(\gamma(\hat{X}(\theta), \theta)|\theta)) - \phi'(p(x|\theta)) \geq \phi'(p(\gamma(\hat{X}(\theta'), \theta')|\theta')) - \phi'(p(x'|\theta)) \tag{725}$$

It follows then by the definition of precision that $p(\theta)$ is more precise about $\gamma(\hat{X}(\theta), \theta)$ than $p(\theta')$ about $\gamma(\hat{X}(\theta'), \theta')$ under ϕ'. □

C.1.7 Proof of Theorem 4

Proof. By Assumption 6, there is a unique efficient stochastic choice rule P^E. Moreover, for any $x, x' \in \mathcal{X}$ and $\theta \in \Theta$ such that $p^E(x|\theta) > 0$ and $p^E(x'|\theta) > 0$, by the same variational arguments used in the Proof of Theorem 1, and exploiting linearity of the aggregator we have that:

$$\begin{aligned} &u(x', X(p^E(\theta)), \theta) - u(x, X(p^E(\theta)), \theta) + [f(x') - f(x)] \int_{\mathcal{X}} u_X(\tilde{x}, X(p^E(\theta)), \theta) p^E(\tilde{x}|\theta) \, \mathrm{d}\tilde{x} \\ &= \lambda(X(p^E(\theta)), \theta) \left(\phi'(p^E(x'|\theta)) - \phi'(p^E(x|\theta)) \right) + [f(x') - f(x)] \lambda_X(X(p^E(\theta)), \theta) \int_{\mathcal{X}} \phi(p^E(\tilde{x}|\theta)) \, \mathrm{d}\tilde{x} \end{aligned} \tag{726}$$

is necessary for optimality of p^E. Moreover, if the efficient stochastic choice rule obtains in equilibrium, we have that (by Equation 705):

$$u(x', X(p^E(\theta)), \theta) - u(x, X(p^E(\theta)), \theta) = \lambda(X(p^E(\theta)), \theta) \left(\phi'(p^E(x'|\theta)) - \phi'(p^E(x|\theta)) \right) \tag{727}$$

These conditions coincide if and only if:

$$[f(x') - f(x)] \int_{\mathcal{X}} u_X(\tilde{x}, X(p^E(\theta)), \theta) p^E(\tilde{x}|\theta) \, \mathrm{d}\tilde{x} = [f(x') - f(x)] \lambda_X(X(p^E(\theta)), \theta) \int_{\mathcal{X}} \phi(p^E(\tilde{x}|\theta)) \, \mathrm{d}\tilde{x} \tag{728}$$

As f is nowhere-constant, $f(x') \neq f(x)$, and this condition reduces to:

$$\int_{\mathcal{X}} u_X(\tilde{x}, X(p^E(\theta)), \theta) p^E(\tilde{x}|\theta) \, \mathrm{d}\tilde{x} = \lambda_X(X(p^E(\theta)), \theta) \int_{\mathcal{X}} \phi(p^E(\tilde{x}|\theta)) \, \mathrm{d}\tilde{x} \tag{729}$$

Substituting in $p^E = p^*$, we obtain the statement in the claim. □

C.1.8 Statement and Proof of Lemma 15

In this appendix, we state and prove a Lemma that specializes several of main results to the case with quadratic payoffs of the form

$$u(x, X, \theta) = \alpha(X, \theta) - \beta(X, \theta)(x - \gamma(X, \theta))^2 \tag{730}$$

In the statement below, we use the definition $\tilde{\beta}(X, \theta) = \beta(X, \theta)/\lambda(X, \theta)$. We also define the bias and dispersion of a stochastic choice rule P in state θ around optimal point $\gamma(X(P), \theta)$ as

$$\begin{aligned} \mathrm{Bias}[P, \theta] &\equiv \int_{\mathcal{X}} (x - \gamma(X(P), \theta)) \ \mathrm{d}P\,(x|\theta) \\ \mathrm{Disp}[P, \theta] &\equiv \left(\int_{\mathcal{X}} (x - \gamma(X(P), \theta))^2 \ \mathrm{d}P\,(x|\theta) \right)^{\frac{1}{2}} \end{aligned} \tag{731}$$

Lemma 15. *Suppose that Assumptions 2 and 3 hold and that payoffs are given by Equation 730. The following properties hold under the additional stated conditions.*

1. ***Uniqueness.*** *There exists a unique equilibrium if the following holds for all $x \in \mathcal{X}, X \in \mathcal{X}$ and $\theta \in \Theta$:*

$$-(1 - \gamma_X(X, \theta)) < \frac{\tilde{\beta}_X(X, \theta)}{\tilde{\beta}(X, \theta)} (x - \gamma(X, \theta)) < \gamma_X(X, \theta) \tag{732}$$

2. ***Monotone actions.*** *The cross-sectional distribution of actions and the aggregate action X are monotone in the fundamental if, in addition to the condition (732), the following holds for all $X \in \mathcal{X}, x \in \mathcal{X}$, and $\theta \in \Theta$:*[1]

$$\frac{\tilde{\beta}_\theta(X, \theta)}{\tilde{\beta}(X, \theta)} (x - \gamma(X, \theta)) < \gamma_\theta(X, \theta) \tag{733}$$

[1] Where, for simplicity, we allow β to be defined for all states in a closed interval that contains Θ, and assume it is differentiable in its second argument.

3. ***Monotone precision.*** *The precision of actions about the optimal action γ under ϕ' is increasing (decreasing) in the strength of fundamentals if, in addition to (732) and (733), $\tilde{\beta}$ is monotone decreasing (increasing) in both arguments.*

4. ***Efficiency.*** *A necessary condition for efficiency of the stochastic choice rule P^* under Assumption 6 is that, for all θ,*

$$\begin{aligned} &\lambda_X(X(P^*(\theta)),\theta)\int_{\mathcal{X}}\phi(p^*(x\mid\theta))\,\mathrm{d}x \\ =&\alpha_X(X(P^*(\theta)),\theta)-\beta_X(X(P^*(\theta)),\theta)(Disp[P^*(\theta),\theta])^2 \\ &+2\,\gamma_X(X(P^*(\theta)),\theta)\,\beta(X(P^*(\theta)),\theta)\,Bias[P^*(\theta),\theta] \end{aligned} \tag{734}$$

Proof. We have directly assumed that Assumptions 2, 3 and 5 hold. The first claim follows so long as condition 732 implies Assumption 1, Supermodularity and Sufficient Concavity, for the outcome-equivalent game with payoff curvature $\tilde{\beta}$ and associated payoff $\tilde{u}$.

For supermodularity, it is sufficient to show that $\tilde{u}_{xX}(x,X,\theta)>0$. We observe that $\tilde{u}_{xX}=-2\tilde{\beta}_X(X,\theta)(x-\gamma(X,\theta))+2\gamma_X(X,\theta)\tilde{\beta}(X,\theta)$ This condition simplifies to $\gamma_X(X,\theta)>\frac{\tilde{\beta}_X(X,\theta)}{\tilde{\beta}(X,\theta)}(x-\gamma(X,\theta))$, which is the second inequality of Equation 732.

For sufficient concavity, it is sufficient to show that $|\tilde{u}_{xx}(X,\theta)|>\tilde{u}_{xX}(x,X,\theta)$. Observe that $|\tilde{u}_{xx}(X,\theta)|=2\tilde{\beta}(X,\theta)$. The condition

$$2\tilde{\beta}(X,\theta)>\tilde{u}_{xX}=-2\tilde{\beta}_X(X,\theta)(x-\gamma(X,\theta))+2\gamma_X(X,\theta)\tilde{\beta}(X,\theta) \tag{735}$$

simplifies to the first inequality of Equation 732: $-(1-\gamma_X(X,\theta))<\frac{\tilde{\beta}_X(X,\theta)}{\tilde{\beta}(X,\theta)}(x-\gamma(X,\theta))$.

The second claim of the Lemma follows so long as condition 733 implies Assumption 4. To see this, as we have already that $\tilde{u}_{xX}(x,X,\theta)>0$ for all x,X,θ, it is sufficient to check that $\tilde{u}_{x\theta}(x,X,\theta)>0$ for all x,X,θ. We note that $\tilde{u}_{x\theta}(x,X,\theta)=-2\tilde{\beta}_\theta(X,\theta)(x-\gamma(X,\theta))+2\gamma_\theta(X,\theta)\tilde{\beta}(X,\theta)$ and re-arrange to the desired expression.

The third claim follows directly by Theorem 3 as the payoffs in Equation 730 satisfy Assumption 5.

The fourth claims follows by Theorem 4. Recall from Theorem 4 that a necessary condition for efficiency of an equilibrium P^* under Assumption 6 is that:

$$\int_{\mathcal{X}}u_X(\tilde{x},X(P^*(\theta)),\theta)\,\mathrm{d}P^*(\tilde{x}|\theta)=\lambda_X(X,\theta)\int_{\mathcal{X}}\phi(p^*(x\mid\theta))\,\mathrm{d}x \tag{736}$$

for all $\theta \in \Theta$. Using the payoff function, we calculate:

$$u_X(x, X, \theta) = \alpha_X(X, \theta) - \beta_X(X, \theta)(x - \gamma(X, \theta))^2 + 2\gamma_X(X, \theta)\beta(X, \theta)(x - \gamma(X, \theta)) \tag{737}$$

Plugging this into the necessary condition and evaluating at the equilibrium aggregate $\hat{X}(\theta) = X(P^*(\theta))$, we obtain:

$$\begin{aligned} &\int_{\mathcal{X}} u_X(\tilde{x}, X(P^*(\theta)), \theta)\, \mathrm{d}P^*(\tilde{x}|\theta) \\ =&\int_{\mathcal{X}} \Big[\alpha_X(X(P^*(\theta)), \theta) - \beta_X(X(P^*(\theta)), \theta)(\tilde{x} - \gamma(X(P^*(\theta)), \theta))^2 \\ &+ 2\gamma_X(X(P^*(\theta)), \theta)\beta(X(P^*(\theta)), \theta)(\tilde{x} - \gamma(X(P^*(\theta)), \theta))\Big]\, \mathrm{d}P^*(\tilde{x}|\theta) \end{aligned} \tag{738}$$

Which can be rewritten in terms of the equilibrium bias and variance with respect to γ as:

$$\begin{aligned} &\int_{\mathcal{X}} u_X(\tilde{x}, X(P^*(\theta)), \theta)\, \mathrm{d}P^*(\tilde{x}|\theta) \\ =&\alpha_X(X(P^*(\theta)), \theta) - \beta_X(X(P^*(\theta)), \theta)\left(\mathrm{Disp}[P^*(\theta), \theta]\right)^2 \\ &+ 2\gamma_X(X(P^*(\theta)), \theta)\beta(X(P^*(\theta)), \theta)\mathrm{Bias}[P^*(\theta), \theta] \end{aligned} \tag{739}$$

as desired. □

C.1.9 Proof of Corollary 9

We first derive the payoff representation of Equation 130. This is a second-order approximation of the payoff function in Equation 129, reprinted here:

$$u(p_i, P, M) = M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left(p_i - M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}\right) p_i^{-\eta} \tag{740}$$

We first calculate

$$\begin{aligned} u_p(p_i, P, M) &= M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left((-\eta + 1)p_i^{-\eta} + \eta M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} p_i^{-\eta-1}\right) \\ u_{pp}(p_i, P, M) &= M^{\frac{1-\sigma}{\sigma}} P^{\eta - \frac{1}{\sigma}} \left(\eta(\eta - 1)p_i^{-\eta-1} - \eta(\eta + 1) M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} p_i^{-\eta-2}\right) \end{aligned} \tag{741}$$

We define $\gamma(P, M)$ as the (unique) solution to $u_p(p_i, P, M)|_{p_i = \gamma(P,M)} = 0$. Re-arranging:

$$\begin{aligned}
0 &= M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \left((-\eta+1)\gamma(P,M)^{-\eta} + \eta M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \gamma(P,M)^{-\eta-1} \right) \\
0 &= \left((-\eta+1) + \eta M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \gamma(P,M)^{-1} \right) \\
\gamma(P,M) &= \frac{\eta}{\eta-1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}}
\end{aligned} \tag{742}$$

We define $\alpha(P,M) = u(\gamma(P,M),P,M)$. We first observe that

$$\gamma(P,M) - M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} = \frac{1}{\eta-1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \tag{743}$$

Then, by direct calculation,

$$\begin{aligned}
\alpha(P,M) &= M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \left(\frac{1}{\eta-1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \right) \left(\frac{\eta}{\eta-1} \right)^{-\eta} M^{-\eta\frac{\chi}{\sigma}} P^{-\eta+\eta\frac{\chi}{\sigma}} \\
&= \frac{1}{\eta-1} \left(\frac{\eta}{\eta-1} \right)^{-\eta} M^{\frac{1-\sigma+\chi(1-\eta)}{\sigma}} P^{\eta-\frac{1}{\sigma}+(1-\eta)(1-\frac{\chi}{\sigma})}
\end{aligned} \tag{744}$$

as desired

We define $\beta(P,M) = -\frac{1}{2} u_{pp}(p_i,P,M)|_{p=\gamma(P,M)}$. We first observe that

$$(\eta-1)\gamma(P,M) - (\eta+1) M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} = -M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \tag{745}$$

Then, by direct calculation,

$$\begin{aligned}
\beta(P,M) &= -\frac{1}{2} \left(M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \eta \gamma(P,M)^{-\eta-2} \left((\eta-1)\gamma(M,P) - (\eta+1) M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \right) \right) \\
&= \frac{1}{2} \left(M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \eta \gamma(P,M)^{-\eta-2} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \right) \\
&= \frac{1}{2} \left(M^{\frac{1-\sigma}{\sigma}} P^{\eta-\frac{1}{\sigma}} \eta \left(\frac{\eta}{\eta-1} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \right)^{-\eta-2} M^{\frac{\chi}{\sigma}} P^{1-\frac{\chi}{\sigma}} \right) \\
&= \frac{\eta}{2} \left(\frac{\eta}{\eta-1} \right)^{-(\eta+2)} M^{\frac{1-\sigma-\chi(\eta+1)}{\sigma}} P^{-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}}
\end{aligned} \tag{746}$$

as required. We finally observe that, since $\lambda(M) = \frac{1}{\pi(M)} = KM^{-\delta}$ (where K is a normalizing constant), we have

$$\tilde{\beta}(P,M) = \frac{\beta(P,M)}{\lambda(P,M)} = \frac{\eta}{2} \left(\frac{\eta}{\eta-1} \right)^{-(\eta+2)} M^{\frac{1-\sigma-\chi(\eta+1)}{\sigma}+\delta} P^{-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}} \tag{747}$$

We now apply the conditions of Lemma 15 to prove the stated result. We first calculate that

$$\gamma_P(P,M) = \left(1-\frac{\chi}{\sigma}\right)\frac{\gamma(P,M)}{P} \qquad \frac{\tilde{\beta}_P(P,M)}{\tilde{\beta}(P,M)} = P^{-1}\left(-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}\right) \tag{748}$$

Applying the condition of Equation 732, we get

$$-(1-\left(1-\frac{\chi}{\sigma}\right)\frac{\gamma(P,M)}{P}) < P^{-1}\left(-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}\right)(p-\gamma(P,M)) < \left(1-\frac{\chi}{\sigma}\right)\frac{\gamma(P,M)}{P} \tag{749}$$

Since $\chi/\sigma < 1$, we divide all three expressions by $\left(1-\frac{\chi}{\sigma}\right)\gamma(P,M)/P$ to get

$$-\left(\frac{P}{\gamma(P,M)\left(1-\frac{\chi}{\sigma}\right)}-1\right) < \frac{-1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma}}{\left(1-\frac{\chi}{\sigma}\right)}\left(\frac{p}{\gamma(P,M)}-1\right) < 1 \tag{750}$$

as desired.

We next verify a condition for monotone aggregates. We first calculate:

$$\gamma_P(P,M) = \frac{\chi}{\sigma}\frac{\gamma(P,M)}{M}, \qquad \frac{\tilde{\beta}_M(P,M)}{\tilde{\beta}(P,M)} = M^{-1}\left(\frac{1-\sigma-\chi(\eta+1)}{\sigma}+\delta\right) \tag{751}$$

We then apply Equation 733:

$$M^{-1}\left(\frac{1-\sigma-\chi(\eta+1)}{\sigma}+\delta\right)(p-\gamma(P,M)) < \frac{\chi}{\sigma}\frac{\gamma(P,M)}{M} \tag{752}$$

Dividing both sides by $\gamma(P,M)\sigma/(M\chi)$, this becomes

$$\frac{1-\sigma-\chi(\eta+1)+\delta\sigma}{\chi}\left(\frac{p}{\gamma(P,M)}-1\right) < 1 \tag{753}$$

We finally derive a condition for monotone precision. For this, we need $\tilde{\beta}$ to decrease in both M and P. This respectively requires:

$$\begin{aligned} 0 &> \frac{1-\sigma-\chi(\eta+1)}{\sigma}+\delta \\ 0 &> -1-\frac{1}{\sigma}+(\eta+1)\frac{\chi}{\sigma} \end{aligned} \tag{754}$$

Re-arranging these inequalities gives the desired condition,

$$\chi(\eta+1) \in (1+\gamma(\delta-1), 1+\gamma) \tag{755}$$

C.1.10 Proof of Corollary 10

Proof. We first derive the payoff representation of Equations 139 and 140. To this end, we begin by deriving the consumer's choices at $t \geq 1$. At $t = 1$, given savings b_{i0} from the first period, each consumer i solves the following program at $t = 1$:

$$\begin{aligned} \max_{\{c_{it}, n_{it}\}_{t=1}^{\infty}} \sum_{t=1}^{\infty} \delta^t \left(c_{it} - \frac{c_{it}^2}{2} - \chi \frac{n_{it}^2}{2} \right) \\ \sum_{t=1}^{\infty} \frac{c_{it}}{R^t} \leq b_{i0} + \sum_{t=1}^{\infty} \frac{w_t n_{it}}{R^t} \end{aligned} \tag{756}$$

where $b_{i0} = y_0 - c_{i0}$ is the agent's savings from $t = 0$. This problem is concave in all arguments. Letting κ denote the Lagrange multiplier in the constraint, we find first-order conditions $\delta^t(1 - c_{it}) = \kappa R^{-t}$ for each c_{it} and $\delta^t \chi n_{it} = w_t \kappa R^{-t}$ for each n_{it}. Using $\delta R = 1$, we transform the former into $\kappa = 1 - c_{it}$ for all t. This implies that consumption is constant. Plugging this into the labor-supply condition, we derive $n_{it} = \frac{1}{\chi} w_t(1 - c_{it})$. This is also constant, if consumption is constant.

We next prove that output is identically equal to $y_t = \bar{y}$ for $t \geq 1$ and solve for $\bar{y}$. Profit maximization for the firm implies that the firm elastically demands labor at the wage $w_t = 1$. Evaluated at this wage, labor demand for each agent i is $n_{it} = \frac{1}{\chi}(1 - c_{it})$. Integrating both sides over i, we get $n_t = \frac{1}{\chi}(1 - c_t)$. Substituting in the production function and market clearing, this becomes $y_t = \frac{1}{\chi}(1 - y_t)$. Therefore, $y_t = \bar{y} = \frac{1}{1+\chi}$.

To derive the household's consumption and labor supply, we return to the budget constraint and simplify it by plugging in constant consumption $c_{it} = c_{i1}$, labor demand, and $w_t = 1$, and by simplifying the sums:

$$\frac{1}{1 - R^{-1}} c_{i1} \leq R b_{i0} + \frac{1}{1 - R^{-1}} \frac{1}{\chi}(1 - c_{i1}) \tag{757}$$

Rearranging, we write

$$c_{i1} \leq \frac{\chi}{1 + \chi} \frac{1 - \delta}{\delta} b_{i0} + \bar{y} \tag{758}$$

This holds at equality if the right-hand-side is less than 1 (the agent's bliss point). This is guaranteed under the maintained assumption that $b_{i0} \leq \bar{c} < \delta/(1 - \delta)$. We finally write the value function from Equation 756 as $V(b_{i0})$. And we observe from the envelope theorem that

$$V'(b_{i0}) = \kappa = 1 - \frac{\chi}{1 + \chi} \frac{1 - \delta}{\delta} b_{i0} - \bar{y} \tag{759}$$

We now return to the payoff of the consumer at time 0, who chooses consumption given rational expectations about this future equilibrium path and their future choices. For notational simplicity, we let $c_{i0} = c$ and $y_0 = y$. The agent's payoff is

$$U(c, y, \theta_d) = (1 + \theta_d)c - \frac{c^2}{2} - \chi \frac{y^2}{2} + V(y - c) \tag{760}$$

Note that all agents are off their labor supply curve and work y labor hours.

We now derive the form in Equation 139. We first observe that $U_c(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} = 0$. Taking the first derivative,

$$\begin{aligned} U_c(c, y, \theta_d) &= (1 + \theta_d) - c - V'(y - c) \\ &= (1 + \theta_d) - c - \left(1 - \frac{\chi}{1 + \chi} \frac{1 - \delta}{\delta}(y - c) - \bar{y}\right) \end{aligned} \tag{761}$$

We next use $U_c(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} = 0$ and rearrange to write

$$\gamma(y, \theta_d) = (1 - m)(\theta_d + \bar{y}) + my \tag{762}$$

where $m = \frac{\chi(1-\delta)}{\chi+\delta}$ is the marginal propensity to consume

We next observe that $-U_{cc}(c, y, \theta_d) = 2\beta(y, \theta_d)$. We calculate, from above,

$$U_{cc} = -\frac{1}{1 - m}, \qquad \beta(y, \theta_d) = \frac{1}{2(1 - m)} \tag{763}$$

Moreover, we have $\tilde{\beta}(y, \theta_d) = \beta(y, \theta_d)/\lambda(y, \theta_d) = \frac{1}{2(1-m)} y^\tau$.

We finally observe that $U(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} = \alpha(y, \theta_d)$. We therefore define

$$\alpha(y, \theta_d) = (1 + \theta_d)\gamma(y, \theta_d) - \frac{\gamma(y, \theta_d)^2}{2} - \chi \frac{y^2}{2} + V(y - \gamma(y, \theta_d)) \tag{764}$$

Having verified the payoff representation, we now prove the claim by applying Lemma 15. First, to show uniqueness, we specialize the condition in Equation 732. We note that

$$\begin{aligned} \gamma_y(y, \theta_d) &= m \\ \frac{\tilde{\beta}_y(y, \theta_d)}{\tilde{\beta}(y, \theta_d)} &= \tau y^{-1} \end{aligned} \tag{765}$$

Using these expressions, we derive the condition that, for all $y, c \in [\underline{c}, \bar{c}]$ and $\theta_d \in \Theta_d$,

$$-(1 - m) < \frac{\tau}{y} \left(c - (1 - m)\left(\bar{y} + \theta_d\right) - my\right) < m \tag{766}$$

We re-arrange this algebraically to

$$0 < m - \frac{\tau}{y}\left(c - (1-m)\left(\bar{y} + \theta_d\right) - my\right) < 1 \tag{767}$$

Next, to show monotonicity, we observe that $\tilde{\beta}_{\theta_d}(y, \theta_d) = 0$, and hence Equation 733 reduces to $\gamma_{\theta_d}(y, \theta_d) > 0$, which is by assumption when $\delta > 0$ and $\chi > 0$.

Next, to show monotone precision, we observe that $\tilde{\beta}(X, \theta_d) = \frac{X^{\tau}}{\delta}$ is monotone decreasing in X, which from Lemma 15 implies that precision in increasing in fundamentals.

Finally, to show efficiency, we plug directly into Equation 734. We first use the definition of α and the envelope theorem to observe that

$$\alpha_y(y, \theta_d) = U_y(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} + \gamma_c(y, \theta_d) U_c(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} = U_y(c, y, \theta_d)|_{c=\gamma(y,\theta_d)} \tag{768}$$

We then calculate

$$\begin{aligned} U_y(c, y, \theta_d) &= -\chi y + V'(y-c) \\ &= -\chi y + 1 - \frac{\chi}{1+\chi}\frac{1-\delta}{\delta}(y-c) - \bar{y} \\ &= -\chi y + 1 - \frac{m}{1-m}(y-c) - \bar{y} \end{aligned} \tag{769}$$

where in the last line we use the definition of m.

We next observe that

$$\lambda_y(y, \theta_d) = -\tau y^{-\tau-1} \qquad \gamma_y(y, \theta_d) = m \qquad \beta(y, \theta_d) = \frac{1}{2(1-m)} \tag{770}$$

Finally, because of linear aggregation,

$$\text{Bias}[P, \theta_d] = \int_{\mathcal{X}} \left(x - \gamma(y(P), \theta_d)\right) \, \mathrm{d}P\left(x|\theta_d\right) = y - \gamma(y, \theta_d) \tag{771}$$

Using all of this, we re-write the condition for efficiency (Equation 734) as

$$\begin{aligned} -\tau y^{-\tau-1} \int_{\mathcal{X}} \phi(p^*(x \mid \theta_d))\,\mathrm{d}x &= -\chi y + 1 - \frac{m}{1-m}(y - \gamma(y, \theta_d)) - \bar{y} + \frac{m}{1-m}\left(y - \gamma(y, \theta_d)\right) \\ &= -\chi y + 1 - \bar{y} \end{aligned} \tag{772}$$

This re-arranges to

$$y = \bar{y} + \frac{\tau}{\chi} y^{-\tau-1} \int_{\mathcal{X}} \phi(p^*(x \mid \theta_d))\, \mathrm{d}x \tag{773}$$

When $\tau = 0$, the solution to the fixed-point equation is $y = \bar{y}$. When $\tau > 0$, then

$$y - \bar{y} = \frac{\tau}{\chi} y^{-\tau-1} \int_{\mathcal{X}} \phi(p^*(x \mid \theta_d))\, \mathrm{d}x \tag{774}$$

The right-hand side is weakly positive under the assumption that cognitive costs in each state are positive. Moreover, an optimal y that solves this condition exists due to Assumption 6 being satisfied. Thus, the introduction of stress weakly increases the optimal level of output. □

C.2 State-Separable *vs.* Mutual Information Costs

In this Appendix, we compare the strategic mistakes model with the rational inattention model of Sims (2003). In Sims' rational inattention model, agents flexibly collect signals about an unknown state subject to a continuous cost or hard constraint monotone in the Shannon mutual information between the signal and the state, and then take actions measurable in this signal. Commonly, researchers assume that agents' information choice is unobserved and restrict focus to testing the model's predictions for behavior. This perspective is apparent in the early applications of Sims (2003, 2006), in the decision-theoretic analysis of Caplin, Dean, and Leahy (2019, 2022), and in many of the applications surveyed by Maćkowiak, Matejka, and Wiederholt (2020). From this perspective, despite their very different motivations—ours from the perspective of costly planning, and Sims (2003)'s from the perspective of costly information acquisition—the strategic mistakes and mutual information models may be each be comparable "candidates" for studying imperfect optimization in a specific equilibrium setting.

We study the similarities and differences between the two models both in theory and practice. We first present an abstract equivalence result which underscores how the models may be equivalent for matching observed data (aggregate and cross-sectional) when the prior distribution is unknown. We then exemplify these differences in a numerical example of a beauty contest, in which the strategic mistakes model has unique predictions and monotone comparative statics while the rational inattention model does not.

C.2.1 Definitions and an Equivalence Result

We first provide abstract conditions under which a version of the strategic mistakes model makes identical equilibrium predictions to the mutual information model, to build intuition about the comparability and differences of the two approaches.

All information acquisition models that have a posterior separable representation, including mutual information, can be recast as a choice over stochastic choice rules in $\mathcal{P}$ subject to some convex cost functional c (Denti, 2018). The mutual information cost of a stochastic choice rule $P \in \mathcal{P}$ can be decomposed into two terms which we label below:[2]

$$c^{MI}(P) = \underbrace{\sum_{\Theta} \int_{\mathcal{X}} p(x|\theta) \log p(x|\theta) \, \mathrm{d}x \ \pi(\theta)}_{\text{State-Separable Term}} - \underbrace{\int_{\mathcal{X}} p(x) \log p(x) \, \mathrm{d}x}_{\text{Cross-State Interactions}} \tag{775}$$

The first term is in fact identical to the state-separable representation (100) with the (quasi-MLRP) kernel $\phi(p) = p \log p$. We label the resulting cost function c^{LSM}, or *logit strategic mistakes*. In a stochastic choice interpretation, this term encodes the agent's desire to *increase* the entropy of the conditional action distributions or play randomly. The second term equals the entropy of the unconditional action distribution and encodes the agents' desire to, on average, anchor toward commonly played actions. This force is absent in the logit strategic mistakes model, and therefore characterizes c^{MI} model compared to its "strategic mistakes cousin" c^{LSM}. Moreover, this decomposition makes clear that there is no conceptual difference in modelling any stochastic choice game with mutual information versus entropic stochastic choice other than that agents have different cost functions, and therefore preferences.

Matějka and McKay (2015) show that the second term ("anchoring") has marginally zero influence on actions when agents' actions are *ex ante* exchangeable, or agents play each action x with equal unconditional probability. From the analyst's perspective, the key free parameter for engineering such exchangeability is the prior $\pi(\cdot)$. We extend this result to show, constructively, that an analyst free to specify the prior can re-construct the equilibrium of a logit strategic mistakes model as an equilibrium of an equivalent game with a mutual information friction provided that a technical condition on payoffs, which ensures that all actions can be made *ex ante* equally attractive, holds:

Lemma 16 (Equilibrium Equivalence)**.** *Suppose that the action space $\mathcal{X}$ is finite. Let*

[2] In this expression, we use the definition of the marginal distribution $p(x) = \sum_{\Theta} p(x|\theta)\pi(\theta)$.

$\Omega = (P^*, \hat{X})$ *be a symmetric equilibrium for the game* $\mathcal{G}^{LSM} = (u(\cdot), \lambda c^{LSM}(\cdot), X(\cdot), \pi'(\cdot), \Theta, \mathcal{X})$. *There exists some* $\pi'(\cdot) \in \triangle(\Theta)$ *such that* Ω *is an equilibrium of* $\mathcal{G}^{LSM}$ *and* $\mathcal{G}^{MI} = (u(\cdot), \lambda c^{MI}(\cdot),$
$X(\cdot), \pi'(\cdot), \Theta, \mathcal{X})$ *if and only if the following linear system has a solution for* $\pi' \in \Delta(\Theta)$*:*

$$\tilde{U}\pi' = \frac{1}{|\mathcal{X}|} \tag{776}$$

where 1 *is a* $|\Theta|$ *length vector, and* $\tilde{U}$ *is a* $|\mathcal{X}| \times |\Theta|$ *matrix with entries:*

$$\tilde{u}_{x_i,\theta_j} = \frac{\exp\{u(x_i, \hat{X}(\theta_j), \theta_j)/\lambda\}}{\sum_{x_k \in \mathcal{X}} \exp\{u(x_k, \hat{X}(\theta_j), \theta_j)/\lambda\}} \tag{777}$$

Proof. To establish that Ω is an equilibrium of the mutual information model, it is sufficient to establish that P^* solves each individual's optimization problem when they take $\hat{X}$ as given. By Corollary 2 in Matějka and McKay (2015), all interior unconditional choice probabilities $p(x) = \sum_{\theta \in \Theta} p(x|\theta)\pi(\theta)$ in the mutual information model satisfy the following first-order condition:

$$p(x \mid \theta) = \frac{p(x)\exp\{u(x, \hat{X}(\theta), \theta)/\lambda\}}{\sum_{\tilde{x} \in \mathcal{X}} p(\tilde{x})\exp\{u(\tilde{x}, \hat{X}(\theta), \theta)/\lambda\}} \tag{778}$$

and the following additional constraint:

$$\sum_{\theta \in \Theta} \frac{\exp\{u(x, \hat{X}(\theta), \theta)/\lambda\}}{\sum_{\tilde{x} \in \mathcal{X}} p(\tilde{x})\exp\{u(\tilde{x}, \hat{X}(\theta), \theta)/\lambda\}} \pi'(\theta) = 1 \tag{779}$$

Observe that, if and only if $p(x) = p(x')$ for all $x, x' \in \mathcal{X}$, then the choice probabilities that solve (778) are

$$p(x \mid \theta) = \frac{\exp\{u(x, \hat{X}(\theta), \theta)/\lambda\}}{\sum_{\tilde{x} \in \mathcal{X}} \exp\{u(\tilde{x}, \hat{X}(\theta), \theta)/\lambda\}} \tag{780}$$

This would verify that the stochastic choice rule P^* is a unique, interior solution to agents' choice problem. Hence it remains only to verify that $p(x) = p(x')$ for all $x, x' \in \mathcal{X}$, or exchangeability, in the agent's optimal program.

It is straightforward to derive such a condition using (779). Stacking equation (779) over all interior $x \in \mathcal{X}$, we obtain the system:

$$\tilde{U}(\{p(x)\}_{x \in \mathcal{X}})\pi' = 1 \tag{781}$$

where:

$$\tilde{u}_{x_i,\theta_j}(\{p(x)\}_{x\in\mathcal{X}}) = \frac{\exp\{u(x_i,\hat{X}(\theta_j),\theta_j)/\lambda\}}{\sum_{x_k\in\mathcal{X}} p(x_k)\exp\{u(x_k,\hat{X}(\theta_j),\theta_j)/\lambda\}} \tag{782}$$

and 1 is a $|\Theta|$ length vector. Thus, there exists a prior consistent with uniform unconditional choice $p(x) = \frac{1}{|\mathcal{X}|}$ if and only if the following linear system has a solution probability vector $\pi' \in \Delta(\Theta)$:

$$\tilde{U}\pi' = |\mathcal{X}|^{-1}1 \tag{783}$$

where 1 is a $|\Theta|$ length vector, and $\tilde{U}$ is as stated in the result. This completes the proof, with π' solving the given system supporting the equilibrium under the mutual information model. □

The proof establishes from first-order conditions that (776) corresponds with a flat unconditional distribution over actions. The condition ensures that there exists a prior such that all actions yield ex-ante equal payoffs. Heuristically, it is likely to fail if some actions in $\mathcal{X}$ are unappealing regardless of the state or the state space does not have many realizations. The intuition for the first idea is clearest in the extreme case in which some actions are dominated by others for all values of X and θ. In this case, there is nothing that an agent could believe that would ever rationalize playing these actions; and the bridge between the control-cost model and the rational-inattention model cannot be crossed. The intuition for the second relates to the fact that our construction varies the prior to make all actions *ex ante* equally plausible. If, for instance, there are only two states but $N > 2$ actions with very different payoffs from one another in each state, then there is likely no belief that will make all of the actions seem equally appealing.

This result has two practical implications. First, an analyst who is unsure of the physical prior distribution can think of the logit strategic mistakes model as a selection criterion for the mutual information model, across *games indexed by different priors* and, within each prior, a *potentially non-singleton set of equilibria*. This is a general-equilibrium analogue of the Matějka and McKay (2015) insight about the relationship between logit and mutual-information models for individual choice: the former approximates the latter when the analyst does not take a specific stand on anchoring toward defaults. Second, comparative statics in the strategic mistakes model which perturb payoffs $u(\cdot)$ or compare across states $\theta \in \Theta$ may be interpreted, under the conditions of Lemma 16, as comparative statics in a mutual information model *jointly* across the aforementioned features and the physical prior and given a specific equilibrium selection rule.

C.2.2 Numerically Revisiting The Beauty Contest

We now return to the beauty contest model to illustrate the differences between the strategic mistakes and mutual information models in a practical scenario that maps to the applications of Section 3.4. Because closed-form solutions are not available for equilibrium action profiles under the mutual information cost, we instead make a feasible approximation of the model on a gridded action space.[3] We will show in this context sharp differences between the predictions of the logit strategic mistakes and mutual information models regarding equilibrium multiplicity and comparative statics, and that these stem from the cross-state interactions embedded in the mutual information cost functional.[4]

Environment and Solution Method

For the simplest exposition and comparison to existing work, we use a version of our model that reduces to the linear beauty contest. We study quadratic payoffs of the form,

$$u(x, X, \theta) = \alpha(X, \theta) - \beta(X, \theta)(x - \gamma(X, \theta))^2 \tag{784}$$

and set $\alpha(X, \theta) \equiv 0$, eliminating the pure externality; $\beta(X, \theta) \equiv 1$, giving constant costs of misoptimization; and $\gamma(X, \theta) = (1-r)\theta + rX$ with $r = 0.85$.[5] The aggregator is the mean. The state space has two points of support, $\Theta = \{\theta_0, \theta_1\} = \{1.0, 2.0\}$. The action space $\mathcal{X}$ is approximated with a 40-point grid between lower endpoint $\underline{x} = 0$ and upper endpoint $\overline{x} = 3$. We use a flat prior with $\pi(\theta_0) = \pi(\theta_1) = \frac{1}{2}$. And we scale both logit and mutual information costs by $\lambda = 0.25$.

Let $p^*(\hat{X}) \in \Delta(\mathcal{X})^2$ return each agent's (unique) optimal stochastic choice rule, expressed as pair of probability mass functions, when they conjecture the equilibrium law of motion $\hat{X} = \left(\hat{X}(\theta_0), \hat{X}(\theta_1)\right)$.[6] As in the proof of our main results, let us

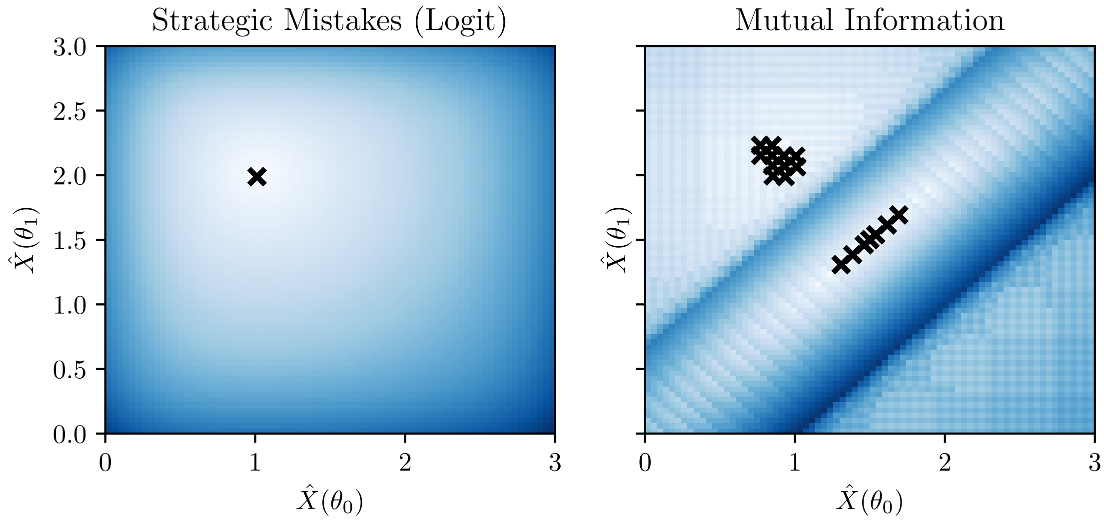
[3] This is due to two reasons, in our application with quadratic preferences: the lack of a Gaussian prior and the bounded action space. Moreover, if we had numerically solved a generalized beauty contest with state-dependent costs of mis-optimization, the non-quadratic payoffs would preclude a closed-form mutual-information solution even with a Gaussian prior and unbounded state space.

[4] Note that using logit strategic mistakes will imply that all actions are played with positive probability. To obtain endogenous consideration sets in the strategic mistakes model, we could have instead used a quadratic kernel.

[5] Hellwig and Veldkamp (2009) remark that, for dynamic beauty contests meant to mimic price-setting in New Keynesian models, that $r = 0.85$ is "commonly used." Finally, observe that these payoffs are jointly supermodular in (x, X, θ) but feature bounded complementarity based on the conditions established in the previous section, provided that $r \in (0, 1)$.

[6] For the logit strategic mistakes model, the optimal action profile is known in closed form. For the mutual information model, we apply the Blahut-Arimoto algorithm as described in Caplin, Dean, and Leahy (2019) which iterates over the first-order condition for optimal stochastic choice and updates the marginal distribution over actions until convergence.

Figure C-1: Equilibria in the Beauty Contest



Notes: Each plot is a 2-D histogram of $||T\hat{X} - \hat{X}||$, where $||\cdot||$ indicates the Euclidean norm. Whiter colors indicate smaller values, and hence "closeness to equilibrium." The cross marks represent equilibria, defined such that $||T\hat{X} - \hat{X}|| < 10^{-6}$.

define the operator $T : \mathbb{R}^2 \to \mathbb{R}^2$ which constructs essentially the "best response" of aggregates to aggregates by composing the best response with the aggregator:

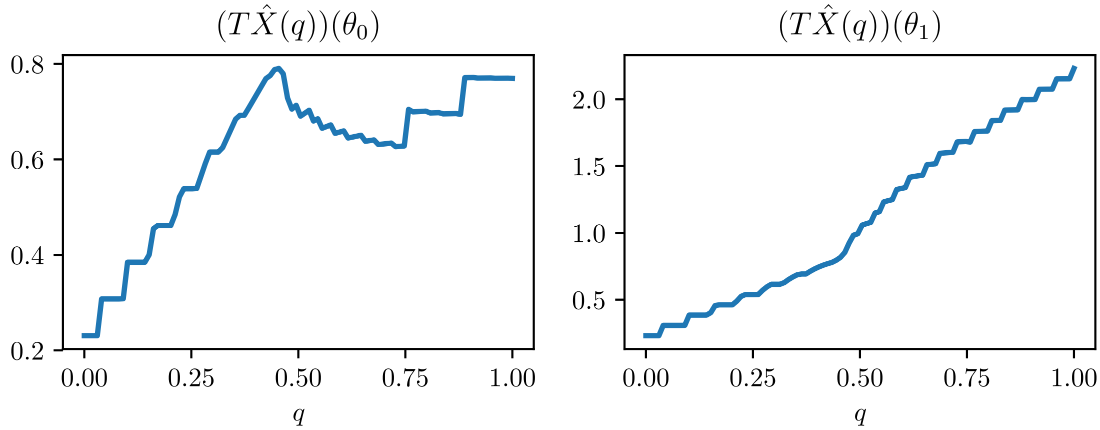
$$T\hat{X} = \Big(X \circ p^*(\theta_0; \hat{X}), X \circ p^*(\theta_1; \hat{X})\Big) \tag{785}$$

We define equilibria by first searching over a grid covering $[\underline{x}, \overline{x}]^2$ for approximate fixed points $\hat{X}$, or low $||T\hat{X} - \hat{X}||$, and then using a numerical fixed-point solving algorithm with fine tolerance to confirm equilibria.

Equilibrium Uniqueness and the Contraction Map

Figure C-1 plots the accuracy of the equilibrium conjecture, $||T\hat{X} - \hat{X}||$, in a heat map or two-dimensional histogram over the grid of candidate conjectures. Whiter areas denote that the equilibrium conjecture is closer to the aggregate best response, bluer areas indicate the opposite, and crosses identify equilibria. The strategic mistakes model, on the left, features a single-peaked surface and a single equilibrium. This is consistent with our theoretical results, and with the fixed-point condition (785) being a contraction. The mutual information model, on the right, features a non-monotone surface and 18 confirmed equilibria.

Figure C-2: Partial Equilibrium Comparative Statics With Mutual Information



Notes: These plots show aggregate best response $T\hat{X}$ in state θ_0 (left pane) and θ_1 (right pane) along the path (786) for the equilibrium conjecture.

We now deconstruct further the failure of the contraction map argument for the mutual information model. Recall, in our proof of Theorem 1, that establishing monotonicity and discounting for the equilibrium operator T required first showing monotone and smooth comparative statics for the single-agent decision problem. To "test" this in the mutual information model, we parameterize a path that increases the equilibrium conjecture of $\hat{X}$ from $(0, 0)$ to one of its equilibrium values.[7] Formally, if we label this chosen equilibrium as $X^*_{MI} = (X^*_{\text{MI}}(\theta_0), X^*_{\text{MI}}(\theta_1))$, we consider points indexed by $q \in [0, 1]$:

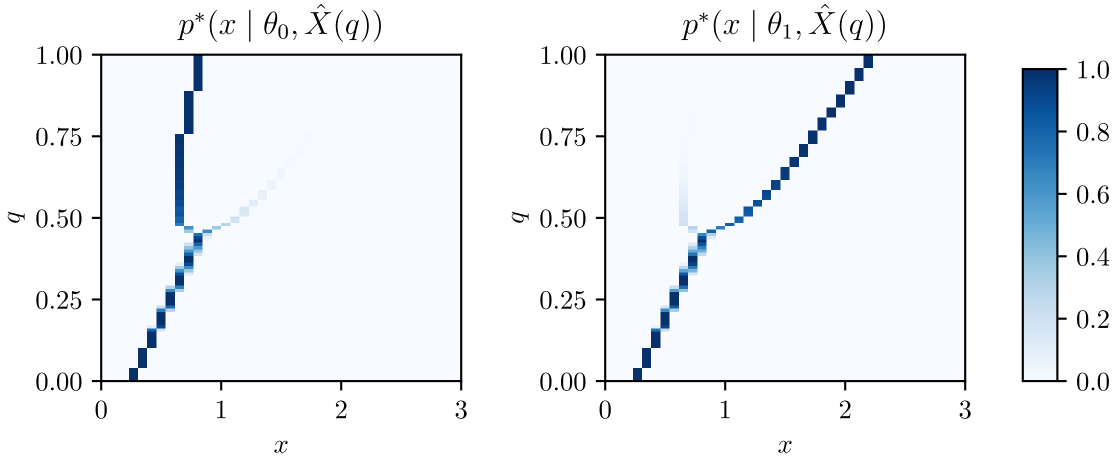
$$\hat{X}(q) = (q \cdot X^*_{\text{MI}}(\theta_0), q \cdot X^*_{\text{MI}}(\theta_1)) \tag{786}$$

and the aggregate best response $T\hat{X}(q)$. Figure C-2 shows each element of $T\hat{X}(q)$ as a function of q. The first element, plotted in the left panel, is (i) non-monotone and (ii) discontinuous in the equilibrium conjecture. In the language of the price-setting application, the mutual information model does not predict that expecting a higher price level increases one's own price, even though the payoff to setting a higher price has globally increased; and when prices increase, they may jump suddenly.

To better understand the agent's behavior along this path, we show in Figure C-3 a two-dimensional histogram of the stochastic choice patterns conditional on each conjecture indexed by q. Equilibrium strategies are mostly supported on either one

[7] We pick the equilibrium with the largest value of $\hat{X}(\theta_1) - \hat{X}(\theta_0)$.

Figure C-3: Stochastic Choice Strategies With Mutual Information

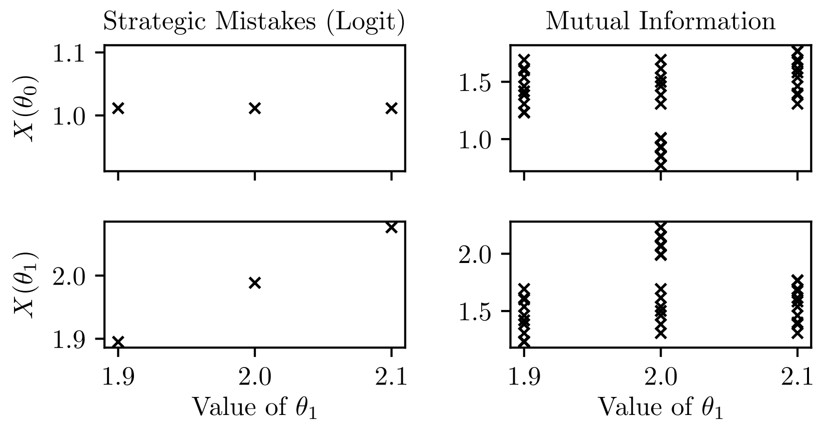


Notes: Each slice on the the vertical axis (q) gives the probability distribution of actions in state 0 (left) or 1 (right), represented via a "heat map" (scale on the right). The path of the equilibrium conjecture corresponds to the same in Figure C-2.

or two points. This sparsity of support is formally described by **?** and Caplin, Dean, and Leahy (2019) in discrete- and continuous-action variants of the mutual information model as a natural consequence of the lowered marginal costs (or, more loosely, "increasing returns to scale") of allocating probability mass to frequently played actions. Sparse behavior is a characteristic feature of the optimal policy in price-setting applications studied by Matějka (2015) and Stevens (2019). In our example, the optimal policy switches between one and two support points around $q = 0.45$. Matějka (2015) refers to such behavior as a *bifurcation* in the optimal policy. As q increases after the bifurcation point, the optimal policy in Figure C-3 pushes the larger and smaller support points away from one another. This violates monotonicity in the sense of first-order stochastic dominance, and therefore can lead to a non-monotone aggregate with respect to some admissible aggregators. Under our chosen aggregator, this behavior causes $X(\theta_0)$ to decrease, as evident in the left panel of Figure C-2.

Our observation is that this force can support multiple equilibria in coordination games because it breaks the contractive properties of the equilibrium map. These multiple equilibria are not, in our reading, very easily interpretable given that choices have an ordinal interpretation, payoffs leverage this interpretation in their definition of complementarity, and agents have a continuum of possible options. This reasoning is quite stark in the price-setting application which Matějka (2015) and Stevens (2019)

Figure C-4: Equilibrium Comparative Statics in the Beauty Contest



Notes: Each cross mark is an equilibrium, under the strategic mistakes (left) and mutual information (right) models, for different values of θ_1. Note the different axis scales in each figure.

study with mutual information. While it is quite reasonable that a single firm wavers between charging \$1.99 and \$2.99 for its product, and indeed Stevens (2019) provides direct evidence for such behavior, it is a much stronger prediction that an entire (symmetric) economy of firms switches between a coordinated equilibrium of charging (\$1.99, \$2.99), respectively in each of two states of nature, to a different equilibrium of charging (\$1.98, \$3.00).

Equilibrium Comparative Statics

A point emphasized in our theory, and in particular the transition from Theorem 1 (existence and uniqueness) to Theorem 2 and Theorem 3 (monotone aggregates and precision), was that the contraction map structure goes hand-in-hand with proving equilibrium comparative statics. We now illustrate the contrast between comparative statics with strategic mistakes and information acquisition in our model. We vary the value of the higher state θ_1 on the grid $\{1.90, 2.00, 2.10\}$ and re-solve for all equilibria of each model. Our main results for the strategic mistakes model suggest that $X^*(\theta_1)$ should monotonically increase in that model while $X^*(\theta_0)$ stays constant, owing to the separability of decisions by state. For the mutual information model, there are no equivalent theoretical results.

Figure C-4 plots the equilibria of each model as a function of the chosen θ_1. In the strategic mistakes model, we verify the predicted comparative statics across unique equilibria. In the mutual information model, we observe non-monotone comparative statics as equilibria move in and out of the set. Thus, while a mutual-information model may be an appealing laboratory to study specific behaviors like discrete pricing, it may not lend itself to straightforward comparative statics analysis conditional on this feature outside of specific numerical calibrations.[8]

C.3 State-Separable Costs in Binary-Action Games

In this Appendix, we adapt our analysis to study binary-action games, which are also common for modeling coordination phenomena in macroeconomics and finance. We first provide results ensuring existence, uniqueness and monotone comparative statics. We next apply our results to study the "investment game," introduced by Carlsson and Van Damme (1993) and studied recently by Yang (2015) and Morris and Yang (2019). Bridging our continuous-action and binary-action analyses, we finally discuss how the the action space can have a large bearing on our model's uniqueness predictions. This may be an important consideration for researchers when the choice of action space is primarily based on analytical convenience and not descriptive realism regarding adjustment on an extensive margin.

C.3.1 Existence, Uniqueness, and Comparative Statics

We now study the same environment as Section 3.2 with the sole change that agents now have a binary action set $\mathcal{X} = \{0, 1\}$.[9] Let $p(\theta)$ denote the probability that a given agent plays action 1 in state θ. It is without loss of generality to restrict to the aggregator $X(p(\theta)) = p(\theta)$, since transformations of this aggregate can be applied within payoffs, and we adopt this convention throughout. Given a conjecture for the law of motion of the aggregate $\hat{p}$ and state $\theta \in \Theta$, we define the cost-adjusted benefit of playing action 1 over action 0 as:

$$\Delta\tilde{u}(\hat{p}(\theta), \theta) \equiv \frac{u(1, \hat{p}(\theta), \theta) - u(0, \hat{p}(\theta), \theta)}{\lambda(\hat{p}(\theta), \theta)} \tag{787}$$

We let $\Delta\tilde{u}_X$ denote this function's derivative in the first argument.

[8] Of course whether this is a "bug" or instead a "feature," reflecting the unstable coordinational nature of activities like price-setting, is an open question that merits additional research. Stevens (2019), for instance, uses a model of coarse pricing with mutual-information costs to match micro-level evidence on pricing strategies and macroeconomic dynamics for aggregates. The microeconomic calibration builds the case that non-uniqueness and ambiguous comparative statics may indeed be features of the "correct" descriptive model of this setting.

[9] Naturally, all integrals are now replaced with summations and density functions by mass functions.

We now provide an existence and uniqueness result. To do so, we place the following regularity condition on the stochastic choice functional:[10]

Assumption 15. *The kernel of the cost functional satisfies the Inada condition* $\lim_{x \to 0} \phi'(x) = -\infty$. *Moreover,* ϕ'' *is globally strictly convex.*[11]

This rules out stochastic choice rule's being concentrated on only one of the two actions in any state. The result follows:[12]

Proposition 43. *Suppose that* ϕ *satisfies assumption 15 and* $\Delta u(p, \theta)$ *is continuously differentiable in its first argument. There exists an equilibrium. All equilibria are symmetric. A sufficient condition for there to be a unique* $p^*(\theta)$ *is that:*

$$\max_{p \in [0,1]} \Delta \tilde{u}_X(p, \theta) < 2\phi'' \left(\frac{1}{2} \right) \tag{788}$$

A sufficient condition for there to be a unique p^* *is that* (788) *holds for all* $\theta \in \Theta$.

Proof. Under Assumption 15, for any θ, we have that $p^*(\theta) \in (0, 1)$. Thus equilibrium is characterized by the first-order condition obtained by moving probability of playing zero to playing one. Thus, the condition characterizing equilibrium is given by:

$$\Delta u(p^*(\theta), \theta) = \lambda(p^*(\theta)) \left(\phi'(p^*(\theta)) - \phi'(1 - p^*(\theta)) \right) \tag{789}$$

To prove uniqueness for a given θ it is sufficient to prove that the minimal slope of the RHS exceeds the maximal slope of the LHS:

$$\max_{p \in [0,1]} \Delta \tilde{u}_X(p, \theta) < \min_{p \in [0,1]} \phi''(p) + \phi''(1 - p) \tag{790}$$

If ϕ'' is strictly convex, then the problem is solved by solving the FOC:

$$\phi'''(p) = \phi'''(1 - p) \tag{791}$$

As ϕ'' is strictly convex, ϕ''' is strictly increasing and is therefore invertible. Thus the unique solution is $p = \frac{1}{2}$ and the minimized value is $2\phi''(\frac{1}{2})$. Applying this argument state by state yields the global condition. □

[10] For existence, this can be weakened in the obvious way: the objective need only be continuous. We present results with this stronger assumption for brevity.

[11] Note that, in view of the Inada condition, it is impossible for ϕ'' to be globally strictly concave.

[12] One can extend this result in the obvious way beyond the differentiability assumption to allow for Lipschitz continuous $\Delta u(p, \theta)$. Naturally, the key property being ruled out is a sudden threshold around which the gains from playing action 1 change discontinuously.

Condition (788) checks the maximum value of complementarity (left-hand-side) against the lowest value for the slope of the marginal cognitive cost of investing (right-hand-side), which is realized at $p = \frac{1}{2}$.[13] We will provide a simple graphical intuition for this condition in the upcoming example.

It is moreover simple to establish when the aggregate $p^*(\theta)$ increases in θ. As in our main analysis, this simply requires supermodularity of payoffs in (x, p, θ), or that higher actions by others and states are complementary with playing $x = 1$:

Assumption 16 (Joint Supermodularity)**.** *The cost-adjusted benefit of playing action 1 over action 0 satisfies, for all $p' \geq p, \theta' \geq \theta$:*

$$\Delta\tilde{u}(p', \theta') \geq \Delta\tilde{u}(p, \theta) \tag{792}$$

Proposition 44. *Suppose that Assumptions 15 and 16 hold, and that the inequality in Equation 788 holds for all $\theta \in \Theta$ so that there is a unique equilibrium p^*. The unique equilibrium $p^*(\theta)$ is monotone increasing in θ.*

Proof. Under Assumption 15, the equilibrium is characterized by Equation 789. Under the assumption that the inequality in Equation 788 holds, there is a unique solution $p^*(\theta)$ for all $\theta \in \Theta$. Note that that this unique equilibrium occurs when $\Delta\tilde{u}(p, \theta)$ intersects $\phi'(p) - \phi'(1 - p)$ from above. Moreover, by Assumption 16 we know that $\Delta\tilde{u}(p, \theta)$ is increasing in (p, θ). Thus, when we take $\theta' \geq \theta$, we know that the unique intersection occurs for $p^*(\theta') \geq p^*(\theta)$. □

Analogous results with general information acquisition or stochastic choice, by contrast, require more extensive analysis (see, *e.g.,* Yang, 2015; Morris and Yang, 2019).

C.3.2 Application: The Investment Game

We now apply these results in a variant of the binary-action investment game introduced by Carlsson and Van Damme (1993), which models coordination motives in financial speculation. Each agent chooses an action $x \in \{0, 1\}$, or "not invest" and "invest." The state of nature $\theta \in \Theta \subseteq \mathbb{R}$ scales the desirability of investing independent of other conditions. Agents' payoffs depend on the action, the total fraction of investing agents, and the state of nature separably and linearly:

$$u(x, p, \theta) = x(\theta - r(1 - p)) \tag{793}$$

[13] That $p = \frac{1}{2}$ is such a point can be derived by noting the symmetry of the state-separable cost around $p = \frac{1}{2}$ and the convexity of ϕ.

where $r \geq 0$ scales the degree of strategic complementarity between investment decisions.

It is straightforward to derive the following fixed-point equation that describes the equilibria of the model when ϕ satisfies the Inada condition in Assumption 15 and $\lambda(p, \theta) \equiv 1$:

$$\theta + rp(\theta) - r = \phi'(p(\theta)) - \phi'(1 - p(\theta)) \tag{794}$$

Equilibrium is guaranteed to be unique by Proposition 44 provided that the following condition holds relating strategic complementarity r with the second derivative of the kernel ϕ:

$$r < 2\phi'' \left(\frac{1}{2}\right) \tag{795}$$

This condition is independent of the state space Θ or the prior. But it does depend on the scale and character of cognitive costs through $\phi''\left(\frac{1}{2}\right)$.

Condition (795) admits the following interpretation about uniqueness with vanishing costs under arbitrary functional forms. For any positive (but arbitrarily small) level of strategic complementarity, and with a sufficiently rich state space, there will be multiple equilibria for a sufficiently small cost of stochastic choice:

Corollary 15. *Consider a family of investment games $\{\mathcal{G}_\lambda : \lambda \in (0, L]\}$ with fixed payoffs, action space, and state space, each with the re-scaled cost functional for some common $\hat{\phi}$ that satisfies Assumption 15, i.e., $\phi_\lambda = \lambda\hat{\phi}$. Then, for all*

$$\lambda > L^* := \frac{r}{2\hat{\phi}''\left(\frac{1}{2}\right)} \tag{796}$$

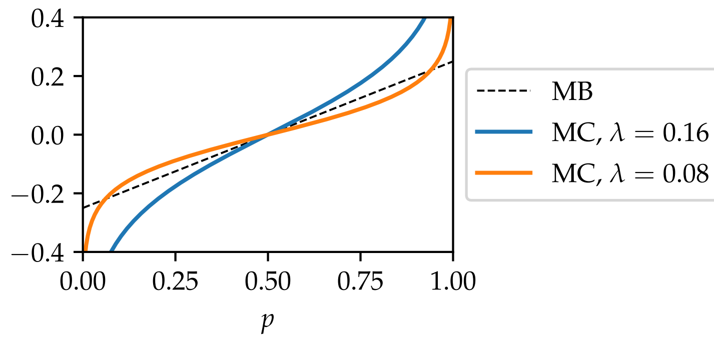
game $\mathcal{G}_\lambda$ has a unique action profile $(p^*(\theta))_{\theta\in\Theta}$. Conversely, when $\lambda < L^*$, there exists at least some $\theta^* \in \mathbb{R}$ such that the equilibrium of $\mathcal{G}_\lambda$ is not unique if $\theta^* \in \Theta$.

Proof. Recall that for any ϕ_λ (owing to $\hat{\phi}$ satisfying Assumption 15), we have that:

$$\theta + rp^*(\theta) - r = \phi'_\lambda(p^*(\theta)) - \phi'_\lambda(1 - p^*(\theta)) \tag{797}$$

Consider state $\theta^* = \frac{r}{2}$. In this state, we have that $p^*(\theta^*) = \frac{1}{2}$ is an equilibrium. Moreover, see that the slope of the LHS in p is given by r and the slope of the RHS in p at $p = \frac{1}{2}$ is given by $2\lambda\hat{\phi}''(\frac{1}{2})$. Hence, when $\lambda < \frac{r}{2\hat{\phi}''(\frac{1}{2})}$, we have that the slope of the LHS exceeds the slope of the RHS. But we know that the RHS is continuous on $(0, 1)$ and that $\lim_{p\to 1} \phi'(p) - \phi'(p) = \infty$. Thus, the RHS must intersect the LHS from below for some other $p \in (\frac{1}{2}, 1)$. Thus, in state θ^*, if $\lambda < \frac{r}{2\hat{\phi}''(\frac{1}{2})}$ there are multiple $p^*(\theta)$

Figure C-5: Multiplicity in the Investment Game



Notes: The dotted line is the marginal benefits of investing more often as a function of others' investment probability, or the right-hand side of (794). The blue and orange lines are the marginal costs of investing more often under respectively more and less severe costs of stochastic choice. Each intersection is an equilibrium.

that can arise in equilibrium. Consequently, if $\theta^* \in \Theta$ and $\lambda < \frac{r}{2\hat{\phi}''(\frac{1}{2})}$, we have that equilibrium is not globally unique. The final claim that we have global uniqueness for $\lambda > \frac{r}{2\hat{\phi}''(\frac{1}{2})}$ follows immediately from Theorem 43. $\square$

The result contrasts with Corollary 1 which showed limit uniqueness in the generalized beauty contest. We will further discuss this issue in Section C.3.4.

To illustrate the uniqueness result, we consider a specialization of the model in which the kernel function is $\hat{\phi}(x) = x \log x$. In this case, $\phi''(0.5) = 2$ and the cost threshold for uniqueness is $L^* = \frac{r}{4}$.[14] Figure C-5 illustrates the scope for multiplicity in a benchmark parameter case of this logit model. We fix $r = 0.50$, and $\theta = 0.25$, the state such that a 50% aggregate investment corresponds with zero payoff. The dotted black line is the "Marginal Benefit," which corresponds with the left-hand-side of (794). The blue and orange lines are the the "Marginal Cost" of increasing the investing probability, or the right-hand-side of (794), with respectively higher and lower values of λ or costs of attention. By construction, there is an equilibrium with $p = \frac{1}{2}$ for any value of λ. Whether or not there are additional equilibria corresponding to more "confident" play, or p closer to 0 or 1, depends on the *slope* of these marginal costs. When λ is high (blue line), it is costly to play more certainly and hence there is only one intersection with the dotted line. When λ is low (orange line), marginal

[14] This is exactly the condition obtained by Yang (2015) for this game with information acquisition costs proportional to mutual information. This foreshadows a deeper connection which we will explore in the next subsection.

costs cross marginal benefits from above at $p = 0.5$. This visualizes a violation of the condition in Proposition 43. As a result there are two more "confident" equilibria near $p = 0$ and $p = 1$.

The right-hand-side of the confident-wavering condition (795) is a well-defined moment which researchers may try to calibrate via laboratory experiments and could interpret in our model without taking a stand on the entire ϕ function. In this way, (795) can be read as a sufficient statistic gauge of the potential for multiplicity and fragility that relies only on one informative aspect of the underlying stochastic choice model.

C.3.3 State-Separable *vs.* Mutual Information Costs

In the vein of our main analysis' comparison of beauty contests with strategic mistakes and mutual information, we now compare the investment game under logit strategic mistakes with the equivalent game under mutual information, as studied by Yang (2015). Observe first that the mutual information model does not always admit an interior solution. Intuitively, if agents place an arbitrarily high prior weight on fundamentals always being very high or very low, they may decide to unconditionally invest or dis-invest without learning anything. These scenarios are ruled out by respectively assuming $\mathbb{E}_\pi\left[\exp\{\lambda^{-1}\theta\}\right] > \exp\{\lambda^{-1}r\}$ and $\mathbb{E}_\pi\left[\exp\{-\lambda^{-1}\theta\}\right] > 1$. No analogue of either is possible in the strategic mistakes model with logistic choice which always features positive probability of playing both actions in all states, so these conditions *a fortiori* rule out an application of Lemma 16. Nonetheless, after ruling out these cases, we can show the following:

Corollary 16. *Compare identical investment games $\mathcal{G}^{LSM}$ and $\mathcal{G}^{MI}$, distinguished by their costs of stochastic choice, scaled by a common scalar λ. Assume*

1. *(Interiority)* $\mathbb{E}_\pi\left[\exp\{\lambda^{-1}\theta\}\right] > \exp\{\lambda^{-1}r\}$ *and* $\mathbb{E}_\pi\left[\exp\{-\lambda^{-1}\theta\}\right] > 1$
2. *(Global uniqueness)* $r < 4\lambda$

Each game has a unique equilibrium $(p^{LSM}(\cdot), p^{MI}(\cdot))$. Moreover,

$$\begin{cases} p^{LSM}(\theta) = p^{MI}(\theta), \forall\theta & \text{if } \sum_\Theta p^{MI}(\theta)\,\pi(\theta) = 1/2, \\ p^{LSM}(\theta) < p^{MI}(\theta), \forall\theta & \text{if } \sum_\Theta p^{MI}(\theta)\,\pi(\theta) > 1/2, \\ p^{LSM}(\theta) > p^{MI}(\theta), \forall\theta & \text{if } \sum_\Theta p^{MI}(\theta)\,\pi(\theta) < 1/2. \end{cases} \tag{798}$$

Proof. It follows from Proposition 2 of Yang (2015), that when $\mathbb{E}_\pi\left[\exp\{\lambda^{-1}\theta\}\right] >$

$\exp\{\lambda^{-1}r\}$ and $\mathbb{E}_\pi\left[\exp\{-\lambda^{-1}\theta\}\right] > 1$, the equilibria of the game with mutual information cost are characterized by:

$$\theta + rp^{\mathrm{MI}}(\theta) - r = \lambda\left[\ln\left(\frac{p^{\mathrm{MI}}(\theta)}{1-p^{\mathrm{MI}}(\theta)}\right) - \ln\left(\frac{\bar{p}^{\mathrm{MI}}}{1-\bar{p}^{\mathrm{MI}}}\right)\right] \tag{799}$$

for all $\theta \in \Theta$ where $\bar{p}^{\mathrm{MI}} = \sum_\Theta p^{\mathrm{MI}}(\theta)\pi(\theta)$. It moreover follows from Proposition 3 of Yang (2015) that when $r < 4\lambda$, this model features a unique equilibrium. Recall that when $r < 4\lambda$ our model with entropic stochastic choice also features a unique equilibrium and this is characterized by:

$$\theta + rp^{L}(\theta) - r = \lambda\left[\ln\left(\frac{p^{L}(\theta)}{1-p^{L}(\theta)}\right)\right] \tag{800}$$

Moreover, when $\bar{p}^{\mathrm{MI}} > \frac{1}{2}$, we have that $\ln\left(\frac{\bar{p}^{\mathrm{MI}}}{1-\bar{p}^{\mathrm{MI}}}\right) > 0$, when $\bar{p}^{\mathrm{MI}} = \frac{1}{2}$, we have that $\ln\left(\frac{\bar{p}^{\mathrm{MI}}}{1-\bar{p}^{\mathrm{MI}}}\right) = 0$ and when $\bar{p}^{\mathrm{MI}} < \frac{1}{2}$, we have that $\ln\left(\frac{\bar{p}^{\mathrm{MI}}}{1-\bar{p}^{\mathrm{MI}}}\right) < 0$. It is then immediate that $p^L(\theta) < p^{\mathrm{MI}}(\theta)$ when $\bar{p}^{\mathrm{MI}} > \frac{1}{2}$, $p^L(\theta) = p^{\mathrm{MI}}(\theta)$ when $\bar{p}^{\mathrm{MI}} = \frac{1}{2}$, and $p^L(\theta) > p^{\mathrm{MI}}(\theta)$ when $\bar{p}^{\mathrm{MI}} < \frac{1}{2}$. □

Conditional on interiority, anchoring in the mutual information model distorts the choice probabilities but perhaps more surprisingly is completely separable from the game's uniqueness properties. More formally, in binary-action games with mutual information, the only difference between the strategic mistakes model with entropy is that log-odds ratio $\log\left(\frac{p(\theta)}{1-p(\theta)}\right)$ in state $\theta \in \Theta$ differs across the models by a state-independent additive constant. In our earlier graphical analysis, this can be seen as a vertical shift of the marginal cost curve. Thus, our confident wavering argument applies directly to the mutual information model and offers an alternative window into the main result of Yang (2015). This separability of anchoring from uniqueness properties with binary actions may be an independently useful insight in other models with mutual information cost.

C.3.4 Discussion: Global *vs.* Local Mistakes

Binary-action settings are sometimes used as a convenient metaphor for underlying environments with many possible actions—for instance, simplifying financial speculation as the choice between extremes of investing and dis-investing instead of a continuous portfolio choice. Our analysis reveals that, in models of stochastic choice, the restriction to two extreme actions may significantly change the character of the game because it removes the possibility of local substitution of actions. The binary-action game allows for "global mistakes," like fully investing when fully disinvesting is instead optimal, that impose discontinuously different externalities and can support multiple equilibria. Our benchmark continuous-action model implies by contrast that agents make "local mistakes" like substituting an optimal action with an alternative that is sub-optimal but nearby in the action space. Whether an analyst should use the binary-action or continuous-action model then depends on the problem at hand and how seriously they take the prediction of global substitution relative to the potential loss in tractability.

Our results also contrast with those in the global games literature in which there is, instead of stochastic choice, vanishing private measurement error in observing the fundamental (Carlsson and Van Damme, 1993; Morris, Rob, and Shin, 1995; Frankel, Morris, and Pauzner, 2003). When combined with the earlier observation linking strategic mistakes with cross-sectional heterogeneity in payoff functions (Section 3.2.1), our results draw a sharp distinction between measurement errors for payoffs (studied here, which do not yield limit uniqueness) and measurement errors for fundamentals (studied in the aforementioned literature, which do yield limit uniqueness). One way of thinking about the difference is that the "contagion" argument formalized in the above references, which shows that having dominant actions in specific states iteratively implies unique rationalizable actions in neighboring states, has no analogue in the present model with no interim beliefs or cross-state reasoning. A different interpretation is that the mere observation that agents have trembling hands is not sufficient to imply the sharp and specific predictions of canonical global games, a point also made by Yang (2015) and Morris and Yang (2019).

Appendix D

Appendix to Priority Design in Centralized Matching Markets

D.1 Omitted Proofs

D.1.1 Proof of Theorem 5

Proof. Fix an arbitrary coarsening Ξ. Divide the schools to two subsets: $c \in \tilde{\mathcal{C}}$ if Ξ_c is finite and $c \in \hat{\mathcal{C}}$ if Ξ_c is the identity function. In what follows, we will construct an alternative coarsening Ξ' that has three indifference classes in some school c and induces the same allocation. Since c was arbitrary, replicating this for all c will yield a trinary coarsening Ξ'.

For any school $c \in \tilde{\mathcal{C}}$, the coarsening Ξ takes any student with score $s_c^{\tilde{\theta}}$ to an equivalence class, hence $\Xi_c(s_c^{\tilde{\theta}}) \in \{P_1^c, P_2^c, \ldots, P_N^c\}$ with $P_1^c < P_2^c < \ldots < P_N^c$ for some N. By Lemma 17 in Appendix D.2, there exists a matching $\tilde{\mu}$ in the coarsened ordinal economy with tie-breakers $\tilde{\Omega}_{\Xi}^{\tau}$ such that for all $(\tilde{\theta}_{\Xi}, \tau) \in \tilde{\Theta}_{\Xi}^{\tau}$, its matching is uniquely given by $\tilde{\mu}(\tilde{\theta}_{\Xi}, \tau)$. Moreover, by Lemma 18 in Appendix D.2, any type $\theta \in \Theta$ whose coarsened ordinal type is given by $\tilde{\theta}_{\Xi}$ has assignment probability at each school c, $g_{\Xi}(c|\theta)$, that is obtained by integrating $\tilde{\mu}(\tilde{\theta}_{\Xi}, \tau)$ over τ.

We will construct an alternative coarsening Ξ' that has only 3 indifference classes for school c and induces the same allocation as Ξ, i.e. $g_{\Xi}(c|\theta) = g_{\Xi'}(c|\theta)$ for all $\theta \in \Theta$ and $c \in \mathcal{C}$. Let P_x^c be the lowest indifference class that has a student placed in school c, i.e. $\tilde{\mu}(\tilde{\theta}_{\Xi}, \tau) = c$ for some $(\tilde{\theta}_{\Xi}, \tau)$ with $s_c^{\tilde{\theta}_{\Xi}} = P_x^c$ and $\tilde{\mu}(\tilde{\theta}_{\Xi}, \tau) \neq c$ for all $(\tilde{\theta}_{\Xi}, \tau)$ with $s_c^{\tilde{\theta}_{\Xi}} = P_y^c$ where $y < x$.[1] Now, define Ξ' by merging all indifference classes above and

[1] Note that this allows $x = 1$ and $x = N$, i.e. this class can be the lowest indifference class or the highest indifference class.

below x for school c:

$$\Xi'_c(s_c^{\tilde{\theta}}) = \begin{cases} P_1^c & , \text{ if } \quad \Xi_c(s_c^{\tilde{\theta}}) = P_z^c, z < x, \\ P_x^c & , \text{ if } \quad \Xi_c(s_c^{\tilde{\theta}}) = P_x^c, \\ P_N^c & , \text{ if } \quad \Xi_c(s_c^{\tilde{\theta}}) = P_z^c, z > x. \end{cases} \tag{801}$$

And $\Xi'_{c'} = \Xi_{c'}$ for all $c' \neq c$. To see that $\tilde{\mu}$ is stable under Ξ', assume (ordinal) student pair i, j (with scores $s_{c'}^i$ and $s_{c'}^j$ and tie-breakers τ_i and τ_j) and school c' blocks $\tilde{\mu}$. Then, $c' \succ_i \tilde{\mu}(i)$, $\tilde{\mu}(j) = c'$ and either $\Xi'_{c'}(s_{c'}^i) > \Xi'_{c'}(s_{c'}^j)$ or $\Xi'_{c'}(s_{c'}^i) = \Xi'_{c'}(s_{c'}^j)$ and $\tau_i > \tau_j$.[2] First, if $c' \neq c$, as the priority for school c' is same under both Ξ and Ξ', (i, j, c') block $\tilde{\mu}$ under Ξ, which is a contradiction. If $c = c'$, $\Xi'_{c'}(s_{c'}^i) > \Xi'_{c'}(s_{c'}^j)$ implies that $\Xi_{c'}(s_{c'}^i) > \Xi_{c'}(s_{c'}^j)$ and (i, j, c') again block $\tilde{\mu}$ under Ξ, which is a contradiction. Next, suppose $\Xi'_{c'}(s_{c'}^i) = \Xi'_{c'}(s_{c'}^j)$ and $\tau_i > \tau_j$. As $\tilde{\mu}(j) = c'$ and $\Xi'_{c'}(s_{c'}^i) = \Xi'_{c'}(s_{c'}^j)$, we have that $\Xi_{c'}(s_{c'}^i) \geq P_x^{c'}$. There are two cases, either $\Xi_{c'}(s_{c'}^i) > P_x^{c'}$ or $\Xi_{c'}(s_{c'}^i) = P_x^{c'}$. In the first case, from the definition of $P_x^{c'}$, there exists k such that $\tilde{\mu}(k) = c'$, $\Xi_{c'}(s_{c'}^k) = P_x^{c'}$. Then, $\Xi_{c'}(s_{c'}^i) > \Xi_{c'}(s_{c'}^k)$, $c' \succ_i \tilde{\mu}(i)$ and $\tilde{\mu}(k) = c'$, which means that (i, k, c') block $\tilde{\mu}$ under Ξ, a contradiction. In the second case, $\Xi_{c'}(s_{c'}^i) = P_x^{c'}$ and $\Xi'_{c'}(s_{c'}^i) = \Xi'_{c'}(s_{c'}^j)$ imply $\Xi_{c'}(s_{c'}^j) = P_x^{c'}$. However, this violates the stability of $\tilde{\mu}$ under Ξ as $\tau_i > \tau_j$, which is a contradiction. Hence $\tilde{\mu}$ is stable under both Ξ and Ξ'. Moreover, the economy under Ξ' with tie-breakers retains the full support property, so there still is a unique stable matching for both economies (see Lemma 17 in Appendix D.2.1). As the stable matching is unique in both economies, we use the same matching in the construction of g_Ξ and $g_{\Xi'}$, so $g_\Xi = g_{\Xi'}$ (applying Lemma 18 in Appendix D.2.1).

Next, take $c \in \hat{\mathcal{C}}$. Let $\tilde{\mu}$ denote the unique stable matching under Ξ. For any school c, let t_c denote the threshold that the students must clear in order to gain admission to that school. Formally,

$$t_c = \inf\{s_c^\theta : \tilde{\mu}(\theta) = c\} \tag{802}$$

Next, define Ξ' in the following way:

$$\Xi'_c(s_c^\theta) = \begin{cases} 0 & , \text{ if } \quad s_c^\theta < t_c, \\ 1 & , \text{ if } \quad s_c^\theta \geq t_c. \end{cases} \tag{803}$$

[2]We abuse notation slightly by evaluating $\tilde{\mu}$ under Ξ' as the set of types changes under Ξ'. However, this is not an issue as we explicitly refer to the uncoarsened ordinal types of students i and j.

Note that from the stability of $\tilde{\mu}$ and the definition of matching (in particular, property (iv) from Footnote 23), there cannot be a student k such that $c' \succ_k \tilde{\mu}(k)$ and $s^k_{c'} \geq t_{c'}$. To see that $\tilde{\mu}$ is stable under Ξ', assume that student pair i, j (with scores $s^i_{c'}$ and $s^j_{c'}$) blocks it in school c'. Then we have $c' \succ_i \tilde{\mu}(i)$, $\tilde{\mu}(j) = c'$ and $\Xi'_{c'}(i) \geq \Xi'_{c'}(j)$. As $\tilde{\mu}(j) = c'$, we have $\Xi'_{c'}(i) = \Xi'_{c'}(j) = 1$. But then this implies that $s^i_{c'} \geq t_{c'}$, which contradicts the stability of $\tilde{\mu}$ under Ξ as $c' \succ_i \tilde{\mu}(i)$. By the same argument as in the first case, it follows that $g_\Xi = g_{\Xi'}$.

Applying this argument for all $c \in \mathcal{C}$ then yields a trinary coarsening Ξ' that replicates Ξ in the sense that $g_\Xi = g_{\Xi'}$. □

D.1.2 Proof of Theorem 6

Proof. First, by Theorem 5, we parameterize coarsenings Ξ by the outcome equivalent $v \in \mathcal{V}$, where we note that $\mathcal{V}$ is compact. Second, define $\tilde{g} : \mathcal{V} \to \mathcal{G}$ such that $\tilde{g}(v) = g_v$. By Lemma 19 in Appendix D.2.2, we have that $\tilde{g}$ is continuous under the appropriate L^1-norm on $\mathcal{G}$. Third, define $\tilde{Z} : \mathcal{V} \to \mathbb{R}$ as $\tilde{Z} = Z \circ \tilde{g}$. By Assumption 19 that Z is continuous under the L^1-norm and the fact that $\tilde{g}$ is continuous, it then follows that $\tilde{Z}$ is continuous. Fourth, observe that we can write Equation 147 as:

$$\max_{v \in \mathcal{V}} \tilde{Z}(v) \tag{804}$$

Finally, by the extreme value theorem as $\mathcal{V}$ is compact and $\tilde{Z}$ is continuous, it follows that $\mathcal{V}^*$ is non-empty. Thus, an optimal trinary coarsening exists. □

D.1.3 Proof of Corollary 11

Proof. This result follows from Theorem 5 specialized to an environment with two schools. However, in this case there is an alternative, simpler proof that we provide below for completeness.

In the first part of the proof, we show that under any stable mechanism, the allocation takes the following form:

$$g^v_i = \begin{cases} 1, v \geq \overline{v}, \\ p_L, v \in [\underline{v}, \overline{v}), \\ 0, v < \underline{v}. \end{cases} \tag{805}$$

for $0 \leq \underline{v} \leq \overline{v} \leq 1$ and $p_L \in [0, 1]$. To see this, consider priority classes defined by: $i \in P_j \iff s_i \in [v_{j-1}, v_j)$ for $j \leq n$ and $i \in P_{n+1} \iff s_i \in [v_n, v_{n+1}]$. Now suppose that $\exists i \in P_j$ and $k \in P_l$ for $l < j$ such that $g_k^v > 0$. If $g_j^v < 1$, then a positive measure of students in P_j will not be assigned to G and a positive measure of students in P_k will be assigned to G. This violates stability. Hence, $g_j^v = 1$. Now suppose that $\exists i \in P_j$ and $k \in P_l$ for $l < j$ such that $g_j^v < 1$. By an identical argument, it must be that $g_k^v = 0$. By the above two conclusions, it must be that if there is any l such that $g_l^v \in (0,1)$ there is a unique P_j such that $g_j^v \in (0,1)$ and that $g_k^v = 1$ for $k > j$ and $g_k^v = 0$ for $k < j$. Taking $\overline{v} = v_j$ and $\underline{v} = v_{j-1}$ thereby proves that the allocation takes the form given in Equation 805.

Given this claim, we can take a (v, n) that induces g^v and construct a $(v', 2)$ that also induces g^v. If there is no j such that $g_j^v \in (0,1)$, then we can simply take the lowest class k for which $g_k^v = 1$ and set $v_1' = 0$ and $v_2' = v_{k-1}$. If there is a j such that $g_j^v \in (0,1)$, then we can take $v_1' = v_{j-1}$ and $v_2' = v_j$. See that v' induces the same allocation as v in both cases.

Having now established that we can replicate any (v, n) with $(v', 2)$, it remains to show that there exists an optimum to establish the result. See that the optimization problem by the replication result can be rewritten as:[3]

$$\begin{aligned} &\max_{v_1, v_2} \frac{Q - (1 - v_2)}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds + \int_{v_2}^{1} W(s)ds \\ &\text{s.t.} \quad 0 \leq v_1 \leq 1 - Q \leq v_2 \leq 1 \end{aligned} \tag{806}$$

If the function $W(s)$ is continuous, then there must exist a solution by the Weierstrass Extreme Value Theorem as we are simply maximizing a continuous function over a compact domain. As B and C are continuous, then so too is W, so a solution exists. This completes the proof. □

D.1.4 Proof of Proposition 11

Proof. The school district's problem is given by:

$$\max_{n, v \in \mathcal{V}^n} \sum_{i=1}^{n} g_i^v \int_{v_i}^{v_{i+1}} W(s)ds \tag{807}$$

Applying the replication argument in Corollary 11, it is without loss of optimality to impose $n = 2$ and to have one class with a zero probability of assignment $[0, v_1)$, one class $[v_1, v_2)$ which faces a lottery of being assigned with probability $0 \leq p_L(v_1, v_2) \leq 1$ and one class $[v_2, 1]$ with a unit probability of assignment. As the planner has measure

[3] See Proposition 2 for the full argument.

Q seats to assign, it must be that:

$$(1 - v_2) + (v_2 - v_1)p_L(v_1, v_2) = Q \tag{808}$$

Or:

$$p_L(v_1, v_2) = \frac{Q - (1 - v_2)}{v_2 - v_1} \tag{809}$$

Thus the objective becomes:

$$V(v_1, v_2) = \frac{Q - (1 - v_2)}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds + \int_{v_2}^{1} W(s)ds \tag{810}$$

We require that $v_1 \geq 0, v_2 \leq 1, p_L(v_1, v_2) \in [0, 1]$. These requirements reduce to:

$$0 \leq v_1 \leq 1 - Q \leq v_2 \leq 1 \tag{811}$$

Thus the planner's problem is:

$$\begin{aligned} \max_{v_1, v_2} & \frac{Q - (1 - v_2)}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds + \int_{v_2}^{1} W(s)ds \\ \text{s.t.} \quad & 0 \leq v_1 \leq 1 - Q \leq v_2 \leq 1 \end{aligned} \tag{812}$$

From the form of the problem, the Lagrangian can be stated as:

$$\begin{aligned} \mathcal{L}(v_1, v_2, \lambda_1, \lambda_2, \mu_1, \mu_2) = & \frac{Q - (1 - v_2)}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds + \int_{v_2}^{1} W(s)ds - \lambda_1\big((v_1 - (1 - Q)\big) \\ & - \lambda_2\big((1 - Q) - v_2\big) + \mu_1 v_1 + \mu_2(1 - v_2) \end{aligned} \tag{813}$$

where the λ_i are the Lagrange multipliers on the constraints that the probability in the lottery zone does not exceed unity or become negative and the μ_i are Lagrange multipliers on the constraints that the cutoffs remain in the unit interval. See that there are five cases of interest.

1. Both v_1 and v_2 are unconstrained: $\lambda_1 = \lambda_2 = \mu_1 = \mu_2 = 0$. In this case, the Lagrangian becomes:

$$\mathcal{L}(v_1, v_2) = \frac{Q - (1 - v_2)}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds + \int_{v_2}^{1} W(s)ds \tag{814}$$

Taking FOCs:

$$\begin{aligned}
p_L(v_1, v_2)W(v_1) &= p_{Lv_1}(v_1, v_2) \int_{v_1}^{v_2} W(s)ds \\
(1 - p_L(v_1, v_2))W(v_2) &= p_{Lv_2}(v_1, v_2) \int_{v_1}^{v_2} W(s)ds
\end{aligned} \tag{815}$$

Noting that:

$$\begin{aligned}
p_{Lv_1} &= \frac{p_L(v_1, v_2)}{v_2 - v_1} \\
p_{Lv_2} &= \frac{1 - p_L(v_1, v_2)}{v_2 - v_1}
\end{aligned} \tag{816}$$

Plugging these relations into the FOCs yields:

$$\begin{aligned}
W(v_1) &= \frac{1}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds \\
W(v_2) &= \frac{1}{v_2 - v_1} \int_{v_1}^{v_2} W(s)ds
\end{aligned} \tag{817}$$

Thus, when either FOC holds, it must be the case that the marginal contribution to social welfare is equal to average utility within the lottery zone. It further follows that when both FOCs hold that:

$$W(v_1) = W(v_2) \tag{818}$$

This reveals that an interior optimum, it must be that the marginal lottery zone student on average contributes as much to welfare as the marginal student in the zone that gets in with probability one.

However, for the above to hold, we must ensure that the SOCs hold. To this end, it is sufficient to show that the Hessian of the Lagrangian is negative definite.

Taking second derivatives and re-arranging:

$$\begin{aligned}
\mathcal{L}_{v_1v_1}(v_1,v_2) &= \frac{2p_L(v_1,v_2)}{v_2-v_1}\Big[\frac{1}{v_2-v_1}\int_{v_1}^{v_2} W(s)ds - W(v_1)\Big] - p_L(v_1,v_2)W'(v_1) \\
\mathcal{L}_{v_1v_2}(v_1,v_2) &= \frac{1-p_L(v_1,v_2)}{v_2-v_1}\Big[\frac{1}{v_2-v_1}\int_{v_1}^{v_2} W(s)ds - W(v_1)\Big] \\
&\quad - \frac{p_L(v_1,v_2)}{v_2-v_1}\Big[\frac{1}{v_2-v_1}\int_{v_1}^{v_2} W(s)ds - W(v_2)\Big] \\
\mathcal{L}_{v_2v_2}(v_1,v_2) &= \frac{2(1-p_L(v_1,v_2))}{v_2-v_1}\Big[W(v_2) - \frac{\int_{v_1}^{v_2} W(s)ds}{v_2-v_1}\Big] - (1-p_L(v_1,v_2))W'(v_2)
\end{aligned} \tag{819}$$

To show that the Hessian of the Lagrangian is negative (semi-)definite, it suffices to show that $\mathcal{L}_{v_1v_1}(v_1,v_2), \mathcal{L}_{v_2v_2}(v_1,v_2) \leq 0$ with one equality strict at a point satisfying the FOCs. See that if the FOC for v_1 holds that:

$$\mathcal{L}_{v_1v_1}(v_1,v_2) = -p_L(v_1,v_2)W'(v_1) \tag{820}$$

And if the FOC for v_2 holds:

$$\mathcal{L}_{v_2v_2}(v_1,v_2) = -(1-p_L(v_1,v_2))W'(v_2) \tag{821}$$

Thus, if both FOCs are satisfied at (v_1,v_2), it suffices that $W'(v_1), W'(v_2) \geq 0$ (with one strict) for (v_1,v_2) to be a local optimum.

2. The lottery probability is equal to unity or zero: $\mu_1 = \mu_2 = \lambda_i = 0$ for one and only one i. In this case, there is an acceptance zone and a rejection zone. As a result, the objective takes the value:

$$V = \int_{1-Q}^{1} W(s)ds \tag{822}$$

which is simply the total value of all students living closest to the school.

3. Only the lower cutoff is zero: $\mu_2 = \lambda_1 = \lambda_2 = 0$ and $\mu_1 > 0$. In this case $v_2 > 1-Q$ and $v_1 = 0$. Thus, the Lagrangian is:

$$\mathcal{L}(v_2) = \frac{Q-(1-v_2)}{v_2}\int_0^{v_2} W(s)ds + \int_{v_2}^{1} W(s)ds \tag{823}$$

Taking the FOC with respect to v_2 as in the first case yields:

$$W(v_2) = \frac{1}{v_2} \int_0^{v_2} W(s)ds \tag{824}$$

which reveals that the marginal student who gets in for sure should contribute as much to social welfare as the average student in the lottery zone. The SOC for this to be an optimum is:

$$-(1 - p_L(0, v_2))W'(v_2) \leq 0 \tag{825}$$

Or simply:

$$W'(v_2) \geq 0 \tag{826}$$

4. Only the upper cutoff is zero: $\mu_1 = \lambda_1 = \lambda_2 = 0$ and $\mu_2 > 0$. Following the same approach as for the lower cutoff being zero. The Lagrangian is:

$$\mathcal{L}(v_1) = \frac{Q}{1 - v_1} \int_{v_1}^{1} W(s)ds \tag{827}$$

Taking the FOC with respect to v_1 yields:

$$W(v_1) = \frac{1}{1 - v_1} \int_{v_1}^{1} W(s)ds \tag{828}$$

again yielding the insight that the marginal student in the lottery zone should contribute as much to social welfare as the average student in the walk-zone.

The SOC for this to be an optimum is given by:

$$-p_L(v_1, 1)W'(v_1) \leq 0 \tag{829}$$

Or simply:

$$W'(v_1) \geq 0 \tag{830}$$

5. Both the upper and lower cutoff are one and zero, respectively: $\lambda_1 = \lambda_2 = 0$ and $\mu_1, \mu_2 > 0$. In this case, the value is simply given by:

$$V = Q \int_0^1 W(s)ds \tag{831}$$

so value is simply the total utility of all students, weighted by the probability that they are admitted to the school.

From the above analysis, it follows that whenever a cutoff is interior and W is differentiable, the following two conditions are necessary and sufficient for a local optimum:

$$\begin{aligned} W(v_i^*) &= \frac{1}{v_2^* - v_1^*} \int_{v_1^*}^{v_2^*} W(s)ds \\ W'(v_i^*) &> 0 \end{aligned} \tag{832}$$

As a result, an optimal cutoff either satisfies the above conditions, lies on the boundary of the constraint set, or lies at a point of non-differentiability of W. This concludes the proof. □

D.1.5 Proof of Proposition 12

Proof. We explicitly provide sufficient conditions on the parameters such that each class of policy is optimal:

1. Full coarsening: if $u^P > u^R$, then there exist $\bar{\alpha}, \bar{\beta}$ such that when $\alpha < \bar{\alpha}$ and $\beta < \bar{\beta}$, $v_1^* = 0$ and $v_2^* = 1$. To show this, we want to prove in the case that $u^p > u^R$ that there exist $\bar{\alpha}, \bar{\beta} > 0$ such that when $\alpha < \bar{\alpha}$ and $\beta < \bar{\beta}$ the optimum is $v_1^* = 0$ and $v_2^* = 1$. To this end, it is sufficient to show that for $\alpha < \bar{\alpha}$ and $\beta < \bar{\beta}$, $W(s)$ is strictly decreasing on $[0, 1]$. See that $W(s)$ is given by:

   $$W(s) = u^P + s(u^R - u^P) - \frac{\alpha}{1 + \delta \exp\{\varepsilon(s - \bar{s})\}} - \mathbb{I}[s \leq \bar{s}]\beta(\bar{s} - s)^\gamma \tag{833}$$

   As $u^R > u^P$, this is clearly the case so long as α, β are sufficiently small.

2. No coarsening: if $u^P < u^R$ then $v_1^* \leq 1 - Q$, $v_2^* = 1 - Q$. Alternatively, if $u^P > u^R$ there exists $\underline{\alpha}, \underline{\beta}$ such that when $\underline{\alpha} < \alpha$ and $\underline{\beta} < \beta$, $v_1^* \leq 1 - Q$ and $v_2^* = 1 - Q$. In this case, it is sufficient to show that $W(s)$ is strictly increasing. This is transparently the case when $u^P < u^R$. Moreover, whenever $\bar{s} > 1 - Q$, when $u^P > u^R$ and α, β are sufficiently large, then for $s < \bar{s}$, W(s) is strictly increasing and it is optimal to set $v_2^* = 1 - Q$ and $v_1^* < 1 - Q$.

3. 'Small' walk-zone: if $u^P > u^R$, $\gamma > 1$ and $\varepsilon \to \infty$ then there exist $\bar{\beta}$ and $\underline{\alpha}$ such that when $\beta < \bar{\beta}$ and $\alpha > \underline{\alpha}$, $v_1^* = 0$ and $v_2^* \in (1 - Q, 1)$. In this case, it is sufficient to show that W(s) is strictly decreasing for $s < \bar{s}$ and $W(s) < W(s')$ for all $s < \bar{s} < s'$. This is clearly the case for β sufficiently small and α sufficiently large.

4. 'Large' walk-zone: it is sufficient to show that there is a $\tilde{s} < 1-Q$ such that for $s < \tilde{s}$, W(s) is increasing and for $s > \tilde{s}$, W(s) is decreasing and that $W(0) < W(1)$. In this case, it follows that $v_1^* \in (0, 1-Q)$ and $v_2^* = 1$. If $u^P > u^R$ and $\gamma > 1$ then there exist $\underline{\beta}, \overline{\beta}$ and $\bar{\alpha}$ such that when $\underline{\beta} < \beta < \overline{\beta}$ and $\alpha < \bar{\alpha}$, the above is true.

5. 'Double' walk-zone: if $u^P > u^R$ and $\gamma > 1$ and $\varepsilon \to \infty$ then there exist $\underline{\beta}, \bar{\beta}, \underline{\alpha} > 0$ such that when $\underline{\beta} < \beta < \bar{\beta}$ and $\underline{\alpha} < \alpha$, $v_1^* \in (0, 1-Q)$ and $v_2^* \in (1-Q, 1)$. In this case, it is sufficient to show that for $W(s) < W(s')$ for all $s < \bar{s} < s'$ and there is a $\tilde{s} < 1-Q$ such that $W(s)$ is increasing for $s < \tilde{s}$ and decreasing over $\bar{s} > s > \tilde{s}$ and $W(0) < W(\bar{s})$. The first condition is satisfied for α sufficiently large. The second is satisfied so long as β is neither too small nor too large.

□

D.1.6 Proof of Proposition 13

Proof. We first show that for any parameter λ in which W is differentiable that:

$$\frac{\partial v_i^*}{\partial \lambda} \lesseqgtr 0 \iff \int_{v_1^*}^{v_2^*} \left(W_\lambda(s) - W_\lambda(v_i^*)\right) ds \lesseqgtr 0 \tag{834}$$

for $i \in \{1, 2\}$. This result allows one to find comparative statics in any parametric environment and is perhaps of independent interest.

First, note that from optimality and interiority of (v_1^*, v_2^*) and from Proposition 11, we have that

$$W(v_2^*) = W(v_1^*) = \frac{1}{v_2^* - v_1^*} \int_{v_1^*}^{v_2^*} W(s)ds \tag{835}$$

Differentiating $(v_2^* - v_1^*)W(v_i^*) = \int_{v_1^*}^{v_2^*} W(s)ds$ implicitly with respect to λ, we obtain

$$\begin{aligned}\left(\frac{\partial v_2^*}{\partial \lambda} - \frac{\partial v_1^*}{\partial \lambda}\right) W(v_i^*) + (v_2^* - v_1^*) \left(W_\lambda(v_i^*) + \frac{\partial v_i^*}{\partial \lambda} W'(v_i^*)\right) = \\ \frac{\partial v_2^*}{\partial \lambda} W(v_2^*) - \frac{\partial v_1^*}{\partial \lambda} W(v_1^*) + \int_{v_1^*}^{v_2^*} W_\lambda(s)ds\end{aligned} \tag{836}$$

Plugging in $W(v_2^*) = W(v_1^*) = W(v_i^*)$ to the RHS of this equation, we obtain

$$\begin{aligned}\left(\frac{\partial v_2^*}{\partial \lambda} - \frac{\partial v_1^*}{\partial \lambda}\right) W(v_i^*) + (v_2^* - v_1^*)\left(W_\lambda(v_i^*) + \frac{\partial v_i^*}{\partial \lambda} W'(v_i^*)\right) = \\ \left(\frac{\partial v_2^*}{\partial \lambda} - \frac{\partial v_1^*}{\partial \lambda}\right) W(v_i^*) + \int_{v_1^*}^{v_2^*} W_\lambda(s) ds\end{aligned} \tag{837}$$

Cancelling the terms and rearranging, we obtain

$$\begin{aligned}(v_2^* - v_1^*) W'(v_i^*) \frac{\partial v_i^*}{\partial \lambda} &= \int_{v_1^*}^{v_2^*} W_\lambda(s) ds - (v_2^* - v_1^*) W_\lambda(v_i^*) \\ &= \int_{v_1^*}^{v_2^*} \left(W_\lambda(s) - W_\lambda(v_i^*)\right) ds\end{aligned} \tag{838}$$

As the solution is interior, we have $1 > v_2^* > 1 - Q > v_1^* > 0$ and from Proposition 11, $W'(v_i^*) > 0$, which proves that $\frac{\partial v_i^*}{\partial \lambda}$ and $\int_{v_1^*}^{v_2^*} \left(W_\lambda(s) - W_\lambda(v_i^*)\right) ds$ have the same sign for $i \in \{1, 2\}$.

To prove the first part of the proposition, note that $W_\beta(s) = -\mathbb{I}[s \leq \bar{s}](\bar{s} - s)^\gamma$, which is increasing in s. Moreover, as $\bar{s} > 1 - Q$, in an interior solution, $v_1^* < \bar{s}$. Then for any $s \in (v_1^*, v_2^*)$, $W_\beta(v_1^*) < W_\beta(s) \leq W_\beta(v_2^*)$. Thus:

$$\int_{v_1^*}^{v_2^*} \left(W_\beta(s) - W_\beta(v_1^*)\right) ds > 0 > \int_{v_1^*}^{v_2^*} \left(W_\beta(s) - W_\beta(v_2^*)\right) ds \tag{839}$$

So $\frac{\partial v_1^*}{\partial \beta} > 0$ and $\frac{\partial v_2^*}{\partial \beta} < 0$ obtains by equation 834.

To prove the second part of the proposition, note that $W_{u_p}(s) = 1 - s$ is decreasing in s. Then for any $s \in (v_1^*, v_2^*)$, $W_{u_p}(v_1^*) > W_{u_p}(s) > W_{u_p}(v_2^*)$. Thus

$$\int_{v_1^*}^{v_2^*} \left(W_{u_p}(s) - W_{u_p}(v_1^*)\right) ds < 0 < \int_{v_1^*}^{v_2^*} \left(W_{u_p}(s) - W_{u_p}(v_2^*)\right) ds \tag{840}$$

As a result, from equation 834 we have $\frac{\partial v_1^*}{\partial u_P} < 0$ and $\frac{\partial v_2^*}{\partial u_P} > 0$. Other comparative statics can be derived in a similar fashion. $\square$

D.1.7 Proof of Proposition 14

Proof. The utility with cut-offs v_1, v_2 is given by:

$$\begin{aligned} Z(v_1, v_2) = &\sum_{\kappa_i} \int_{v_2}^{1} s dF_{\kappa_i}(s) + \frac{v_2 - r}{v_2 - v_1} \sum_{\kappa_i} \int_{v_1}^{v_2} s dF_{\kappa_i}(s) \\ &+ h\Big(\int_{v_2}^{1} dF_{\kappa_1}(s) + \frac{v_2 - r}{v_2 - v_1} \int_{v_1}^{v_2} dF_{\kappa_1}(s) \Big) \end{aligned} \tag{841}$$

As $f_{\kappa_1}(s) + f_{\kappa_2}(s) = 1$, by computing the score integrals and rearranging, we have that:

$$Z(v_1, v_2) = \frac{1}{2}\Big[1 - v_2^2 + (v_2 + v_1)(v_2 - r)\Big] + h\Big(\int_{v_2}^{1} dF_{\kappa_1}(s) + \frac{v_2 - r}{v_2 - v_1} \int_{v_1}^{v_2} dF_{\kappa_1}(s) \Big) \tag{842}$$

Similarly, utility without coarsening is given by:

$$Z^{NC} = \sum_{\kappa_i} \int_{r}^{1} s dF_{\kappa_i}(s) + h\Big(\int_{r}^{1} dF_{\kappa_1}(s) \Big) = \frac{1}{2}(1 - r^2) + h\Big(\int_{r}^{1} dF_{\kappa_1}(s) \Big) \tag{843}$$

Thus:

$$\begin{aligned} Z(v_1, v_2) - Z^{NC} = &\frac{1}{2}[(v_2 - r)(v_1 - r)] \\ &+ h\Big(\int_{v_2}^{1} dF_{\kappa_1}(s) + \frac{v_2 - r}{v_2 - v_1} \int_{v_1}^{v_2} dF_{\kappa_1}(s) \Big) - h\Big(\int_{r}^{1} dF_{\kappa_1}(s) \Big) \end{aligned} \tag{844}$$

For $Z(v_1, v_2) > Z^{NC}$ to hold, it must be that:

$$\frac{1}{2}[(v_2 - r)(r - v_1)] < h\Big(\frac{1}{2} - F_{\kappa_1}(v_2) + (F_{\kappa_1}(v_2) - F_{\kappa_1}(v_1))\frac{v_2 - r}{v_2 - v_1}\Big) - h\Big(\frac{1}{2} - F_{\kappa_1}(r)\Big) \tag{845}$$

Which proves the first part of the result. Plugging in $h(x) = \alpha x$, we have:

$$[(v_2 - r)(r - v_1)] < 2\alpha\Big(F_{\kappa_1}(r) - F_{\kappa_1}(v_2) + (F_{\kappa_1}(v_2) - F_{\kappa_1}(v_1))\frac{v_2 - r}{v_2 - v_1} \Big) \tag{846}$$

Dividing both sides by $(v_2 - r)$:

$$(r - v_1) < 2\alpha\Big(\frac{F_{\kappa_1}(v_2) - F_{\kappa_1}(v_1)}{v_2 - v_1} - \frac{F_{\kappa_1}(v_2) - F_{\kappa_1}(r)}{v_2 - r} \Big) \tag{847}$$

Yielding the second part of the result. □

D.1.8 Proof of Corollary 12

Proof. This is an immediate consequence of Proposition 49 in Appendix D.5. □

D.1.9 Proof of Proposition 15

Proof. First, let us compute social welfare function under the induced trinary coarsening of any policy, v. In the $\gamma \to \infty$ limit, we have that social welfare is given by (up to constant of proportionality that we omit):

$$Z(v) = \min_{\kappa \in \mathcal{K}, s \in [0,1]} \bar{v}(\kappa, P(s, v)) \tag{848}$$

Note that agents have outside options given by $\kappa = h(s)$ where h is decreasing. Thus, within any equivalence class, the agent with the lowest expected utility is the agent at the upper threshold for each equivalence class. Moreover, we know that the allocation probabilities for these agents are given by 0, $p_L(v)$ and 1, respectively. Thus:

$$Z(v) = \min \{\bar{v}(h(v_1), 0), \bar{v}(h(v_2), p_L(v)), \bar{v}(h(1), 1)\} \tag{849}$$

We now prove both parts of the proposition. In the first case, suppose that all agents have positive assignment probability. It follows that the coarsening is given by a single number $v = v_2$. Under this policy, welfare is given by:[4]

$$Z(v) = \min \{\bar{v}(h(v_2), p_L(v)), \bar{v}(h(1), 1)\} \tag{850}$$

Now consider an alternative policy $v' = (\varepsilon, v_2)$ for $0 < \varepsilon < v_2$. Under this policy, welfare is given by:

$$Z(v') = \min \{\bar{v}(h(\varepsilon), 0), \bar{v}(h(v_2), p_L(v')), \bar{v}(h(1), 1)\} \tag{851}$$

If h is sufficiently steep, then $\bar{v}(h(\varepsilon), 0) > \bar{v}(h(v_2), 1)$. Moreover, $p_L(v') > p_L(v)$. Thus, $Z(v') \geq Z(v)$. Thus an optimal coarsening features a lower cutoff and consequently some agents who have zero assignment probability.

For the second part of the proposition, simply take h to be the constant function. The optimal policy is $v = (0, 1)$ and all agents have interior assignment probability. This completes the proof. □

[4] In the knife-edge case with full coarsening, $Z(v) = \bar{v}(h(1), p_L(v))$. The steps below still follow.

D.2 Additional Technical Results

In this Appendix we collect technical results that are necessary for the proofs of our main results but that have been omitted from the main text for clarity and brevity.

D.2.1 Construction of the Allocation

In this section, we establish that we can characterize the allocation of students to schools under a coarsening Ξ as a conditional probability mass function $g_\Xi(c|\theta)$ that assigns each type θ a probability of being assigned to each school c. This allocation will serve as the object over which the planner has preferences and its construction and properties are important for the analysis.

Under a coarsening Ξ, a school prioritizes student i to student j if i is in a higher indifference class than j or i and j are in the same indifference class and i has higher tie-breaker. Hence the priorities of school c become lexicographic with order (p, τ), where $p = \Xi(s)$ for some score s. By Definition 7, any admissible coarsening map partitions students into a finite number of partitions or retains a strict priority structure. Moreover, there is no aggregate uncertainty in the economy: the same measure of students of each type is assigned to each school with probability one. This is stated rigorously in Lemma 17, which also establishes that the coarsened economy after tie-breaking inherits the full support condition that we assumed for the initial economy.

Lemma 17. *Let Ξ be a coarsening and $\tilde{F}_\Xi^{\tau^*}$ the distribution of students corresponding to a strict ordinal economy for a given realized distribution of tiebreakers.*

1. *$\tilde{F}_\Xi^{\tau^*} = \tilde{F}_\Xi^\tau$ almost surely where $\tilde{F}_\Xi^\tau$ is given by $\tilde{f}_\Xi^\tau(\tilde{\theta}_\Xi, \tau) = \tilde{f}_\Xi(\tilde{\theta}_\Xi)$ for all $\tilde{\theta}_\Xi \in \tilde{\Theta}_\Xi$, $\tau \in [0, 1]$. Moreover, $\tilde{F}_\Xi^\tau$ has full support.*

2. *In the coarsened ordinal economy with tie-breakers $\tilde{\Omega}_\Xi^\tau = (\tilde{F}_\Xi^\tau, Q, \tilde{\Theta}_\Xi^\tau)$, for any student of any induced type $\tilde{\theta}_\Xi = (\succ_{\tilde{\theta}_\Xi}, s_c^{\tilde{\theta}_\Xi}) \in \tilde{\Theta}_\Xi$, there is a unique mapping $\tilde{\mu}_{\tilde{\theta}_\Xi} : [0, 1] \to \mathcal{C}$ that determines the assignment of that student as a function of her realized tie-breaker.*

Proof. Part 1: Each induced type, $\tilde{\theta}_\Xi = (\succ_{\tilde{\theta}_\Xi}, s_c^{\tilde{\theta}_\Xi})$, draws a random number from $U[0, 1]$, by the law of large numbers, for any type $\tilde{\theta}_\Xi$, the distribution of realized tie-breakers $\tilde{F}_\Xi^{\tau^*}(\tau|\tilde{\theta}_\Xi)$ is $U[0, 1]$ almost surely, which gives a unique strict economy type space almost surely.[5] As the interim economy (with induced types) before tie-

[5] This unique economy can be constructed by the following method: for any school, let the number

breaking $\tilde{\Theta}$ has full support by Assumption 7, and $\tilde{f}^{\tau}_{\Xi}(\tilde{\theta}_{\Xi}, \tau) = \tilde{f}_{\Xi}(\tilde{\theta}_{\Xi})$, it follows that $(\tilde{F}^{\tau}_{\Xi}, Q, \tilde{\Theta}^{\tau}_{\Xi})$ has full support.

Part 2: As $\tilde{\Omega}^{\tau}_{\Xi}$ has full support, it has a unique stable matching (see Theorem 1 in Azevedo and Leshno (2016)), denoted by $\tilde{\mu}$. Thus for any type $\tilde{\theta}_{\Xi}$ its assigned school is determined directly as a function of its tie-breaking number τ, $\mu(\tilde{\theta}_{\Xi}, \tau)$. Now define $\tilde{\mu}_{\tilde{\theta}_{\Xi}} : [0,1] \to \mathcal{C}$ so that $\tilde{\mu}_{\tilde{\theta}_{\Xi}}(\tau) = \mu(\tilde{\theta}_{\Xi}, \tau)$. □

Using this fact, under any coarsening Ξ one can construct the *allocation* $g_{\Xi} : \Theta \times \mathcal{C} \to [0,1]$, with the probability that type θ is assigned to school c given by $g_{\Xi}(c|\theta)$. This is stated formally in Lemma 18.

Lemma 18. *For any coarsening Ξ the probability that any student of type $\theta \in \Theta$ is assigned to a school $c \in \mathcal{C}$ is well-defined and can be represented by a conditional probability mass function $g_{\Xi}(c|\theta)$, which is given by:*

$$g_{\Xi}(c|\theta) = \int_0^1 \mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_{\Xi}(\theta)}(\tau) = c\right] d\tau \tag{852}$$

where $\tilde{\theta}_{\Xi}(\theta)$ is the induced ordinal type of type $\theta \in \Theta$ under the coarsening Ξ, and $\tilde{\mu}_{\tilde{\theta}_{\Xi}}$ is the constructed mapping from Lemma 17 for the coarsened ordinal economy with tie-breakers $\tilde{\Omega}^{\tau}_{\Xi}$ corresponding to Ξ.

Proof. As a result of Lemma 17, in the coarsened ordinal economy with tie-breakers $\tilde{\Omega}^{\tau}_{\Xi}$ corresponding to Ξ, for any ordinal type $\tilde{\theta}_{\Xi} \in \tilde{\Theta}_{\Xi}$ we know that their assigned school is uniquely determined by their tie-breaker τ according to $\tilde{\mu}_{\tilde{\theta}_{\Xi}}(\tau)$. The probability that type $\tilde{\theta}_{\Xi}$ is matched to a school c is then:

$$\mu_{\Xi}(c|\tilde{\theta}_{\Xi}) = \int_0^1 \mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_{\Xi}}(\tau) = c\right] d\tau \tag{853}$$

Moreover, given a type $\theta = (u^{\theta}, s^{\theta})$, we can deduce their induced type under a coarsening Ξ, or $\tilde{\theta}_{\Xi}(\theta) = (\succ^{\tilde{\theta}_{\Xi}(\theta)}, s^{\tilde{\theta}_{\Xi}(\theta)})$ by taking the ordinal representation of their cardinal utility and coarsening their priority according to Ξ. Thus, for a type θ, the probability

of equivalence classes that the school has be k and enumerate them in increasing priority, so the i'th class has measure $m(i)$. Divide the interval $[0,1]$ to k ordered sub intervals (where the intervals are $[a_0, a_1), [a_1, a_2), \ldots, [a_{k-1}, a_k]$ with $a_0 = 0$ and $a_k = 1$) with each having measure $m(i)$ and score distribution $U[a_i, a_{i+1}]$ (in other words, with density $m(i)/(a_i - a_{i-1})$). In the unique economy after tie-breaking, any student who is in priority class i has a uniform probability of having a priority of any number in the i'th sub interval. So long as the initial economy has full support, the economy after coarsening and tie-breakers also has full support.

that they are assigned to school c is simply $\mu_\Xi(c|\tilde{\theta}_\Xi(\theta))$. Consequently, we can define:

$$g_\Xi(c|\theta) = \mu_\Xi(c|\tilde{\theta}_\Xi(\theta)) \tag{854}$$

or, as claimed in the Lemma:

$$g_\Xi(c|\theta) = \int_0^1 \mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_\Xi(\theta)}(\tau) = c\right] d\tau \tag{855}$$

If the coarsening map is the identity function, then there is no need for tie breaking and the economy has full support, so $\tilde{\mu}(\tilde{\theta}, \tau) = c$ is still unique. As a result, an analogous argument yields the result. □

This construction simply takes the function $\tilde{\mu}$ for any student from Lemma 17 and averages over the distribution of tie-breakers to compute the assignment probabilities.

D.2.2 Continuity of Allocations

To establish the existence of an optimum, we show that allocations are continuous in the cutoff vectors corresponding to coarsenings. We first use Theorem 5 to represent coarsenings as vectors $v \in \mathcal{V}$. We refer to $g_v : \Theta \times \mathcal{C} \to [0, 1]$ as the corresponding allocation. As highlighted in Footnote 25, the set of potential allocations $\mathcal{G}$ is a subset of the space $L^1(\Theta \times \mathcal{C})$ of functions measurable in the measure space $(\Theta \times \mathcal{C}, \Sigma, \check{F})$, where Σ is is the Borel σ-algebra generated by the product topology on $\Theta \times \mathcal{C}$ and $\check{F}$ extends the domain of the measure F from Θ to $\Theta \times \mathcal{C}$ by stacking $|\mathcal{C}|$ copies of F. We now prove that g_v is continuous in v in the sense that for any sequence $\{v_n\}_{n\in\mathbb{N}} \subset \mathcal{V}$ such that $v_n \to v$, we have that $g_{v_n} \to g_v$ in the L^1-norm.

Lemma 19. *g_v is continuous in v in the L^1-norm.*

Proof. As shown in Lemma 18, we may represent:

$$g_v(c|\theta) = \int_0^1 \mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_v(\theta)}(\tau) = c\right] d\tau \tag{856}$$

Fix any sequence $\{v_n\}_{n\in\mathbb{N}} \subset \mathcal{V}$ such that $v_n \to v$. If $g_{v_n}(c|\theta) \to g_v(c|\theta)$ for any such sequence for $\check{F}$-almost all (c, θ), then g_v is continuous in v in the L^1-norm. By the dominated convergence theorem, this is equivalent to showing the following for

$\check{F}$-almost all (c, θ) (as $L^1(\Theta \times \mathcal{C})$ is complete):

$$0 = \lim_{n\to\infty} \left[g_{(v_n)}(c|\theta) - g_v(c|\theta) \right] = \lim_{n\to\infty} \int_0^1 \left[\mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_{v_n}(\theta)}(\tau) = c \right] - \mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_v(\theta)}(\tau) = c \right] \right] d\tau \tag{857}$$

Now fix (c, θ), to show the above, once again by the dominated convergence theorem, it is sufficient to show that:

$$\lim_{n\to\infty} \left[\mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_{v_n}(\theta)}(\tau) = c \right] - \mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_v(\theta)}(\tau) = c \right] \right] = 0, \text{ for almost all } \tau \in [0,1] \tag{858}$$

Now fix a $\tau \in [0,1]$ and consider the sequence of economies with the priority structures induced by v_n. As $v_n \to v$, the priority structures converge pointwise. By Lemma 17, the full support assumption in the ex-ante economy implies that interim economy has full support for all n, hence the interim economy has a unique stable matching for all v_n. Thus, by Theorem 2.3 of Azevedo and Leshno (2016), we know that the stable matching is continuous in the induced economy with coarsening vector $v \in \mathcal{V}$ within the set of economies $\mathcal{O}$. Thus, for each $\tilde{\theta}_v(\theta)$ and almost all τ, there must exist an N such that for all $n > N$, we have that $\mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_{v_n}(\theta)}(\tau) = c \right] = \mathbb{I}\left[\tilde{\mu}_{\tilde{\theta}_v(\theta)}(\tau) = c \right]$ for almost all $\theta \in \Theta$ and $\tau \in [0,1]$. Thus, we have shown that Equation 857 holds for $\check{F}$-almost all (c, θ) and completed the proof. □

D.3 Priority Design with Aggregate Uncertainty

In the main text we assumed that the mechanism designer knows the distribution of students' types. We start with a simple discrete example that shows this is necessary for the optimality of trinary coarsenings.

Example 7. *There are six students,* $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6\}$ *and one school with capacity* 2, $\mathcal{C} = \{c\}$ *and* $Q_c = 2$. *The scores of the students are given by their index number, i.e.* $s_c^{\theta_i} = i$. *The designer prefers to admit* θ_6, θ_4 *and to a lesser extent,* θ_2 *and has the following utility function:*

$$Z(\mu) = \sum_{i \in \{4,6\}} \mathbb{I}\{\theta_i \in \mu(c)\} + \frac{1}{2}\mathbb{I}\{\theta_2 \in \mu(c)\} \tag{859}$$

There are two states denoted by γ *and* γ' *and both states have strictly positive probability. The ordinal preferences of student* θ_i *in state* $j \in \{\gamma, \gamma'\}$ *is denoted by* $\succ_j^i$. *The preferences are given by:*

- $\succ_j^i = c, \theta_i$ *for* $i \in \{1, 2, 3, 6\}$ *and* $j = \{\gamma, \gamma'\}$.
- $\succ_\gamma^i = c, \theta_i$ *and* $\succ_{\gamma'}^i = \theta_i, c$ *for* $i \in \{4, 5\}$.

Under state γ, *the unique optimal trinary coarsening is given by:*

$$\Xi_\gamma(i) = \begin{cases} 3, i = 6 \\ 2, i \in \{4, 5\} \\ 1, i \in \{1, 2, 3\} \end{cases} \tag{860}$$

This coarsening admits 6 *with probability 1 and maximizes the probability of admission for* 4. *Note that any other trinary coarsening will result in a strictly lower utility for the designer. Under state* γ', Ξ_γ *is not optimal as it admits* 6 *with probability 1 but admits* 2 *with probability* $1/3$. *As* Ξ_γ *is the unique trinary coarsening that is optimal under* γ, *no trinary coarsening can attain the optimum under both realizations of uncertainty. Moreover, the following is an optimal trinary coarsening under* γ':

$$\Xi_{\gamma'}(i) = \begin{cases} 3, i = 6 \\ 2, i \in \{2, 3, 4, 5\} \\ 1, i \in \{1\} \end{cases} \tag{861}$$

$\Xi\gamma'$ admits 6 with probability 1 and admits 2 with probability 1/2, and thus performs strictly better than Ξ_γ. However, there exists a coarsening with 4 indifference classes that attains the optimum under both states:

$$\Xi^*(i) = \begin{cases} 4, i = 6 \\ 3, i \in \{4,5\} \\ 2, i \in \{2,3\} \\ 1, i \in \{1\} \end{cases} \tag{862}$$

Ξ^* replicates Ξ_γ under γ and $\Xi_{\gamma'}$ under γ', and thus attains the optimal allocation in both cases. This example shows that trinary coarsenings are not without loss of optimality under uncertainty.

We now depart from that assumption by considering a mechanism designer who believes finitely many distributions are possible and has a prior belief over those probability distributions. Formally, let $\mathcal{F}$ denote the set of finitely many possible distributions with a particular element f. Let $p(f)$ denote the probability measure over this finite set and g_f denote the allocation under f. In this environment, we introduce a new assumption (in place of Assumption 9) on the objective function of the planner that they are a subjective expected utility maximizer with respect to their prior as to the market they face:

Assumption 17. *Under uncertainty, the social planner has subjective V.N-M utility Z with a Bernoulli utility function $z : \mathcal{G} \to \mathbb{R}$ that is continuous in g_f.*

Notice that in any realization of the student distribution, the same coarsening rule must be applied. Hence schools must use the same priority structure across realizations of different market structures. This draws a parallel between uncertainty and our analysis of homogeneous coarsening in Appendix D.4: under homogeneity multiple schools must use the same coarsening rule; under uncertainty, the same rule must be used across multiple states.

First, and analogously to the previous analysis, we provide a result on the maximum number of partitions needed to achieve an optimal coarsening under uncertainty. This is stated as Proposition 45:

Proposition 45. *There exists an optimal coarsening Ξ with at most $2|\mathcal{F}|$ cutoffs at every school.*

Proof. Fix an arbitrary coarsening Ξ. Divide the schools to two subsets: $c \in \tilde{\mathcal{C}}$ if Ξ_c is finite and $c \in \hat{\mathcal{C}}$ if Ξ_c is the identity function. Take $c \in \tilde{\mathcal{C}}$, the coarsening Ξ takes any student with score $s_c^{\tilde{\theta}}$ to an equivalence class, hence $\Xi_c(s_c^{\tilde{\theta}}) \in \{P_1^c, P_2^c, \ldots, P_n^c\}$ with $P_1^c < P_2^c, \ldots, < P_n^c$. By Theorem 5, for any $f_i \in \mathcal{F}$, there can be at most one lottery region for school c and denote that region by $P_{f_i}^c$ and enumerate f_i in a way that $P_{f_i}^c$ is increasing in i. Let $P_n < P_{n+1} < \ldots < P_m$ denote the indifference classes that are between two consecutive lottery regions $P_{f_i}^c$ and $P_{f_{i+1}}^c$. Define Ξ' by merging $P_n < P_{n+1} < \ldots < P_m$.[6] Note that the allocation under Ξ and Ξ' are the same under any f_i.[7] Repeating this for all schools $c \in \tilde{\mathcal{C}}$ and lottery regions results in a coarsening rule Ξ' that induces same allocation as the arbitrary coarsening Ξ. Moreover, there is only 1 partition between any two lottery regions for a schools in $\tilde{\mathcal{C}}$, hence there are at most $2|\mathcal{F}|$ cut-offs and $2|\mathcal{F}|+1$ partitions per schools in $\tilde{\mathcal{C}}$ under Ξ'.

Next, take $c \in \hat{\mathcal{C}}$, so Ξ_c is the identity function. As the economy has full support, there is a unique stable matching under any $f_i \in \mathcal{F}$. Let μ_i denote this stable matching. For any school c and distribution f_i, let t_c^i denote the threshold that the students must clear in order to gain admission to that school under f_i. Formally,

$$t_c^i = \min\{s_c^{\tilde{\theta}} : \mu_i(\tilde{\theta}) = c\} \tag{863}$$

Let $T_c = \cup_{i \le |\mathcal{F}|}\{t_c^i\}$. Next, define Ξ' in the following way:

$$\Xi'_c(s_c^{\tilde{\theta}}) = \max\{t_c^i \in T_c : s_c^{\tilde{\theta}} \ge t_c^i\} \tag{864}$$

Note that from the stability of μ_i under the distribution f_i and the definition of matching (in particular, property (iv) from Footnote 23)) there cannot be a student k such that $c' \succ_k \mu_i(k)$ and $s_{c'}^k \ge t_{c'}^i$. To see that μ_i is stable under Ξ', assume that student pair l, j (with scores $s_{c'}^l$ and $s_{c'}^j$) blocks it in school c'. Then we have $c' \succ_l \mu_i(l)$, $\mu_i(j) = c'$ and $\Xi'_{c'}(l) \ge \Xi'_{c'}(j)$. As $\mu_i(j) = c'$, we have $\Xi'_{c'}(j) \ge t_c^i$. As $\Xi'_{c'}(l) \ge \Xi'_{c'}(j)$ and $\Xi'_{c'}(j) \ge t_c^i$, we have that $s_{c'}^l \ge t_{c'}^i$, which contradicts the stability of μ_i under Ξ as $c' \succ_l \mu_i(l)$. By the same argument as in the first case, it follows that $g_\Xi = g_{\Xi'}$. Then non-emptiness follows from the same argument in Theorem 6 under Assumption 17. $\square$

This result generalizes Theorems 5 and 6 to the case with aggregate uncertainty. Instead of an optimum being attainable with at most 2 cutoffs at each school, now up to $2|\mathcal{F}|$ cutoffs are required: 2 for each state of the world. This demonstrates

[6] This can be done by setting $\Xi'(P_n) = \ldots = \Xi'(P_m) = \Xi(P_n)$.

[7] The argument for this is exactly same as the one made in the proof of Theorem 5.

how the presence of aggregate uncertainty can substantially complicate the problem of priority design.

To understand the normative implications of aggregate uncertainty, we characterize when the presence of uncertainty induces a welfare loss to the planner. We denote the lottery region cut-offs of school c under distribution f by $[\underline{P^c_f}, \overline{P^c_f}]$. We can define a no-overlap condition in the case with uncertainty:

Definition 19. *For any problem, the set of distributions $\{f_1, f_2, \ldots, f_n\} = \mathcal{F}$ satisfies no overlap across distributions if for any f_i, there exists coarsenings:*

$$\Xi_i = [\underline{P^1_i}, \ldots, \underline{P^{|\mathcal{C}|}_i} | \overline{P^1_i}, \ldots, \overline{P^{|\mathcal{C}|}_i}] \tag{865}$$

such that Ξ_i is optimal under f_i and for all c, k and j, $[\underline{P^c_k}, \overline{P^c_k}] \cap [\underline{P^c_j}, \overline{P^c_j}] = \emptyset$ or $[\underline{P^c_k}, \overline{P^c_k}] = [\underline{P^c_j}, \overline{P^c_j}]$.

We can now prove that there is a loss from uncertainty if and only if the no-overlap condition is not satisfied. This is stated as Proposition 46:

Proposition 46. *There is no welfare loss from uncertainty if and only if $\mathcal{F}$ satisfies no overlap across distributions.*

Proof. Assume that the problem satisfies no overlap across distributions and fix optimal trinary coarsenings Ξ_{f_i} that satisfy no overlap across distributions. Let $|\mathcal{F}| = n$. For any school c, Ξ_{f_i} has two cut-offs v_{f_i}: denote them by $L^c_{f_i}$ and $U^c_{f_i}$ with $L^c_{f_i} < U^c_{f_i}$. As the trinary coarsening satisfies no overlap across distributions, we can enumerate them as $L^c_{f_1} < U^c_{f_1} < \ldots < L^c_{f_n} < U^c_{f_n}$. Define a new set of cut-offs $w^c = L^c_{f_1}, U^c_{f_1}, \ldots, L^c_{f_n}, U^c_{f_n}$. Define a new coarsening v' by setting the cut-offs as w^c for all c. Notice that for any realization of f_i, the coarsenings v_{f_i} and v' induce the same allocation.[8] Hence the coarsening v' induces the same allocation as v_{f_i} for any f_i. As each v_{f_i} maximizes the mechanism designer's utility under f_i, the maximum utility can be replicated by using v' for any f_i and there is no loss from uncertainty.

For the converse result, fix an arbitrary coarsening w and suppose that the problem does not satisfy the no overlap across distributions condition. Consider the set of coarsenings that uses $v_f = w$ for all $f \in \mathcal{F}$. Notice that this set trivially has no overlap. Clearly, w and $\{v_f\}$ attain the same utility value for the mechanism designer. As the problem does not satisfy the no overlap across distributions assumption, there

[8] To see why this is true, notice that v_{f_i} can be obtained from v' by merging indifference classes that are above and below each lottery class for any school, using the construction in the proof of Theorem 5.

must be a set of coarsenings $\{v'_f\}$ that attains a strictly higher expected utility than $\{v_f\}$ and w. As w is arbitrary, there is welfare loss from uncertainty regardless of its choice. $\square$

As no overlap is a very strenuous condition, the above result clarifies that generally uncertainty will lead to welfare losses for the planner. It moreover makes clear that this loss in welfare stems from the inability of the planner to preserve the lottery classes from the optimal *ex post* designs in each state.

D.4 Priority Design with Homogeneous Coarsening

In many applications, the assumption that the policymaker can set a school-specific coarsening may be unrealistic. For example, in cases where the underlying score represents exam scores or some measure of student achievement, it is possible that the schools are constrained to use the same grading criteria. Hence, in this section, we examine optimal coarsenings when the coarsening rule is constrained to be the same across all schools. Formally, we find the optimal coarsening in the set $\mathcal{H}$ of homogeneous coarsenings:

Definition 20. *A homogeneous coarsening rule* Ξ *is a coarsening rule such that*

$$\Xi(s) = (\Xi_1(s_1), \ldots, \Xi_1(s_n)) \tag{866}$$

Let the set of all homogeneous coarsening maps be $\mathcal{B}_h$. When only homogeneous coarsenings are allowed, the planner faces the general problem that satisfies Assumption 9:

$$\mathcal{B}_h^* = \arg\max_{\Xi \in \mathcal{B}_h} Z(g_\Xi) \tag{867}$$

where g_Ξ is the allocation induced by Ξ. We first prove a proposition which shows that appropriately revised versions of Theorems 5 and 6 continue to hold in this environment:

Proposition 47. $\mathcal{B}_h^*$ *is non-empty. There exists a homogeneous coarsening with* $2|\mathcal{C}|$ *cutoffs such that* $\Xi \in \mathcal{B}_h^*$.

Proof. Fix an arbitrary finite homogeneous coarsening Ξ. The coarsening Ξ takes any student with score $s_c^{\tilde{\theta}}$ to a partition, hence $\Xi_c(s_c^{\tilde{\theta}}) \in \{P_1^c, P_2^c, \ldots, P_n^c\}$ with $P_1^c < P_2^c < \ldots < P_n^c$ for all $\tilde{\theta} \in \tilde{\Theta}$. By Theorem 5, for any $c \in \tilde{\mathcal{C}}$, there can be at most one lottery class.

Without loss of generality, assume the lottery regions are enumerated as $P_{x(1)} < P_{x(2)} < \ldots < P_{x(|\mathcal{C}|)}$. Let $P_n < P_{n+1} < \ldots < P_m$ denote the indifference classes that are between two consecutive lottery regions $P_{x(z)}$ and $P_{x(z+1)}$. Define Ξ' by merging $P_n < P_{n+1} < \ldots < P_m$.[9] Note that $g_\Xi(c|\theta) = g_\Xi(c|\theta)$ for all $\theta \in \Theta$.[10] This process can be repeated for all consecutive lottery regions without changing the allocation. This results in a coarsening rule Ξ' where there is only 1 partition between any two lottery

[9] This can be done by setting $\Xi_c'(P_n) = \ldots = \Xi_c'(P_m) = \Xi_c(P_n)$.

[10] The argument for this is exactly same as the one made in the proof of Theorem 5.

regions for all $c \in \mathcal{C}$. Hence the total number of partitions are at most $2|\mathcal{C}|+1$ and cut-offs are $2|\mathcal{C}|$ for all $c \in \mathcal{C}$.

Next, assume the homogeneous coarsening is the identity function, in other words, there is no coarsening.[11] As the economy has full support, there is a unique stable matching. Let μ denote this stable matching. For any school c and distribution, let t_c denote the threshold that the students must clear in order to gain admission to that school. Formally,

$$t_c = \min\{s_c^{\tilde{\theta}} : \mu(\tilde{\theta}) = c\} \tag{868}$$

Let $T = \cup_{c \in \mathcal{C}} t_c$. Next, define Ξ' in the following way:

$$\Xi'_c(s_c^{\tilde{\theta}}) = \max\{t \in T : s_c^{\tilde{\theta}} \geq t\} \tag{869}$$

Note that from the stability of μ and the definition of matching (in particular, property (iv) from Footnote 23)) there cannot be a student k such that $c' \succ_k \mu_i(k)$ and $s_{c'}^k \geq t_{c'}$. To see that μ is stable under Ξ', assume that student pair l, j (with scores $s_{c'}^l$ and $s_{c'}^j$) blocks it in school c'. Then we have $c' \succ_l \mu(l)$, $\mu(j) = c'$ and $\Xi'_{c'}(l) \geq \Xi'_{c'}(j)$. As $\mu(j) = c'$, we have $\Xi'_{c'}(j) \geq t_c^i$. As $\Xi'_{c'}(l) \geq \Xi'_{c'}(j)$ and $\Xi'_{c'}(j) \geq t_c^i$, we have that $s_{c'}^l \geq t_{c'}^i$, which contradicts the stability of μ under Ξ as $c' \succ_l \mu(l)$. By the same argument as in the first case, it follows that $g_{\Xi} = g_{\Xi'}$. This shows that a homogeneous coarsening Ξ' with $|\mathcal{C}|$ cut-offs attains the same allocation as no coarsening.

Then any homogeneous coarsening Ξ can be represented by a vector of cut-offs $v = P_1, P_2, \ldots, P_{|2\mathcal{C}|}$. The existence of an optimum then follows from the same argument as in Theorem 6. $\square$

Like Theorems 5 and 6, Proposition 47 reduces a possibly infinite dimensional problem to a $2|\mathcal{C}|$-dimensional problem and allows us to compute optimal homogeneous coarsenings.

We next study the welfare implications of imposing homogeneity in coarsenings. We show that there is a loss from imposing homogeneity whenever in the optimal trinary coarsening the lottery regions of two schools intersect without being equal. To this end, we define a subclass of problems:

[11] Note that the case where some schools have a finite coarsening while others have no coarsening is ruled out by assuming the coarsening must be homogeneous.

Definition 21. *A problem satisfies no overlap across schools if there exists an optimal coarsening such that for all lottery regions P^c and $P^{c'}$ either $[\overline{P^c}, \underline{P^c}] \cap [\overline{P^{c'}}, \underline{P^{c'}}] = \emptyset$ or $[\overline{P^c}, \underline{P^c}] = [\overline{P^{c'}}, \underline{P^{c'}}]$.*

We can now characterize exactly the case when imposing homogeneity results in a welfare loss. This is stated formally in Proposition 48:

Proposition 48. *There is no loss from imposing homogeneity if and only if the problem satisfies no overlap across schools.*

Proof. In this environment we can define an object similar to the vector we defined earlier to represent the coarsenings as a vector when there are n_i cut-offs for each school i:

$$\begin{aligned} w = [P^1_1, P^1_2, \ldots, P^1_{n_1} | P^2_1, \ldots, P^2_{n_2} | \ldots | P^{|\mathcal{C}|}_1, \ldots, P^{|\mathcal{C}|}_{n_{|\mathcal{C}|}}] \\ \text{s.t.} \quad w \in \mathcal{W} = \{w \in [0,1]^{\sum_i n_i}, P^c_j > P^c_l, \forall c, j > l\} \end{aligned} \tag{870}$$

Let v be an optimal (i.e. non homogeneous) trinary coarsening that satisfies no overlap across schools. As above, v is of the form $v = \{v^1_1, v^1_2 | \ldots | v^{|\mathcal{C}|}_2, v^{|\mathcal{C}|}_2\}$ as v is a trinary coarsening. Enumerate the $2|\mathcal{C}|$ elements of v in increasing order and let z be the vector obtained by doing this, hence $z = \{z_1, \ldots, z_{2|\mathcal{C}|}\}$ and it is increasing. Therefore it is a well-defined set of coarsening cut-offs. Now, form the alternative coarsening v' by using the cut-offs in z for all schools:

$$v' = [z_1, \ldots, z_{2|\mathcal{C}|} | \ldots | z_1, \ldots, z_{2|\mathcal{C}|}] \tag{871}$$

v' is a homogeneous coarsening as it uses the same cut-offs. Furthermore, v and v' induce the same allocation.[12] Hence there is a homogeneous coarsening v' that induces the same allocation as the optimal trinary coarsening v and there is no loss from imposing homogeneity.

For the converse result, fix an arbitrary coarsening w and suppose that the problem does not satisfy the no overlap across schools condition. Consider the set of coarsenings that uses $v_c = w$ for all $c \in \mathcal{C}$. Notice that this set trivially has no overlap. Clearly, w and $\{v_c\}$ attain the same utility value for the mechanism designer. As the problem does not satisfy the no overlap across schools assumption, there must be a set of coarsenings $\{v'_c\}$ that attains a strictly higher utility than $\{v_c\}$ and w. As w is arbitrary, there is welfare loss from homogeneity regardless of its choice. □

[12] To see why this is true, notice that v can be obtained from v' by merging indifference classes that are above and below each lottery class for any school, using the construction in the proof of Theorem 5.

As no-overlap is an extremely strenuous condition, Proposition 48 would seem to indicate that in practical contexts, there will very likely be a welfare loss from imposing homogeneity. As a result, rationalizing the observed homogeneity of many systems is challenging in this environment. Indeed, it may be the case that one has to appeal to concerns for simplicity in order to understand observed homogeneity.

D.5 Solving the Dynamic Model of Housing Assignment

In this Appendix, we provide the prerequisites for determining the structure of the optimal priority design by first solving the dynamic matching model. Relative to the static matching models we have so far considered, this features two complications. First, we have to determine the agents' optimal stopping rule for when to accept a given house as a function of their individual characteristics and the allocation. Second, we have to find the steady-state distribution of agents who are unmatched (which is endogenous to the stopping rule) to find the steady-state allocation.

We first fix an allocation and solve for the agent's optimal stopping policy. Denoting by $P(s, v)$ the steady-state probability that an agent with income s is matched to a house in any given period given a priority design v, we have that the value function of an unmatched agent is given by:

$$\begin{aligned} V(\kappa, P(s, v)) = \kappa + \beta\delta P(s, v)\mathbb{E}_{\tilde{v}}\left[\max\left\{\frac{\tilde{v}}{1-\beta\delta}, V(\kappa, P(s, v))\right\}\right] \\ + \beta\delta(1 - P(s, v))V(\kappa, P(s, v)) \end{aligned} \tag{872}$$

where the first term is simply the agents value of going unmatched this period: their outside option. The second term is the discounted expected value of being matched to a house, weighted by the chance that that the agent is matched. The final term simply says that if the agent goes unmatched, they receive their discounted value of being unmatched.

The agent's optimal stopping rule is a threshold strategy indexed by $\bar{v}(\kappa, P(s, v))$ such that agents accept any house with $v \geq \bar{v}(\kappa, P(s, v))$ and reject otherwise. It follows that this reservation value is the unique solution to the following equation:

$$\begin{aligned} \bar{v}(\kappa, P(s, v)) &= (1 - \beta\delta)\kappa \\ &\quad + \beta\delta P(s, v)\left[\bar{v}(\kappa, P(s, v))\Lambda(\bar{v}(\kappa, P(s, v))) + \int_{\bar{v}(\kappa, P(s,v))}^{v_{\max}} \tilde{v} d\Lambda(\tilde{v})\right] \\ &\quad + \beta\delta(1 - P(s, v))\bar{v}(\kappa, P(s, v)) \end{aligned} \tag{873}$$

In general this cannot be solved in closed form.[13] However, it is simple to see that $\bar{v}(\kappa, P(s, v))$ features the following comparative statics: it is increasing in the agent's

[13] One exception to this is the case where $\tilde{v} \sim U[v_{\min}, v_{\max}]$. In this case, Equation 873 becomes quadratic in $\bar{v}(\kappa, P(s, v))$ and can be solved analytically.

outside option κ; it is increasing in the agent's patience β; and it is increasing in the probability that an agent is matched to a house in any given period $P(s, v)$.

Having solved for the agent's optimal stopping rule as a function of the assignment probability, it remains to solve for the assignment probability as a function of a given stopping rule. To do this, one first needs to find the steady-state distribution of unmatched agents and the steady state measure of occupied housing. To this end, define the steady-state measure of unmatched agents as the joint density $k(\kappa, s; v)$ with marginal cumulative measure over incomes given by $K_s(s; v)$ and let $M(v)$ be the steady-state measure of matched agents.

Using our more general analysis, we can once again greatly simplify the problem: Theorem 5 directly applies to any steady state of this model. We can therefore restrict to priority designs $v = (v_1, v_2)$ that feature three tiers. Given a CDF of incomes in the population in the steady state, the allocation probability for an agent with income s is given by:

$$P(s, v) = \begin{cases} 1, s \geq v_2, \\ \frac{(Q-\delta M(v))-(1-K_s(v_2;v))}{K_s(v_2;v)-K_s(v_1;v)}, s \in [v_1, v_2), \\ 0, s < v_1 \end{cases} \tag{874}$$

Moreover, the stationary joint distribution at the end of a period $k^l(\kappa, s; v)$ solves the following fixed point equation which balances inflows of new agents and outflows of agents due to accepting a house and death.

$$\begin{aligned} (1-\delta)f(\kappa, s) &= (1-\delta)k^l(\kappa, s; v) \\ &\quad + (1 - \Lambda(\bar{v}(\kappa, P(s, v))))P(s, v)\left(\delta k^l(\kappa, s; v) + (1-\delta)f(\kappa, s)\right) \end{aligned} \tag{875}$$

Thus, the stationary joint distribution of agents at the time of matching is given by:

$$k(\kappa, s; v) = \delta k^l(\kappa, s; v) + (1-\delta)f(\kappa, s) \tag{876}$$

Finally, the stationary measure of matched agents is given by:

$$(1-\delta)M(v) = (1 - \Lambda(\bar{v}(\kappa, P(s, v))))P(s, v)\left(\delta k^l(\kappa, s; v) + (1-\delta)f(\kappa, s)\right) \tag{877}$$

Having now solved the stopping rule as a function of the allocation and the allocation as a function of the stopping rule, we have characterized the matching of agents to houses in this economy. From the steady-state policy function it is moreover simple to compute welfare from the relationship:

$$V(\kappa, P(s,v)) = \frac{\bar{v}(\kappa, P(s,v))}{1-\beta\delta} \tag{878}$$

This discussion is formalized in Proposition 49:

Proposition 49. *For any coarsening policy v', there exists a $v = (v_1, v_2)$ that is outcome equivalent. Moreover, under this v, welfare is given by:*

$$Z(v) = \int_{\mathcal{U}} \int_0^1 \left(\frac{\frac{\bar{v}(\kappa, P(s,v))}{1-\beta\delta}}{1-\gamma} \right)^{1-\gamma} dF(\kappa, s) \tag{879}$$

where the reservation value policy solves Equation 873, the allocation is given by Equation 874, the stationary measure of unmatched agents of each type is given by Equations 875 and 876 and the stationary measure of matched agents is given by Equation 877.

Proof. In a steady state of the model, there is a time-invariant assignment probability as a function of the policy v and one's score s, $P(s,v)$. Given this, the agent has Bellman equation given by:

$$\begin{aligned} V(\kappa, P(s,v)) = \kappa + \beta\delta P(s,v)\mathbb{E}_{\tilde{v}}\left[\max\left\{\frac{\tilde{v}}{1-\beta\delta}, V(\kappa, P(s,v))\right\}\right] \\ + \beta\delta(1-P(s,v))V(\kappa, P(s,v)) \end{aligned} \tag{880}$$

The first claim is that the optimal stopping solution, where 1 corresponds to accepting a house and 0 to rejecting is given by:

$$x^*(\kappa, \tilde{v}, P(s,v)) = \begin{cases} 1, \tilde{v} \geq \bar{v}(\kappa, P(s,v)), \\ 0, \tilde{v} < \bar{v}(\kappa, P(s,v)), \end{cases} \tag{881}$$

for some function $\bar{v}(\kappa, P(s,v))$. This is immediate: suppose an agent accepts $\tilde{v}$ and rejects $\tilde{v}' > \tilde{v}$. As $\tilde{v}' > \tilde{v}$, either accepting $\tilde{v}$ or rejecting $\tilde{v}'$ must be suboptimal. Moreover, we know that at this reservation value $\bar{v}(\kappa, P(s,v))$, the agent must be indifferent between accepting and rejecting. It follows that:

$$V(\kappa, P(s,v)) = \frac{\bar{v}(\kappa, P(s,v))}{1-\beta\delta} \tag{882}$$

Combining Equations 880 and 882, we obtain:

$$\begin{aligned}\frac{\bar{v}(\kappa, P(s,v))}{1-\beta\delta} &= \kappa + \beta\delta P(s,v)\mathbb{E}_{\tilde{v}}\left[\max\left\{\frac{\tilde{v}}{1-\beta\delta}, \frac{\bar{v}(\kappa, P(s,v))}{1-\beta\delta}\right\}\right] \\ &\quad + \beta\delta(1-P(s,v))\frac{\bar{v}(\kappa, P(s,v))}{1-\beta\delta}\end{aligned} \tag{883}$$

Re-arranging this equation yields:

$$\begin{aligned}\bar{v}(\kappa, P(s,v)) &= (1-\beta\delta)\kappa + \beta\delta P(s,v)\mathbb{E}_{\tilde{v}}\left[\max\left\{\tilde{v}, \bar{v}(\kappa, P(s,v))\right\}\right] \\ &\quad + \beta\delta(1-P(s,v))\bar{v}(\kappa, P(s,v))\end{aligned} \tag{884}$$

where:

$$\mathbb{E}_{\tilde{v}}\left[\max\left\{\tilde{v}, \bar{v}(\kappa, P(s,v))\right\}\right] = \bar{v}(\kappa, P(s,v))\Lambda(\bar{v}(\kappa, P(s,v))) + \int_{\bar{v}(\kappa, P(s,v))}^{v_{\max}} \tilde{v}\, d\Lambda(\tilde{v}) \tag{885}$$

yielding the claimed equation for the reservation value function:

$$\begin{aligned}\bar{v}(\kappa, P(s,v)) = (1-\beta\delta)\kappa + \beta\delta P(s,v)&\left[\bar{v}(\kappa, P(s,v))\Lambda(\bar{v}(\kappa, P(s,v))) + \int_{\bar{v}(\kappa, P(s,v))}^{v_{\max}} \tilde{v}\, d\Lambda(\tilde{v})\right] \\ + \beta\delta(1-P(s,v))\bar{v}(\kappa, P(s,v))&\end{aligned} \tag{886}$$

As was claimed in the text, this has a unique solution. To see this simply observe that the slope of the RHS of Equation 886 is given by (suppressing all constants):

$$RHS'(\bar{v}) = \beta\delta\left[1 - P(1-\Lambda(\bar{v}))\right] < 1 \tag{887}$$

while the slope of the LHS is 1, and $RHS(0) > LHS(0) = 0$. We additionally claimed in the text that the case with $\tilde{v} \sim U[v_{\min}, v_{\max}]$ has a closed form solution for Equation 886. To this end, see that we have:

$$\begin{aligned}\bar{v}(\kappa, P(s,v)) = (1-\beta\delta)\kappa + \beta\delta P(s,v)&\left[\bar{v}(\kappa, P(s,v))\frac{\bar{v}(\kappa, P(s,v)) - v_{\min}}{v_{\max} - v_{\min}} + \frac{v_{\max}^2 - \bar{v}(\kappa, P(s,v))^2}{2(v_{\max} - v_{\min})}\right] \\ + \beta\delta(1-P(s,v))\bar{v}(\kappa, P(s,v))&\end{aligned} \tag{888}$$

which can be rewritten as a quadratic in $\bar{v}(\kappa, P(s,v))$:

$$0 = a\bar{v}(\kappa, P(s,v))^2 + b\bar{v}(\kappa, P(s,v)) + c \tag{889}$$

where:

$$\begin{aligned}
a &= \frac{\beta\delta P(s,v)}{2(v_{\max}-v_{\min})} \\
b &= -1+\beta\delta(1-P(s,v))-\beta\delta P(s,v)\frac{v_{\min}}{v_{\max}-v_{\min}} \\
c &= (1-\beta\delta)\kappa+\beta\delta P(s,v)\frac{v_{\max}^2}{2(v_{\max}-v_{\min})}
\end{aligned} \tag{890}$$

yielding the standard solution of a quadratic for $\bar{v}(\kappa, P(s,v))$.

As a function of the steady-state assignment probability, we have solved the agents' optimal stopping problems. We now solve for the steady-state assignment probability as a function of the agents' optimal stopping problems. To this end, observe that the joint density of agents at the end of a period must match the inflow of agents to the outflow of agents:

$$\begin{aligned}
\underbrace{(1-\delta)f(\kappa,s)}_{\text{New Agents}} = \underbrace{(1-\delta)k^l(\kappa,s;v)}_{\text{Dead Agents}} \\
+ \underbrace{(1-\Lambda(\bar{v}(\kappa,P(s,v))))P(s,v)}_{\text{Agents Who Stop}}\underbrace{\left(\delta k^l(\kappa,s;v)+(1-\delta)f(\kappa,s)\right)}_{\text{Agents Available to Match This Period}}
\end{aligned} \tag{891}$$

While the joint density of agents at the time of matching is given by:

$$k(\kappa,s;v) = \delta k^l(\kappa,s;v)+(1-\delta)f(\kappa,s) \tag{892}$$

Moreover, the steady state measure of occupied housing balances outflows from death and inflows from new matches:

$$\underbrace{(1-\delta)M(v)}_{\text{Death of Old Matches}} = \underbrace{(1-\Lambda(\bar{v}(\kappa,P(s,v))))P(s,v)\left(\delta k^l(\kappa,s;v)+(1-\delta)f(\kappa,s)\right)}_{\text{New Matches}} \tag{893}$$

Given these equations, each period there are $Q-\delta M(v)$ houses available. Given the policy v, agents are allocated houses tier by tier until the first tier such that not all agents can be allocated housing. Let the first such tier be l where agents are in l if $s \in [v_{l-1}, v_l)$. Clearly all agents with $s < v_{l-1}$ have zero assignment probability and all agents with $s \geq v_l$ have unit assignment probability. In the steady state there is a mass $1-K_s(v_l;v)$ of agents with priority above v_l. Moreover, there is a mass $K_s(v_l;v)-K_s(v_{l-1};v)$ who have score $s \in [v_{l-1}, v_l)$. Thus agents with scores

$s \in [v_{l-1}, v_l)$ are allocated a house with probability given by:

$$p_L(v) = \frac{(Q - \delta M(v)) - (1 - K_s(v_l; v))}{K_s(v_l; v) - K_s(v_{l-1}; v)} \tag{894}$$

It follows that the allocation is given by:

$$P(s, v) = \begin{cases} 1, s \geq v_l, \\ p_L(v), s \in [v_{l-1}, v_l), \\ 0, s < v_{l-1} \end{cases} \tag{895}$$

Consider now a new coarsening policy $\hat{v} = (v_{l-1}, v_l)$. The allocations induced by $\hat{v}$ and v are equivalent. Thus all other equations describing the steady state are equivalent under the two policies. This proves the claim that for any v there exists a $v' = (v_1, v_2)$ that is outcome equivalent. Indeed, this is given by $v' = \hat{v}$. We have now proved all claims in the proposition. □

This highlights again the usefulness of Theorem 5 in providing a revealing structure to the optimal design: one can employ a tiered design, but it need contain at most three tiers.

D.5.1 An Example with Two Income Cutoffs

The main text uses the above result to explore the optimal design, giving sufficient conditions for there to be at least two income tiers in the optimum. Here, we additionally provide an explicit example that shows how three income tiers can be optimal when there is sufficient heterogeneity in outside options.

Example 8. *Suppose that there are three levels of outside option κ_P, κ_M and κ_R where an agent with score s (or equivalently income $1 - s$) has outside option given by:*

$$\kappa = \begin{cases} \kappa_R, s < \bar{s}_1, \\ \kappa_M, s \in [\bar{s}_1, \bar{s}_2), \\ \kappa_P, s \geq \bar{s}_2. \end{cases} \tag{896}$$

where $\bar{s}_1 < \bar{s}_2$. Moreover, we suppose that outside options are such that:

$$\begin{aligned} \bar{v}(\kappa_R, 0) > \bar{v}(\kappa_M, 1) \\ \bar{v}(\kappa_M, 0) > \bar{v}(\kappa_P, 1) \end{aligned} \tag{897}$$

That is to say, an agent with a better outside option who is never assigned public housing still has higher welfare than an agent with a lower outside option when they are certain of receiving public housing each period. Such an environment is perhaps natural when the extremely poor have much worse outside options than the moderately poor and the wealthy have very good outside options.

For intermediate values of Q,[14] *in the $\gamma \to \infty$ limit the optimal policy is such that:*

$$v_1^* = \bar{s}_1, \quad v_2^* = \bar{s}_2 \tag{898}$$

Moreover, agents with score $e < \bar{s}_1$ have zero probability of assignment, agents with score $s \in [\bar{s}_1, \bar{s}_2)$ have intermediate probability of assignment and agents with score $s \geq \bar{s}_2$ are assigned with certainty. This is optimal because assigning agents below $\bar{s}_1$ is always welfare reducing as they are the richest agents with the best outside options. Similarly, the planner must give agents above $\bar{s}_2$ unit probability of assignment. Finally, they assign the remaining houses uniformly among those with intermediate incomes by virtue of the egalitarian motive.

[14] Formally, we require H is such that in the steady state corresponding to the described optimum, there are fewer than $Q - \delta M$ agents with score above $\bar{s}_1$ in the stationary distribution of searching agents and there are more than $Q - \delta M$ agents with score above $\bar{s}_2$ in the stationary distribution of searching agents, where M is the stationary measure of matched agents.

D.6 Relaxing the Full Support Assumption

Despite being a fairly weak condition, it is interesting to consider the implications for our analysis when the full support assumption (Assumption 7) fails. In particular, in this Appendix, we proceed under the assumption only that there are no mass points in the distribution of students.[15]

In this case, the interim economy after tie-breaking may have multiple stable matchings.[16] Thus, restricting the mechanism designer to use a stable matching mechanism does not pin down the allocation as there is a one-to-one correspondence between stable matchings and allocations. In general, under a stable mechanism, the selected matching may depend on the entire priority structure, i.e. a mechanism can even return two different stable matchings for two different priority structures even if the set of stable matchings is the same under those two priority structures.

To deal with these issues, we focus on two different selection rules: one is a selection rule that selects the optimal stable matching with respect to the preferences of the mechanism designer and the other is the student optimal stable matching for any given economy. We show that with mechanism-designer optimal selection, suitably modified versions of Theorems 5 and 6 continue to hold and that with student-optimal selection, Theorem 1 continues to hold but Theorem 6 fails.

D.6.1 The Mechanism Designer Optimal Selection

In this section, we assume whenever multiple stable matchings exist, the mechanism will implement the best matching with respect to the preferences of the mechanism designer. We show that suitably modified versions of Theorems 5 and 6 continue to hold without a full support assumption. First, we prove a modified version of Theorem 5:[17]

Theorem 13. *For any coarsening Ξ with allocation g_Ξ corresponding to the mechanism designer optimal stable matching, there is an alternative trinary coarsening Ξ' with allocation $g_{\Xi'}$ corresponding to the mechanism designer optimal stable matching such that $Z(g_{\Xi'}) \geq Z(g_\Xi)$.*

Proof. For any Ξ and matching μ that is selected by the mechanism designer optimal selection, construct a trinary coarsening $\Xi'(\mu)$ as in the proof of Theorem 5. Notice that the matching μ is still stable under $\Xi'(\mu)$. Hence the mechanism designer optimal

[15] We give an example that shows the necessity of the no mass points assumption in this section.

[16] See Azevedo and Leshno (2016) for examples.

[17] Recall that Z is the objective function of the mechanism designer.

selection picks a matching that is weakly better than μ under the coarsening $\Xi'(\mu)$, thus $Z(g_{\Xi'}) \geq Z(g_{\Xi})$. $\square$

Hence, even if it is not possible to replicate each allocation via a trinary coarsening, there always exists an alternative trinary coarsening that is preferred by the mechanism designer under the mechanism designer optimal selection. Thus, restricting attention to trinary coarsenings is without loss of optimality.

The following result states that under mechanism designer optimal selection, an optimal coarsening exists, so Theorem 6 holds under this selection.

Theorem 14. *Under a stable matching mechanism with mechanism designer optimal selection, there exists an optimal trinary coarsening* Ξ*.*

Proof. From Theorem 13, we can restrict attention to trinary coarsenings, which we continue to represent by the coarsenings $v \in \mathcal{V}$, which is a compact set. Let $\psi : \mathcal{V} \rightrightarrows \mathbb{R}$ be the correspondence mapping each coarsening to the set of utility values of the mechanism designer under stable selections. We prove first that ψ is upper hemicontinuous (UHC) and then that it is closed-valued; establishing existence of an upper semi-continuous maximum selection $\mathcal{H}(v) = \max \psi(v)$. It then follows by compactness of $\mathcal{V}$ that an optimum exists.

Denoting the set of stable matchings (hence the possible allocations) under coarsening v by $\mathcal{S}(v)$, we can write $\psi = Z \circ \mathcal{S}(v)$. Azevedo and Leshno (2016) prove that the set of stable matchings has a cut-off structure and is an upper hemicontinuous function of the economy (Proposition B1 in the appendix of Azevedo and Leshno (2016)). Moreover, by Assumption 9, Z is a continuous function of the allocation (which is the stable matching selected by the mechanism), hence ψ is a UHC correspondence of a coarsening as it is a continuous function of a UHC correspondence.[18]

We now establish that ψ is closed valued. To this end, fix a coarsening v. Now consider the set of market clearing cutoffs $\mathcal{P}$ that correspond to the set of stable matchings under v. Consider a sequence $\{z_n\}$ such that $z_n \in \psi(v)$ for all n that converges to some z^*. We wish to prove that $z^* \in \psi(v)$. For any z_n, the utility value from an admissible stable matching under v, there must exist a corresponding vector of market clearing cutoffs $P_n \in \mathbb{R}^{|\mathcal{C}|}$. Then we can find a subsequence P_{n_k} that is element-wise monotone in $\mathbb{R}^{|\mathcal{C}|}$. As the set of market clearing cut-offs forms a complete lattice (see Azevedo and Leshno (2016) Theorem A1), $\lim_{n_k} P_{n_k} = P^*$ and $P^* \in \mathcal{P}$. Define ν_n as the measure of students whose assignment is different under

[18] See Border (1989) for a proof.

P_n to P^*. In the coarsened economy, there are no mass points in the distribution of students. Thus, for any $\varepsilon > 0$ there exists an N such that $\nu_{n_k} < \varepsilon$ for all $n_k > N$. Hence, as $P_{n_k} \to P^*$, we have $Z(P_{n_k}) \to Z(P^*) = z^*$. Thus, as $P^* \in \mathcal{P}$, we have shown that $z^* \in \psi(v)$. Hence, ψ is closed-valued.

Finally, as ψ is closed-valued, $\mathcal{H}(\psi(v)) = \max \psi(v)$ is well-defined. Moreover, $\mathcal{H}(v)$ corresponds to the utility value of the mechanism designer under the mechanism designer optimal selection. By the UHC property of ψ, $\mathcal{H}$ is upper-semicontinuous. Thus, as $\mathcal{V}$ is a compact set, by the extreme value theorem it follows that $\max_{v \in \mathcal{V}} \mathcal{H}(\psi(v))$ exists. □

D.6.2 The Student Optimal Selection

In this section we assume whenever multiple stable matchings exist, the mechanism will implement the student optimal selection. In this case, while it is without loss of optimality to restrict attention to trinary coarsenings (Theorem 5 continues to hold), we show via an example that an optimum can fail to exist in the absence of the full support assumption (Theorem 6 does not hold). First, we prove that Theorem 5 holds under the student optimal stable selection whether or not the full support assumption holds:

Theorem 15. *Under a stable matching mechanism with student optimal selection, for any coarsening Ξ that induces the allocation g_Ξ, there is an alternative trinary coarsening Ξ' such that $g_{\Xi'} = g_\Xi$.*

Proof. Fix an arbitrary coarsening Ξ. Divide the schools to two subsets: $c \in \tilde{\mathcal{C}}$ if Ξ_c is finite and $c \in \hat{\mathcal{C}}$ if Ξ_c is the identity function. For any school $c \in \tilde{\mathcal{C}}$, the coarsening Ξ takes any student with score $s_c^{\tilde{\theta}}$ to an equivalence class, hence $\Xi(s_c^{\tilde{\theta}}) \in \{P_1^c, P_2^c, \ldots, P_N^c\}$ with $P_1^c < P_2^c < \ldots < P_N^c$ for some N.

Let $\tilde{\mu}$ denote the student optimal stable matching and P_x^c be the lowest indifference class that has a student placed in school c, i.e. $\tilde{\mu}(\tilde{\theta}_\Xi, \tau) = c$ for some $(\tilde{\theta}_\Xi, \tau)$ with $s_c^{\tilde{\theta}_\Xi} = P_x^c$ and $\tilde{\mu}(\tilde{\theta}_\Xi, \tau) \neq c$ for all $(\tilde{\theta}_\Xi, \tau)$ with $s_c^{\tilde{\theta}_\Xi} = P_y^c$ where $y < x$. For all $c \in \tilde{\mathcal{C}}$, define Ξ'_c as in the proof of Theorem 5,

$$\Xi'_c(s_c^{\tilde{\theta}}) = \begin{cases} P_1^c & \text{, if} \quad \Xi_c(s_c^{\tilde{\theta}}) = P_z^c, z < x, \\ P_x^c & \text{, if} \quad \Xi_c(s_c^{\tilde{\theta}}) = P_x^c, \\ P_N^c & \text{, if} \quad \Xi_c(s_c^{\tilde{\theta}}) = P_z^c, z > x. \end{cases} \tag{899}$$

Next, for all $c \in \hat{\mathcal{C}}$, as we did in the proof of Theorem 5, let t_c denote the threshold that the students must clear in order to gain admission to that school. Formally,

$$t_c = \inf\{s_c^\theta : \tilde{\mu}(\theta) = c\} \tag{900}$$

Then, for all $c \in \hat{\mathcal{C}}$, define Ξ'_c in the following way,

$$\Xi'_c(s_c^\theta) = \begin{cases} 0 & , \text{ if } \quad s_c^\theta < t_c, \\ 1 & , \text{ if } \quad s_c^\theta \geq t_c. \end{cases} \tag{901}$$

Note that Ξ' is a trinary coarsening. The stability of $\tilde{\mu}$ under Ξ' follows exactly the same arguments in Theorem 1 and is therefore omitted.

Let μ' denote the student optimal stable matching under Ξ'. To see why μ' is stable under Ξ, assume student pair i, j (with scores s_c^i and s_c^j and tie-breakers τ_i and τ_j) blocks it in school c', i.e. $c' \succ_i \mu'(i)$, $\mu'(j) = c'$ and $\Xi_{c'}(s_{c'}^i) > \Xi_{c'}(s_{c'}^j)$ or $\Xi_{c'}(s_{c'}^i) = \Xi_{c'}(s_{c'}^j)$ and $\tau_i > \tau_j$.

First, assume $c' \in \tilde{\mathcal{C}}$ and $\Xi_{c'}(s_{c'}^i) > \Xi_{c'}(s_{c'}^j)$. Then $\mu'(j) = c'$ and $\Xi_{c'}(s_{c'}^i) > \Xi_{c'}(s_{c'}^j)$ imply that $\Xi'_{c'}(s_{c'}^i) = P_N^{c'}$, which means that i belongs to a class strictly higher than $P_x^{c'}$. From the definition of $P_x^{c'}$, we know that there exists k such that $\Xi'_{c'}(k) = P_x^{c'}$ and $\mu'(k) = c'$, which is a contradiction as i and c' will block μ' under Ξ'.

Next, assume $c' \in \tilde{\mathcal{C}}$, $\Xi_{c'}(s_{c'}^i) = \Xi_{c'}(s_{c'}^j)$ and $\tau_i > \tau_j$. There are two cases, either $\Xi'_{c'}(s_{c'}^i) > P_x^{c'}$ or $\Xi'_{c'}(s_{c'}^i) = P_x^{c'}$. In the first case, from the definition of $P_x^{c'}$, there exists k such that $\tilde{\mu}(k) = c'$ and $\Xi'_{c'}(s_{c'}^k) = P_x^{c'}$, which is a contradiction as i and c' will block μ' under Ξ'. In the second case, as the lottery classes are same under Ξ and Ξ', $\Xi'_{c'}(s_{c'}^i) = P_x^{c'}$ and $\Xi_{c'}(s_{c'}^i) = \Xi_{c'}(s_{c'}^j)$ imply that $\Xi'_{c'}(s_{c'}^j) = P_x^{c'}$. However, as $\tau_i > \tau_j$, this is a contradiction: i and c' will block μ' under Ξ'.

Finally, if $c' \in \hat{\mathcal{C}}$, then Ξ is the identity function. There are two cases, $s_{c'}^i = s_{c'}^j$ and $\tau_i > \tau_j$ or $s_{c'}^i > s_{c'}^j$. In the first, case, $\mu'(j) = c'$ implies that $s_{c'}^i = s_{c'}^j \geq t_{c'}$, which implies $\Xi'_{c'}(s_{c'}^i) = \Xi'_{c'}(s_{c'}^j) = 1$. As $\tau_i > \tau_j$ and $c' \succ_i \mu'(i)$, this contradicts the stability of μ' under Ξ'. In the second case, as $s_{c'}^i > t_{c'}$, it must be that $\tilde{\mu}(i) \succeq_i c'$ (where $c \succeq_i c'$ if $c \succ_i c'$ or $c = c'$), as otherwise $\tilde{\mu}$ will not be stable under Ξ. Moreover, we know that $\tilde{\mu}$ is stable under Ξ'. As μ' is the student optimal stable matching under Ξ', all students must weakly prefer their matching under μ' to $\tilde{\mu}$ which is a contradiction as $\tilde{\mu}(i) \succeq_i c' \succ_i \mu'(i)$. Thus, μ' is stable under Ξ.

As μ' is stable under Ξ, $\tilde{\mu}$ is stable under Ξ' and both are student optimal stable matchings, they must be equivalent. Thus the same matching is used in the construction of g_Ξ and $g_{\Xi'}$ under the student optimal selection and $g_\Xi = g_{\Xi'}$. This proves the result as Ξ' is a trinary coarsening. $\square$

This result implies that even without the full support assumption, if the student optimal stable mechanism is used to compute the stable matching in the interim economy, then it is still without loss of optimality to restrict attention to trinary coarsenings.

However, in the absence of the full support assumption, an optimal coarsening can fail to exist. The intuition for why this is true is that in the limit as the mechanism designer changes the coarsening thresholds, there can exist multiple stable matchings and the student optimal selection from the limit set of stable matchings may be worse from the mechanism designer's perspective than the matchings attained with arbitrarily close coarsening thresholds. Example 9 provides a concrete example of this phenomenon.

Example 9. *There are four schools $\mathcal{C} = \{c_0, c_1, c_2, c_3\}$, capacities $Q = (4, 1, 1, 1)$, and four ordinal types of students, each of measure one, with preferences:*

$$\begin{aligned} &\succ_{t_1}: c_1, c_2, c_0 \\ &\succ_{t_2}: c_2, c_1, c_0 \\ &\succ_{t_3}: c_3, c_1, c_0 \\ &\succ_{t_4}: c_3, c_0 \end{aligned} \tag{902}$$

The marginal score distributions in the acceptable schools are given by:

$$\begin{aligned} s^{c_1}(t_2) \sim U[2,3],\ s^{c_1}(t_3) \sim U[1,2],\ &s^{c_1}(t_1) \sim U[0,1] \\ s^{c_2}(t_1) \sim U[1,2],\ &s^{c_2}(t_2) \sim U[0,1] \\ s^{c_3}(t_3) \sim U[1,2],\ &s^{c_3}(t_4) \sim U[0,1] \end{aligned} \tag{903}$$

Assume the mechanism designer maximizes the following utility function:

$$Z(g) = \sum_t \sum_{c \in \{c_1, c_2, c_3\}} u_t(c) g(c|t) \tag{904}$$

with:

$$u_{t_1}(c_i) = \begin{cases} 1 & \text{if} \quad i = 2 \\ 0 & \text{if} \quad i = 1 \end{cases} \tag{905}$$

$$u_{t_2}(c_i) = \begin{cases} 1 & \text{if} \quad i = 1 \\ 0 & \text{if} \quad i = 2 \end{cases} \tag{906}$$

$$u_{t_3}(c_i) = 0 \quad \text{for} \quad i = 1, 3 \tag{907}$$

$$u_{t_4}(c_3) = -1 \tag{908}$$

This utility function assumes that the designer prefers assigning type t_1 students to school c_2 and type t_2 students to the school c_1, while assigning the minimal amount of type t_4 students to school c_3. The first best outcome for the designer is to match all t_1 students to c_2, all t_2 students to c_1 and all t_3 students to c_3. This yields a payoff of $Z^{FB} = 2$. Note that without priority design, the student optimal stable matching assigns all students of types t_1 to c_1, t_2 to c_2, t_3 to c_3, and t_4 to c_0. The yields a payoff of $Z^{SOSM} = 0$.

We note that the first-best matching is not attainable as the student-optimal stable matching under any coarsening. To match all t_1 students to c_2 and all t_2 students to c_1 in a stable matching, a necessary and sufficient condition is to reject a positive measure of t_3 students from school c_3 (and assign a positive measure of type t_4 students to c_3). But this necessarily leads to a payoff strictly smaller than 2 as some positive measure t_4 students are assigned to c_3.

However, the payoff from the first-best matching can be arbitrarily well approximated through coarsening. Define for any $\epsilon \geq 0$, the coarsening:

$$\Xi_{c_3}(s_{c_3}; \epsilon) = \begin{cases} 0, s_{c_3} < 1 - \epsilon \\ 1, s_{c_3} \in [1-\epsilon, 1+\epsilon), \\ 2, s_{c_3} \geq 1 + \epsilon \end{cases} \tag{909}$$

and $\Xi_c(s_c; \epsilon) = s_c$ for all $c \neq c_3$.

Under this coarsening for $\epsilon > 0$, the unique stable matching assigns $\frac{\epsilon}{2}$ measure type t_4 students are assigned to c_3. Thus, $\frac{\epsilon}{2}$ measure t_3 students are assigned to either c_1 or c_0. Given that t_2 types have the highest priority at c_1 and t_1 types have the highest priority at c_2, it is not possible under any stable matching that any type t_1 or t_2 student is assigned to c_0. Therefore, $\frac{\epsilon}{2}$ type t_3 students are assigned to c_0. Hence, it is not possible that any t_1 student is assigned to c_1 or c_1 and any t_3 type would form a blocking pair, as t_3 students have higher priority at c_1. Thus, in the unique stable matching, all t_1 types are assigned to c_2, all t_2 types are assigned to c_1, and measure $\frac{\epsilon}{2}$ of t_4 types are assigned to c_3. For $\epsilon = 0$, the outcome is the outcome without coarsening. Thus, the payoff of the mechanism designer is given by:

$$Z(\epsilon) = \begin{cases} 2 - \frac{\epsilon}{2}, \epsilon > 0, \\ 0, \epsilon = 0. \end{cases} \tag{910}$$

To conclude, the first-best is not attainable but can be arbitrarily well approximated by coarsening. Thus, the problem of the designer is not upper semi-continuous and an optimum does not exist under the student-optimal selection.

As a result, an optimum can cease to exist in the absence of full support under the student optimal stable matching. This notwithstanding, given the earlier results, it remains true that if the mechanism-designer optimal and student optimal selections coincide, then Theorem 6 holds. This is, for instance, true in the natural case with a utilitarian mechanism designer.

D.6.3 The Necessity of the No Mass Points Assumption for Existence of an Optimum

While we relaxed the full support assumption in this appendix, we retained the no mass points assumption. To see the necessity of the no mass points assumption for the existence of an optimum, we introduce the following example. Intuitively, with mass points, changing the coarsening cutoffs by an arbitrarily small amount can alter the allocation of students to schools in a way that adversely impacts the objective function of the mechanism designer.

Example 10. *Suppose that there is one school c with capacity of unit measure and an outside option. Suppose all students prefer c to the outside option. There are three types of students, each of unit measure. Students with score $s \in [0.5, 1]$ who give the mechanism designer utility of 1 when they are assigned to the school (Type 1), those with score $s \in [0, 0.5)$, who give the mechanism designer utility of 2 (Type 2), and those with score $s = 0$ who give utility -M to the mechanism designer (Type 3). The first two types are uniformly distributed over their respective domains. See that there is a mass point of students with score zero.*

It is clear here that under any coarsening, if there exists another coarsening such that more type 2 students attend the school while no type 3 students are admitted, then that initial coarsening must not be optimal. In particular, suppose that those with scores $s \geq \underline{v} > 0$ are coarsened into the same indifference class. The mechanism designer's utility is:

$$U(\underline{v}) = \frac{0.5}{1 - \underline{v}} \times 1 + \frac{0.5 - \underline{v}}{1 - \underline{v}} \times 2 = \frac{1.5 - 2\underline{v}}{1 - \underline{v}}, \ \underline{v} \in (0, 1] \tag{911}$$

which is decreasing in $\underline{v}$ for all $\underline{v} > 0$ so no $\underline{v} > 0$ can be optimal. Now suppose that *the mechanism designer fully coarsens the students so that* $\underline{v} = 0$*. Their utility is:*

$$U(0) = \frac{1}{3} \times 1 + \frac{1}{3} \times 2 - \frac{1}{3}M \tag{912}$$

Hence, a sufficient condition that there will not exist an optimal coarsening is that setting $\underline{v} = 0.5$ *(i.e. not coarsening) is better than full coarsening. That is:*

$$U(0) = 1 - \frac{1}{3}M < 1 = U(0.5) \tag{913}$$

which is satisfied for any $M > 0$*. Hence, whenever* $M > 0$*, there does not exist an optimal coarsening.*

Appendix E

Appendix to Adaptive Priority Mechanisms

E.1 Omitted Proofs

E.1.1 Proof of Proposition 16

Proof. Part (i): In state ω the payoff from admitting the highest-scoring minority students of measure $x(\omega)$ is:

$$q\omega + (1+\gamma-\omega)x(\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)x(\omega)^2 \tag{914}$$

Thus, the $x(\omega)$ that solves the FOC is given by:

$$x(\omega) = \frac{\kappa(1+\gamma-\omega)}{1+\kappa\gamma\beta} \tag{915}$$

Under our maintained assumptions, we have that:

$$x(\omega) = \frac{\kappa(1+\gamma-\omega)}{1+\kappa\gamma\beta} \leq \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega}-\underline{\omega}) < \min\{\kappa, q\} \tag{916}$$

and:

$$x(\omega) = \frac{\kappa(1+\gamma-\omega)}{1+\kappa\gamma\beta} \geq \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta} > \kappa(1-\underline{\omega}) \geq 0 \tag{917}$$

Thus, this level of minority admissions is feasible. Substituting, we have that:

$$V^* = q\mathbb{E}[\omega] + \frac{1}{2}\frac{\mathbb{E}[\kappa(1+\gamma-\omega)^2]}{1+\kappa\gamma\beta} \tag{918}$$

Consider now the APM $A(y) = \gamma(1 - \beta y)$. Agents are allocated the resource if their modified scores exceed ω, with a uniform lottery over students with score exactly ω. Thus, in state ω, this policy admits measure $y(\omega)$ minorities that solve the fixed point equation:

$$y(\omega) = \min\left\{\kappa \int_0^1 \mathbb{I}[s + A(y(\omega)) \geq \omega]\mathrm{d}s, q\right\} = \min\{\kappa\left(1 - \max\{\omega - A(y(\omega)), 0\}\right), q\} \tag{919}$$

Denote the RHS of this fixed point equation by the function $\mathrm{RHS}(y, \omega)$, which is continuous and decreasing in y. Moreover, $\mathrm{RHS}(0, \omega) = \min\{\kappa(1 - \max\{\omega - \gamma, 0\}), q\} > 0$ and $\mathrm{RHS}(\min\{\kappa, q\}, \omega) < \min\{\kappa, q\}$. The second of these inequalities is true because the condition $\underline{\omega} > \gamma(1 - \beta \min\{q, \kappa\})$ follows from our assumption that $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$. Thus, there exists a unique $y(\omega)$ implemented by the APM. Moreover, let $y_A(\omega)$ denote the unique solution to the equation 919, which gives the measure of admitted minority students under APM A at state ω.

$$\begin{aligned}
y_A(\omega) &= \kappa(1 - (\omega - \gamma(1 - \beta y_A(\omega)))) \\
&= \kappa(1 - \omega + \gamma) - \kappa\gamma\beta y_A(\omega) \\
&= \frac{\kappa(1 - \omega + \gamma)}{1 + \kappa\gamma\beta}
\end{aligned} \tag{920}$$

Thus, A implements the optimal level of minority admissions characterized in equation 915 and $V_A = V^*$.

Part (ii): First, if we admit all minority students over some threshold $\hat{s}$, the total score of admitted minority students is $\kappa \int_{\hat{s}}^1 s\mathrm{d}s$. Moreover, when we admit measure x minority students where $x \leq \min\{\kappa, q\}$, this admissions threshold is defined by $x = \kappa \int_{\hat{s}}^1 \mathrm{d}s = \kappa(1 - \hat{s})$. Thus, we have that $\hat{s} = 1 - \frac{x}{\kappa}$. Finally, the residual measure $q - x$ admitted majority students all score ω. Thus, the total score is given by $\bar{s} = q\omega + (1 - \omega)x - \frac{1}{2\kappa}x^2$ for $0 \leq x \leq \min\{\kappa, q\}$. As both quota and priority policies always admit the highest-scoring minority students, the authority's utility is given by:

$$\mathcal{U} = q\mathbb{E}[\omega] + \mathbb{E}[(1 + \gamma - \omega)x] - \frac{1}{2}\left(\frac{1}{\kappa} + \gamma\beta\right)\mathbb{E}[x^2] \tag{921}$$

We now derive the admitted measure of minority students. In the absence of a priority or quota policy, $\alpha = 0$ or $Q = 0$, we have that $x = \kappa(1 - \omega)$ measure minority students is admitted. Thus, under a quota policy Q, measure $x = \max\{Q, \kappa(1 - \omega)\}$ minority students are admitted. Under a priority policy, the measure of admitted minority students is $x = \kappa \int_{\omega-\alpha}^{1} \mathrm{d}x = \kappa(1+\alpha-\omega)$. In each case x is capped by $\min\{\kappa, q\}$ and floored by 0.

The expected utility function over quotas is given by one of four cases. First, $Q > \min\{\kappa, q\}$ and:

$$\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\min\{\kappa, q\} - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\min\{\kappa, q\}^2 \tag{922}$$

Second, $Q \in [\kappa(1-\underline{\omega}), \min\{\kappa, q\})$ and:[1]

$$\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])Q - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^2 \tag{923}$$

Third, $Q \in (\kappa(1-\overline{\omega}), \kappa(1-\underline{\omega}))$ and:

$$\begin{aligned}\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] &+ \int_{1-\frac{Q}{\kappa}}^{\overline{\omega}} \left((1+\gamma-\omega)Q - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^2\right)\mathrm{d}\Lambda(\omega) \\ &+ \int_{\underline{\omega}}^{1-\frac{Q}{\kappa}} \left((1+\gamma-\omega)\kappa(1-\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)(\kappa(1-\omega))^2\right)\mathrm{d}\Lambda(\omega)\end{aligned} \tag{924}$$

Finally, $Q \leq \kappa(1-\overline{\omega})$ and:

$$\mathcal{U}_Q(Q) = q\mathbb{E}[\omega] + \mathbb{E}\left[(1+\gamma-\omega)\kappa(1-\omega)\right] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}\left[(\kappa(1-\omega))^2\right] \tag{925}$$

We claim that the optimum lies in the second case. See that in case two the strict maximum is attained at $Q^* = \frac{1+\gamma-\mathbb{E}[\omega]}{\frac{1}{\kappa}+\gamma\beta} \in (\kappa(1-\underline{\omega}), \min\{\kappa, q\})$, by our assumptions that $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega}-\underline{\omega})$ and $\kappa(1-\underline{\omega}) < \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta}$. Moreover, in case three, the first derivative of the payoff is given by:

$$\mathcal{U}_Q'(Q) = \int_{1-\frac{Q}{\kappa}}^{\overline{\omega}} \left((1+\gamma-\omega) - \left(\frac{1}{\kappa}+\gamma\beta\right)Q\right)\mathrm{d}\Lambda(\omega) \tag{926}$$

Thus, checking that the sign of this is positive amounts to verifying that for all $Q \in (\kappa(1-\overline{\omega}), \kappa(1-\underline{\omega}))$, we have that:

$$Q < \frac{1+\gamma-\mathbb{E}[\omega|\omega \geq 1-\frac{Q}{\kappa}]}{\frac{1}{\kappa}+\gamma\beta} \tag{927}$$

[1] By our maintained assumptions we have that this interval has non-empty interior.

As the RHS is an increasing function of Q, it suffices to show that:

$$\kappa(1-\underline{\omega}) < \frac{1+\gamma-\overline{\omega}}{\frac{1}{\kappa}+\gamma\beta} \tag{928}$$

which we have assumed. Moreover, the expected utility in the first case equals $\mathcal{U}_Q(\kappa(1-\underline{\omega}))$, thus is lower than the optimum of the second case. The expected utility in the fourth case equals $\mathcal{U}_Q(\kappa(1-\overline{\omega}))$, thus is lower than the optimum of the third case. We therefore have that:

$$V_Q = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])Q^* - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^{*2} \tag{929}$$

We now turn to characterizing the value of priorities. There are three cases to consider. First, when $\kappa(1+\alpha-\overline{\omega}) \geq \min\{\kappa, q\}$ we have that $x = \min\{\kappa, q\}$ and:

$$\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\min\{\kappa,q\} - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\min\{\kappa,q\}^2 \tag{930}$$

Second, when $\kappa(1+\alpha-\underline{\omega}) \geq \min\{\kappa, q\} \geq \kappa(1+\alpha-\overline{\omega})$ we have that:

$$\begin{aligned}\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] &+ \int_{\underline{\omega}}^{1+\alpha-\min\{\frac{q}{\kappa},1\}} \left((1+\gamma-\omega)\min\{\kappa,q\} - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\min\{\kappa,q\}^2\right) \mathrm{d}\Lambda(\omega) \\ &+ \int_{1+\alpha-\min\{\frac{q}{\kappa},1\}}^{\overline{\omega}} \left((1+\gamma-\omega)\kappa(1+\alpha-\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)[\kappa(1+\alpha-\omega)]^2\right) \mathrm{d}\Lambda(\omega)\end{aligned} \tag{931}$$

Finally, when $\min\{\kappa, q\} \geq \kappa(1+\alpha-\underline{\omega})$, we have that:

$$\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] + \mathbb{E}[(1+\gamma-\omega)\kappa(1+\alpha-\omega)] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}[(\kappa(1+\alpha-\omega))^2] \tag{932}$$

We claim that the optimum under our assumptions lies only in the third case. First, we argue that there is a unique local maximum in the third case. Second, we show the value in the second case is decreasing in α. By continuity, the unique optimum then lies in the third case.

First, it is helpful to write $\bar{x}(\alpha) = \kappa(1+\alpha-\mathbb{E}[\omega])$ and $\varepsilon = \kappa\left(\mathbb{E}[\omega]-\omega\right)$. The value in the third case can then be re-expressed as:

$$\begin{aligned}\mathcal{U}_P(\alpha) &= q\mathbb{E}[\omega] + \mathbb{E}[(1+\gamma-\omega)\left(\bar{x}(\alpha)+\varepsilon\right)] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}[(\bar{x}(\alpha)+\varepsilon)^2] \\ &= q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\bar{x}(\alpha) - \mathbb{E}[\omega\varepsilon] - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\bar{x}(\alpha)^2 - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\mathbb{E}[\varepsilon^2]\end{aligned} \tag{933}$$

Finally, we have that $\mathbb{E}[\varepsilon^2] = \kappa^2\mathrm{Var}[\omega]$ and $\mathbb{E}[\omega\varepsilon] = \mathrm{Cov}[\omega, \varepsilon] = -\kappa\mathrm{Var}[\omega]$. Thus:

$$\mathcal{U}_P(\alpha) = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])\bar{x}(\alpha) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)\bar{x}(\alpha)^2 + \frac{\kappa}{2}\left(1-\kappa\gamma\beta\right)\mathrm{Var}[\omega] \tag{934}$$

We then see that the optimal α^* in this range sets $\bar{x}(\alpha^*) = Q^* < \min\{\kappa, q\}$. It remains only to check that this optimal α^* indeed lies within this case, or equivalently that $\kappa(1+\alpha^*-\underline{\omega}) \leq \min\{\kappa, q\}$. To this end, see that $\kappa(1+\alpha^*-\mathbb{E}[\omega]) = Q^*$, and:

$$\begin{aligned}\kappa(1+\alpha^*-\underline{\omega}) &= Q^* + \kappa(\mathbb{E}[\omega]-\underline{\omega}) \leq Q^* + \kappa(\overline{\omega}-\underline{\omega}) \\ &\leq \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega}-\underline{\omega}) < \min\{\kappa, q\}\end{aligned} \tag{935}$$

where the final inequality follows by our assumption that $\min\{\kappa, q\} > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$.

Second, in the second case we have that the first derivative of the payoff in α is given by:

$$\begin{aligned}\mathcal{U}_P'(\alpha) &= \int_{1+\alpha-\min\{\frac{q}{\kappa},1\}}^{\overline{\omega}} \frac{d}{d\alpha}\left((1+\gamma-\omega)\kappa(1+\alpha-\omega) - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)[\kappa(1+\alpha-\omega)]^2\right)\mathrm{d}\Lambda(\omega) \\ &= \kappa\int_{1+\alpha-\min\{\frac{q}{\kappa},1\}}^{\overline{\omega}}\left((1+\gamma-\omega) - \left(\frac{1}{\kappa}+\gamma\beta\right)(\bar{x}(\alpha)+\varepsilon(\omega))\right)\mathrm{d}\Lambda(\omega)\end{aligned} \tag{936}$$

Checking that the sign of this is negative for all α such that $\kappa(1+\alpha-\underline{\omega}) \geq \min\{\kappa, q\} \geq \kappa(1+\alpha-\overline{\omega})$ then amounts to checking that:

$$\bar{x}(\alpha) > \frac{1+\gamma-\mathbb{E}[\omega|\omega \geq 1+\alpha-\min\{\frac{q}{\kappa},1\}]}{\frac{1}{\kappa}+\gamma\beta} - \mathbb{E}\left[\varepsilon(\omega)|\omega \geq 1+\alpha-\min\{\frac{q}{\kappa},1\}\right] \tag{937}$$

for all $\bar{x}(\alpha) \in [\min\{\kappa, q\} - \kappa(\mathbb{E}[\omega]-\underline{\omega}), \min\{\kappa, q\} - \kappa(\mathbb{E}[\omega]-\overline{\omega})]$. So it suffices to check that the minimal possible value of the LHS exceeds the maximal possible value of the RHS. A sufficient condition for this is that:

$$\min\{\kappa, q\} - \kappa(\mathbb{E}[\omega] - \underline{\omega}) > \frac{1+\gamma-\underline{\omega}}{\frac{1}{\kappa}+\gamma\beta} - \kappa(\mathbb{E}[\omega] - \overline{\omega}) \tag{938}$$

Which holds as we assumed that $\min\{\kappa, q\} > \frac{1+\gamma-\omega}{\frac{1}{\kappa}+\gamma\beta} + \kappa(\overline{\omega} - \underline{\omega})$. Substituting the optimal priority policy $\bar{x}(\alpha) = Q^*$ in equation 932, we obtain

$$V_P = q\mathbb{E}[\omega] + (1+\gamma-\mathbb{E}[\omega])Q^* - \frac{1}{2}\left(\frac{1}{\kappa}+\gamma\beta\right)Q^{*2} + \frac{\kappa}{2}(1-\kappa\gamma\beta)\operatorname{Var}[\omega] \tag{939}$$

We have now established that:

$$\Delta = V_P - V_Q = \frac{\kappa}{2}(1-\kappa\gamma\beta)\operatorname{Var}[\omega] \tag{940}$$

Part (iii): We have V^*, V_Q, V_P. Thus, we can compute the loss from restricting to quota policies:

$$\mathcal{L}_Q = \frac{1}{2}\frac{\kappa\operatorname{Var}[\omega]}{1+\kappa\gamma\beta} \tag{941}$$

To find the loss from restricting to priority policies, we compute:

$$\mathcal{L}_P = \mathcal{L}_Q - \Delta = \frac{1}{2}(\kappa\gamma\beta)^2\frac{\kappa\operatorname{Var}[\omega]}{1+\kappa\gamma\beta} \tag{942}$$

Enveloping over these losses yields the claimed formula. □

E.1.2 Proof of Proposition 17

Proof. Adapting Definition 13 to single object setting, we say that a matching μ admits a cutoff structure if there exists $S(\omega) = \{S_m(\omega)\}_{m\in\mathcal{M}}$ such that $\mu(s, m; \omega) = 1$ if and only if $s \geq S_m(\omega)$. A mechanism admits a cutoff structure if it admits a cutoff structure at every ω. We will first prove that any monotone APM admits a cutoff structure.

Lemma 20. *A monotone APM admits a cutoff structure.*

Proof. For a contradiction, assume it does not. Then there exists ω and matching μ implemented by the monotone APM such that for some $m \in \mathcal{M}$, $s > s'$, $\mu(s, m; \omega) = 0$ but $\mu(s', m; \omega) = 1$. Let x_m denote the measure of group m agents allocated the resource at μ. Since A is a monotone APM and $s > s'$, we have that $A(x_m, s) > A(x_m, s')$, which contradicts that A implements μ. □

We now use Lemma 20 to show that a monotone APM implements a unique allocation. Assume for a contradiction that $A_m(y_m, s)$ is monotone and implements two different allocations, μ and μ'. Let x_m and x'_m denote the measure of type m students assigned the resource at μ and μ'. First, we prove that if μ and μ' admit the same measure of students from each group, *i.e.*, $x_m = x'_m$ for all m, then the average score of admitted students are the same. Let s_m and s'_m denote the score of the lowest-scoring type m students assigned the resource at μ and μ'.

Claim 1. *If $x_m = x'_m$ for all $m \in \mathcal{M}$, then $\bar{s}_h(\mu, \omega) = \bar{s}_h(\mu', \omega)$*

Proof. Fix an m. Without loss of generality, let $s_m \geq s'_m$. First, since APM has cutoff structure and $x_m = x'_m$, we have that

$$\int_\Theta \mathbb{I}\{s(\theta) \in [s'_m, s_m], m(\theta) = m\} \mathrm{d}F_\omega(s, m) = 0 \tag{943}$$

Note that this holds regardless of $m' \in \mathcal{M}$ and whether $s_m \geq s'_m$ or $s'_m \geq s_m$. Therefore,

$$\begin{aligned}
\bar{s}_h(\mu, \omega) &= \int_\Theta \mu(s, m) h(s) \mathrm{d}F_\omega(s, m) \\
&= \sum_{m \in \mathcal{M}} \int_\Theta \mathbb{I}\{s(\theta) \geq s_m, m(\theta) = m\} h(s(\theta)) \mathrm{d}F_\omega(s, m) \\
&= \sum_{m \in \mathcal{M}} \int_\Theta \mathbb{I}\{s(\theta) \geq s'_m, m(\theta) = m\} h(s(\theta)) \mathrm{d}F_\omega(s, m) \\
&= \int_\Theta \mu(s, m) h(s) \mathrm{d}F_\omega(s, m) \\
&= \bar{s}_h(\mu', \omega)
\end{aligned} \tag{944}$$

where line equation holds from Equation 943 and all others are by definition. This finishes the proof of the claim. □

Therefore, if μ and μ' do not yield identical measures, then there are m and n such that $x_m > x'_m$ and $x'_n > x_n$. Since $x_m > x'_m$, it follows that $s'_m > s_m$. Likewise $x'_n > x_n$ implies that $s_n > s'_n$. Note that these imply: (i) $\mu'(s'_n, n) = 1$ while $\mu'(s'_m, n) = 0$ and (ii) $\mu(s_m, m) = 1$ while $\mu(s'_n, n) = 0$. Thus, the following inequalities hold:

$$A_n(s'_n, x'_n) > A_m(s'_m, x'_m) \geq A_m(s_m, x_m) > A_n(s'_n, x_n) \geq A_n(s'_n, x'_n) \tag{945}$$

where the first inequality follows from (i), the second inequality follows from the fact that $x'_m < x_m$ and A is monotone, the third inequality follows from (ii) and the fourth inequality follows from the fact that $x_n < x'_n$ and A is monotone. This equation yields $A_n(s'_n, x'_n) > A_n(s'_n, x'_n)$, which is a contradiction. Therefore, all allocations implemented by A yield the same x. Thus, from Lemma 20, if a monotone APM A implements μ and μ', both allocations admit the highest-scoring measure x_m agents from group m and can differ (at most) on a measure 0 set, proving the essential uniqueness of A. □

E.1.3 Proof of Theorem 7

Proof. We characterize the optimal allocation for each $\omega \in \Omega$ and show that the claimed adaptive priority mechanism implements the same allocation. Fix an $\omega \in \Omega$ and suppress the dependence of F_ω and f_ω thereon, and define the utility index of a score as $\tilde{s} = h(s)$ with induced densities over $\tilde{s}$ given by $\tilde{f}_m$ for all $m \in \mathcal{M}$. Let the measure of agents from any group $m \in \mathcal{M}$ that is allocated the resource be $x_m \in [0, \bar{x}_m]$ where $\bar{x}_m = \int_{h(0)}^{h(1)} \tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s}$. Observe that, conditional on fixing the measures of agents from each group that are allocated the resource $x = \{x_m\}_{m\in\mathcal{M}}$, there is a unique optimal allocation (*i.e.,* ξ-maximal μ up to measure zero transformations). In particular, as g and h are continuous and strictly increasing, the optimal allocation conditional on x satisfies $\mu^*(\tilde{s}, m; x) = 1 \iff \tilde{s} \geq \underline{\tilde{s}}_m(x_m)$ for some thresholds $\{\underline{\tilde{s}}_m(x_m)\}_{m\in\mathcal{M}}$ that solve:

$$\int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} = x_m \tag{946}$$

We can then express the problem of choosing the optimal $x = \{x_m\}_{m\in\mathcal{M}}$ as:

$$\max_{x_m \in [0,\bar{x}_m],\, \forall m \in \mathcal{M}} \sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} + \sum_{m\in\mathcal{M}} u_m(x_m) \quad \text{s.t.} \sum_{m\in\mathcal{M}} x_m \leq q \tag{947}$$

where a solution exists by compactness of the constraint sets and continuity of the objective. We can derive necessary and sufficient conditions on the solution(s) to this problem by considering the Lagrangian:

$$\begin{aligned}\mathcal{L}(x, \lambda, \overline{\kappa}, \underline{\kappa}) = &\sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} + \sum_{m\in\mathcal{M}} u_m(x_m) \\ &+ \lambda\left(q - \sum_{m\in\mathcal{M}} x_m\right) + \sum_{m\in\mathcal{M}} \overline{\kappa}_m(\bar{x}_m - x_m) + \sum_{m\in\mathcal{M}} \underline{\kappa}_m x_m\end{aligned} \tag{948}$$

The first-order necessary conditions to this program are given by:

$$\frac{\partial \mathcal{L}}{\partial x_m} = -\underline{\tilde{s}}'_m(x_m)\underline{\tilde{s}}_m(x_m)\tilde{f}_m(\underline{\tilde{s}}_m(x_m)) + u'_m(x_m) - \lambda - \overline{\kappa}_m + \underline{\kappa}_m = 0 \tag{949}$$

$$\lambda\frac{\partial \mathcal{L}}{\partial \lambda} = \lambda\left(q - \sum_{m\in\mathcal{M}} x_m\right) = 0 \tag{950}$$

$$\overline{\kappa}_m\frac{\partial \mathcal{L}}{\partial \overline{\kappa}_m} = \overline{\kappa}_m(\bar{x}_m - x_m) = 0 \tag{951}$$

$$\underline{\kappa}_m\frac{\partial \mathcal{L}}{\partial \underline{\kappa}_m} = \underline{\kappa}_m x_m = 0 \tag{952}$$

for all $m \in \mathcal{M}$. By implicitly differentiating Equation 946, we obtain that:

$$-\underline{\tilde{s}}'_m(x_m)\tilde{f}_m(\underline{\tilde{s}}_m(x_m)) = 1 \tag{953}$$

Thus, we can simplify Equation 949 to:

$$\frac{\partial \mathcal{L}}{\partial x_m} = \underline{\tilde{s}}_m(x_m) + u'_m(x_m) - \lambda - \overline{\kappa}_m + \underline{\kappa}_m = 0 \tag{954}$$

Observe that all constraints are linear. Thus, if the objective function is strictly concave, the first-order conditions are also sufficient. Observe by Equation 953 that $\underline{\tilde{s}}_m(x_m)$ is a strictly decreasing function of x_m, and all cross-partial derivatives are zero. Therefore, the first summation is strictly concave. Moreover u'_m is a decreasing function of x_m by virtue of the assumption that u_m is concave for all $m \in \mathcal{M}$. Therefore, the second summation is concave. Thus, the objective function is strictly concave and the optimal allocation is unique.

Thus, to verify that our claimed adaptive priority mechanism is a first-best mechanism, it suffices to show that the allocation it implements satisfies Equations 949 to 952. The adaptive priority mechanism $A_m(y_m, s) = h^{-1}\left(h(s) + u'_m(y_m)\right)$ in the transformed score space yields transformed scores $h\left(A_m(y_m, s)\right) = \tilde{s} + u'_m(y_m)$. Define x_m as the admitted measure of agents from group m under this mechanism. Agents in group $m \in \mathcal{M}$ are allocated the resource if and only if $\tilde{s} + u'_m(x_m) \geq s^C$ for some threshold s^C that solves:

$$\sum_{m\in\mathcal{M}} \int_{\max\{\min\{s^C - u'_m(x_m), h(1)\}, h(0)\}}^{h(1)} \tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} = q \tag{955}$$

We can therefore partition $\mathcal{M}$ into three sets that are uniquely defined: (i) interior

$\mathcal{M}_I = \{m \in \mathcal{M} | s^C - u'_m(x_m) \in (h(0), h(1))\}$; (ii) no allocation $\mathcal{M}_0 = \{m \in \mathcal{M} | s^C - u'_m(x_m) \geq h(1)\}$; (iii) full allocation $\mathcal{M}_1 = \{m \in \mathcal{M} | s^C - u'_m(x_m) \leq h(0)\}$. For all $m \in \mathcal{M}_0$, we implement $x_m = 0$. For all $m \in \mathcal{M}_1$, we implement $x_m = \bar{x}_m$. For all $m \in \mathcal{M}_I$, we implement $x_m \in (0, \bar{x}_m)$. For any $m \in \mathcal{M}_I$, the allocation threshold is $\underline{\tilde{s}}_m(x_m) = s^C - u'_m(x_m)$. For any $m \in \mathcal{M}_0$, the allocation threshold is $h(1)$. For any $m \in \mathcal{M}_1$, the allocation threshold is $h(0)$.

We now verify that this outcome satisfies the established necessary and sufficient conditions. For all $m \in \mathcal{M}_I$, by the complementary slackness conditions we have that $\underline{\kappa}_m = \overline{\kappa}_m = 0$. Substituting the above into Equation 949 for all $m \in \mathcal{M}_I$ we obtain that:

$$s^C - \lambda = 0 \tag{956}$$

which is satisfied for $\lambda = s^C$. As $q = \sum_{m \in \mathcal{M}} x_m$, the complementary slackness condition for λ is then satisfied. For all $m \in \mathcal{M}_0$, by complementary slackness we have that $\overline{\kappa}_m = 0$ and Equation 949 is satisfied by:

$$\underline{\kappa}_m = \lambda - h(1) - u'_m(0) \tag{957}$$

For all $m \in \mathcal{M}_1$, by complementary slackness we have that $\underline{\kappa}_m = 0$ and Equation 949 is satisfied by:

$$\overline{\kappa}_m = h(0) + u'_m(\bar{x}_m) - \lambda \tag{958}$$

This completes the proof of first-best optimality of A^*. Moreover, as the optimal allocation is unique for all ω, any allocation that differs from the allocation implemented by the optimal APM at any ω would not be first-best optimal. Therefore, any first-best-optimal mechanism must implement essentially the same allocation as A^*. □

E.1.4 Proof of Theorem 8

Proof. First, we prove the if parts of the results.

Part (i): When u_m is linear, u'_m is constant and the first-best optimal adaptive priority mechanism is a priority mechanism $P(s, m) = h^{-1}(h(s) + u'_m)$. Part (ii): When $\tilde{u}'_m(x_m) \geq k$ for $x_m \leq x_m^{\text{tar}}$ and $\tilde{u}'_m(x_m) = 0$ for $x_m > x_m^{\text{tar}}$ and $\sum_{m \in \mathcal{M}} x_m^{\text{tar}} < q$, observe that the optimal mechanism admits $x_m \geq x_m^{\text{tar}}$ for all $m \in \mathcal{M}$ in all states of the world, but conditional on $x_m \geq x_m^{\text{tar}}$ for all $m \in \mathcal{M}$ admits the highest-scoring set of agents. A quota $Q_m = x_m^{\text{tar}}$ and $Q_R = q - \sum_{m \in \mathcal{M}} x_m^{\text{tar}}$, with $D(R) = |\mathcal{M}| + 1$ implements this allocation and is first-best optimal for any authority that is extremely risk-averse.

Second, we prove the only if parts of the results.

Part (i): Assume the utility functions are not linear and let m denote a group where u'_m is not constant in $[0, q]$. We say that a state ω has full support if f_w has full support. A state ω has *full support in m and n* if $f_w(\cdot, m) > 0$ and $f_w(\cdot, n) > 0$ for some m and n and positive measures of only m and n. Let ω be a state that has full support in m and n. Moreover, assume both groups have a measure q of agents. We first establish that in any optimal allocation, agents from both groups are allocated the resource.

Claim 2. *If preferences are non-trivial, then the optimal allocation has $x_n, x_m > 0$.*

Proof. Toward a contradiction, suppose without loss of generality that $x_n = 0$. This implies that $x_m = q$. By the necessary first-order condition from Theorem 7 (combing Equations 949 and 952), we have that:

$$u'_m(q) + h(0) = u'_n(0) + h(1) + \underline{\kappa}_n \geq u'_n(0) + h(1) \tag{959}$$

where the inequality follows as $\underline{\kappa}_n \geq 0$. Thus, we have that:

$$u'_m(q) - u'_n(0) \geq h(1) - h(0) > u'_m(q) - u'_n(0) \tag{960}$$

where the first inequality follows by rearranging Equation 959 and the second follows by the definition of non-triviality of preferences. This is a contradiction, thus $x_n, x_m > 0$ in any optimal allocation. □

We now establish an equation relating x_n and x_m that will be useful in the steps to come.

Claim 3. *Let ω have full support in m and n, μ denote a cutoff matching with cutoffs s_m and s_n. Let x_m and x_n denote the measures of agents who are allocated the object at μ. μ is optimal if and only if $u'_m(x_m) + h(s_m) = u'_n(x_n) + h(s_n)$ and $x_n + x_m = q$.*

Proof. By the necessary and sufficient first-order conditions from Theorem 7, we again have that:

$$u'_m(x_m) + h(s_m) - \overline{\kappa}_m + \underline{\kappa}_m = u'_n(x_n) + h(s_n) - \overline{\kappa}_n + \underline{\kappa}_n \tag{961}$$

By Claim 2, we have $x_m, x_n > 0$. Thus, by the complementary slackness conditions

(Equations 951 and 952), we have that $\overline{\kappa}_m = \underline{\kappa}_m = \overline{\kappa}_n = \underline{\kappa}_n = 0$. Thus, we obtain:

$$u'_m(x_m) + h(s_m) = u'_n(x_n) + h(s_n) \tag{962}$$

together with $x_n + x_m = q$, we have characterized the optimal allocation as claimed. $\square$

Continue to let x_m and x_n denote the measures of group m and n agents at the optimal allocation under ω, and s_m and s_n denote the cutoff scores for admission. There are now two cases to consider: (i) $u'_m(x_m)$ and $u'_n(x_n)$ are locally constant. (ii) $u'_m(x_m)$ or $u'_n(x_n)$ are not locally constant. If we are in case (i), we will construct an ω' with a unique optimal allocation x'_m and x'_n where $u'_m(x'_m)$ or $u'_n(x'_n)$ is not locally constant, and then show jointly how we arrive at a contradiction in both cases (i) and (ii).

To this end, suppose that we are in case (i). Let x^*_m and x^*_n denote the measures that are closest to x_m and x_n such that $u'_m(x_m)$ and $u'_n(x_n)$ are not locally constant, *i.e.,*:

$$\begin{aligned} x^*_k = \arg\min_{x'_k} \Big\{ |x_k - x'_k| \Big| u'_k(x_k) = u'_k(x'_k) \text{ and for all } \varepsilon > 0 \\ \text{either } u'_k(x'_k - \varepsilon) > u'_k(x_k) \text{ or } u'_k(x'_k + \varepsilon) > u'_k(x_k) \Big\} \end{aligned} \tag{963}$$

As u'_k is continuous, this minimum is attained and x^*_k is well-defined. Without loss of generality, assume $|x_m - x^*_m| \leq |x_n - x^*_n|$ and define both $\hat{x}_m = x^*_m$ and $\hat{x}_n = q - x^*_m$. We now construct a state ω' such that $\hat{x}$ is optimal:

Claim 4. *Define ω' where $F_m(1) - F_m(s_m) = \hat{x}_m$, $F_n(1) - F_n(s_n) = \hat{x}_n$ and ω' has full support in m and n. The allocation that admits the highest-scoring $\hat{x}_m$ group m agents and the highest-scoring $\hat{x}_n$ group n agents is the unique optimal allocation.*

Proof. By Claim 3, as $\hat{x}_m + \hat{x}_n = q$ by construction, $\hat{x}$ is optimal if and only if Equation 962 holds. To this end, observe that if we admit $\hat{x}$, then the cutoff scores are the same as under x as $F_m(1) - F_m(s_m) = \hat{x}_m$ and $F_n(1) - F_n(s_n) = \hat{x}_n$, by construction. Thus, we have that:

$$\begin{aligned} u'_m(\hat{x}_m) + h(s_m) &= u'_m(x_m) + h(s_m) \\ &= u'_n(x_n) + h(s_n) \\ &= u'_n(\hat{x}_n) + h(s_n) \end{aligned} \tag{964}$$

where the first equality holds by construction as $\hat{x}_m = x_m^*$ and $u_m'(x_m^*) = u_m'(x_m)$, the second equality holds by optimality of x, and the third equality holds as $|x_m - x_m^*| \leq |x_n - x_n^*|$, which implies that $u_n'(\hat{x}_n) = u_n'(x_n^*)$. Thus, Equation 962 holds, and $\hat{x}$ is optimal, as claimed. $\square$

Observe that this construction also applies trivially in case (ii) with $x_m^* = x_m$. Thus, using this construction, we can now study cases (i) and (ii) together. In state ω', to implement this optimal allocation, we must have that $P(s,m) < P(s_n,n)$ for all but a measure zero set of s such that $s < s_m$. We will now construct another state ω'' such that any priority mechanism with this property is suboptimal.

First, suppose that $x_m^* \leq x_m$ and fix some $\varepsilon \in (0, x_m^*)$. Define $\tilde{s}_m$ as solving the following equation:

$$u_m'(\hat{x}_m - \varepsilon) + h(\tilde{s}_m) = u_n'(\hat{x}_n + \varepsilon) + h(s_n) \tag{965}$$

We then have that:

$$\begin{aligned}
\tilde{s}_m &= h^{-1}\left(h(s_n) + u_n'(\hat{x}_n + \varepsilon) - u_m'(\hat{x}_m - \varepsilon)\right) \\
&< h^{-1}\left(h(s_n) + u_n'(\hat{x}_n) - u_m'(\hat{x}_m)\right) \\
&= s_m
\end{aligned} \tag{966}$$

where the first equality rearranges Equation 965 and the second inequality uses the facts that $u_n'(\hat{x}_n) \leq u_n'(\hat{x}_n + \varepsilon)$ and $u_m'(\hat{x}_m) < u_m'(\hat{x}_m - \varepsilon)$. We now construct a state ω'' such that $(\hat{x}_m - \varepsilon, \hat{x}_n + \varepsilon)$ is optimal.

Claim 5. *Define ω'' where $1 - F_m(\tilde{s}_m) = \hat{x}_m - \varepsilon$, $1 - F_n(s_n) = \hat{x}_n + \varepsilon$ with full support in m and n. The allocation that admits the highest-scoring $(\hat{x}_m - \varepsilon, \hat{x}_n + \varepsilon)$ agents is the unique optimal allocation.*

Proof. Following the same steps as Claim 4, and the fact that Equation 965 holds by construction, we have that the claim holds. $\square$

Observe that to implement this optimal allocation a priority mechanism must set $P(s,m) \geq P(s_n,n)$ for all but zero measure $s > \tilde{s}_m$. However, since $\tilde{s}_m < s_m$, this contradicts the optimality condition for state ω' that $P(s,m) < P(s_n,n)$. This is because for all but measure zero $s \in (\tilde{s}_m, s_m)$, which we have established is non-empty, we have that:

$$P(s,m) \geq P(s_n,n) > P(s,m) \tag{967}$$

which is a contradiction. To complete the proof, we need only consider the case that $x^*_m > x_m$. In this case, we can apply essentially the same steps and the result follows. Concretely, instead increasing $\hat{x}_m$ by ε and following the same steps yields the required contradiction.

We have now constructed three states $\omega, \omega', \omega''$ such that no priority mechanism can be optimal in each state when the authority is not risk-neural, completing the proof.

Part (ii): Assume that a quota policy is optimal, we now show that the authority's preferences must be extremely risk-averse. For each group $m \in \mathcal{M}$, let $c_m \in [0, 1]$ and $c_m \neq c_n$ if $m \neq n$. Let ω be such that the scores of agents from each group m are uniformly distributed between $[c_m, c_m + \epsilon]$, where ϵ is chosen to be small so that there is no overlap of these supports and each group has measure q agents. Let m_ω denote the group with the highest c_m at ω. Now, compute the optimal allocation at ω and denote the measure of admitted agents from each group at the optimal allocation by $\{x^*_m(\omega)\}_{m \in \mathcal{M}}$. We first show that under any optimal quota policy, the level of the quotas must be set equal to the optimal allocation for all but the highest-scoring group:

Claim 6. *If a quota policy Q attains the optimal allocation, then for each $m \neq m_\omega$, $Q_m = x^*_m(\omega)$.*

Proof. If $Q_m > x^*_m(\omega)$, then we admit $x_m \geq Q_m > x^*_m(\omega)$, which is suboptimal as there is a unique optimal allocation by Theorem 7. If $Q_m < x^*_m(\omega)$ and $m \neq m_\omega$, then $x_m = Q_m$ as $c_{m_\omega} > c_m + \varepsilon$ and no agent from group m can claim a merit slot. This is suboptimal. Thus, $Q_m = x^*_m(\omega)$ for all $m \neq m_\omega$. □

Next, create ω' by changing the highest-scoring group, *i.e.*, $m_\omega \neq m_{\omega'}$. Let $x^*_{m_\omega}(\omega')$ denote the measure of admitted agents from group m_ω under ω'. Applying Claim 6, If Q attains the optimal allocation, then it must be that $Q_{m_\omega} = x^*_{m_\omega}(\omega')$. Define Q^*_m by $Q^*_m = x^*_m$ for all $m \in \mathcal{M} \setminus \{m_\omega\}$ and $Q^*_{m_\omega} = x^*_{m_\omega}(\omega')$.

Now, we have proved that if Q is an optimal policy, then $Q_m = Q^*_m$ for all $m \in \mathcal{M} \setminus \{m_\omega\}$ and $Q_{m_\omega} = Q^*_{m_\omega}$. We now establish that merit slots must be processed after any positive measure quota slots if the merit slots are of positive measure:

Claim 7. *If there is a quota policy that attains the first-best, Q, then $Q_m = Q^*_m$ and either $\sum_{m \in \mathcal{M}} Q^*_m = q$, i.e., there are no merit slots (thus the order of processing the merit slots does not matter), or $\sum_{m \in \mathcal{M}} Q^*_m < q$ and merit slots are processed after any positive measure quota slots.*

Proof. We have already proved $Q_m = Q_m^*$. If $\sum_{m\in\mathcal{M}} Q_m^* = q$, there are no merit slots and any processing order yields the same result. If $\sum_{m\in\mathcal{M}} Q_m^* < q$, for a contradiction, assume merit slots are processed before quota slots for group m and $Q_m^* > 0$. There are two cases, $m \neq m_\omega$ and $m = m_\omega$. We start with the first case. Note that there is a cutoff s_m for group m with $s_m < c_m + \epsilon$ and all agents from group m who score above s_m are allocated the resource. Create ω'' by taking measure $x_m/2$ of these agents who are allocated the resource and give them scores above c_{m_ω} (the highest-scoring group at ω). The scores of the remaining $x_m/2$ agents are distributed uniformly at $[c_m, c_m + \epsilon]$.

We now observe that the optimal allocations at ω and ω'' are the same. This is because increasing the scores of already admitted agents does not change the preferences of the authority of whom to admit. Moreover, the optimal allocation at ω'' cannot be attained if the quota slots for group m are processed after the merit slots. This follows as, if merit slots are processed before quota slots for group m, a strictly positive measure of them would go to group m agents at ω'' since now they have a measure of agents with the highest scores, which violates optimality.

This proves the claim for $m \neq m_\omega$. To prove the result for $m = m_\omega$, replicate the above steps with ω' where m_ω is not the highest-scoring group. □

We now use these claims to establish that if quotas are first-best optimal, then (u, h) must agree with $(\tilde{u}, \tilde{h})$ on optimal allocations.

Claim 8. *The quota first policy with $Q_m = x_m^{tar}$ maximizes the utility with respect to $\tilde{u}, \tilde{h}$.*

Proof. This is clear since for $\tilde{u}, \tilde{h}$, diversity utility dominates until x_m^{tar} and has no effect after. □

This proves the result since if there exists a first-best optimal quota policy, then it is rationalized by $(\tilde{u}, \tilde{h})$ with $x_m^{tar} = Q_m^*$. Hence, if there is a first-best quota mechanism, the authority is extremely risk-averse. □

E.1.5 Proof of Proposition 18

Let x_m^* denote the measure of group m agents in the optimal allocation, with $x^* = \{x_m^*\}_{m\in\mathcal{M}}$. A priority policy $P(s, m) = h^{-1}(h(s) + u_m'(x_m^*)) = A_m(x_m^*, s)$ implements the same allocation as the optimal adaptive priority mechanism and by Theorem 7, is optimal. A quota mechanism with (Q, D) where $Q_m = x_m^*$ implements x^* for all D, and is therefore optimal.

E.1.6 Proof of Theorem 9

Proof. We first prove the following lemma.

Lemma 21. *Any stable matching is a cutoff matching.*

Proof. Assume that μ is a stable matching. Let $S_{m,c} = \inf_\theta\{s_c(\theta) : m(\theta) = m, \mu(\theta) = c\}$. Since μ satisfies within-group fairness, for all m and $s' > S_{m,c}$, if $m(\theta) = m$ and $s_c(\theta) = s'$, $\mu(\theta) \succeq_\theta c$. Moreover, from part (iv) of the definition of matching, this extends to the case where $s' = S_{m,c}$. Concretely, suppose that $\mu(\theta) \neq c$, $c \succ_\theta \mu(\theta)$ and $s_c(\theta) = S_{m,c}$. Consider a sequence of types $\{\theta_k\}_{k\in\mathbb{N}}$ with common group m and scores $\{s_c(\theta_k)\}_{k\in\mathbb{N}}$ such that $s_c(\theta_k) > S_{m,c}$ for all $k \in \mathbb{N}$ and $s_k(\theta) \to S_{m,c}$. Define the set $\Theta^E = \{\theta \in \Theta : c \succ \mu(\theta)\}$, which must be open by part (iv) of the definition of a matching. We have that $\theta_k \notin \Theta^E$ for all $k \in \mathbb{N}$ but $\lim_{k\to\infty} \theta_k \in \Theta^E$, which contradicts that Θ^E is open. Thus, if μ is stable, then it is also a cutoff matching. $\square$

Therefore, to characterize stable matchings, it is enough to characterize cutoffs that induce a stable matching, which we call *stable cutoffs*.

Definition 22. *A vector S is a market-clearing cutoff if it satisfies the following:*

1. $D_c(S) \leq q_c$ *for all* c.
2. $D_c(S) = q_c$ *if* $S_{m,c} > 0$ *for some* $m \in \mathcal{M}$.

Since an authority can admit different measures of agents from different groups, there is a continuum of cutoffs that clear the market given S_{-c}, as long as $\{(0, \ldots, 0)\}$ is not the only market-clearing cutoff. Let $I(S_{-c})$ denote the set of market-clearing cutoffs. Let $I^*(S_{-c}) \subseteq I(S_{-c})$ denote the unique (by Lemma 20) cutoffs that implement the outcome under APM A^*_c when the authority faces the induced type distribution over the set $\tilde{D}_c(S_{-c})$. Define the map $T_c : [0,1]^{|\mathcal{M}|\times|\mathcal{C}|} \to [0,1]^{|\mathcal{M}|}$ as $T_c(S) = I^*_c(S_{-c})$ with $T : [0,1]^{|\mathcal{M}|\times|\mathcal{C}|} \to [0,1]^{|\mathcal{M}|\times|\mathcal{C}|}$ given by $T = \{T_c\}_{c\in\mathcal{C}}$. We first show that the set of fixed points of T equals the set of stable cutoffs:

Claim 9. *The set of fixed points of T equals the set of stable cutoffs.*

Proof. If S^*, with corresponding matching μ^* (by Lemma 21), is a fixed point of T, then each $c \neq c_0$ admits their most preferred measure q_c agents in $\tilde{D}_c(S^*_{-c})$ (by Theorem 7). Note that any $\hat{\Theta}$ that can block the matching must prefer c to their allocation at μ^* and therefore $\hat{\Theta} \subset \tilde{D}_c(S^*_{-c})$. Then there cannot be a $\hat{\Theta}$ that blocks μ^* at c since c already attains the first-best utility under $\tilde{D}_c(S^*_{-c})$ from the definition of $T_c(S)$ and Theorem 7. Conversely, if S^*, with corresponding matching μ^*, is a not fixed point of T, then there exists c such that $T_c(S^*) \neq S^*_c$. Let $\hat{\Theta}$ denote the set of agents who are not matched to c at μ^* but have scores greater than $T_c(S^*)$, and $\tilde{\Theta}$ denote the set of agents who are matched to c at μ^* but have scores lower than $T_c(S^*)$. From optimality of A^*_c (by Theorem 7), $\hat{\Theta}$ blocks μ^* at c with $\tilde{\Theta}$. $\square$

We now show that T is increasing.

Claim 10. *T is increasing.*

Proof. Fix an arbitrary $c \in \mathcal{C}$ and suppose that $S'_{-c} \geq S_{-c}$. Toward a contradiction suppose that there exists $m \in \mathcal{M}$ such that $t'_{c,m} = T_{c,m}(S') = I^*_c(S'_{-c}) < I^*_c(S_{-c}) = T_{c,m}(S) = t_{c,m}$, *i.e.,* the admissions threshold for group m at authority c goes down. Let f and f' be the induced joint densities of agents over scores at c and groups by the sets $\tilde{D}_c(S_{-c})$ and $\tilde{D}_c(S'_{-c})$, respectively. Let $\{x_{m,c}\}_{m\in\mathcal{M}}$ and $\{x'_{m,c}\}_{m\in\mathcal{M}}$ denote the measure agents who score above $t_{m,c}$ for their group (*i.e.,* admitted under A^*_c) under $\tilde{D}_c(S_{-c})$ and $\tilde{D}_c(S'_{-c})$, respectively. As $S'_{-c} \geq S_{-c}$, we have that $D^c(S_{-c}, S_c) \subseteq D^c(S'_{-c}, S_c)$ for all $S_c \in [0,1]^{|\mathcal{M}|}$. It follows that $f'(\theta_c) \geq f(\theta_c) > 0$ for all $\theta_c = (s_c, m_c) \in [0,1] \times \mathcal{M}$. As $t'_{c,m} < t_{c,m}$, f' has full support, and $f' \geq f$, we have that the measure of admitted group m agents under increases $x'_{c,m} > x_{c,m}$. But as $\sum_{k\in\mathcal{M}} x'_k = \sum_{k\in\mathcal{M}} x_k = q$, we know that there exists an $m' \in \mathcal{M}$ such that $x'_{c,m'} < x_{c,m'}$. It follows that $t'_{c,m'} > t_{c,m'}$, otherwise, if $t'_{c,m'} \leq t_{c,m'}$, then $x'_{c,m'} \geq x_{c,m'}$. But now we have shown the following:

$$h_c(t'_{c,m'})+u'_{m',c}(x'_{c,m'}) > h_c(t_{c,m'})+u'_{m',c}(x_{c,m'}) \geq h_c(t_{c,m})+u'_{m,c}(x_{c,m}) > h_c(t'_{c,m})+u'_{m,c}(x'_{c,m}) \tag{968}$$

where the first inequality follows by $t_{c,m'} < t'_{c,m'}$, $x_{c,m'} > x'_{c,m'}$, concavity of u_m and strictly increasing h_c. The second inequality follows by optimality. This is because the facts that $t_{c,m} > 0$ and $t'_{c,m'} < 1$ imply that $\overline{\kappa}_m = \underline{\kappa}_{m'} = 0$ and so Equation 954 implies that:

$$h_c(t_{c,m'}) + u'_{m',c}(x_{c,m'}) - \overline{\kappa}_{m'} = h_c(t_{c,m}) + u'_{m,c}(x_{c,m}) + \underline{\kappa}_m \tag{969}$$

with $\overline{\kappa}_{m'}, \underline{\kappa}_m \geq 0$. The final inequality follows as $t'_{c,m} < t_{c,m}$ and $x'_{c,m} > x_{c,m}$. But this contradicts the optimality condition for APM (Theorem 7), which implies that $T_c \neq I^*_c$, which is a contradiction. Hence, for all c and $m \in \mathcal{M}$, $T_{c,m}$ is an increasing function. $\square$

As $T : [0,1]^{|\mathcal{M}|\times|\mathcal{C}|} \to [0,1]^{|\mathcal{M}|\times|\mathcal{C}|}$ is monotone and $[0,1]^{|\mathcal{M}|\times|\mathcal{C}|}$ is a lattice under the elementwise order $\geq$, Tarski's fixed point theorem implies that the set of stable matching cutoffs is a non-empty lattice.

Finally, we use the fact that the set of stable cutoffs is a complete lattice to argue that there is a unique cutoff consistent with stability.

Claim 11. *The stable matching cutoffs are unique.*

Proof. Assume that there are multiple stable cutoffs. As the set of stable cutoffs is a lattice, there exists a largest (S^+) and smallest (S^-) stable cutoffs, where $S^+ \geq S^-$, with strict inequality for some $m \in \mathcal{M}$, $c \in \mathcal{C}$ as $S^+ \neq S^-$. But then, as there is full support of agent types and authority c fills the capacity under stable cutoffs S^+, it must exceed its capacity under S^-, which is a contradiction. Hence, we have shown that there exists a unique stable matching cutoff. □

The combination of Lemma 21 and Claim 11 completes the proof. □

E.1.7 Proof of Theorem 10

Proof. If ϕ is equivalent to A_c^*, Claim 9 implies that ϕ is consistent with stability.

We prove consistency with stability implies that ϕ is equivalent to A_c^* by the contrapositive. To this end, suppose that ϕ is not equivalent to A_c^*. It follows that there exists a full-support density $\{\tilde{f}(s_c, m)\}_{s_c \in [0,1], m \in \mathcal{M}}$ such that ϕ yields a different allocation than A_c^* under $\tilde{f}$. The rest of the proof constructs a full-support measure F with unique stable matching μ_F such that $\tilde{f}$ is the induced density of scores and groups of the agents who demand authority c at μ_F. Given such an F, we will have that ϕ cannot be consistent with stability as it yields a different allocation than A_c^*, which itself yields $\mu_F(c)$, the set of students c is matched to in the unique stable matching.

We first define some notation. Given a density f, for any set of types $\check{\Theta} \subseteq \Theta$, we define the marginal density of agents with score $s_c \in [0,1]$ at authority c in group $m \in \mathcal{M}$ as:

$$f_{\mathrm{marg}(\check{\Theta})}(s_c, m) = \int_{\check{\Theta}} \mathbb{I}[s_c(\theta) = s_c, m(\theta) = m] \, \mathrm{d}F(\theta) \tag{970}$$

To construct such an F, we proceed in three steps.

1. Take a full-support density f^0 that satisfies the following two conditions: i) Define $\hat{S}_c \in [0,1]^{|\mathcal{M}|}$ as the cutoff vector that obtains by applying A_c^* to $\tilde{f}$.[2] We assume that f^0 is such that authority c's cutoff vector that is consistent with the unique stable matching, μ_{F_0}, coincides with $\hat{S}_c$; ii) for all $m \in \mathcal{M}$ and

[2]Which exists as any monotone APM admits a cutoff structure (Lemma 21) and the optimal APM is monotone (Theorem 7).

$s_c < \hat{S}_{m,c}$, $f^0_{\text{marg}(\Theta)}(s,m) < \tilde{f}(s,m)$; and iii) all authorities have strictly positive cutoffs for all groups at the unique stable matching.

2. We transform f^0 into a new density f^1 that differs from f^0 on the set of types that is matched with c under μ_{F_0}, which we call Θ_c.[3] We define the scaling factor $\iota^1(s_c, m)$ as:
$$\iota^1(s_c, m) = \frac{\tilde{f}(s_c, m)}{f^0_{\text{marg}(\Theta_c)}(s_c, m)} \tag{971}$$
Given this scaling factor, we define:
$$f^1(\theta) = \begin{cases} f^0(\theta)\iota^1(s_c(\theta), m(\theta)) & \text{if} \quad \theta \in \Theta_c, \\ f^0(\theta) & \text{otherwise.} \end{cases} \tag{972}$$
Observe that this changes the scores of the types who are allocated to c under μ_{F^0} but it does not change their total measure, their composition, or their scores at any other authority. Thus, the unique stable matching under f^1, μ_{F^1}, coincides with μ_{F^0}. Moreover, by assumption i) in the construction of f^0 in step 1, we have that $f^1_{\text{marg}(\Theta_c)}(s_c, m) = \tilde{f}(s_c, m)$ for all m and $s_c \geq \hat{S}_{m,c}$.

3. We transform f^1 into a new density f^2 that differs on the set of unmatched agents under f^0 (and also therefore f^1 by step 2), $\tilde{\Theta}$.[4] Define the set of types who strictly prefer c to their assignment under μ_{F^0} (and also therefore μ_{F^1} by step 2), $\hat{\Theta}_c$.[5] We define a new scaling factor $\iota^2(s_c, m)$ as:
$$\iota^2(s_c, m) = \frac{\tilde{f}(s_c, m) - f_{\text{marg}(\hat{\Theta}_c)}(s_c, m)}{f_{\text{marg}(\tilde{\Theta})}(s_c, m)} \tag{973}$$
which is strictly positive by assumption ii) of step 1. We then define f^2 as:
$$f^2(\theta) = \begin{cases} f^1(\theta)(1 + \iota^2(s_c(\theta), m(\theta))) & \text{if} \quad \theta \in \tilde{\Theta}, \\ f^1(\theta) & \text{otherwise.} \end{cases} \tag{974}$$
By construction, $f^2_{\text{marg}(\hat{\Theta}_c)}(s_c, m) = \tilde{f}(s_c, m)$ for all m and $s_c < \hat{S}_{m,c}$. Moreover, $\mu_{F^2} = \mu_{F^1} = \mu_{F^0}$ as all $\theta \in \tilde{\Theta}$ remain unmatched.

[3] Formally, $\Theta_c = \{\theta : \theta \in D_c(\mu_{F^0}), s_c(\theta) \geq \hat{S}_{m(\theta),c}\}$.

[4] Formally, $\tilde{\Theta} = \{\theta : \theta \in D_c(\mu_{F^1}), s_c(\theta) < \hat{S}_{m(\theta),c}, s_{c'}(\theta) < S^{\mu_{F^1}}_{m(\theta),c'}$ for all $c' \neq c\}$, where $S^{\mu_{F^1}}_{m,c'}$ denotes the group m cutoff at school c' at the stable matching μ_{F^1}, which is strictly positive by assumption iii) of step 2.

[5] Formally, $\hat{\Theta}_c = \{\theta : \theta \in D_c(\mu_{F^1}), s_c(\theta) < \hat{S}_{m(\theta),c}\}$.

We have now constructed a full-support density f^2 with unique stable matching μ_{F^2} (by Theorem 9) such that the density over $D_c(\mu_{F^2})$ coincides with $\tilde{f}$. Moreover, by Claim 9, A_c^* selects $\mu_{F^2}(c)$ from $D_c(\mu_{F^2})$. As ϕ selects a different allocation from $D_c(\mu_{F^2})$ (as it has density of types $\tilde{f}$), it is inconsistent with stability. □

E.1.8 Proof of Theorem 11

Proof. We prove that APM A_c^* implements a dominant strategy for all authorities in all stages by backward induction. Consider the terminal time $t = |\mathcal{C}| - 1$. Some measure of agents λ applies to the authority. Regardless of the measure λ, by Theorem 7 we have that the APM A_c^* is first-best optimal (to see this more concretely, simply index λ by an arbitrary $\omega \in \Omega$ and apply Theorem 7). Thus, A_c^* is dominant. Moreover, from Theorem 7, any strategy that differs from A_c^* on a strictly positive measure set cannot be optimal. Thus any dominant strategy implements essentially the same allocation as A_c^*. Consider now any time $t < |\mathcal{C}| - 1$, precisely the same argument applies and A_c^* is (essentially uniquely) dominant. □

E.1.9 Proof of Proposition 19

Proof. We first prove the following claim.

Claim 12. *μ_{Σ^*} is (almost surely) a deterministic allocation that corresponds to a cutoff matching μ^*.*

Proof. Since there is a continuum of agents, under any Σ^*, with probability 1, any authority c faces a given set of agents who apply Θ_c^{A,Σ^*} with induced measure $\lambda_c^{\Sigma^*}$. As c uses APM A_c^*, with probability 1, any agent θ is admitted to an authority if and only if $s_c(\theta) \geq S_{m,c}^{\Sigma^*}$, where $S_{m,c}^{\Sigma^*}$ denotes the cutoffs when APM A_c^* is applied to agent measure $\lambda_c^{\Sigma^*}$. Since the agents have strict preferences, in any equilibrium, each agent applies to the $\succeq_\theta$ −maximal authority in $\{c : s_c(\theta) \geq S_{m,c}^{\Sigma^*}\}$, and is admitted, which establishes that μ_{Σ^*} is (almost surely) deterministic allocation that corresponds to a cutoff matching with cutoffs $S_{m,c}^{\Sigma^*}$. □

We now establish that μ_{Σ^*} is the unique stable matching of the economy.

Claim 13. *μ^* is the unique stable matching of this economy.*

Proof. For a contradiction, assume μ_{Σ^*} is not stable. Let S denote the unique cutoffs associated with μ_{Σ^*}. Since μ_{Σ^*} is not stable, by Claim 9, S is not a fixed point of T. Let $t_c = T_c(S)$. Since S is not a fixed point of T, there exists $m \in \mathcal{M}$ and $c \in \mathcal{C}$ such that $t_{m,c} \neq S_{m,c}$. Moreover, let $\{x_{m,c}^t\}_{m \in \mathcal{M}}$ and $\{x_{m,c}^s\}_{m \in \mathcal{M}}$ denote the measure of agents in $\tilde{D}_c(S_{-c})$ who are above the admission thresholds for authority c under t_c and S_c. As in Claim 10, note that if there exists m, c such that $t_{m,c} > S_{m,c}$, then from full support, we have that $x^s_{m,c} > x^t_{m,c}$. Since the authority fills its capacity in both cases, there must exists m' such that $x^t_{m',c} > x^s_{m',c}$ which is only possible if $S_{m,c} > t_{m,c}$. By an identical argument, if there is m, c such that $t_{m,c} < S_{m,c}$, then there exists m' $S_{m',c} < t_{m',c}$. Therefore, whenever $t_{m,c} \neq S_{m,c}$, there exists c and m, m' such that $t_{m,c} > S_{m,c}$ and $S_{m',c} > t_{m',c}$. But now we have shown the following:

$$h_c(S_{c,m'})+u_{m'}(x^s_{c,m'}) > h_c(t_{c,m'})+u_{m'}(x^t_{c,m'}) \geq h_c(t_{c,m})+u_m(x^t_{c,m}) > h_c(S_{c,m})+u_m(x^s_{c,m})$$

where the first inequality follows by $t_{c,m'} < S_{c,m'}$, $x^s_{c,m'} < x^t_{c,m'}$, and concavity of u_m. The second inequality follows by optimality. This is because the facts that $t_{c,m} > 0$ and $t'_{c,m'} < 1$ imply that the Lagrange multipliers in the proof of Theorem 7 $\overline{\kappa}_m = \underline{\kappa}_{m'} = 0$. The final inequality follows since $t_{m,c} > S_{m,c}$ and $x^t_{m,c} < x^s_{m,c}$. However, this is a contradiction since $h_c(S_{c,m'}) + u_{m'}(x^s_{c,m'}) > h_c(S_{c,m}) + u_m(x^s_{c,m})$ implies that there exists $\varepsilon > 0$, an agent θ with score $s_c(\theta) = S_{c,m'} - \epsilon$ and type $m(\theta) = m$ has higher score under A^* than the agent θ' with score $s_c(\theta') = S_{c,m}$ and type $m(\theta') = m$. Since θ' is admitted to c, θ would be if it applied to c. Moreover, from full support, there is such θ whose top choice is c and the strategy of this agent is not a best response, which is a contradiction. □

The combination of Claims 12 and 13 completes the proof. □

E.1.10 Proof of Proposition 20

Proof. We prove the result by explicitly constructing an economy in which the optimal APMs lead to inefficiency. There are two authorities, c and c', both with capacity 1/2 and two groups of agents, m and m'. Both agent groups have a measure of 1 and their scores are uniformly distributed in $[1/2, 1]$. Authorities' utility functions are given by

$$\xi_c(\bar{s}_h, x) \equiv \bar{s}_h + \frac{1}{4}\sqrt{x_m} + \frac{1}{8}\sqrt{x_{m'}} \tag{975}$$

$$\xi_{c'}(\bar{s}_h, x) \equiv \bar{s}_h + \frac{1}{4}\sqrt{x_{m'}} + \frac{1}{8}\sqrt{x_m} \tag{976}$$

with $h(x) \equiv x$ while all agents of type m prefer authority c' to c while all agents of type m' prefer authority c to c'.[6]

[6] This assumption on the preferences and the distribution of scores violate our full support assumption, but adding an arbitrarily small full support density to all types makes arbitrarily small changes in the utility under the stable matching and optimal allocation but complicates the calcu-

We will now derive the stable outcome of this economy, which is (up to measure zero transformation) the unique outcome implemented when the authorities use the optimal APM. Let x^c_m denote the measure of type m agents at authority c. First, note that higher-scoring agents from the same group go to the more preferred authority. To see why this is true, note that if $m(\theta) = m(\theta') = m$, $s(\theta) > s(\theta')$ and $\mu(\theta) = c$ while $\mu(\theta') = c'$, c and θ would violate within group fairness since θ has higher priority at c than θ' regardless of the allocation. As a result, in any stable allocation μ, the highest-scoring $x^c_{m'}$ type m' agents are assigned to c and the next highest-scoring $x^{c'}_{m'}$ agents are assigned to authority c', while rest of the type m' agents are not assigned to any authority. The allocation for type m agents is analogous. Moreover, since $q = 1/2$ for both authorities, $x^{c'}_{m'}$=1/2-$x^c_{m'}$ and x^c_m=1/2-$x^{c'}_m$ and the allocation is completely determined by the measures $x^c_{m'}$ and $x^{c'}_m$.

Next, note that at μ, the adaptive priority of the lowest-scoring type m and m' agents must be equal at both authorities. To see why this is true, take authority c without loss of generality. Let $s^c_{m'} = 1 - x^c_{m'}$ and $s^c_m = 1 - x^c_{m'} - x^c_m$ denote the scores of the lowest-scoring type m and m' agents and A_m denote the optimal APM. For a contradiction, assume $A_m(x^c_m, s^c_m) > A_{m'}(x^c_{m'}, s^c_{m'})$. Since agents of type m' with scores lower than s^c_m are unassigned at μ, for small enough ϵ, a type m agent with score $s^c_m - \epsilon$ and authority c blocks the matching. Similarly, assume $A_m(x^c_m, s^c_m) < A_{m'}(x^c_{m'}, s^c_{m'})$. Since agents of type m' with scores lower than $s^{c'}_m$ are assigned to authority c or unmatched at μ, a type m' agent with score $s^c_{m'} - \epsilon$ and authority c blocks the matching μ. Thus, the following equations must be satisfied:

$$A_m(x^c_m, s^c_m) = A_{m'}(x^c_{m'}, s^c_{m'}) \text{ and } A_m(x^{c'}_m, s^{c'}_m) = A_{m'}(x^{c'}_{m'}, s^{c'}_{m'}) \tag{977}$$

As the optimal APM in this setting is given by:

$$A^*_{\hat{m},\hat{c}}(y_{\hat{m}}, s) \equiv s + u'_{\hat{m},\hat{c}}(y_{\hat{m}}) \tag{978}$$

for all $\hat{m} \in \{m, m'\}$ and $\hat{c} \in \{c, c'\}$, we have that:

$$1 - x^c_{m'} + \frac{1}{8}\frac{1}{\sqrt{x^c_{m'}}} = 1 - x^c_{m'} - x^c_m + \frac{1}{4}\frac{1}{\sqrt{1/2 - x^c_{m'}}} \tag{979}$$

and:

$$1 - x^{c'}_m + \frac{1}{8}\frac{1}{\sqrt{x^{c'}_m}} = 1 - x^{c'}_m - x^{c'}_{m'} + \frac{1}{4}\frac{1}{\sqrt{1/2 - x^{c'}_m}} \tag{980}$$

lation, so we omit it for expositional clarity.

These equations are identical up to relabelling and so $x_{m'}^{c} = x_{m'}^{c'} = x^*$ for some x^*. Thus, we need to find the solution to the following single equation to characterize the allocation:

$$1 - x^* + \frac{1}{8}\frac{1}{\sqrt{x^*}} = \frac{1}{2} + \frac{1}{4}\frac{1}{\sqrt{1/2 - x^*}} \tag{981}$$

Observe that this equation can be rewritten as the fixed point equation:

$$x^* = \frac{1}{2} + \frac{1}{8}\frac{1}{\sqrt{x^*}} - \frac{1}{4}\frac{1}{\sqrt{1/2 - x^*}} \tag{982}$$

We observe that the RHS satisfies the following properties: (i) $\lim_{x^* \to 0} \mathrm{RHS}(x^*) = \infty$, (ii) $\lim_{x^* \to \frac{1}{2}} \mathrm{RHS}(x^*) = -\infty$, and (iii) $\mathrm{RHS}'(x^*) < 0$ for all $x^* \in (0, \frac{1}{2})$. Thus, there exists a unique solution. Moreover, we can guess-and-verify that this solution is $x^* = \frac{1}{4}$.

In summary, if both authorities use the optimal APM, then the outcome is

$$\mu(\theta) = \begin{cases} c & \text{if } m(\theta) = m, s(\theta) \in [1/2, 3/4) \text{ or } m(\theta) = m', s(\theta) \in [3/4, 1] \\ c' & \text{if } m(\theta) = m', s(\theta) \in [1/2, 3/4) \text{ or } m(\theta) = m, s(\theta) \in [3/4, 1] \\ \theta & \text{otherwise} \end{cases} \tag{983}$$

In this outcome, both authorities have an average score of 3/4 and admit measure 1/4 agents from both groups, giving them a utility of 15/16. Thus, total utilitarian welfare is 15/8 under the decentralized outcome.

We now show that this does not attain the efficient benchmark. A necessary condition for the (utilitarian) efficient outcome is that for c:

$$\frac{1}{4}\frac{1}{\sqrt{x_m^c}} = \frac{1}{8}\frac{1}{\sqrt{1/2 - x_m^c}} \tag{984}$$

and for c':

$$\frac{1}{4}\frac{1}{\sqrt{x_{m'}^{c'}}} = \frac{1}{8}\frac{1}{\sqrt{1/2 - x_{m'}^{c'}}} \tag{985}$$

This implies that $x_m^c = x_{m'}^{c'} = 4/10$ and $x_{m'}^c = x_m^{c'} = 1/10$. In this case, the same set of agents is admitted overall, so the score contribution to utility remains 3/4 on average across the authorities. Total utilitarian welfare is now:

$$3/2 + 1/2 \times \sqrt{4/10} + 1/4 \times \sqrt{1/10} \approx 1.895 > 1.875 = 15/16 \tag{986}$$

Completing the proof. $\square$

E.1.11 Proof of Theorem 12

Proof. First, we define a fictitious *composite authority* with utility function defined over vectors of total scores $\bar{s}_h = \{\bar{s}_h^c\}_{c\in\mathcal{C}}$, and aggregate allocation to each group $x = \{x_m\}_{m\in\mathcal{M}}$. To do this, we define:

$$\begin{aligned}\tilde{u}(\{x_m\}_{m\in\mathcal{M}}) = & \max_{\{x_{m,c}\}_{c\in\mathcal{C}}} \sum_{c\in\mathcal{C}} \sum_{m\in\mathcal{M}} u_{m,c}(x_{m,c}) \\ & \text{s.t.} \sum_{c\in\mathcal{C}} x_{m,c} \le x_m, \sum_{m\in\mathcal{M}} x_{m,c} \le q_c, \forall m \in \mathcal{M}, c \in \mathcal{C}\end{aligned} \tag{987}$$

and $\tilde{\bar{s}}_h = \sum_{c\in\mathcal{C}} \bar{s}_h^c$. We write the utility function of this composite authority as

$$\tilde{\xi}\left(\tilde{\bar{s}}_h, x\right) = \tilde{\bar{s}}_h + \tilde{u}(x) \tag{988}$$

We first establish that $\tilde{u}$ satisfies the properties necessary to invoke Proposition 50, which establishes the optimality of the claimed APM for the fictitious authority.

Claim 14. *The function $\tilde{u}$ is concave and partially differentiable in each argument.*

Proof. First, we establish concavity. That is, for all $\lambda \in [0,1]$ and $x, x' \in \mathbb{R}_+^{|\mathcal{M}|}$, we have that $\tilde{u}(\lambda x' + (1-\lambda)x) \ge \lambda\tilde{u}(x') + (1-\lambda)\tilde{u}(x)$. Let $\{x^*_{m,c}\}_{m\in\mathcal{M},c\in\mathcal{C}}$ and $\{x^{*'}_{m,c}\}_{m\in\mathcal{M},c\in\mathcal{C}}$ correspond to optimal values under x and x'. Under $\tilde{x} = \lambda x' + (1-\lambda)x$, we have that $\lambda x^{*'}_{m,c} + (1-\lambda)x^*_{m,c}$ is feasible for all $m \in \mathcal{M}$ and $c \in \mathcal{C}$. Thus, we have that:

$$\begin{aligned}\tilde{u}(\tilde{x}) &\ge \sum_{m\in\mathcal{M}} \sum_{c\in\mathcal{C}} u_{m,c}(\lambda x^{*'}_{m,c} + (1-\lambda)x^*_{m,c}) \\ &\ge \sum_{m\in\mathcal{M}} \sum_{c\in\mathcal{C}} \lambda u_{m,c}(x^{*'}_{m,c}) + (1-\lambda)u_{m,c}(x^*_{m,c}) \\ &= \lambda\tilde{u}(x') + (1-\lambda)\tilde{u}(x)\end{aligned} \tag{989}$$

where the second inequality is by concavity of $u_{m,c}$ for all $m \in \mathcal{M}, c \in \mathcal{C}$.

Second, we establish partial differentiability in each argument. That is, for all $x \in \mathbb{R}_{++}^{|\mathcal{M}|}$, $\frac{\partial}{\partial x_m}\tilde{u}(x) = \tilde{u}^{(m)}(x)$ exists. This follows by Corollary 5 in Milgrom and Segal (2002). Concretely, the domain of optimization can be taken to be a compact and convex subset of a normed vector space – a sufficiently large cube in $\mathbb{R}_+^{|\mathcal{M}|\times|\mathcal{C}|}$ equipped with the Euclidean norm, for example. The objective function does not depend on x, and constraints are linear in x (and therefore both continuous and continuously differentiable). Moreover, as $x \gg 0$, there exists a $\{x_{m,c}\}$ that satisfies all constraints with strict inequality.

□

It follows that the objective function of the composite authority satisfies Assumption 18, and so Proposition 50 implies that the non-separable APM $\tilde{A}_m(y,s) = h^{-1}\left(h(s)+\tilde{u}^{(m)}(y)\right)$ uniquely implements the first-best optimal allocation for the composite authority.

It remains to establish that the quota functions implement the optimal allocation $\{x_{m,c}\}$ conditional on $\{x_m\}$. Let λ_m be the Lagrange multiplier on the x_m constraint, γ_c be the Lagrange multiplier on the q_c constraint and $\underline{\kappa}_{m,c}$ be the Lagrange multiplier on the positivity constraint. Under our maintained Inada condition, we have that $\underline{\kappa}_{m,c} = 0$. Moreover, by Corollary 5 in Milgrom and Segal (2002), we have that $\tilde{u}^{(m)}(x) = \lambda_m$, $\tilde{u}_{q_c}(x) = \gamma_c$, and $u'_{m,c}(x^*_{m,c}) = \lambda_m + \gamma_c - \underline{\kappa}_{m,c}$. Hence, we obtain that:

$$x^*_{m,c} = \left(u'_{m,c}\right)^{-1}\left(\tilde{u}^{(m)}(x)+\tilde{u}_{q_c}(x)\right) \tag{990}$$

Thus, the following profile of quota functions implements the optimal cross-sectional allocation:

$$Q_{m,c}(x) = \left(u'_{m,c}\right)^{-1}\left(\tilde{u}^{(m)}(x)+\tilde{u}_{q_c}(x)\right) \tag{991}$$

Completing the proof. □

E.2 Additional Results for the Example (Section 5.2)

E.2.1 Formal Equivalence Between Prices *vs.* Quantities and Priorities *vs.* Quotas

The structure of the comparative advantage of priorities over quotas from Section 5.2 hints at a more formal relationship between our analysis of affirmative action policies and Weitzman's analysis of price and quantity regulation. In Weitzman's model, there is a single firm producing a quantity of a single good $x \in \mathbb{R}$ with production costs $C(x,\zeta)$ and benefits $B(x,\zeta')$:

$$\begin{aligned} C(x,\zeta) &= a_0(\zeta) + (C' + a_1(\zeta))(x - \hat{x}) + \frac{C''}{2}(x-\hat{x})^2 \\ B(x,\zeta') &= b_0(\zeta') + (B' + b_1(\zeta'))(x - \hat{x}) + \frac{B''}{2}(x-\hat{x})^2 \end{aligned} \tag{992}$$

where $B', C', C'' > 0$, $B'' < 0$, and ζ and ζ' are random variables. The regulator can either set a price that the firm must charge (after which the firm chooses its optimal production quantity) or mandate the production of a given quantity. The comparative advantage of prices over quantities Δ^{Weitzman} is then defined as the difference between expected benefits net of costs under the optimal price regime minus the corresponding net benefits under the optimal quantity regime. This comparative advantage is given by:

$$\Delta^{\text{Weitzman}} = \frac{C''^{-1}}{2}\left(1 + C''^{-1}B''\right)\operatorname{Var}[a_1(\zeta)] \tag{993}$$

The intuition for this formula is that when benefits are more curved than costs $|B''| > C''$, reducing variability in production is more valuable than the gain of having producers minimize costs. Thus, quantities are preferred. On the other hand, when costs are more curved than benefits, prices are preferred as there is greater production when producers have the lowest marginal costs of production.

These trade-offs are, in a certain sense, formally analogous to those that we have highlighted between priorities and quotas. In particular, under the mapping $C''^{-1} \mapsto \kappa$, $B'' \mapsto -\gamma\beta$, $\operatorname{Var}[\omega] \mapsto \operatorname{Var}[a_1(\zeta)]$, we have that $\Delta^{\text{Weitzman}} = \Delta$. The intuition for this is that C''^{-1} in the Weitzman framework determines how sensitive production is to changes in marginal cost, while κ in our framework determines how sensitive the admitted measure of minority students is to the relative scores. Moreover, B'' corresponds to curvature in the benefits of production while $\gamma\beta$ corresponds to curvature in the benefits of admitting more minority students. Finally, $\operatorname{Var}[a_1(\zeta)]$

corresponds to the authority's uncertainty in the level of marginal costs of production while Var[ω] corresponds to the authority's uncertainty regarding the marginal cost of admitting more minority students in terms of lost total score. Thus, the positive selection effect whereby priorities admit more minority students in the states of the world where they score highest is directly analogous to the effect that price regulation gives rise to the greatest production in states where the firm's marginal cost is lowest. Moreover, the guarantee effect whereby quotas prevent variation in the measure of admitted minority students across states of the world is analogous to the ability of quantity regulation to stabilize the level of production. Importantly, our results therefore allow one to apply established price-theoretic intuition for the benefits of price *vs* quantity choice to matching markets without an explicit price mechanism.

E.2.2 Beyond Affirmative Action: Medical Resource Allocation

The lessons of this paper apply not only to affirmative action in academic admissions, but also more broadly to settings in which centralized authorities must allocate resources to various groups. One prominent such context is the allocation of medical resources during the Covid-19 pandemic. An important issue faced by hospitals is how to prioritize health workers (doctors, nurses and other staff) in the receipt of scarce medical resources: hospitals wish to both treat patients according to clinical need and ensure the health of the frontline workers needed to fight the pandemic. To map this setting to our example, suppose that the score s is an index of clinical need for a scarce medical resource available in amount q, the measure of frontline health workers is κ, and ω indexes the level of clinical need in the patients currently (or soon to be) treated by the hospital, which is unknown. The risk aversion of the authority $\gamma\beta$ corresponds to both a fear of not treating sufficiently many frontline workers and excluding too many clinically needy members of the general population.

In practice, both priority systems and quota policies have been used, as detailed extensively by Pathak, Sönmez, Unver, and Yenmez (2021).[7] The primary concern that has been voiced is that if a priority system is used, some groups (or characteristics) may be completely shut out of allocation of the scarce resource and that this is unethical, so quotas should be preferred. Our framework can be used to understand this argument: if there is an unusually high draw of ω, a priority system would lead to the allocation of very few resources to frontline workers, and vice-versa. Our

[7]Some other papers that study the allocation of scarce medical resources are Akbarpour, Budish, Dworczak, and Kominers (2021), Grigoryan (2021) and Dur, Thayer, and Phan (2022).

Proposition 16 implies that if the authority is very averse to such outcomes ($\gamma\beta$ is high), quotas will be preferred and for exactly the reasons suggested. However, we also highlight a fundamental benefit of priority systems in inducing positive selection in allocation: when ω is high, it is beneficial that fewer resources go to the less sick medical workers and more to the relatively sicker general population. More generally, we argue that an adaptive priority mechanism that awards frontline workers a score subsidy that depends on the number of more clinically needy frontline workers could further improve outcomes.

An important additional consideration in this context arises if the hospital or authority must select a regime (priorities or quotas) before it understands the clinical need of its frontline workers κ, after which it can decide exactly how to prioritize these workers or set quotas, but before ultimate demand for medical resources ω is known. It follows from Proposition 16 that the comparative advantage of priorities over quotas is:

$$\mathbb{E}[\Delta] = \frac{1}{2}\left(\mathbb{E}[\kappa] - (\mathrm{Var}[\kappa] + \mathbb{E}^2[\kappa])\gamma\beta\right)\mathrm{Var}[\omega] \tag{994}$$

Thus, an increase in uncertainty $\mathrm{Var}[\kappa]$ regarding the need of frontline workers leads to a greater preference for quotas. This highlights a further advantage of quotas in settings where a clinical framework must be adopted in the face of uncertainty regarding the clinical needs of frontline workers, as was the case at the onset of the Covid-19 pandemic.

E.2.3 Optimal Precedence Orders

Thus far we have modelled quotas by first allocating minority students to quota slots and then allocating all remaining students according to the underlying score. However, we could have instead allocated $q - Q$ places to all agents according to the underlying score and then allocated the remaining Q places to minority students. The order in which quotas are processed is called the *precedence order* in the matching literature and their importance for driving outcomes has been the subject of a growing literature (see *e.g.,* Dur, Kominers, Pathak, and Sönmez, 2018; Dur, Pathak, and Sönmez, 2020; Pathak, Rees-Jones, and Sönmez, 2020). Our framework can be used to understand which precedence order is optimal, a question that has not yet been addressed.

In this example, the same factors that determine whether one should prefer priorities or quotas determine whether one should prefer processing quotas second or first. By virtue of uniformity of scores, it can be shown in the relevant parameter range that a priority subsidy of α is equivalent to a quota policy of $\kappa\alpha$ when quotas are processed second. Thus, the comparative advantage of priorities over quotas is exactly equal to the comparative advantage of processing quotas second over first. The intuition is analogous: processing quotas second allows for positive selection while processing quotas first fixes the number of admitted minority students. Thus, on the one hand, when the authority is more risk-averse, they should process quota slots first to reduce the variability in the admitted measure of minority students. On the other hand, when they are less risk-averse, they should process quotas second to take advantage of the positive selection effect such policies induce.

Corollary 17. *The optimal quota-second policy achieves the same value as the optimal priority policy; quota-second policies are preferred to quota-first policies if and only if* $\frac{1}{\kappa} \geq \gamma\beta$.

Proof. We show that a quota-second policy Q is equivalent to a priority subsidy of $\alpha(Q) = \frac{Q}{\kappa}$. A quota-second policy admits the highest-scoring $x = \kappa(1-\omega) + Q$ minority students, floored by zero and capped by $\min\{\kappa, q\}$. A priority policy $\alpha(Q) = \frac{Q}{\kappa}$ admits the highest-scoring $x = \kappa(1 + \alpha(Q) - \omega) = \kappa(1-\omega) + Q$ minority students, floored by zero and capped by $\min\{\kappa, q\}$. Thus, state-by-state, quota-second policy Q and priority subsidy $\alpha(Q) = \frac{Q}{\kappa}$ yield the same allocation. The claims then follow from Proposition 16. $\square$

We emphasize that this equivalence is a result of the uniform distribution of scores and merely illustrates the similarity between priority policies and processing quotas second. This result does not hold in the more general model we study in the remainder of the paper. Indeed, in Theorem 8, we show that for any quota policy to be optimal in the presence of uncertainty, it must process quotas first.

E.3 Extension of the Main Results to Discrete Economies

In this Appendix, we extend the results in the main text to discrete economies and thereby establish that our analysis generalizes from the continuum framework. Concretely, we show that appropriate analogs of Theorems 7, 8, 11 and Proposition 18 carry over to discrete economies. As discrete economies do not necessarily admit a unique stable matching, Theorems 9 and 10 do not generalize. As Theorem 12 relies on convex optimization techniques, it also does not generalize as written to a discrete setting.

E.3.1 Primitives

An authority has q resources to allocate. At each state ω, the economy the authority faces corresponds to agents $\Theta^\omega = \{\theta_1, \ldots, \theta_{N(\omega)}\}$ where $q \leq |N(\omega)|$. As in the continuum case, $\theta \in [0,1] \times \mathcal{M}$ denotes the type of an agent who has score s and belongs to group m. We denote the score and group of any type θ by $s(\theta)$ and $m(\theta)$, respectively. For simplicity, we assume that no two agents have the same score at any ω, formally, if $\{\theta, \theta'\} \subseteq \Theta^\omega$, then $s(\theta) \neq s(\theta')$.

An allocation $\mu : \Theta \to \{0,1\}$ specifies for any type $\theta \in \Theta$ whether they are assigned to the resource. The set of possible allocations is $\mathcal{U}$ and Ω is the set of all possible economies. An allocation is feasible if it allocates no more than measure q of the resource. A mechanism is a function $\phi : \Omega \to \mathcal{U}$ that returns a feasible allocation for any possible Θ^ω.

The authority believes ω has distribution $\Lambda \in \Delta(\Omega)$. $x(\mu, \omega) = \{x_m(\mu, \omega)\}_{m \in \mathcal{M}}$ denotes the number of agents of each group allocated the resource at matching μ, while $\bar{s}_h(\mu, \omega) = \sum_{\theta \in \Theta^\omega} \mu(\theta) h(s(\theta))$ denotes the utility the authority derives from scores at μ. The preferences of the authority are given by $\xi : \mathbb{R}^{|\mathcal{M}|+1} \to \mathbb{R}$:

$$\xi\left(\bar{s}_h, x\right) \equiv \bar{s}_h + \sum_{m \in \mathcal{M}} u_m(x_m) \tag{995}$$

where h is continuous and strictly increasing and u_m is concave for all $m \in \mathcal{M}$.

E.3.2 Optimal Mechanisms in Discrete Economies

We adapt our definition of the Adaptive Priority Mechanisms to the discrete setting. An *adaptive priority policy* $A = \{A_m\}_{m \in \mathcal{M}}$, where $A_m : \mathbb{R} \times [0,1] \to \mathbb{R}$. The adaptive priority policy assigns priority $A_m(y_m, s)$ to an agent with score s in group m when y_m of agents of the same group is allocated the object. Given an adaptive priority policy, an APM implements allocations in the following way:

Definition 23 (Adaptive Priority Mechanism)**.** *An adaptive priority mechanism, induced by an adaptive priority A, implements an allocation μ in state ω if the following are satisfied:*

1. *Allocations are in order of priorities: $\mu(\theta) = 1$ if and only if*

   (i) for all θ' with $m(\theta') \neq m(\theta)$ and $\mu(\theta') = 0$,

$$A_{m(\theta)}(x_{m(\theta)}(\mu, \omega), s(\theta)) \geq A_{m(\theta')}(x_{m(\theta')}(\mu, \omega) + 1, s(\theta')) \tag{996}$$

   (ii) for all θ' with $m(\theta') = m(\theta)$ and $\mu(\theta') = 0$, $s(\theta) > s(\theta')$

2. *The resource is fully allocated:*

$$\sum_{m \in \mathcal{M}} x_m(\mu, \omega) = q \tag{997}$$

Definition 23 makes two modifications relative to the continuum model. First, the measures of agents from each group are replaced by the number of agents from each group. Second, when $m(\theta) \neq m(\theta')$, the adaptive priority of θ' is now evaluated in the case where an extra agent from $m(\theta')$ is assigned the resource.[8] Unlike the continuum case, it is possible for a monotone APM to implement two different allocations, since it can assign the same priority to two different agents, which could happen only for a zero-measure set of agents in the continuum model.

Define $A^*_m(y_m, s) \equiv h(s) + u_m(y_m) - u_m(y_m - 1)$, which will turn out to be the optimal APM. We first show that A^* is monotone, and all allocations that A^* implements give the authority the same utility.

Lemma 22. *A^* is monotone. Moreover, if A^* implements μ and $\mu' \neq \mu$ in state ω, then $\xi(\mu, \omega) = \xi(\mu', \omega)$.*

Proof. Monotonicity is immediate from the definition of A^* and concavity of u_m. Assume that A^* implements two different allocations, μ and μ' at ω. Let x_l and x'_l denote the number of group $l \in \mathcal{M}$ agents assigned the resource at μ and μ'. Since A^* is monotone and $\mu \neq \mu'$, there are m and n such that $x_m > x'_m$ and $x'_n > x_n$. Let $\tilde{\theta}_l$ and $\tilde{\theta}'_l$ denote the lowest-scoring type l agent assigned the resource at μ and μ', respectively. Similarly, let $\hat{\theta}_l$ and $\hat{\theta}'_l$ denote the highest-scoring type l agents who is not assigned the resource at μ and μ', respectively. Let $\tilde{\mu}$ denote the matching given

[8]This was not the case in the continuum model since all types of agents have measure 0 and therefore replacing θ with θ' has no effect the evaluation of diversity.

by: $\tilde{\mu}(\theta) = \mu(\theta)$ if $\theta \notin \{\tilde{\theta}_m, \hat{\theta}'_n\}$, $\tilde{\mu}(\tilde{\theta}_m) = 0$ while $\mu(\hat{\theta}_n) = 1$. $\tilde{\mu}$ starts with μ, takes the resource away from the lowest-scoring group m agent who has it, $\tilde{\theta}_m$, and allocates it to the highest-scoring group n agent who does not have it, $\hat{\theta}_n$. Note that since A^* is monotone, from $x_m > x'_m$ and $x'_n > x_n$, under μ', $\hat{\theta}_n$ is already allocated the resource while $\tilde{\theta}_m$ is not.

Claim 15. *$\tilde{\mu}$ is implemented under A^* in state ω and $\xi(\mu, \omega) = \xi(\tilde{\mu}, \omega)$.*

Proof. Since A^* implements μ and $\mu(\hat{\theta}_n) = 0$, we have that $A^*_m(s(\tilde{\theta}_m), x_m) \geq A^*_n(s(\hat{\theta}_n), x_n+1)$. Conversely, since A^* also implements μ' and $\mu'(\hat{\theta}'_m) = 0$, we have that $A^*_n(s(\tilde{\theta}'_n), x'_n) \geq A^*_m(s(\hat{\theta}'_m), x'_m+1)$. Moreover, since $x_m > x'_m$ and $x'_n > x_n$, we have that $s(\hat{\theta}'_m) \geq s(\tilde{\theta}_m)$ and $s(\hat{\theta}_n) \geq s(\tilde{\theta}'_n)$. From this, it follows that:

$$\begin{aligned} A^*_n(s(\hat{\theta}_n), x_n + 1) &\geq A^*_n(s(\tilde{\theta}'_n), x_n + 1) \geq A^*_n(s(\tilde{\theta}'_n), x'_n) \\ &\geq A^*_m(s(\hat{\theta}'_m), x'_m + 1) \geq A^*_m(s(\hat{\theta}'_m), x_m) \geq A^*_m(s(\tilde{\theta}_m), x_m) \end{aligned} \tag{998}$$

where the first inequality holds as $s(\hat{\theta}_n) \geq s(\tilde{\theta}'_n)$, the second inequality holds as $x'_n > x_n$ (which implies $x'_n \geq x_n + 1$) and A^*_n is decreasing in its second argument, the third inequality holds as A^* also implements μ' (as stated above), the fourth inequality holds as $x'_m < x_m$ (which implies $x'_m + 1 \leq x_m$) and A^*_n is decreasing in its second argument, and the fifth inequality holds as $s(\hat{\theta}'_m) \geq s(\tilde{\theta}_m)$. Thus, $A^*_m(s(\tilde{\theta}_m), x_m) \leq A^*_n(s(\hat{\theta}_n), x_n + 1)$. This shows that $A^*_m(s(\tilde{\theta}_m), x_m) = A^*_n(s(\hat{\theta}_n), x_n + 1)$, which implies that $\tilde{\mu}$ is implemented under A^* and $\xi(\mu, \omega) = \xi(\tilde{\mu}, \omega)$. □

Note that Claim 15 shows that starting from a matching μ which is implemented by A^*, taking away the object from a particular agent who does not have it in μ' and allocating it to a particular agent who has it in μ', we arrive at another matching $\tilde{\mu}$ that is implemented under A^* and gives the authority the same payoff. Therefore, starting from any μ that is implemented by A^* and repeating this construction (by replacing μ at step i with $\tilde{\mu}$ at step $i - 1$) where at each step we take the resource from an agent who is not allocated the resource at μ' and assign it to an agent who is, in finitely many steps we arrive at μ'. Since the payoff stays the same at each step, μ' gives the authority the same payoff as μ. □

Theorem 16. *If μ is implemented by A^*, then μ is an optimal matching.*

Proof. First, note that an optimal matching exists since the economy (and therefore the set of matchings) is finite. We first show the following lemma.

Lemma 23. *If μ is not implemented by A^*, then there exists μ' that gives the authority a strictly higher payoff.*

Proof. If μ is not implemented by A^*, then there exists θ and θ' such that $\mu(\theta) = 0$, $\mu(\theta') = 1$ and either $m(\theta) = m(\theta')$ and $s(\theta) > s(\theta')$ or $m(\theta) \neq m(\theta')$ and

$$\begin{aligned} &h(s(\theta)) + u_{m(\theta)}(x_{m(\theta)}(\mu) + 1) - u_{m(\theta)}(x_{m(\theta)}(\mu)) > \\ &h(s(\theta')) + u_{m(\theta')}(x_{m(\theta')}(\mu)) - u_{m(\theta')}(x_{m(\theta')}(\mu) - 1) \end{aligned} \tag{999}$$

However, in both cases, a μ' that allocates the resource to θ instead of θ' (while not changing any other agent's matching) strictly improves the utility of the authority. □

Lemma 23 proves that the optimal matching cannot be a matching that is not implemented by A^*. Since the optimal matching exists, then it is implemented by A^*. From Lemma 22, all matchings implemented by A^* give the authority the same payoff, proving the result. □

Note that Lemma 22 and Theorem 16 imply that any mechanism that is defined by an arbitrary singleton selection from the set of matchings that A^* implements would achieve the optimal matching under any ω and therefore would be first-best optimal.

E.3.3 Priorities *vs.* Quotas in Discrete Economies

Now, we define Priority and Quota Mechanisms in the discrete model and extend our (sub)optimality results to discrete economies.

A *priority policy* $P : \Theta \to [0, 1]$ awards an agent of type $\theta \in \Theta$ a priority $P(\theta)$.

Definition 24 (Priority Mechanisms)**.** *A priority mechanism, induced by a priority policy P, allocates the resource in order of priorities until measure q has been allocated, with ties broken uniformly and at random.*

A *quota policy* is given by (Q, D), where $Q = \{Q_m\}_{m\in\mathcal{M}}$ and $D : \mathcal{M} \cup \{R\} \to \{1, 2, \ldots, |\mathcal{M}| + 1\}$ is a bijection. The vector Q reserves Q_m objects for agents in group m, with residual capacity $Q_R = q - \sum_{m\in\mathcal{M}} Q_m$ open to agents of all types. The bijection D (often called the precedence order) determines the order in which the groups are processed.

Definition 25 (Quota Mechanisms)**.** *A quota mechanism, induced by a quota policy (Q, D), proceeds by allocating $Q_{D^{-1}(k)}$ objects to agents from group $D^{-1}(k)$ (if there are sufficient agents from this group) to the resource in ascending order of k, and in* *descending order of score within each k. If there are insufficiently many agents of any group to fill the quota, the residual capacity is allocated to a final round in which all agents are eligible.*

We also extend the definitions of risk-neutrality and high risk aversion to the discrete setting. Authority preferences are *non-trivial* if for all $m, n \in \mathcal{M}$:

$$h(1) + (u_n(1) - u_n(0)) > h(0) + (u_m(q) - u_m(q-1)) \tag{1000}$$

The authority is *risk-neutral* if for all $m \in \mathcal{M}$, $u_m(x) = c_m x$ for some $c_m \geq 0$ and all $x \in \{0, 1, \ldots, q\}$. Define $\tilde{u}$ and $\tilde{h}$ as follows: there exists x_m^{tar} such that $\tilde{u}_m(x_m + 1) - \tilde{u}_m(x_m) = 0$ for all $x_m \geq x_m^{\text{tar}}$ and $\tilde{u}_m(x_m + 1) - \tilde{u}_m(x_m) \geq h(1) - h(0)$ for $x_m < x_m^{\text{tar}}$ and where $\sum_{m \in \mathcal{M}} x_m^{\text{tar}} \leq q$. Let $\tilde{\xi}$ denote the preferences of the authority under $\tilde{u}$ and $\tilde{h}$. The authority with preferences ξ is *extremely risk-averse* if the set of optimal allocations under ξ and $\tilde{\xi}$ coincide for all ω.

Theorem 17. *The following statements are true:*

1. *If there is no uncertainty, then there exist first-best priority and quota mechanisms.*

2. *Suppose that the authority has non-trivial preferences. There exists a first-best priority mechanism if and only if the authority is risk-neutral. This mechanism is given by $P(s, m) = s + u_m(1) - u_m(0)$.*

3. *Suppose that the authority has non-trivial preferences. There exists a first-best quota mechanism if and only if the authority is extremely risk-averse. This mechanism is given by $Q_m = x_m^{tar}$ and $D(R) = |\mathcal{M}| + 1$.*

Proof. Part (1):

Claim 16. *Let μ denote an optimal allocation at ω. Then μ is a cutoff matching.*

Proof. If μ is not a cutoff matching, then there exists (s, m) and (s', m) where $\mu(s, m) = 1$, $\mu(s', m) = 0$ and $s' > s$. Define μ' by setting: $\mu'(s, m) = 0$, $\mu(s', m) = 1$ and $\mu(\tilde{s}, \tilde{m}) = \mu(\tilde{s}, \tilde{m})$ for all $(\tilde{s}, \tilde{m})$ such that $(\tilde{s}, \tilde{m}) \notin \{(s, m), (s', m)\}$. Observe that, $\xi(\mu', \omega) - \xi(\mu, \omega) = s' - s > 0$. Therefore, μ is not an optimal allocation, which is a contradiction. $\square$

Let μ denote an optimal allocation under ω, $\{\hat{s}_m(\mu, \omega)\}_{m \in \mathcal{M}}$ denote the cutoff scores at μ and s^* denote an arbitrary number. Any priority policy that assigns

$P(\hat{s}_m(\omega), m) = s^*$ for all $m \in \mathcal{M}$ and is strictly increasing in the first argument allocates the resource to any agent who has a higher score than the cutoff for their group and implements the optimal allocation.

Let x_m denote the number of group m agents who are allocated the resource at an optimal allocation under ω. Then a quota policy that sets $Q_m = x_m$ allocates the resource to any agent who has a higher score than the cutoff for their group and implements the optimal allocation.

Part (2): The if part of the result follows from observing the priority policy $P(s, m) = s + u_m(1) - u_m(0)$ is equivalent to the optimal APM A^* under risk neutrality since $u_m(1) - u_m(0) = u_m(y_m + 1) - u_m(y_m)$ for all m, y_m. Thus, by Theorem 16, $P(s, m) = s + u_m(1) - u_m(0)$ is first-best optimal.

To prove the only if part, assume risk neutrality does not hold and let m denote a group such that u_m does not satisfy risk neutrality. For a contradiction, assume that P is an optimal priority policy. First, we observe that $P(s, m)$ must be strictly increasing in s for all m. To see why, assume $P(s, m) = P(s', m)$ where $s > s'$ and just consider an ω where there are $q - 1$ group m agents with scores strictly higher than s, and no other agents. Clearly, the optimal allocation would be to allocate the resource to all agents but (s', m), while P allocates the resource to (s', m) with at least probability 1/2.

Second, let m denote a group such that u_m does not satisfy risk neutrality. Take another arbitrary group n. We have the following:

Claim 17. *Either (i) there exists $t < q$, s_m, s_n such that*

$$u_m(t + 1) - u_m(t) + h(s_m) = u_n(q - t) - u_n(q - t - 1) + h(s_n) \tag{1001}$$

or (ii) there exists $t < q$ such that

$$u_m(t + 1) - u_m(t) + h(1) < u_n(q - t) - u_n(q - t - 1) + h(0) \tag{1002}$$

$$u_m(t) - u_m(t - 1) + h(0) > u_n(q - t + 1) - u_n(q - t) + h(1) \tag{1003}$$

Proof. From non-triviality, we know that $u_m(1) - u_m(0) + h(1) > u_n(q) - u_n(q - 1) + h(0)$ and $u_n(1) - u_n(0) + h(1) > u_m(q) - u_m(q - 1) + h(0)$. The result then follows from the fact that h is continuous and strictly increasing and u_m and u_n are concave. □

We first prove the result under case (ii). Fix two agents with scores $s_m \in (0, 1)$, who belong to group m and $s_n \in (0, 1)$, who belong to group n. Assume that there are

$t-1$ group m agents and $q-t$ group n agents with higher scores than $\max\{s_n, s_m\}$, so a total of t group m agents and $q-t+1$ group n agents. Note that in this case, only one agent will not be allocated the resource in the optimal allocation, and that would be either (s_m, m) or (s_n, n). From equation 1003, (s_m, m) is more preferred than (s_n, n) and therefore it must be that $P(s_n, n) < P(s_m, m)$, as otherwise P would not be optimal. Next, assume that there are t group m agents and $q-t-1$ group n agents with higher scores than $\max\{s_n, s_m\}$. From equation 1002, (s_n, n) is more preferred than (s_m, m) and therefore it must be that $P(s_m, m) < P(s_n, n)$, which is a contradiction.

We now prove the result under case (i).

Claim 18. *In case (i), any optimal priority policy P must satisfy $P(s_m+\epsilon, m) > P(s_n, n)$ for all $\epsilon > 0$ and $P(s_m-\epsilon, m) < P(s_n, n)$ for all $\epsilon > 0$*

Proof. From Equation 1001, we see that when there are t group m agents and $q-t-1$ group n agents with higher scores, $(s_m+\epsilon, m)$ is strictly preferred to (s_n, n), which is strictly preferred to $(s_m-\epsilon, m)$. □

Since u_m is not linear, there exists an l such that $u_m(l+1)-u_m(l) < u_m(l)-u_m(l-1)$. There are two possibilities: $l \leq t$ or $l > t$. First, suppose that $l \leq t$. We have that:

$$u_m(l)-u_m(l-1)+h(s_m) > u_m(l+1)-u_m(l)+h(s_m) \geq u_n(q-l)-u_n(q-l+1)+h(s_n) \tag{1004}$$

where the first inequality follows from $u_m(l+1)-u_m(l) < u_m(l)-u_m(l-1)$ and the second inequality follows as $u_m(t+1)-u_m(t)+h(s_m) = u_n(q-t)-u_n(q-t-1)+h(s_n)$, u_m and u_n are concave, and $l \leq t$. Thus, for sufficiently small $\epsilon > 0$, we have that:

$$u_m(l)-u_m(l-1)+h(s_m-\epsilon) > u_n(q-l)-u_n(q-l+1)+h(s_n) \tag{1005}$$

Given this inequality, we see that when there are $l-1$ group m agents and $q-l$ group n agents with higher scores, $(s_m-\epsilon, m)$ is strictly preferred to (s_n, n). Thus, to implement the optimal allocation, it must be that $P(s_m-\epsilon, m) \geq P(s_n, n)$, which is a contradiction to Claim 18.

Second, suppose that $l > t$. We know that:

$$u_m(t+1)-u_m(t)+h(s_m) = u_n(q-t)-u_n(q-t-1)+h(s_n) \tag{1006}$$

As $l > t$, from concavity of u_m and u_n,

$$u_m(l) - u_m(l-1) + h(s_m) \leq u_n(q-l+1) - u_n(q-l) + h(s_n) \tag{1007}$$

From concavity of u_n and u_m:

$$u_m(l+1) - u_m(l) + h(s_m) < u_n(q-l) - u_n(q-l-1) + h(s_n) \tag{1008}$$

Thus, for sufficiently small $\epsilon > 0$, we have that:

$$u_m(l+1) - u_m(l) + h(s_m+\epsilon) < u_n(q-l) - u_n(q-l-1) + h(s_n) \tag{1009}$$

Given this inequality, we see that when there are l group m agents and $q-l-1$ group n agents with higher scores, (s_n, n) is strictly preferred to $(s_m+\epsilon, m)$. Thus, to implement the optimal allocation, it must be that $P(s_m+\epsilon, m) \leq P(s_n, n)$, which is a contradiction to Claim 18.

Part (3): To prove the if part, fix an ω and let μ^* denote the optimal allocation under ω. Let x_m^* denote the number of group m agents allocated the resource at μ^* and $x_m(\omega)$ denote the total number of group m agents under ω.

Claim 19. *If the authority is extremely risk-averse, then* $x_m^* \geq \min\{x_m(\omega), x_m^{tar}\}$

Proof. Assume for a contradiction this is not the case. Then $x_m^* < x_m(\omega)$ and $x_m^* < x_m^{tar}$. Since $\sum_{m\in\mathcal{M}} x_m^{tar} \leq q$ and $x_m^* < x_m^{tar}$, there exists $n \in \mathcal{M}$ such that $x_n^* > x_n^{tar}$. Let s_n denote the score of the lowest-scoring group n agent who is allocated the resource, and let s_m denote the score of any group m agent who is not allocated the resource, which exists as $x_m^* < x_m(\omega)$. Since the authority is extremely risk-averse, we have the following:

$$h(s_m) + u_m(x_m^*+1) - u_m(x_m^*) > h(s_n) - u_n(x_n^*) + u_m(x_n^*-1) \tag{1010}$$

However, this contradicts the optimality of μ^* and proves the claim. □

Claim 20. *If the authority is extremely risk-averse,* $x_m^* > x_m^{tar}$ *and* $x_n^* > x_n^{tar}$*,* $\mu^*(s,m) = 0$ *and* $\mu^*(s',n) = 1$*, then* $s' > s$*.*

Proof. Assume for a contradiction that $s > s'$.[9] The difference in the utility of the

[9] Remember that $s' = s$ was ruled out by assumption.

authority when allocating the resource to (s, m) rather than (s', n) is given by

$$h(s)+u_m(x_m^*+1)-u_m(x_m^*)-(h(s')-u_n(x_n^*)+u_m(x_n^*-1)) = h(s)-h(s') > 0 \quad (1011)$$

which is a contradiction to optimality of μ^*. □

The previous two claims show that under any ω, the optimal allocation admits (i) the highest-scoring x_m^{tar} agents from each group (provided that they exist) and (ii) highest-scoring agents who are not in (i), until the capacity is exhausted. Clearly, the quota policy $Q_m = x_m^{\text{tar}}$ and $D(R) = |\mathcal{M}| + 1$ implements this outcome at every ω.

To prove the only if part, assume that $\{Q_m\}_{m\in\mathcal{M}}$ is part of an optimal quota policy.

Claim 21. *For and each $m, n \in \mathcal{M}$ and any t, l such that $t \leq Q_m$, $Q_m > 0$ and $l \geq Q_n$, we have that:*

$$u_m(t) - u_m(t-1) + h(0) \geq u_n(l+1) - u_n(l) + h(1) \quad (1012)$$

Proof. Assume that at ω, there are t group m agents, one of which one has score 0 and $l+1$ group n agents with scores higher than $1-\epsilon_1$ and q agents from other groups who have scores higher than $1-\epsilon_2$, where $\epsilon_1 > \epsilon_2 > 0$. As $t \leq Q_m$ and $Q_n < l+1$, t group m agents and $Q_n < l+1$ group n agents are admitted under Q. Since Q is optimal for all ϵ_1, we must have that:

$$u_m(t) - u_m(t-1) + h(0) \geq u_n(l+1) - u_n(l) + h(1-\epsilon_1) \quad (1013)$$

The statement then follows from continuity of h by taking the limit $\epsilon_1 \to 0$. □

Claim 22. *Merit slots are processed last at the optimal quota policy.*

Proof. For a contradiction, assume there is a merit slot that is processed before a quota slot. Let l denote the last merit slot that precedes a quota slot. Let m denote a group that has a quota slot after l. We consider a state in which: (i) there are q group n agents with scores $\hat{s} - \epsilon_i$, where $\epsilon_i > 0$ for all $i \in \{1, \ldots, q\}$ (let $\underline{\hat{s}}$ denote the score of the highest-scoring agent from this group), (ii) there are Q_m group m agents with scores $\hat{s} + \epsilon_j$ for $j \in \{1, \ldots, Q_m\}$ (let $\overline{\hat{s}}$ denote the score of lowest-scoring agent from this group) and one with score $\hat{s}/2$, and (iii) q agents from other groups with scores in $(\underline{\hat{s}}, \overline{\hat{s}})$. A group m agent with score $\hat{s} + \epsilon_k$ for some k is matched to l, thus $(\hat{s}/2, m)$ is matched to a later quota slot, while some agents with type $(\hat{s} - \epsilon_j, n)$ are rejected for some j. Let $\hat{s} - \epsilon_{j'}$ be the score of the highest-scoring such agent. From the optimality of the quota policy we have that

$$u_m(Q_m + 1) - u_m(Q_m) + h(\hat{s}/2) \geq u_n(Q_n + 1) - u_n(Q_n) + h(\hat{s} - \epsilon_{j'}) \tag{1014}$$

Let s^* be the score of the lowest-scoring group n agent (*i.e.*, $s^* = \min_{i \in \{1,\ldots,q\}} \hat{s} - \epsilon_i$). Next, consider the modified version of the above state, all group n agents are the same, but all of the other Q_m group m agents as well as q agents from other groups now have scores in $(s^* - \hat{\epsilon}, s^*)$ and the group m agent who had a score of $\hat{s}/2$ now has a score of $\hat{s}/2 + \hat{\epsilon}$ for $\hat{\epsilon} > 0$. Note that now the group n agent with score $\hat{s} - \epsilon_{j'}$ is allocated the slot l or an earlier slot, while the agent $(\hat{s}/2 + \hat{\epsilon}, m)$ is not allocated to any slot. Thus

$$u_m(Q_m + 1) - u_m(Q_m) + h(\hat{s}/2 + \hat{\epsilon}) \leq u_n(Q_n + 1) - u_n(Q_n) + h(\hat{s} - \epsilon_{j'}) \tag{1015}$$

which, since h is strictly increasing, implies that $u_m(Q_m + 1) - u_m(Q_m) + h(\hat{s}/2) < u_n(Q_n+1) - u_n(Q_n) + h(\hat{s} - \epsilon_{j'})$. This contradicts Equation 1014, proving the claim. □

Given the previous two claims, the following claim proves the result.

Claim 23. *If merit slots are processed last, then for all $l \geq Q_m$ and $j \geq Q_n$*

$$u_m(l + 1) - u_m(l) = u_n(j + 1) - u_n(j) \tag{1016}$$

Proof. Assume for a contradiction this does not hold. Without loss of generality, assume $u_m(l + 1) - u_m(l) > u_n(j + 1) - u_n(j)$ and define δ as

$$\delta = (u_m(l + 1) - u_m(l)) - (u_n(j + 1) - u_n(j)) \tag{1017}$$

Consider a state with $q - 1$ agents with scores higher than s^*, of which exactly Q_m are group m agents and Q_n are group n agents. Moreover, there is one more group m agent with score $s' < s^*$ (denote this agent by θ_m) and one more group n agent with score $s'' \in (s', s^*)$ where $h(s'') - h(s') < \delta$ (denote this agent by θ_n). Note that all agents apart from θ_m and θ_n are allocated the resource before the final merit slot. Moreover, since θ_n has a higher score, she obtains the final merit slot. However, this is a contradiction to the optimality of Q as $h(s'') - h(s') < \delta$ and allocating that resource to θ_m gives the authority higher utility. This proves the claim. □

Taken together, claims 21 and 23 prove that a fictitious authority that is extremely risk-averse with $x_m^{\text{tar}} = Q_m$ agrees with the authority on the optimal allocation, for all ω. To see this, observe that claim 21 implies that diversity preferences dominate any concern for scores when a group is allocated less than Q_m. Moreover, conditional on being allocated at least Q_m, it is as if there is no residual diversity preference, by claim 23. This proves the only if part of (3), which finishes the proof of the result. $\square$

E.3.4 Dominance of APM in Discrete Economies

In this section, we extend our discrete model to the multiple authority case and show that the dominance of the optimal APM in the decentralized admissions setting studied in Section 5.4.3 can be extended to this setting. Let Θ_0 denote the set of agents. $\mathcal{C} = \{c_0, c_1, \ldots, c_{|\mathcal{C}|-1}\}$ denote the set of authorities. q_c denotes the capacity of authority c and $q_{c_0} \geq |\Theta_0|$. $\theta = (s, m, \succ) \in [0,1]^{|\mathcal{C}|} \times \mathcal{M} \times \mathcal{R} = \Theta$, where $\mathcal{R}$ is set of all complete, transitive, and strict preference relations over $\mathcal{C}$ such that c_0 is less preferred than all $c \in \mathcal{C}$. For each type θ, $s_c(\theta)$ denotes the score of θ at authority c and $m(\theta)$ denotes the group of θ.

A matching in this environment is a function $\mu : \mathcal{C} \cup \Theta \to 2^{\Theta} \cup \mathcal{C}$ where $\mu(\theta) \in \mathcal{C}$ is the authority any type θ is assigned and $\mu(c) \subseteq \Theta$ is the set of agents assigned to authority c, which satisfies $|\mu(c)| \leq q_c$ for all c. $x_c(\mu) = \{x_{m,c}(\mu)\}_{m \in \mathcal{M}}$ denotes the number of agents of each group assigned to school c at μ while $\bar{s}_{h_c}(\mu) = \sum_{\theta \in \mu(c)} h(s(\theta))$ denotes the score utility the authority derives from μ. The preferences of the authority are given by:

$$\xi_c(\bar{s}_{h_c}, x_c) = \bar{s}_{h_c} + \sum_{m \in \mathcal{M}} u_{m,c}(x_{m,c}) \tag{1018}$$

where h_c is continuous and strictly increasing and $u_{m,c} : \mathbb{R} \to \mathbb{R}$ is concave for all $m \in \mathcal{M}$ and $c \in \bar{\mathcal{C}}$.

Agents apply to the authorities sequentially, who decide which agents to admit. We index the stage of the game by $t \in \mathcal{T} = \{1, \ldots, |\mathcal{C}| - 1\}$. Each stage corresponds to an authority $I(t)$, where $I : \mathcal{T} \to \mathcal{T}$. At each stage t, any unmatched agents choose whether apply to authority $I(t)$. Given the set of applicants, authority $I(t)$ chooses to admit a subset of these agents. Given this, histories are indexed by the path of the remaining of agents who have not yet matched, $h^{t-1} = (\Theta_0, \Theta_1, \ldots, \Theta_{t-1}) \in \mathcal{H}^{t-1}$. Given each history h^{t-1} and set of applicants $\Theta_c^A \subseteq \Theta$, a strategy for an authority returns a set of agents $\Theta_c^G \subseteq \Theta$ whom they will admit such that $\Theta_c^G \subseteq \Theta_c^A$ and $|\Theta_c^G| \leq q_c$ for each time at which they could move $t \in \mathcal{T}$, $a_{c,t} : \mathcal{H}^{t-1} \times \mathcal{P}(\Theta) \to \mathcal{P}(\Theta)$, where $\mathcal{P}(\Theta)$ is the power set over Θ. A strategy for an agent returns a choice of whether to apply to authorities at each history and time for all agent types $\theta \in \Theta$,

$\sigma_{\theta,t} : \mathcal{H}^{t-1} \to [0,1]$. We moreover say that a strategy $a_{\tilde{c},t}$ for an authority $\tilde{c}$ at time t is *dominant* if it maximizes authority utility regardless of $\{\{a_{c,t}\}_{c\in\mathcal{C}/\{\tilde{c}\}}, \{\sigma_{\theta,t}\}_{\theta\in\Theta}\}_{t\in\mathcal{T}}$ and I.

Theorem 18. *The APM A_c^* is a dominant strategy for all authorities.*

Proof. We prove that APM A_c^* implements a dominant strategy for all authorities in all stages by backward induction. Consider the terminal time $t = |\mathcal{C}| - 1$. Some set of agents $\hat{\Theta} \subseteq \Theta$ applies to the authority. Regardless of $\hat{\Theta}$, by Theorem 16 we have that the set of agents chosen under any selection from APM A_c^* is first-best optimal. Thus, A_c^* is dominant. Consider now any time $t < |\mathcal{C}| - 1$, precisely the same argument applies and A_c^* is dominant. □

E.3.5 Discrete Model: Example under Imperfect Information

We develop a simple example to show how the qualitative trade-offs between priorities and quotas we have identified are those present in discrete matching markets. There are 4 students, $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$ and one authority c with capacity two. Students θ_3 and θ_4 belong to an underrepresented minority. The scores of minority students are distributed independently and uniformly on $[0,1]$, so that $s_3, s_4 \sim U[0,1]$. For simplicity, we assume there is no uncertainty over the scores of other students: $s_1 = s_2 = 1$.[10] We further specify that the authority has the following utility function:[11]

$$W(\beta, \mu) = \beta \mathbb{I}\{\mu(c) \cap \{\theta_3, \theta_4\} \neq \emptyset\} + \sum_{i:\mu(\theta_i)=c} s_i \tag{1019}$$

This function embodies the main trade-off we have studied: the trade-off between scores and diversity. The first term indicates that whenever the authority admits at least one minority student, the utility of authority increases by β, which denotes the strength of affirmative action or diversity preferences. The second term simply indicates that the authority cares about scores and wants to admit the highest-scoring students they can. An alternative interpretation in this context, where allocating to agents with low scores is perceived as unfair, is that the authority wants to ensure outcomes that are fair in this sense.

The authority implements a stable matching and has two different policies at their disposal to influence the outcome of the matching mechanism. The first is a

[10]The qualitative result here does not change as long as non-minority students draw their scores from a distribution that FOSD $U[0,1]$.

[11]μ denotes a matching, where $\mu(c) \subset \Theta$ and $|\mu(c)| = 2$. We assume μ is stable, which uniquely determines the allocations.

priority subsidy, denoted by $\alpha \in [0,1]$. A subsidy α simply increases the scores of minority students by α and moves the distribution of scores of minority students to $U[\alpha, 1+\alpha]$. The second is a minority quota $Q \in \{0,1,2\}$ that reserves Q seats for minority students.

We start by characterizing the first-best where the authority can choose the matching they most prefer in each state of the world. Intuitively, if the score of the highest-scoring minority student is sufficiently high, then the designer prefers to admit her and one of the non-minority students. Otherwise, it is optimal for the designer to admit the two highest-scoring students, that is the two non-minority students. The first-best matching μ^* is therefore given by:

$$\mu^*(c) = \begin{cases} \{\theta_1, \theta_2\} & , \text{ if } \quad \max_{i\in\{3,4\}} s_i < 1-\beta, \\ \{\theta_1, \theta_k\} & , \text{ if } \quad \max_{i\in\{3,4\}} s_i > 1-\beta \text{ and } s_k = \max_{i\in\{3,4\}} s_i. \end{cases} \tag{1020}$$

See that if the authority had perfect information and knew s_3 and s_4 that both a priority subsidy and a quota can implement the first-best.[12] In particular, if $\max_{i\in\{3,4\}} s_i > 1-\beta$, then both a quota $Q = 1$ that reserves one seat for minority students and a priority subsidy for minority students of $\alpha \in (1-\min_{i\in\{3,4\}} s_i, 1-\max_{i\in\{3,4\}} s_i)$ implement the first-best. When $\max_{i\in\{3,4\}} s_i \leq 1-\beta$, then both a quota of $Q = 0$ and a subsidy $\alpha = 0$ implement the first-best. Thus, with perfect information, both policies yield the first-best and there is no trade-off for the authority.

We now consider the second best where an authority is constrained to implement a priority or a quota before the realization of uncertainty. Note that implementing the first-best is impossible with both quotas and priorities as neither can be adapted to the underlying realized scores of the minority students. We now solve for the optimal quota and priority designs and compare their values. In order to characterize the optimal reserve policy, one first notes that reserving both seats for minority students is always strictly dominated by reserving only one. Thus the designer only needs to compare a policy with a quota of one against a policy with no quotas. With no quota, no minority student is admitted and the utility of the designer $W_{nr}(\beta) = 2$ with probability one. On the other hand, if a quota of one is used, only the highest-scoring minority student is admitted. Thus, the expected utility of the designer is:

$$\mathbb{E}[W(\beta)] = \beta + 1 + \mathbb{E}\left[\max\{s_3, s_4\}\right] = \frac{5}{3} + \beta \tag{1021}$$

[12] Formally speaking, this is only true outside of the knife-edge case where $s_3 = s_4$, which is probability zero. In this case, there is no subsidy that can implement the first-best.

The optimal quota policy is therefore to reserve one seat for minority students if $\beta > 1/3$ and reserve no seats otherwise. Moreover, the utility of the designer under the optimal quota policy is:

$$V_Q(\beta) = \begin{cases} 2 & ,\text{if} \quad \beta \leq \frac{1}{3}, \\ \dfrac{5}{3} + \beta & ,\text{if} \quad \beta > \frac{1}{3}. \end{cases} \tag{1022}$$

We now compute the optimal priority design and the authority's value thereof. To this end, we first calculate the utility of the designer under subsidy α. We start by calculating the matching conditional on the realized scores. There are three cases to consider. If both minority students have scores above $1-\alpha$, then both of them are admitted to the authority. If both minority students score below $1-\alpha$, then neither of them is admitted. Lastly, if one of them scores above $1-\alpha$ while the other scores below $1-\alpha$, then only one minority student is admitted. The following equation gives the utility of the designer as a function of β and α:

$$\begin{aligned} \mathbb{E}[W(\beta, \alpha)] = & \int_{1-\alpha}^{1} \int_{1-\alpha}^{1} (s_3 + s_4 + \beta) ds_3 ds_4 + 2 \int_{1-\alpha}^{1} \int_{0}^{1-\alpha} (1 + s_4 + \beta) ds_3 ds_4 \\ & + \int_{0}^{1-\alpha} \int_{0}^{1-\alpha} 2 ds_3 ds_4 \\ = & -(1+\beta)\alpha^2 + 2\beta\alpha + 2 \end{aligned} \tag{1023}$$

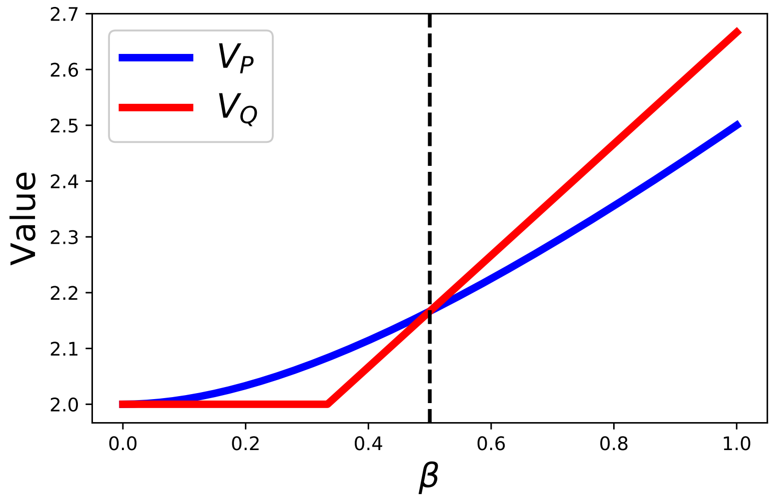
A quick calculation shows that the optimal subsidy is always interior and $\alpha^* = \frac{\beta}{1+\beta}$. Plugging the optimal subsidy policy into the authority's payoff function, we obtain:

$$V_P(\beta) = 1 + \beta + \frac{1}{1+\beta} \tag{1024}$$

We now compare the value of the optimal quota and priority designs as the strength of the affirmative action motive changes. Comparing $V_P(\beta)$ and $V_Q(\beta)$ shows that the optimal policy depends on the strength of affirmative action preferences of the authority. Figure E-1 plots these two values as a function of the affirmative action motive with the dotted line giving the value $\beta = \frac{1}{2}$ at which the two value functions cross. Importantly, we see that a quota policy is optimal whenever $\beta > 1/2$ and a priority subsidy policy is optimal whenever $\beta < 1/2$.

This example highlights the main differences between priorities and quotas under uncertainty and suggests when we might expect to prefer one over the other. When

Figure E-1: Comparative Statics for the Preference Between Priorities and Quotas



Notes: Values of the optimal priority policy, V_P, and optimal quota policy, V_Q, as a function of the strength of the diversity preference β. The dashed black line corresponds to $\beta = 1/2$ and is the point at which both policies yield the same value.

the preference for diversity is low, the authority only wants to admit a minority student if her score is high enough. In this case, a subsidy is a better policy as its outcome can depend on the relative scores of the students. In particular, it only admits minority students if they obtain sufficiently high scores while a quota admits minority students equally across states of the world. Consequently, priority designs generate a desirable positive selection of minority students which tends to improve scores. However, the drawback of a subsidy policy is that it applies to all students and can therefore cause either the admission of a second minority student with a lower average score or fail to admit any minority students. On the other hand, if the preference for diversity is sufficiently high, then the authority wants to admit one minority student for sure, regardless of her score. In this case, the subsidy policy is undesirable as even under the optimal subsidy, there are many realizations where neither or both minority students are admitted, while the reserve policy ensures that one minority student is admitted in all states of the world.

E.4 Extension to More General Authority Preferences

In this Appendix, we relax Assumption 10 to allow for (i) non-separable diversity preferences, (ii) non-separable score and diversity preferences, (iii) non-differentiable preferences, and (iv) non-concave diversity preferences. We show how these changes in assumptions lead to certain modified APM mechanisms becoming first-best optimal.

E.4.1 Non-Separable Diversity Preferences

First, we relax Assumption 10 and instead suppose that the authority's preferences satisfy the following assumption:

Assumption 18. *The authority's utility function can be represented as:*

$$\xi\left(\bar{s}_h, x\right) \equiv g\left(\bar{s}_h + u(x)\right) \tag{1025}$$

for some continuous, strictly increasing function $g : \mathbb{R} \to \mathbb{R}$ and a concave, partially differentiable u in each argument.

In this environment, we define a *non-separable APM* $\tilde{A} = \{\tilde{A}_m\}_{m\in\mathcal{M}}$ where $\tilde{A}_m : \mathbb{R}^{|\mathcal{M}|} \times [0,1] \to \mathbb{R}$. This implements allocation μ in state ω as per Definition 10 (under the modification of point 1 in Definition 10 to allow A_m to depend on x rather than just x_m).

We generalize Theorem 7 to show that the following non-separable APM uniquely implements the first-best optimal allocation:

Proposition 50. *The non-separable APM $\tilde{A}^*_m(y,s) \equiv h^{-1}(h(s)+u^{(m)}(y))$ and uniquely implements the first-best optimal allocation.*[13]

Proof. Follow every step in the proof of Theorem 7 with $\sum_{m\in\mathcal{M}} u_m(x_m)$ replaced by $u(x)$ and $u'_m(x_m)$ replaced by $u^{(m)}(x)$. □

Thus, allowing for non-separable diversity preferences does not substantially change the analysis of adaptive priority mechanisms. One must simply adapt the APM to be non-separable to allow cross-group diversity concerns to shape the marginal benefits of admitting agents from various groups. The main difference is that this a non-separable APM does not necessarily allow the greedy implementation of Algorithm 1. This is because, in the presence of cross-group adaptive priorities, it is no

[13] Where we define $u^{(m)}(y) = \frac{\partial}{\partial y_m} u(y)$.

longer enough to rank agents within their own group. A small adaptation to this algorithm that dynamically admits agents, starting from the highest-scoring agents in each group, would naturally implement the unique first-best optimal allocation.

E.4.2 Non-Separable Score and Diversity Preferences

Second, we relax Assumption 10 and instead suppose that the authority's preferences are represented by:

Assumption 19. *The authority's Bernoulli utility function can be represented as:*

$$\xi\left(\bar{s}_h, x\right) \tag{1026}$$

where ξ is monotone, differentiable, and concave.

We define a state-dependent APM $\hat{A} = \{\hat{A}_m\}_{m\in\mathcal{M}}$ where $\hat{A}_m : \mathbb{R}^{|\mathcal{M}|} \times [0,1] \times \Omega \to \mathbb{R}$. This implements allocation μ in state ω as per Definition 10 (where point 1 in Definition 10 is modified to allow A_m to depend on both x and ω).

In this more general setting, we now find a state-dependent APM that implements the optimal allocation.

Proposition 51. *The following state-dependent APM implements a first-best optimal allocation:*

$$A_m(y, s, \omega) \equiv h^{-1}\left(h(s) + \frac{\xi_{x_m}\left(\bar{s}_h(y,\omega), y\right)}{\xi_{\bar{s}_h}\left(\bar{s}_h(y,\omega), y\right)}\right) \tag{1027}$$

where $\bar{s}_h(y,\omega)$ is the score index in state ω when the highest-scoring $y = \{y_m\}_{m\in\mathcal{M}}$ agents of each attribute are allocated.

Proof. Follow every step in Theorem 7 with $\sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} + \sum_{m\in\mathcal{M}} u_m(x_m)$ replaced with $\xi\left(\bar{s}_h(y,\omega), x\right)$ where $\bar{s}_h(y,\omega) = \sum_{m\in\mathcal{M}} \int_{\underline{\tilde{s}}_m(x_m)}^{h(1)} \tilde{s}\tilde{f}_{m,\omega}(\tilde{s})\mathrm{d}\tilde{s}$. $\square$

There are two substantial differences in this optimal policy from our baseline APM. First, the policy depends on the joint distribution of agents in the population. Thus, specifying it *ex ante* is likely to be extremely challenging in any practical setting. This is necessary because the marginal rate of substitution between diversity and scores depends on the level of scores, which depends on the distribution of agents. Second, without assumptions on the shape of the distribution of agents, there is no guarantee that this policy is monotone and thus no guarantee that it implements a unique policy.

Thus, while Proposition 50 showed that cross-group separability is largely inessential for our main conclusions, separability between score and diversity preferences is key to the power of APM.

E.4.3 Non-Differentiable Preferences

In this section, we retain the majority of Assumption 10, where we instead suppose that the authority's diversity preferences $\{u_m\}_{m\in\mathcal{M}}$ are potentially non-differentiable at finitely many points.

As u_m is concave, the left and right derivatives of u_m, u_m^- and u_m^+, exist. The definition of our first-best APM is not applicable to this case since u_m' might not exist. Therefore, we define the following generalized optimal APM $A_m^*(y_m, s) \equiv h^{-1}(h(s)+u_m^-(y_m))$, which simply replaces u_m' with u_m^- in the definition. By concavity of u_m, u_m^- is monotone decreasing. Thus, this generalized optimal APM (as it is a montone APM) implements a unique allocation by Proposition 17. Moreover, the unique allocation that it implements is an optimal allocation:

Proposition 52. *Let μ^* denote the allocation implemented by the generalized optimal APM. μ^* is an optimal allocation.*

Proof. We first prove a claim. An allocation in this setting is a cutoff allocation if there exists cutoffs $\{s_m\}_{m\in\mathcal{M}}$ such that an agent θ is assigned the resource if and only if $s(\theta) \geq s_m$ and $m(\theta) = m$.

Claim 24. *There exists a unique optimal allocation μ' in the sense that all other allocations that attain the optimal payoff differ from μ' on at most a measure zero set of types. Moreover, there exists an optimal allocation that is a cutoff allocation.*

Proof. In the setting of Theorem 7, observe that $\tilde{\underline{s}}_m(x_m)$ is strictly decreasing in x_m. This, together with the concavity of u implies that the objective is strictly concave and constraints are linear. Therefore an optimal allocation exists and is unique up to measure zero transformations. Given this allocation μ' (with measures x_m), an optimal cutoff allocation is obtained by the cutoff scores s_m^* that satisfy

$$s_m' = \sup\left\{s_m \in [0,1] : \int_{s_m}^1 \tilde{f}_m(\tilde{s})\mathrm{d}\tilde{s} = x_m\right\} \tag{1028}$$

□

Using this claim, toward a contradiction, assume there exists another allocation μ', which gives the authority a strictly higher utility. Moreover, take μ' to be an optimal cutoff allocation (which must exist by the claim). As μ' differs from μ^* and both are cutoff allocations, we have that there exist two groups $m, n \in \mathcal{M}$ such that: (i) $s'_m > s^*_m$ and $x'_m < x^*_m$ and (ii) $s'_n < s^*_n$ and $x'_n > x^*_m$. We have that:

$$A^*_m(x, s) \geq A^*_m(x^*_m, s) > A^*_m(x^*_m, s^*_m) \geq A^*_n(x^*_n, \hat{s}) \geq A^*_n(\hat{x}, \hat{s}) \tag{1029}$$

for all $s \in (s^*_m, s'_m)$, $\hat{s} \in (s'_n, s^*_n)$, $x \leq x^*_m$, $\hat{x} \geq x^*_n$. The first inequality follows by concavity of u_m, the second follows by the fact that h is strictly increasing, the third follows by the definition of APM and the fact that $\mu^*(s^*_m, m) = 1$ and $\mu^*(\hat{s}, n) = 0$, and the fourth follows from concavity of u_n. Thus, we have that, for all $s \in (s^*_m, s'_m)$, $\hat{s} \in (s'_n, s^*_n)$, $x \leq x^*_m$, $\hat{x} \geq x^*_n$:

$$u^-_m(x) + h(s) > u^-_n(\hat{x}) + h(\hat{s}) \tag{1030}$$

Thus, the total marginal utility obtained by replacing any positive measure type m students with scores $s \in (s^*_m, s'_m)$ with an identical measure of type n students with scores $\hat{s} \in (s'_n, s^*_n)$ is positive. But this contradicts the optimality of μ'. Thus, if $\tilde{\mu}$ is optimal, then $\tilde{\mu} = \mu^*$ (up to a measure zero set). □

E.4.4 Non-Concave Preferences

In this section, we relax the assumption that the u_m are concave.

Proposition 53. *If μ is an optimal allocation, then μ is implemented by A^*.*

Proof. Without concavity, the optimal allocation characterized in the proof of Theorem 7 is no longer unique. However, the Lagrangian conditions we have derived are still necessary for any optimal allocation $x = \{x_m\}_{m \in \mathcal{M}}$. Thus, any optimal allocation is implemented by A^*. □

This result shows that any optimal allocation is implemented by the optimal APM. However, when $\{u_m\}_{m \in \mathcal{M}}$ are not concave, A^* is not necessarily monotone. Therefore, A^* does not necessarily implement a unique allocation. Indeed, it is possible that A^* implements suboptimal allocations, as it will implement any locally optimal allocation. Therefore, a mechanism defined by an arbitrary selection from the allocations implemented by A^* would not be first-best optimal. However, A^* may still help decision-making in this setting as it implements any optimal allocation.

E.5 Implementation, Precedence Orders, and an Illustration from H1-B Visa Allocation

In this appendix, we show that (with no uncertainty) priority and quota policies can implement the same set of allocations. We apply this insight to study the effect of precedence orders in US H1-B visa allocation.

E.5.1 Equivalence of Priorities and Quotas for Implementation

In Proposition 18, we showed that if there is no uncertainty, both priorities and quotas can achieve the optimal allocation. We say that a priority policy P is monotone if $P(s,m)$ is strictly increasing in s. Note that since the authority prefers higher-scoring agents to lower-scoring ones, monotone policies perform better non-monotone policies. We will now show that, in the setting of Section 5.3, for a given ω (which we suppress for the rest of this section), these quota and monotone priority policies are equivalent in the sense that any allocation that is achieved by one can also be achieved by the other.

Proposition 54. *μ is implemented by a quota policy if and only if it is also implemented by a monotone priority policy.*

Proof. Assume that μ is implemented by a quota policy. Then μ is a cutoff allocation since the resource is allocated in descending order of score. Let s_m denote the lowest-scoring agent from group m who is allocated the object at μ for $m \in \mathcal{M}$. Let $\bar{s} = \max_{m\in\mathcal{M}} s_m$. Define the priority policy as $P(s,m) = s + (s_m - \bar{s})$. Note that if $\mu(s,m) = 1$ and $\mu(s',m') = 0$, then $s + s_m - \bar{s} > s' + s_{m'} - \bar{s}$ and therefore $P(s,m) > P(s',m')$. As P allocates the resource to measure q highest-scoring agents under P and measure q of agents who are allocated the resource under the quota policy has higher priorities than those who are not, P implements the same allocation as the quota policy.

Conversely, assume that μ is implemented by a monotone priority policy. Let x_m denote the measure of agents from group m allocated the object at μ for $m \in \mathcal{M}$. Let Q denote a quota policy where $Q_m = x_m$. Under any processing order, Q implements the same allocation and allocates the resource to the highest-scoring measure x_m agents from group m, for all m. This is the allocation under $P(s,m)$ since P is a monotone priority policy and allocates the resource to the highest-scoring measure x_m agents from group m, which proves the result. $\square$

In the next section, we use this result to provide a diagnostic test for evaluating quota policies with different precedence orders by the strength of the equivalent priority policy.

E.5.2 Application of a Diagnostic Test for the Effect of Precedence Orders to US H1-B Visa Allocation

In this Appendix, we argue that in light of Proposition 54, priorities can be used as a diagnostic test for the effect of precedence orders in the context of US H1-B Visa allocation, which has had historical issues in implementation arising from the choice of precedence order. The American H1-B visa program enables American companies to temporarily employ educated foreign workers in high-skill occupations.[14] The statutory law enacted by the U.S. Congress mandates the total number of visas to be granted and The U.S. Customs and Immigration Service (USCIS) implements this mandate. The visa allocation is governed by the H-1B Visa Reform Act of 2004 that established an annual system in which 65,000 visas were made available for all eligible applicants and an additional 20,000 visas were reserved for applicants with advanced degrees. Until 2009, USCIS used the arrival time of the application to determine priorities. Since then, the priorities are determined according to a uniform lottery.

As we have emphasized, under quota policies, specifying the processing order is critical. Between 2009 and 2019, USCIS used a Reserve-Initiated processing rule. In 2020, in accordance with the 2017 *Buy American and Hire American Executive Order*, USCIS switched to a Unreserved-Initiated rule, in order to award visas to the most-skilled workers. Pathak, Rees-Jones, and Sönmez (2020) document this switch and give a detailed account of the consequences. In particular, they calculate the effect of this change on visa allocation between 2013-2017. They find that from (on average) 55,900 applicants with advanced degrees, 33,495 of them obtain a visa under the Reserve-Initiated rule, while 38,843 of them obtain a visa under the Unreserved-Initiated rule. This fact underscores how the complexity of quota policies can lead to issues in implementation, even in the simplest case with two groups.

We now use the structure of Proposition 54 to provide a diagnostic test that authorities can use to see the degree of effective affirmative action when employing a quota policy and apply it to the H1-B and Boston Public Schools settings.[15] Con-

[14] See Pathak, Rees-Jones, and Sönmez (2020) for a detailed account of H-1B policies and reforms.

[15] We note that the H-1B lottery is a setting where the perfect information assumption is justified. First, the only object that is allocated is the visa and all applicants prefer obtaining the visa to not obtaining it. This removes any uncertainty over the preferences of the individuals. Second, the priorities are determined according to a uniform random lottery and the market is large. In

Table E.1: Equivalent Priorities for Different Precedence Orders

	# Applicants		# Reserve-eligible Visas	
	General	Reserve-eligible	R-I Rule	NR-I Rule
5-yr Average (2013-2017)	137,017	55,900	33,495	38,834
Equivalent Subsidy (α)			23	35

Notes: Allocation of H-1B visas under Reserve-Initiated and Nonreserve-Initiated rules along with the equivalent subsidies that would induce these allocations.

cretely, when an authority is considering designing its precedence order, it can simply compute the implied priority subsidy being afforded to each group. In the context of H1-B allocation, we assume the uniform random lottery of USCIS is implemented by drawing a number uniformly from the interval $[0, 100]$. In the counterfactual priority mechanism, there are no quotas for reserve category applicants, but they get a score subsidy of α, i.e. their random numbers are distributed uniformly on $[\alpha, 100 + \alpha]$. Computing the implied α under both processing orders to compare the policies, we obtain Table E.1. Note that even though both quota policies correspond to 20,000 visas being reserved for applicants with advanced degrees, there is an important difference in the number of visas allocated to advanced degree applicants and therefore in the subsidy required to achieve that allocation. In particular, the Unreserved-Initiated order leads to a 12-point subsidy increase relative to the Reserve-Initiated benchmark.

particular, between 2013 and 2017, each year, 85,000 visas are allocated to an average of 137,017 reserve ineligible and 55,900 reserve eligible applicants. Thus, although there is uncertainty at the individual level, the distribution of lottery numbers conditional on reserve eligibility is essentially fixed.

E.6 Additional Quantitative Results

In this Appendix, we describe both the methodology and results of the two robustness exercises that are not discussed in full detail in the main text. First, we estimate the gains from APM when we assume that CPS sets one tier size for all tiers rather than separately optimizing the sizes of the four tiers. Second, we estimate the gains from APM under alternative utility functions that differentially penalize underrepresentation and overrepresentation.

E.6.1 Estimation with Homogeneous Reserves

As we have motivated, in this section we estimate an alternative model, where CPS chooses a single reserve size, r, instead of separate reserve sizes for all tiers. Formally, we replace the vector of reserve sizes of the four socioeconomic tiers, $r = (r_1, r_2, r_3, r_4)$ by $r = (r, r, r, r)$. In this setting, we define the marginal benefit of increasing reserve size as

$$G(r, \Lambda; \beta, \gamma) = \frac{\partial}{\partial r} \Xi(r, \Lambda; \beta, \gamma) \tag{1031}$$

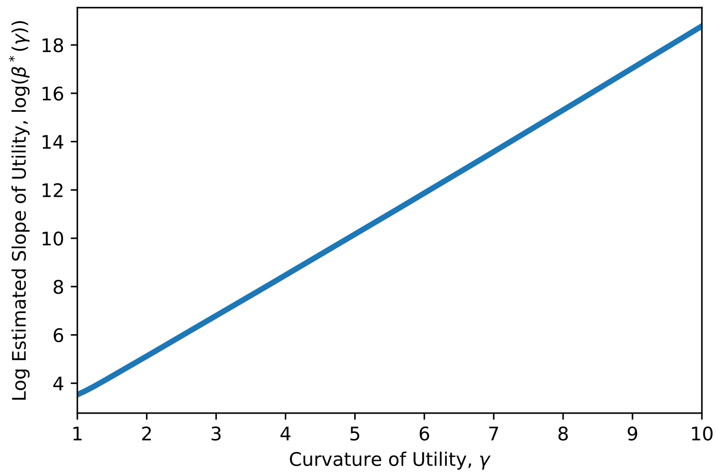
As in the general model, any (interior) reserve policy r^* must satisfy $G(\hat{r}^*, \hat{\Lambda}; \beta, \gamma) = 0$. This first-order condition yields one moment, and so we can estimate one parameter. To this end, we fix γ, and for each $\gamma \in [1, 10]$, and we estimate $\beta^*(\gamma)$ as the exact solution to the following empirical moment condition:

$$G(\hat{r}^*, \hat{\Lambda}; \beta^*(\gamma), \gamma) = 0 \tag{1032}$$

Figure E-2 plots the logarithm of the estimated $\beta^*(\gamma)$. The estimated $\beta^*(\gamma)$ is increasing in γ. As the loss term $|x_t - 0.25|$ is in $(0, 1)$, $\beta^*(\gamma)$ is increasing and convex in γ, where $\beta^*(1) = 34$ and $\beta^*(10) = 1.436 \times 10^8$. In Figure E-3, we plot the gains as a function of γ, which shows that even though the estimated value for β moves quite a lot, the empirical gains range from 2 to 4 points. This also shows that the estimated gain from APM of 2.1 under our benchmark specification is close to the lower bound of the estimated gains under the alternative specification with homogeneous reserves.

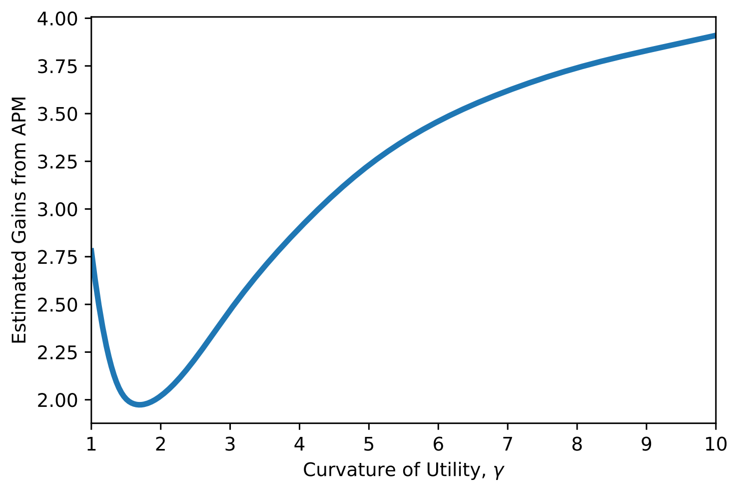
Finally, we benchmark these gains as a fraction of loss from underrepresentation under the CPS policy, where the loss of underrepresentation is calculated under the estimated parameter values. In Figure E-4, we plot the gains under APM as a percentage of diversity loss under the CPS policy. These range from 26% to 300%. Our baseline percentage gain estimate of 37.5% is again close to the lower bound that we

Figure E-2: Estimated Slope of Utility Under Homogeneous Reserves



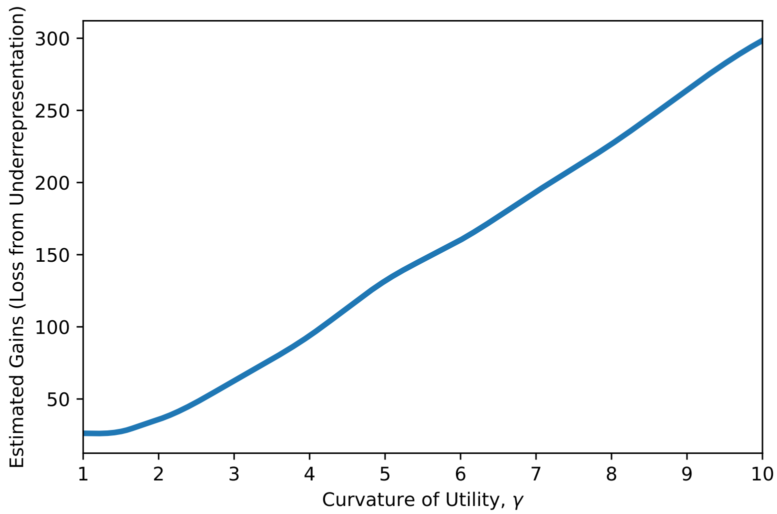
Notes: This graph plots the estimated logarithm of the slope of utility $\log \beta^*(\gamma)$ in the homogeneous reserve case as we vary the curvature of utility $\gamma \in [1, 10]$.

Figure E-3: Payoff Gains from APM Under Homogeneous Reserves



Notes: This graph plots the estimated difference in payoffs in the homogeneous reserves case between the optimal APM and the CPS policy as we vary the curvature of utility $\gamma \in [1, 10]$.

Figure E-4: The Gains from APM as a Fraction of the Loss From Underrepresentation Under Homogeneous Reserves



Notes: This graph plots the estimated difference between the payoffs under the optimal APM under homogeneous reserves as a fraction of loss from underrepresentation as we vary the curvature of utility $\gamma \in [1, 10]$.

estimate under the alternative specification with homogeneous reserves.

E.6.2 Gains from APM Under Different Utility Functions

In this section, as we have motivated, we estimate alternative objective functions to investigate the robustness of our findings.

First, we analyze a setting that includes a loss term only for underrepresented tiers (and does not penalize overrepresentation of any tier). To this end, we replace the term $|0.25 - x_t|$ with $\min\{0, (0.25 - x_t)\}$ and perform the same estimation with the following parametric utility function:

$$\xi(\bar{s}, x; \beta, \gamma) = \bar{s} + \beta \sum_{t=1}^{4} (\min\{0, (0.25 - x_t)\})^{\gamma} \tag{1033}$$

The estimated parameter values are $\beta^* = 52058$ and $\gamma^* = 3.87467$. We compute the difference between the empirical payoffs under APM and the CPS reserve policy to be 0.262, which is significantly lower than our estimate of 2.1. However, the reason for this is that the diversity domain is estimated to be less important under this specification, and the diversity loss under the CPS policy is 2.71. Thus, improvements from APM correspond to 9.6% of the loss from underrepresentation, which is attenuated relative to our baseline specification, but remains non-negligible.

Second, we allow CPS to care differentially about underrepresentation and overrepresentation by considering a utility function with separate coefficients for underrepresented and overrepresented tiers. To this end, we define the following loss function:

$$f(x_t, \beta_l, \beta_h, \gamma) = \begin{cases} \beta_l (0.25 - x_t)^\gamma & \text{if } x_t \leq 0.25 \\ \beta_h (x_t - 0.25)^\gamma & \text{if } x_t > 0.25 \end{cases} \tag{1034}$$

where β_l indexes the loss from underrepresentation of a tier, while β_h indexes the loss from overrepresentation. We then perform the same estimation with the following parametric utility function:

$$\xi(\bar{s}, x; \beta, \gamma) = \bar{s} + \sum_{t=1}^{4} f(x_t, \beta_l, \beta_h, \gamma) \tag{1035}$$

This yields the following estimated values: $\beta_l^* = 1362270$, $\beta_h^* = 12278$, $\gamma^* = 5.28021$. We compute the difference between the empirical payoffs under APM and the CPS reserve policy to be 0.195 and the loss from underrepresentation under the CPS policy to be 2.24. Thus, we conclude that improvement from APM corresponds to 8.7% of loss from underrepresentation under the CPS policy, which is similar to what we obtain under the specification in which there is no loss from overrepresentation.

Appendix F

Appendix to Nonlinear Pricing with Under-Utilization: A Theory of Multi-Part Tariffs

F.1 Proofs of Main Results

In this appendix, we provide the proofs of the main results. In Section F.1.1, we define and characterize implementable consumption functions under free disposal and characterize optimal contracts, proving Proposition 21. In Section F.1.2, we characterize the occurrence of multi-part tariffs by proving Proposition 22 and the corresponding corollaries. Finally, in Section F.1.3, we derive comparative statics for welfare, proving Propositions 23 and 24.

F.1.1 Implementation and Proof of Proposition 21

We say that consumption function ϕ is implementable if there exist a purchase function ξ and a price schedule T such that (ϕ, ξ, T) jointly satisfy the constraints (O), (IC), and (IR) of Problem 198. In this case, we say that ϕ is supported by (ξ, T). The following intermediate results fully characterize implementable consumption functions in terms of their functional properties. We say real functions are *monotone* when they are monotone non-decreasing.

Lemma 24. *Fix a consumption function ϕ that is monotone and such that $\phi \leq \phi^A$. Define the transfer function $t : \Theta \to \mathbb{R}$ as*

$$t(\theta) = C + u(\phi(\theta), \theta) - \int_0^\theta u_\theta(\phi(s), s)\,\mathrm{d}s \tag{1036}$$

for some $C \leq 0$, and define the price schedule $T : X \to \overline{\mathbb{R}}$ as

$$T(x) = \inf_{\theta' \in \Theta} \{t(\theta') : x \leq \phi(\theta')\} \tag{1037}$$

Then t and T are monotone.

Proof. Fix $\theta', \theta \in \Theta$ such that $\theta' \geq \theta$. Given that ϕ is monotone, it is almost everywhere differentiable with derivative denoted by ϕ' when defined. By the Fundamental Theorem of calculus, we can re-write the transfer function as

$$t(\theta) = C + u(\phi(0), 0) + \int_0^\theta (u_x(\phi(s), s)\phi'(s) + u_\theta(\phi(s), s))\, \mathrm{d}s - \int_0^\theta u_\theta(\phi(s), s)\, \mathrm{d}s \tag{1038}$$

Subtracting $t(\theta)$ from $t(\theta')$, we get

$$t(\theta') - t(\theta) = \int_\theta^{\theta'} u_x(\phi(s), s)\phi'(s)\, \mathrm{d}s \tag{1039}$$

Given that $\phi \leq \phi^A$, and that u is strictly quasiconcave in x, it follows that $u_x(\phi(s), s) \geq 0$ for all $s \in [0, \theta']$. Moreover, given that ϕ is monotone, it follows that $\phi'(s) \geq 0$ for almost all $s \in [0, \theta']$. Given that $\theta' \geq \theta$, Equation 1039 implies that $t(\theta') \geq t(\theta)$. Given that θ', θ were arbitrarily chosen, it follows that t is monotone.

Next, fix $x, y \in X$ such that $y \leq x$. Given that ϕ is monotone, the definition of T implies that $T(y) \leq T(x)$. We then conclude that T is monotone. □

Lemma 25. *A consumption function ϕ is implementable if and only if ϕ is monotone and such that $\phi \leq \phi^A$. In this case, ϕ is supported by (ϕ, T), where T is defined as in Equation (1037) for some $C \leq 0$.*[1]

Proof. (Only if). If ϕ is implementable, then there exists (ξ, T) that support ϕ. By Incentive Compatibility and by the taxation principle, there exists a transfer function $t : \Theta \to \mathbb{R}$ such that $u(\phi(\theta), \theta) - t(\theta) \geq u(\phi(\theta'), \theta) - t(\theta')$ for all $\theta, \theta' \in \Theta$. By a standard implementation result (see, e.g., Proposition 1 in Rochet, 1987), this implies that ϕ is monotone. Finally, if there exists $\theta \in \Theta$ such that $\phi(\theta) > \phi^A(\theta)$, then we would contradict Obedience for type θ since $u(\phi^A(\theta), \theta) > u(\phi(\theta), \theta)$ and $\phi^A(\theta)$ would be feasible given $\phi(\theta)$ by construction.

(If). Now suppose that ϕ is monotone and such that $\phi(\theta) \leq \phi^A(\theta)$ for all $\theta \in \Theta$. Define t and T given ϕ as in Equations 1036 and 1037 respectively. We next prove that (ϕ, ϕ, T) satisfies (O), (IC), and (IR).

[1] Observe that here the purchase function is $\xi = \phi$.

First, for every $\theta \in \Theta$, we have

$$u(\phi(\theta), \theta) - T(\phi(\theta)) \geq u(\phi(\theta), \theta) - t(\theta) = \int_0^\theta u_\theta(\phi(s), s)\,\mathrm{d}s - C \geq 0 \qquad (1040)$$

where the first inequality follows from the definition of T and the last inequality follows from $C \leq 0$ and $u_\theta(\phi(\theta), \theta) \geq 0$ for all $\theta \in \Theta$ (u is monotone increasing in θ). This proves Individual Rationality.

Next, assume toward a contradiction that Obedience does not hold. That is, there exist $\theta \in \Theta$ and $y < \phi(\theta) \leq \phi^A(\theta)$ such that $u(y, \theta) > u(\phi(\theta), \theta)$. However, this yields a contradiction with strict quasiconcavity of u in x. Therefore, Obedience holds.

We are left to prove that (ϕ, ϕ, T) satisfy Incentive Compatibility. Fix $\theta', \theta \in \Theta$ such that $\theta' \neq \theta$. We first prove that, for all θ, θ', we have

$$u(\phi(\theta), \theta) - t(\theta) \geq \max_{x \leq \phi(\theta')} u(x, \theta) - t(\theta') \qquad (1041)$$

This is a variation of the standard reporting problem under consumption function ϕ and transfers t, where each agent, on top of misreporting their type, can freely dispose of the allocated quantity. Violations of this condition can take two forms. First, an agent of type θ could report type θ' and consume $x = \phi(\theta')$. We call this a single deviation. Second, an agent of type θ could report type θ' and consume $x < \phi(\theta')$. We call this a double deviation. Under our construction of transfers t and monotonicity of ϕ, by a standard mechanism-design argument (e.g., Proposition 1 in Rochet, 1987), there is no strict gain to any agent of reporting θ' and consuming $x = \phi(\theta')$. Thus, there are no profitable single deviations under (ϕ, t).

We now must rule out double deviations. Define the value function $V : \Theta \to \mathbb{R}$ under ϕ and t as

$$V(\theta) = u(\phi(\theta), \theta) - t(\theta) = \int_0^\theta u_\theta(\phi(s), s)\,\mathrm{d}s - C \qquad (1042)$$

Suppose, toward a contradiction, that there exists a double deviation in which type θ reports type θ'. We separate the argument by various cases comparing $(\theta, \phi(\theta), \phi^A(\theta))$ and $(\theta', \phi(\theta'), \phi^A(\theta'))$.

1. $\theta' < \theta$: Given that ϕ is monotone, it must be that $\phi(\theta') \leq \phi(\theta)$. Moreover, as (O) holds, we have that $\phi(\theta') < \phi(\theta)$. For the same reason, we have that $\phi(\theta')$ is optimal for type θ' when they could choose any $x \leq \phi(\theta')$. Moreover, by strict single-crossing of u and strict quasiconcavity of $u(\cdot, \theta)$, it is optimal for type

θ to consume some $x \geq \phi(\theta')$. But, we know that $x \leq \phi(\theta')$; thus $x = \phi(\theta')$ is optimal. Hence, if there is a double deviation with $\theta' < \theta$, there is also a single deviation. This is a contradiction as we already showed that there are no strictly profitable single deviations.

2. $\theta' > \theta$ and $\phi^A(\theta) \geq \phi(\theta')$: the optimal choice of consumption for agent θ in $[0, \phi(\theta')]$ is given by $\phi(\theta')$ by strict quasiconcavity of u. Thus, there is a profitable single deviation, which is a contradiction.

3. $\theta' > \theta$ and $\phi^A(\theta) < \phi(\theta')$: We know $x = \phi^A(\theta)$ is most attractive following the misreport θ'. Suppose that there exists some $\hat{\theta} \in (\theta, \theta']$ such that $\phi(\hat{\theta}) = \phi^A(\theta)$. Given that t is monotone by Lemma 24, we know that a single deviation to $\hat{\theta}$ is weakly more attractive than a double deviation to $x \leq \phi(\theta')$. As no single deviations exist, this is a contradiction. As $\phi(\theta) \leq \phi^A(\theta) < \phi(\theta')$, it follows that no type $\hat{\theta} \in (\theta, \theta']$ receives $\phi^A(\theta)$. We know that the most attractive misreport is the smallest type θ' such that $\phi(\theta') \geq \phi^A(\theta)$. It follows that $\phi^A(\theta) \leq \phi(\theta') \leq \phi^A(\theta')$ and therefore that there exists some $\hat{\theta}$ such that $\phi^A(\hat{\theta}) = \phi(\theta')$, by continuity of ϕ^A.

   We now work toward a contradiction. By the hypothesis of a double deviation for type θ:
   $$u(\phi^A(\theta), \theta) - t(\theta') > u(\phi(\theta), \theta) - t(\theta) \tag{1043}$$

   Define for any type θ, the value of optimal autarkic consumption as $V^*(\theta) = u(\phi^A(\theta), \theta)$. We can write $V^*(\theta) - V(\theta) > t(\theta')$. As we have ruled out single deviations, we know that:

   $$u(\phi^A(\hat{\theta}), \hat{\theta}) - t(\theta') \leq u(\phi(\hat{\theta}), \hat{\theta}) - t(\hat{\theta}) \tag{1044}$$

   Thus $V^*(\hat{\theta}) - V(\hat{\theta}) \leq t(\theta')$. Together, we then have that $V(\hat{\theta}) - V(\theta) > V^*(\hat{\theta}) - V^*(\theta)$. From the definition of V in Equation 1042, the left-hand-side is $V(\hat{\theta}) - V(\theta) = \int_\theta^{\hat{\theta}} u_\theta(\phi(s), s)\, \mathrm{d}s$. From the envelope theorem applied to the autarkic consumption problem, the right-hand-side is $V^*(\hat{\theta}) - V^*(\theta) = \int_\theta^{\hat{\theta}} u_\theta(\phi^A(s), s)\, \mathrm{d}s$. Combining these substitutions with the original inequality,

   $$\int_\theta^{\hat{\theta}} u_\theta(\phi(s), s)\, \mathrm{d}s > \int_\theta^{\hat{\theta}} u_\theta(\phi^A(s), s)\, \mathrm{d}s \tag{1045}$$

   But we know that $\phi^A(s) \geq \phi(s)$ for all $s \in [\theta, \hat{\theta}]$, and this implies by single-crossing of u that $u_\theta(\phi^A(s), s) \geq u_\theta(\phi(s), s)$, which contradicts the inequality above.

We have ruled out double deviations in all cases and thereby completed the proof of the claim in Equation 1041. We next prove that Equation 1041 implies that (ϕ, ϕ, T) satisfy Incentive Compatibility. For all $\theta \in \Theta$, we have

$$\begin{aligned}
&u(\phi(\theta), \theta) - T(\phi(\theta)) \geq u(\phi(\theta), \theta) - t(\theta) \\
&\geq \sup_{\theta' \in \Theta} \left\{ \sup_{x \in X : x \leq \phi(\theta')} \{u(x, \theta)\} - t(\theta') \right\} = \sup_{x \in X} \left\{ \sup_{\theta' \in \Theta : x \leq \phi(\theta')} \{u(x, \theta) - t(\theta')\} \right\} \\
&= \sup_{x \in X} \{u(x, \theta) - T(x)\}
\end{aligned} \tag{1046}$$

yielding Incentive Compatibility. This concludes the proof of the implication.

The second part of the statement directly follows from the proof of sufficiency. $\square$

We now show that optimizing over the set of implementable allocations is equivalent to maximizing virtual surplus subject to the implementation constraints from Lemma 25.

Lemma 26. *A consumption function ϕ^* is part of a solution to Problem 198 if any only if it solves*

$$\begin{aligned}
\max_{\phi} \quad & \int_\Theta J(\phi(\theta), \theta) \, \mathrm{d}F(\theta) \\
s.t. \quad & \phi(\theta') \geq \phi(\theta), \quad \phi(\theta) \leq \phi^A(\theta), \quad \theta, \theta' \in \Theta : \theta' \geq \theta
\end{aligned} \tag{1047}$$

Proof. We begin by eliminating the proposed allocation and transfers from the objective function of the seller. From the proof of Lemma 25, we have that every implementable ϕ is supported by $\xi = \phi$ and by a price schedule T defined as in Equation 1037 where the transfer function t is defined in Equation 1036 for some constant $C \leq 0$. Given that any ξ that supports ϕ leads to the same seller payoff, we can then set $\xi = \phi$ without loss of optimality. Moreover, we know that ϕ being implementable is equivalent to ϕ being monotone increasing and $\phi \leq \phi^A$ (given that $C \leq 0$). Finally, it is not optimal for the seller to exclude any agent from the mechanism as it is without loss to allocate any agent $x = 0$ rather than exclude them owing to the fact that $\pi(0, \cdot) = 0$, $u(0, \cdot) = 0$, $u(x, \cdot)$ is monotone increasing over Θ, and u has strict single-crossing in (x, θ). In particular, for any incentive compatible allocation that excludes some type θ, it is without loss of optimality to set $\phi(\theta) = \xi(\theta) = t(\theta) = 0$.

Each agent is indifferent between participation and not, and this does not change the principal's payoff.

Plugging in the expression (1036), we can simplify the expression for the seller's total transfer revenue as the following:

$$\begin{aligned}\int_\Theta t(\theta)\,\mathrm{d}F(\theta) &= \int_\Theta \left(C + u(\phi(\theta),\theta) - \int_0^\theta u_\theta(\phi(s),s)\,\mathrm{d}s\right)\mathrm{d}F(\theta) \\ &= \int_\Theta \left(C + u(\phi(\theta),\theta) - \frac{(1-F(\theta))}{f(\theta)}u_\theta(\phi(\theta),\theta)\right)\mathrm{d}F(\theta)\end{aligned} \tag{1048}$$

where the final equality follows by applying the standard integration-by-parts argument.

Plugging into the seller's objective, we find that the principal solves:

$$\begin{aligned}&\max_{\phi,C} \int_\Theta (J(\phi(\theta),\theta)+C)\,\mathrm{d}F(\theta) \\ &\text{s.t.} \quad C \le 0,\ \phi(\theta') \ge \phi(\theta),\ \phi(\theta) \le \phi^A(\theta) \quad \forall \theta,\theta' \in \Theta : \theta' \ge \theta\end{aligned} \tag{1049}$$

It follows that it is optimal to set $C = 0$, completing the proof. □

Proof of Proposition 21. By Lemma 26, any optimal consumption function must solve Problem 1047. Consider now the family of problems $\max_{x\in[0,\phi^A(\theta)]} J(x,\theta)$, indexed by $\theta \in \Theta$. As J is strictly quasiconcave in x, there is a unique maximum in this problem, which we call $\phi^*(\theta)$. Moreover, whenever $\phi^P(\theta) < \phi^A(\theta)$, we know that $\phi^*(\theta) = \phi^P(\theta)$. Otherwise $\phi^*(\theta) = \phi^A(\theta)$, by strict quasiconcavity of J in x. Thus, the solution of this pointwise problem is $\phi^*(\theta) = \min\left\{\phi^A(\theta), \phi^P(\theta)\right\}$. As ϕ^A and ϕ^P are monotone, ϕ^* is monotone and is therefore the unique solution to Problem 1047.

We next prove the remaining parts of the statement by explicitly constructing the claimed supporting price schedules and purchases. From Lemma 25, we can construct the claimed formula for the price schedule directly. Because ϕ^* is invertible over $(\phi^*(0), \phi^*(1))$ and using its extension on the boundaries (see Footnote 14), we have that for all $x \in X^* = [\phi^*(0), \phi^*(1)]$:

$$T^*(x) = t({\phi^*}^{-1}(x)) = u(x, {\phi^*}^{-1}(x)) - \int_0^{{\phi^*}^{-1}(x)} u_\theta(\phi^*(s),s)\,\mathrm{d}s \tag{1050}$$

As T^* is monotone, it is almost everywhere differentiable. Moreover, whenever it is differentiable, by differentiating Equation 1050 we obtain ${T^*}'(x) = u_x(x, {\phi^*}^{-1}(x))$. Integrating, we obtain the price schedule in Equation 203 on X^*.

Finally, we show that the optimal level of consumption is supported by any selection from Ξ_{ϕ^*}, and only by selections from Ξ_{ϕ^*}. To this end, consider the selection $\overline{\xi} \in \Xi_{\phi^*}$ defined as $\overline{\xi} = \max \Xi_{\phi^*}$. We want to show that the triple $(\phi^*, \overline{\xi}, T^*)$ satisfies Obedience, Incentive Compatiblility, and Individual Rationality. We now define $t = T^* \circ \overline{\xi}$.

Consider first the Obedience constraint that $\phi^*(\theta) \in \arg\max_{x \in [0, \overline{\xi}(\theta)]} u(x, \theta)$, for all $\theta \in \Theta$. Observe that $\phi^* \leq \overline{\xi}$ by construction of $\overline{\xi}$. Moreover, toward a contradiction, suppose that there exists $\theta \in \Theta$ and $x \leq \overline{\xi}(\theta)$ such that $u(\phi^*(\theta), \theta) < u(x, \theta)$. There are two cases:

1. If $\phi^*(\theta) < \phi^A(\theta)$, then by construction $x \leq \overline{\xi}(\theta) = \phi^*(\theta) < \phi^A(\theta)$ implying that $u(\phi^*(\theta), \theta) \geq u(x, \theta)$ by strict quasiconcavity of $u(\cdot, \theta)$, hence yielding a contradiction.

2. If $\phi^*(\theta) = \phi^A(\theta)$, then by construction $u(\phi^*(\theta), \theta) \geq u(x, \theta)$ yielding a contradiction.

Consider now the Incentive Compatibility constraint that for all $\theta \in \Theta$:

$$\overline{\xi}(\theta) \in \arg\max_{y \in X} \left\{ \max_{x \in [0,y]} u(x, \theta) - T^*(y) \right\} \tag{1051}$$

and define $g(y, \theta) = \max_{x \in [0,y]} u(x, \theta)$. Toward a contradiction, suppose that there exist $\theta, \theta' \in \Theta$ such that $g(\overline{\xi}(\theta), \theta) - t(\theta) < g(\overline{\xi}(\theta'), \theta) - t(\theta')$. There are two cases to consider:

1. If $\phi^*(\theta') < \phi^A(\theta')$, then $\Xi_{\phi^*}(\theta') = \{\phi^*(\theta')\}$. Thus, $\overline{\xi}(\theta') = \phi^*(\theta')$. Hence:

$$g(\phi^*(\theta), \theta) - t(\theta) = g(\overline{\xi}(\theta), \theta) - t(\theta) < g(\overline{\xi}(\theta'), \theta) - t(\theta') = g(\phi^*(\theta'), \theta) - t(\theta') \tag{1052}$$

   where the first equality follows by Obedience, the inequality follows by hypothesis, and the last equality follows as $\overline{\xi}(\theta') = \phi^*(\theta')$.

2. If $\phi^*(\theta') = \phi^A(\theta')$, then define $\theta'' = \inf \left\{ \hat{\theta} \in \Theta : \hat{\theta} \geq \theta', \phi^*(\hat{\theta}) < \phi^A(\hat{\theta}) \right\}$. Note that, by monotonicity of ϕ^* and by construction of $\overline{\xi}$, we have $\phi^*(\theta'') = \overline{\xi}(\theta'') = \overline{\xi}(\theta')$. Moreover, by construction we necessarily have that $[\theta', \theta''] \subseteq \left\{ \hat{\theta} \in \Theta : \phi^*(\hat{\theta}) = \phi^A(\hat{\theta}) \right\}$. Therefore, by Equation 1039 in Lemma 24, we have that:

$$t(\theta'') - t(\theta') = \int_{\theta'}^{\theta''} u_x(\phi^A(s), s) \left(\phi^A\right)'(s) \, \mathrm{d}s = 0 \tag{1053}$$

by optimality of $\phi^A(s)$ for all $s \in [0,1]$. Thus, $t(\theta') = t(\theta'')$ and we have that:

$$g(\phi^*(\theta),\theta) - t(\theta) = g(\overline{\xi}(\theta),\theta) - t(\theta) < g(\overline{\xi}(\theta'),\theta) - t(\theta') = g(\phi^*(\theta''),\theta) - t(\theta'') \tag{1054}$$

where the first equality is by Obedience, the inequality is by hypothesis and the second equality follows as $\phi^*(\theta'') = \overline{\xi}(\theta')$ and $t(\theta') = t(\theta'')$.

In both cases, there exists $\theta'' \in \Theta$ such that $g(\phi^*(\theta),\theta) - t(\theta) < g(\phi^*(\theta''),\theta) - t(\theta'')$ (in case 1, $\theta'' = \theta'$). This contradicts the fact that (ϕ^*, ϕ^*, T^*) is implementable, which we established in Lemma 25. Thus, Incentive Compatibility is satisfied.

Finally, consider the Individual Rationality constraint that $u(\phi^*(\theta),\theta) - T^*(\overline{\xi}(\theta)) \geq 0$ for all $\theta \in \Theta$. Observe that $T^*(\overline{\xi}(\theta)) = T^*(\phi^*(\theta))$ for all θ such that $\phi^*(\theta) < \phi^A(\theta)$. When $\phi^*(\theta) = \phi^A(\theta)$, we have that $T^*(\overline{\xi}(\theta)) - T^*(\phi^*(\theta)) = \int_{\phi^*(\theta)}^{\overline{\xi}(\theta)} u_x(z, {\phi^*}^{-1}(z))\,\mathrm{d}z = 0$ as all types that consume between $\phi^*(\theta) = \phi^A(\theta)$ and $\overline{\xi}(\theta)$ consume their bliss point, by construction. Thus, $T^* \circ \overline{\xi} = T^* \circ \phi^*$ and by implementability of (ϕ^*, ϕ^*, T^*) (see Lemma 25), Individual Rationality holds.

This proves that $(\phi^*, \overline{\xi}, T^*)$ is implementable and therefore optimal. We now argue that for any other selection $\xi \in \Xi_{\phi^*}$, the triple (ϕ^*, ξ, T^*) is necessarily implementable and therefore optimal. Indeed, by way of contradiction, suppose that the latter is not implementable. It follows that $(\phi^*, \overline{\xi}, T^*)$ is not implementable either as all feasible deviations under purchase function ξ are still feasible under $\overline{\xi}$ and $T^* \circ \overline{\xi} = T^* \circ \xi$. However, this contradicts our demonstration that $(\phi^*, \overline{\xi}, T^*)$ is implementable.

We finally show that if $\xi \notin \Xi_{\phi^*}$, then it is not part of an optimal contract. We will use the observation that all agents' payments to the seller are pinned down by the envelope formula for t. There are two cases to consider. First, suppose that there exists a $\theta \in \Theta$ such that $\xi(\theta) \neq \phi^*(\theta)$ and $\phi^*(\theta) < \phi^A(\theta)$. If $\xi(\theta) < \phi^*(\theta)$, then $\phi^*(\theta) \notin [0, \xi(\theta)]$, which makes ϕ^* infeasible. If $\xi(\theta) > \phi^*(\theta)$, then, as $\phi^*(\theta) < \phi^A(\theta)$, $t(\theta)$ is strictly increasing at θ. Thus $T^*(\xi(\theta)) > t(\theta)$, which is a contradiction. Second, suppose that there exists a $\theta \in \Theta$ such that $\xi(\theta) \notin [\phi^A(\theta), \inf_{\theta' \in [\theta,1]} \left\{\phi^*(\theta') : \phi^*(\theta') < \phi^A(\theta')\right\}]$ and $\phi^*(\theta) = \phi^A(\theta)$. Once again if $\xi(\theta) < \phi^*(\theta)$, then $\phi^*(\theta) \notin [0, \xi(\theta)]$, which makes ϕ^* infeasible. If $\xi(\theta) > \inf_{\theta' \in [\theta,1]} \left\{\phi^*(\theta') : \phi^*(\theta') < \phi^A(\theta')\right\}$, then as before $T^*(\xi(\theta)) > t(\theta)$, which is a contradiction. $\square$

F.1.2 Proof of Proposition 22, Corollary 13, and Corollary 14

Proof of Proposition 22. We first prove that $H(x) > 0$ implies that $T^*(x)$ is flat at x, for any $x \in X^*$. By the definition of a multi-part tariff, this will also prove that T^* is a multi-part tariff. Consider first any $x \in \mathrm{Int}(X^*)$, where $\mathrm{Int}(\cdot)$ denotes the interior of a set. Observe that $\phi^* = \min\{\phi^P, \phi^A\}$ is invertible over $\text{Int}(X^*)$. Suppose now that $H(x) = J_x(x, (\phi^A)^{-1}(x)) > 0$ and define $\theta(x) = (\phi^*)^{-1}(x)$. It is either the case that $x = \overline{x}$ (which is not in $\text{Int}(X^*)$), or $\phi^A(\theta(x)) < \phi^P(\theta(x))$. Thus, when $H(x) > 0$, $\phi^A(\theta(x)) < \phi^P(\theta(x))$, so $\phi^*(\theta(x)) = \phi^A(\theta(x))$. As ϕ^A and ϕ^P are continuous functions by the Theorem of the Maximum and invertible at x, we can find a neighborhood $O(x)$ of x, such that for all $x' \in O(x)$, and corresponding $\theta' = (\phi^*)^{-1}(x')$, we have that $\phi^*(\theta') = \phi^A(\theta')$. To see that prices are constant on $O(x)$, take any two points $x_1, x_2 \in O(x)$, and observe that (by Equation 203 of Proposition 21):

$$T(x_1) - T(x_2) = \int_{x_2}^{x_1} u_x(z, \phi^{A^{-1}}(z))\, \mathrm{d}z = 0 \tag{1055}$$

by optimality of z for type $\phi^{A^{-1}}(z)$, which implies the necessary optimality condition, for all $z \in [x_2, x_1]$, $u_x(z, \phi^{A^{-1}}(z)) = 0$. It remains to consider all $x \notin \text{Int}(X^*)$. Continuity of H implies the result for the boundary points of X^*.[2] Thus, we have shown that, if $H(x) > 0$, then there exists a neighborhood of x such that prices are constant, proving the claim.

We now prove that, for every $x \in X^*$, if T^* is a multi-part tariff that is flat at x, then $H(x) \geq 0$. First, consider $x \in \text{Int}(X^*)$. If T is flat at x, then there exists a neighborhood $O(x)$ such that for all $x_1, x_2 \in O(x)$, we have that $T(x_1) - T(x_2) = 0$. Thus, by Equation 203 of Proposition 21, we have that $\int_{x_2}^{x_1} u_x(z, \phi^{*^{-1}}(z))\, \mathrm{d}z = 0$ for all $x_1, x_2 \in O(x)$. Thus, we have that $u_x(z, \phi^{*^{-1}}(z)) = 0$ (as $u_x(z, \phi^{*^{-1}}(z)) \geq 0$ by Obedience) for almost all $z \in O(x)$. By strict quasiconcavity of u, this implies that $\phi^{*^{-1}}(z) = \phi^{A^{-1}}(z)$ for almost all $z \in O(x)$. Toward a contradiction, suppose that $H(x) < 0$. By continuity of H, there exists a neighborhood $O'(x) \subseteq O(x)$ such that $\phi^{*^{-1}}(z) = \phi^{P^{-1}}(z) < \phi^{A^{-1}}(z)$ for all $z \in O'(x)$. But we have already shown that $\phi^{*^{-1}}(z) = \phi^{A^{-1}}(z)$ for almost all $z \in O(x)$. This is a contradiction, and so $H(x) \geq 0$. It remains to consider all $x \notin \text{Int}(X^*)$. As before, continuity of H implies the result for the boundary points of X^*.[3] □

Proof of Corollary 13. Immediate from Proposition 22 and the pricing-scheme definitions. □

Proof of Corollary 14. By Proposition 22, if $H(x) > 0$ at $\phi^A(1)$, then T^* is flat at

[2] A neighborhood at $\max X^*$ is of the form $(\max X^* - \varepsilon, \max X^*]$ for some $\varepsilon > 0$, and at $\min X^*$ of the form $[\min X^*, \min X^* + \varepsilon)$.

[3] A careful reader may ask why it is not true that T^* being flat at $x \in X^*$ implies $H(x) > 0$. Toward a counter-example to this claim, suppose that $\phi^P \equiv \phi^A$. We have that T^* is flat everywhere but $H(x) \equiv 0$.

$\phi^*(1) = \phi^A(1)$. Moreover, at $x = \phi^A(1)$, we have $H(\phi^A(1)) = \pi_x(\phi^A(1), 1)$. It follows that, when $\pi_x(\phi^A(1), 1) > 0$, $H(\phi^A(1)) > 0$ and T^* features an unlimited subscription. Likewise, if $H(x) > 0$ at $\phi^A(0)$, then T^* is flat at $\phi^*(0) = \phi^A(0)$. Moreover, at $x = \phi^A(0)$, we have $H(\phi^A(0)) = \pi_x(\phi^A(0), 0) - \frac{1}{f(0)} u_{x\theta}(\phi^A(0), 0)$. It follows that, when $f(0)\pi_x(\phi^A(0), 0) - u_{x\theta}(\phi^A(0), 0) > 0$, $H(\phi^A(0)) > 0$ and T^* features a trial. $\square$

F.1.3 Proofs of Propositions 23 and 24

Proof of Proposition 23. We first prove that $V(\theta; T) \geq V_N(\theta; T)$, for all $\theta \in \Theta$. We compare the values with and without free disposal to each type $\theta \in \Theta$ under any T:

$$V(\theta; T) = \sup_{y \in X, x \in [0,y]} \{u(x, \theta) - T(y)\} \geq \sup_{y \in X} \{u(y, \theta) - T(y)\} = V_N(\theta; T) \tag{1056}$$

because any payoff in the problem on the right of the inequality is attainable in the problem on the left of the inequality.

We now show that $V^*(\theta) \leq V_N^*(\theta)$, for all $\theta \in \Theta$. Without free disposal, the optimal allocation is $\phi_N^*(\theta) = \phi^P(\theta)$. With free disposal, the optimal allocation is $\phi^*(\theta) = \min\{\phi^A(\theta), \phi^P(\theta)\}$. It follows that $\phi^*(\theta) \leq \phi_N^*(\theta)$ for all $\theta \in \Theta$. Using the formula for agent welfare under the optimal mechanism (see Equation 1042), we can then see that:

$$V^*(\theta) = \int_0^\theta u_\theta(\phi^*(s), s)\, \mathrm{d}s \leq \int_0^\theta u_\theta(\phi_N^*(s), s)\, \mathrm{d}s = V_N^*(\theta) \tag{1057}$$

for all $\theta \in \Theta$, where the inequality follows as u is strictly single-crossing in (x, θ) and $\phi^* \leq \phi_N^*$.

For the seller, by Proposition 21, we have that for all $\theta \in \Theta$:

$$\Pi^*(\theta) = \max_{x \in [0, \phi^A(\theta)]} J(x, \theta) \leq \max_{x \in X} J(x, \theta) = \Pi_N^*(\theta) \tag{1058}$$

The inequality follows because the problem without free disposal allows for more choices of $x \in X$. $\square$

Proof of Proposition 24. We first show how J and ϕ^P change when (i) π changes to $\tilde{\pi}$ such that $\tilde{\pi}_x \geq \pi_x$ and (ii) F changes to $\tilde{F}$ such that F dominates $\tilde{F}$ in the hazard-rate order. Observe that $J(\cdot, \theta)$ increases pointwise for all $\theta \in \Theta$ as we may write (noting that $J(0, \theta) = 0$ by the properties that $u(0, \theta) \equiv \pi(0, \theta) \equiv 0$):

$$J(x, \theta) = \int_0^x \left[\pi_x(z, \theta) + u_x(z, \theta) - \frac{1 - F(\theta)}{f(\theta)} u_{x\theta}(z, \theta)\right] \mathrm{d}z \tag{1059}$$

and see that the integrand, $J_x(z, \theta)$, increases pointwise under (i) and (ii). As J_x increases pointwise and J is strictly quasiconcave, we moreover have that ϕ^P increases pointwise while ϕ^A remains unchanged. Let ϕ^P, J, V^*, and Π^* be evaluated at the original π and/or F, and $\tilde{\phi}^P$, $\tilde{J}$, $\tilde{V}^*$, and $\tilde{\Pi}^*$ be the same objects evaluated at the new $\tilde{\pi}$ and/or $\tilde{F}$.

We first study consumer welfare and establish that $\tilde{V}^* \geq V^*$. See that (by Equation 1042):

$$\tilde{V}^*(\theta) = \int_0^\theta u_\theta(\tilde{\phi}^*(s), s)\,\mathrm{d}s \geq \int_0^\theta u_\theta(\phi^*(s), s)\,\mathrm{d}s = V^*(\theta) \tag{1060}$$

for all $\theta \in \Theta$, where the inequality follows as u is strictly single-crossing in (x, θ) and $\tilde{\phi}^* = \min\{\tilde{\phi}^P, \phi^A\} \geq \min\{\phi^P, \phi^A\} = \phi^*$. Showing that $\tilde{V}_N^* - V_N^* \geq \tilde{V}^* - V^*$ is equivalent to showing that $\tilde{V}_N^* - \tilde{V}^* \geq V_N^* - V^*$, or (by Equation 1042):

$$\int_0^\theta \left[\left(u_\theta(\tilde{\phi}_N^*(s), s) - u_\theta(\tilde{\phi}^*(s), s)\right) - \left(u_\theta(\phi_N^*(s), s) - u_\theta(\phi^*(s), s)\right)\right] \mathrm{d}s \geq 0\,, \forall \theta \in \Theta \tag{1061}$$

There are three possible cases for each $s \in \Theta$ to compute the integrand:

i $\phi^P(s) < \phi^A(s)$ and $\tilde{\phi}^P(s) < \phi^A(s)$. Hence: $\phi^*(s) = \phi^P(s) = \phi_N^*(s)$ and $\tilde{\phi}^*(s) = \tilde{\phi}^P(s) = \tilde{\phi}_N^*(s)$. In this case, the value of the integrand is zero.

ii $\phi^P(s) < \phi^A(s)$ and $\tilde{\phi}^P(s) \geq \phi^A(s)$. Hence: $\phi^*(s) = \phi^P(s) = \phi_N^*(s)$ and $\tilde{\phi}^*(s) = \phi^A(s)$. Thus, the integrand is $u_\theta(\tilde{\phi}^P(s), s) - u_\theta(\phi^A(s), s) \geq 0$ by strict single-crossing of u.

iii $\phi^P(s) \geq \phi^A(s)$ and $\tilde{\phi}^P(s) \geq \phi^A(s)$, so $\phi^*(s) = \phi^A(s)$ and $\tilde{\phi}^*(s) = \phi^A(s)$. Thus, the value of the integrand is $u_\theta(\tilde{\phi}^P(s), s) - u_\theta(\phi^P(s), s) \geq 0$ by strict single-crossing of u.

Thus, the integrand is positive for all $s \in \Theta$ and the claimed inequality holds.

We now study producer welfare and establish that $\tilde{\Pi}^* \geq \Pi^*$. By Proposition 21, we have:

$$\tilde{\Pi}^*(\theta) = \tilde{J}(\tilde{\phi}^*(\theta), \theta) \geq \tilde{J}(\phi^*(\theta), \theta) \geq J(\phi^*(\theta), \theta) = \Pi^*(\theta) \tag{1062}$$

where the first inequality is by feasibility of ϕ^* before and after the change (as ϕ^A is unchanged), and the second inequality follows as $\tilde{J} \geq J$. Showing that $\tilde{\Pi}_N^*(\theta) - \Pi_N^*(\theta) \geq \tilde{\Pi}^*(\theta) - \Pi^*(\theta)$ is equivalent to showing that $\tilde{\Pi}_N^*(\theta) - \tilde{\Pi}^*(\theta) \geq \Pi_N^*(\theta) - \Pi^*(\theta)$, or:

$$\left(\tilde{J}(\tilde{\phi}_N^*(\theta),\theta) - \tilde{J}(\tilde{\phi}^*(\theta),\theta)\right) - \left(J(\phi_N^*(\theta),\theta) - J(\phi^*(\theta),\theta)\right) \geq 0\,, \forall\theta \in \Theta \tag{1063}$$

We have that $\phi_N^*(\theta) = \phi^P(\theta)$ and $\tilde{\phi}_N^*(\theta) = \tilde{\phi}^P(\theta)$, and there are three cases for each $\theta \in \Theta$:

i $\phi^P(\theta) < \phi^A(\theta)$ and $\tilde{\phi}^P(\theta) < \phi^A(\theta)$, so $\phi^*(\theta) = \phi^P(\theta) = \phi_N^*(\theta)$ and $\tilde{\phi}^*(\theta) = \tilde{\phi}^P(\theta) = \tilde{\phi}_N^*(\theta)$. In this case, the value of the expression is zero.

ii $\phi^P(\theta) < \phi^A(\theta)$ and $\tilde{\phi}^P(\theta) \geq \phi^A(\theta)$, so $\phi^*(\theta) = \phi^P(\theta) = \phi_N^*(\theta)$ and $\tilde{\phi}^*(\theta) = \phi^A(\theta)$. In this case, the value of the expression is $\tilde{J}(\tilde{\phi}^P(\theta),\theta) - \tilde{J}(\phi^A(\theta),\theta) \geq 0$ as $\tilde{\phi}^P$ is maximal for $\tilde{J}$.

iii $\phi^P(\theta) \geq \phi^A(\theta)$ and $\tilde{\phi}^P(\theta) \geq \phi^A(\theta)$, so $\phi^*(\theta) = \phi^A(\theta)$ and $\tilde{\phi}^*(\theta) = \phi^A(\theta)$. In this case, the value of the expression is $\left(\tilde{J}(\tilde{\phi}^P(\theta),\theta) - \tilde{J}(\phi^A(\theta),\theta)\right) - \left(J(\phi^P(\theta),\theta) - J(\phi^A(\theta),\theta)\right)$, and we wish to show that this is positive. Now observe that we can write this inequality as:

$$\int_{\phi^P(\theta)}^{\tilde{\phi}^P(\theta)} \tilde{J}_x(z,\theta)\,\mathrm{d}z + \int_{\phi^A(\theta)}^{\phi^P(\theta)} \left(\tilde{J}_x(z,\theta) - J_x(z,\theta)\right)\mathrm{d}z \geq 0 \tag{1064}$$

As $\tilde{J}$ is strictly quasiconcave and $\tilde{\phi}^P$ is $\tilde{J}$ maximal, we know that $\int_{\phi^P(\theta)}^{\tilde{\phi}^P(\theta)} \tilde{J}_x(z,\theta)\,\mathrm{d}z \geq 0$. Moreover, as $\tilde{J}_x \geq J_x$, we have that $\int_{\phi^A(\theta)}^{\phi^P(\theta)} \left(\tilde{J}_x(z,\theta) - J_x(z,\theta)\right)\mathrm{d}z \geq 0$. The claimed inequality follows.

Thus, the expression in (1063) is positive for all $\theta \in \Theta$ and the claimed inequality follows. □

F.2 Additional Results

F.2.1 Optimal Bunching and Free Disposal

This appendix extends our main analysis to cover cases in which the virtual surplus function J does not satisfy single-crossing and thereby allows for the possibility that multiple buyer types bunch on the same level of consumption. We apply techniques from Nöldeke and Samuelson (2007) to study the inverse problem of assigning types to consumption. For this reason, we make make the additional assumptions that J is concave and that both π_{xx} and $u_{xx\theta}$ exist and are continuous.

Denote an inverse consumption function by $\psi : X \to \Theta$. This corresponds to an inverse of the standard consumption function ϕ. For any monotone ψ, define the correspondence:

$$\Psi(x) = \left[\lim_{y \to^- x} \psi(y), \lim_{y \to^+ x} \psi(y)\right] \tag{1065}$$

which "fills in" discontinuities in the inverse consumption function.[4] Moreover, define the generalized inverse of ϕ^A as $\left(\phi^A\right)^{[-1]}(x) = \min\left\{\theta \in [0,1] : \phi^A(\theta) = x\right\}$. Our first result concerns implementation in this setting.

Lemma 27. *A consumption function ϕ is implementable and supported by (ϕ, T) if and only if there exists a monotone inverse consumption $\psi : X \to \Theta$ such that $\psi(x) \geq \left(\phi^A\right)^{[-1]}(x)$ for all $x \in X$, $\theta \in \Psi(\phi(\theta))$ for all $\theta \in \Theta$, and $T(x) = C + \int_0^x u_x(z, \psi(z))\,\mathrm{d}z$ with $C \leq 0$.*

Proof. By construction, (T, ψ) are consistent as defined in Equation 5 in Nöldeke and Samuelson (2007). Moreover, $\phi \leq \phi^A$ if and only if $\psi \geq (\phi^A)^{[-1]}$. Therefore, the statement immediately follows from Lemma 1 and Lemma 2 in Nöldeke and Samuelson (2007) and the proof of Lemma 25 in this paper. $\square$

We now provide the solution to the seller's screening problem. Toward simplifying the seller's problem, we define the following function:

$$\hat{J}(x, \theta) = u_x(x, \theta)(1 - F(\theta)) + \int_\theta^1 \pi_x(x, s)\,\mathrm{d}F(s) \tag{1067}$$

Using this function as well as Lemma 27 in this paper and Remark 1 and Lemma 5 in Nöldeke and Samuelson (2007), we can re-express the seller's problem as:

$$\begin{aligned} \max_{\psi} \quad & \int_0^{\overline{x}} \hat{J}(x, \psi(x))\,\mathrm{d}x \\ \text{s.t.} \quad & \psi(x') \geq \psi(x),\ \psi(x) \geq \left(\phi^A\right)^{[-1]}(x),\ \forall x', x \in X : x' \geq x \end{aligned} \tag{1068}$$

The following result solves this problem and uses the solution to solve Problem 198.

Proposition 55. *Problem 1068 is solved by the inverse consumption $\psi^* : X \to \Theta$ given by*

$$\psi^*(x) = \max\left\{\arg \max_{\theta \in [(\phi^A)^{[-1]}(x), 1]} \hat{J}(x, \theta)\right\} \tag{1069}$$

[4]Where we follow the convention from Nöldeke and Samuelson (2007) that:

$$\lim_{y \to^- 0} \psi(y) = 0, \quad \lim_{y \to^+ \overline{x}} \psi(y) = 1 \tag{1066}$$

Moreover, Problem 198 is solved by $\phi^*(\theta) = \inf\{x \in X : \psi^*(x) \geq \theta\}$ *for all* $\theta \in \Theta$.

Proof. We first show that $\hat{J}$ is supermodular. We follow Lemma 6 in Nöldeke and Samuelson (2007) and observe that the cross partial derivative of $\hat{J}$ is:

$$\hat{J}_{x\theta}(x,\theta) = -\left[u_{xx}(x,\theta) + \pi_{xx}(x,\theta)\right] f(\theta) + \left[1 - F(\theta)\right] u_{xx\theta}(x,\theta) = -J_{xx}(x,\theta) f(\theta) \geq 0 \tag{1070}$$

where the last inequality uses the concavity of J and $f > 0$. Next, we argue that the correspondence $x \mapsto [(\phi^A)^{[-1]}(x), 1]$ is monotone in the strong set order (SSO). This immediately follows from the fact that $(\phi^A)^{[-1]}$ is increasing. We then apply Theorem 4' in Milgrom and Shannon (1994) to argue that ψ^* is monotone. Since $\psi^* \geq (\phi^A)^{[-1]}$, the inverse consumption function ψ^* is implementable and therefore optimal in Problem 1068.

We now prove the optimality of ϕ^* in Problem 198. Given that ψ^* is monotone and such that $\psi^* \geq (\phi^A)^{[-1]}$, it follows that ϕ^* is also monotone and such that $\phi^* \leq \phi^A$. Hence, by Lemma 25, it is implementable. Next, suppose there exists an implementable consumption function ϕ such that $\int_0^1 J(\phi(\theta),\theta)\,\mathrm{d}F(\theta) > \int_0^1 J(\phi^*(\theta),\theta)\,\mathrm{d}F(\theta)$. Given that ϕ is implementable, there exist (ξ, T) that support it. By the proof of Lemma 1 in Nöldeke and Samuelson (2007) it follows that there exists an inverse consumption function ψ such that $T(x) = T(0) + \int_0^x u(z,\psi(z))\,\mathrm{d}z$. In turn, Lemma 3 in Nöldeke and Samuelson (2007) implies that

$$\int_0^{\overline{x}} \hat{J}(x,\psi^*(x))\,\mathrm{d}x = \int_0^1 J(\phi^*(\theta),\theta)\,\mathrm{d}F(\theta) < \int_0^1 J(\phi(\theta),\theta)\,\mathrm{d}F(\theta) = \int_0^{\overline{x}} \hat{J}(x,\psi(x))\,\mathrm{d}x \tag{1071}$$

contradicting the optimality of ψ^* in Problem 1068. Therefore, ϕ^* solves Problem 198. $\square$

As in Nöldeke and Samuelson (2007), bunching manifests in the solution to this problem as a discontinuity in the resulting inverse consumption function ψ. In particular, whenever ψ is discontinuous the outcome at the discontinuity is assigned to a positive measure of types.

As explained in Remark 3, these bunching regions in the type space do not generate flat regions of the optimal price schedule. Similarly to Proposition 22, we can fully characterize the regions where T^* is flat. These are the quantities x where the local constraint $\theta \in [(\phi^A)^{[-1]}(x), 1]$ in (1069) binds at the optimum. However, here we do not assume strict concavity of $\hat{J}$ since this would be equivalent to assuming strict supermodularity of J. Therefore, we cannot rely on first-order conditions to replicate

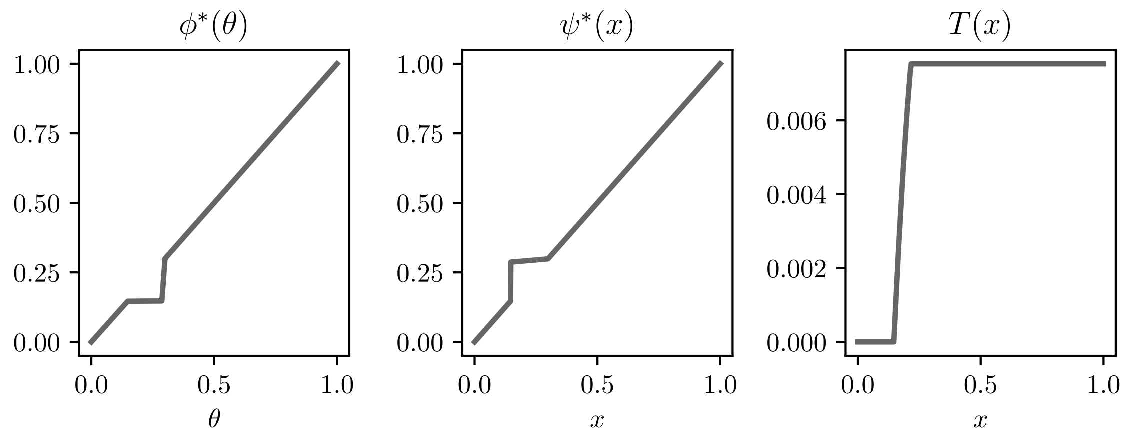


Figure F-1: Multi-part tariff with bunching in Example 11.

the local characterization of Proposition 22. We conclude with an example in which the optimal contract features both bunching and a multi-part tariff:

Example 11. *Consumer preferences, the outcome space, and the external revenue function are identical to those in Example 4. The type distribution has density*

$$f(\theta) = 1 + \frac{k}{2\pi\omega}(\cos(2\pi\omega) - 1) + k\sin(2\pi\omega\theta) \tag{1072}$$

for some $k > 0$ and $\omega > 0$. We solve the example for $\alpha = 1$, $\beta = 0$, $k = \frac{1}{2}$, and $\omega = 3$. In Figure F-1, we plot $\phi^*(\theta)$, $\psi^*(x)$, and $T(x)$ in the optimal contract. In the price schedule, there is both an unlimited subscription and a free trial. A mass of types, approximately between 0.15 and 0.29, is bunched at the allocation $\phi^* = 0.15$. These types all consume the maximum amount possible in the free trial. Anecdotally, this is a common occurrence for free trials in practice (e.g., the free allotment of online Wall Street Journal articles).

F.2.2 Competition and Free Disposal

In this appendix, we study the relationship between competition and optimal pricing under free disposal. We do this by comparing our monopoly screening benchmark with one model of perfect competition. We show that our results are robust to this extension by demonstrating that zero marginal pricing is in fact more prevalent under perfect competition.

The nature of perfect competition we consider is that our monopolist faces a perfectly competitive fringe of firms that can enter and displace them to serve the entire market. In this case (as in, e.g., Grubb, 2009), the equilibrium contract maximizes expected consumer surplus subject to our usual implementation constraints and a new constraint that the monopolist actually wishes to serve the market. That is, the screening problem becomes:

$$\begin{aligned}\sup_{\phi,\xi,T} \quad & \int_{\Theta} \left(u(\phi(\theta),\theta) - T(\xi(\theta))\right) \mathrm{d}F(\theta) \\ \text{s.t.} \quad & \text{(O),(IC),(IR)} \\ & \int_{\Theta} \left(\pi(\phi(\theta),\theta) + T(\xi(\theta))\right) \mathrm{d}F(\theta) \geq 0\end{aligned} \tag{1073}$$

The last constraint, which we call "Monopolist's IR," encodes the requirement that the monopolist wishes to serve the market compared to an outside option of earning nothing.

Toward characterizing the solution of this problem, define the total surplus function as $S(x,\theta) = \pi(x,\theta) + u(x,\theta)$. In analogy to our assumptions that J is strictly single-crossing and strictly quasiconcave, we assume that S is strictly single-crossing in (x,θ) and strictly quasiconcave in x. We further define the total surplus maximizing consumption level as $\phi^O(\theta) = \arg\max_{x\in X} S(x,\theta)$.

Proposition 56. *The equilibrium consumption under perfect competition is* $\phi^{PC} = \min\{\phi^A, \phi^O\}$.

Proof. As in the proof of Lemma 26, we have that agents' transfers under any locally incentive compatible menu are given by Equation 1036 for some $C \in \mathbb{R}$. We can therefore rewrite the objective (using the same integration-by-parts argument as Lemma 26) as:

$$-C + \int_{\Theta} \frac{1-F(\theta)}{f(\theta)} u_\theta(\phi(\theta),\theta) \, \mathrm{d}F(\theta) \tag{1074}$$

By integrating over types, we can then express the monopolist's IR constraint as:

$$\int_{\Theta} \left(\pi(\phi(\theta),\theta) + u(\phi(\theta),\theta) - \frac{1-F(\theta)}{f(\theta)} u_\theta(\phi(\theta),\theta)\right) \mathrm{d}F(\theta) + C \geq 0 \tag{1075}$$

Thus, the optimal C sets this inequality tight. Substituting, we obtain that the objective function becomes $\int_\Theta S(\phi(\theta),\theta)\,\mathrm{d}F(\theta)$. Moreover, by the same arguments as in Lemma 25, the remaining implementation constraints are that $\phi(\theta) \leq \phi^A(\theta)$ for all $\theta \in \Theta$, ϕ is monotone increasing and $u(\phi(0),0) - t(0) \geq 0$. By identical arguments to Proposition 21 (as S is strictly single-crossing and quasiconcave), it follows that the optimal consumption levels satisfy $\phi^{PC}(\theta) = \min\{\phi^A(\theta), \phi^O(\theta)\}$, which is monotone. Moreover, $t(0) = C + u(\phi^{PC}(0),0) \leq 0$ as C is negative and $u(\phi^{PC}(0),0) \geq 0$ as $\phi^{PC}(0) \in [0, \phi^A(0)]$. $\square$

We now show that zero marginal pricing is more prevalent under perfect competition:

Corollary 18. *The set of outcomes at which there is flat pricing under perfect competition includes the set of outcomes at which there is flat pricing under monopoly pricing.*

Proof. Define $H^{PC}(x) = S_x\left(x, \left(\phi^A\right)^{-1}(x)\right)$ for all $x \in X^*$ and observe that:

$$
\begin{aligned}
&H^{PC}(x) = S_x\left(x, \left(\phi^A\right)^{-1}(x)\right) = u_x\left(x, \left(\phi^A\right)^{-1}(x)\right) + \pi_x\left(x, \left(\phi^A\right)^{-1}(x)\right) \\
&\geq u_x\left(x, \left(\phi^A\right)^{-1}(x)\right) + \pi_x\left(x, \left(\phi^A\right)^{-1}(x)\right) - \frac{1 - F\left(\left(\phi^A\right)^{-1}(x)\right)}{f\left(\left(\phi^A\right)^{-1}(x)\right)} u_{x\theta}\left(x, \left(\phi^A\right)^{-1}(x)\right) \\
&= J_x\left(x, \left(\phi^A\right)^{-1}(x)\right) = H(x)
\end{aligned}
\tag{1076}
$$

Thus, $H(x) \geq 0 \implies H^{PC}(x) \geq 0$. Hence, by an identical argument to Proposition 22, whenever T^* is flat, so is T^{PC}. □

The intuition for this result is that there are no quantity distortions from information rents under the competitive solution. Thus, total-surplus-maximizing consumption is greater than virtual-surplus-maximizing consumption, and the constraint $\phi \leq \phi^A$ binds more often.

F.3 Microfoundations of Revenue from Usage

F.3.1 Network Effects from Platform Externalities

Sellers may value usage because it makes the platform more valuable for other end users. That is, usage generates network effects. Examples include networking services (e.g., LinkedIn), matching services (e.g., Tinder, Match.com, or OK Cupid), online games (e.g., Fortnite, Candy Crush Saga, or World of Warcraft), and content-streaming platforms with social rating systems (e.g., Hulu or Netflix).

The function $W : X \times \Theta \to \mathbb{R}_+$ maps each agent's consumption to a positive externality for every agent. Agents' payoffs if they participate, given a consumption function ϕ, are:

$$
v\left(x, \theta, (\phi(s))_{s\in[0,1]}\right) = u(x, \theta) + \int_0^1 W(\phi(s), s)\, \mathrm{d}F(s) \tag{1077}
$$

with the maintained assumption of a zero outside option otherwise. The rest of the model is as in Section 6.2. The externality of others' usage is obtained by an agent whenever they use the platform at the extensive margin. This makes the model amenable to settings where an agent may gain from participating, even if they do not regularly use the platform. For example, having a LinkedIn profile may generate the "passive" benefit of being findable by job recruiters, even if the user spends essentially zero time using the website. In analogy to the main analysis, we assume that the modified virtual surplus function

$$J^{\dagger}(x,\theta) = \pi(x,\theta) + u(x,\theta) + W(x,\theta) - \frac{1-F(\theta)}{f(\theta)} u_{\theta}(x,\theta) \tag{1078}$$

is strictly quasiconcave in x and strictly single-crossing in (x,θ). We now show how this setting maps to our baseline setting of Section 6.2.

Lemma 28. *Optimal consumption is given by* $\phi^*(\theta) = \min\{\phi^A(\theta), \phi^P(\theta)\}$*, where* $\phi^A(\theta) = \arg\max_{x\in X} u(x,\theta)$ *and* $\phi^P(\theta) = \arg\max_{x\in X} J^{\dagger}(x,\theta)$*.*

Proof. Observe first that the externality cannot affect the Obedience or Incentive Compatibility constraints since it has no dependence on consumer choice. The Individual Rationality constraint becomes $v\left(\phi(\theta),\theta,(\phi(s))_{s\in[0,1]}\right) \geq 0$. The same arguments from the proof of Lemma 26 imply that, without loss of optimality, we can restrict attention to allocations in which all agents participate (as $W \geq 0$), but now where $C = \int_{\Theta} W(\phi(\theta),\theta)\,\mathrm{d}F(\theta)$. Thus, by Equation 1049, the objective of the monopolist is now $\int_{\Theta} J^{\dagger}(x,\theta)\,\mathrm{d}F(\theta)$ and the constraints are the same as those in Equation 1047. The result then follows by application of the arguments in the proof of Proposition 21. □

Intuitively, since the externality is excludable, or not available to agents that do not participate in the mechanism, the seller can extract the full value of the externality as part of a "participation fee." Thus, each agent's marginal contribution to the externality, $W(x,\theta)$, is "as if" additional usage-derived revenue.

F.3.2 Irrational Addiction

Addicted users are commonly cited as a major source of revenue for digital goods (see, e.g., Allcott, Gentzkow, and Song, 2022). In this appendix, we describe a simple microfoundation of how external revenue could be derived from irrational addiction of consumers.

Suppose that agents live for two periods but are myopic. Let $x \in X$ be the agent's consumption today ($t = 0$) and $\tilde{x} \in X$ their consumption tomorrow ($t = 1$).

An agent of type $\theta \in \Theta$ believes they have lifetime payoff from consumption x given by $u(x, \theta)$, where u satisfies our running assumptions. In reality, however, the agent also values consumption tomorrow. Moreover, the more (or less) that they consumed today the more (or less) they value consumption tomorrow. Thus, at $t = 1$, the agent has utility function $\tilde{u} : X^2 \times \Theta \to \mathbb{R}$, where $u(x, \tilde{x}, \theta)$ is their payoff. This *complete myopia* can be thought of as an extreme form of the inattention toward habit formation that Allcott, Gentzkow, and Song (2022) find is necessary to empirically rationalize the total demand for six ubiquitous mobile apps (Facebook, Instagram, Twitter, Snapchat, web browsers, and YouTube).

Observe that given a full-revelation mechanism (or equivalently under observation of agent consumption under an implementable mechanism), the seller knows the agent's type tomorrow. Thus, when agents consume x today and their type is θ, tomorrow the monopolist sells them $\tilde{x}^*(x, \theta) \in \arg\max_{\tilde{x} \in X} \tilde{u}(x, \tilde{x}, \theta)$ and charges a transfer of $\pi(x, \theta) = \tilde{u}(x, \tilde{x}^*(x, \theta), \theta)$ to extract full surplus. Thus, from the perspective of today, the monopolist faces the non-linear pricing problem we study in the main text, with an external revenue function π that captures the gains from addicting a user through contemporaneous consumption and extracting this surplus from them in the future.

F.3.3 Overconfidence

A natural reason why a seller may allocate more of a good than an agent wants *ex post* is that the agent expected to want something different *ex ante*. This story is at the heart of Grubb (2009)'s analysis of selling to overconfident consumers and his leading example of pricing cell phone plans, a context in which individuals regularly (based on anecdotes and empirical exploration) underestimate the variance of their future demand (see also Grubb and Osborne, 2015). We now illustrate how over-confidence at the participation stage can be mapped to our framework as a particular external revenue function.

The Grubb (2009) model is a monopoly pricing model, with continuous, increasing, and convex production costs $K(x)$ and no additional revenue from usage. The twist relative to the standard model is that agents decide whether to participate *ex ante* without knowing their type θ, but with a prior belief $\theta \sim \check{F}$ which may differ from the objective truth $\theta \sim F$ (see Grubb (2009) for the full details of the model). The common individual rationality constraint for all consumers is that the expected payoff at the allocation $(\phi(\theta), \xi(\theta), t(\theta))_{\theta \in \Theta}$ exceeds the outside option 0, or

$$\int_0^1 (u(\phi(\theta), \theta) - t(\theta))\, \mathrm{d}\check{F}(\theta) \geq 0 \tag{1079}$$

We derive the following mapping of the Grubb (2009) model into ours:

Lemma 29. *The optimal consumption in the monopoly problem of Grubb (2009) is equal to the consumption that solves Problem 198, with* $\pi(x, \theta) = \frac{1-\check{F}(\theta)}{f(\theta)} u_\theta(x, \theta) - K(x)$

Proof. This follows immediately from our Lemma 26 and Proposition 1 in Grubb (2009). ☐

Observe first that, in a classical model with correctly specified expectations $\check{F} = F$, the first term in π cancels with the information rents and the Obedience constraint is always slack in the optimum. With mis-specified $\check{F} \neq F$, the first term and information rents do not cancel. When the first term stemming from overconfidence dominates both production costs and information rents on the margin at $x \in X$, the model generates $H(x) > 0$ and multi-part tariffs. Grubb (2009) applies this model to understand the occurrence of trial tiers (in our language) in cell phone pricing.

List of Figures

List of Tables

Bibliography

Abdulkadiroğlu, A., J. D. Angrist, Y. Narita, and P. A. Pathak (2017): "Research design meets market design: Using centralized assignment for impact evaluation," *Econometrica*, 85(5), 1373–1432.

Abdulkadiroğlu, A., Y.-K. Che, and Y. Yasuda (2015): "Expanding 'choice' in school choice," *American Economic Journal: Microeconomics*, 7(1), 1–42.

Abdulkadiroğlu, A., P. A. Pathak, and A. E. Roth (2005): "The new york city high school match," *American Economic Review*, 95(2), 364–367.

Abdulkadiroğlu, A. (2005): "College Admission with Affirmative Action," *International Journal of Game Theory*, 33, 535–549.

Abdulkadiroğlu, A., and T. Sönmez (2003): "School Choice: A Mechanism Design Approach," *American Economic Review*, 93, 729–747.

Abel, A. B., J. C. Eberly, and S. Panageas (2013): "Optimal Inattention to the Stock Market With Information Costs and Transactions Costs," *Econometrica*, 81(4), 1455–1481.

Acemoglu, D., V. Chernozhukov, and M. Yildiz (2016): "Fragility of asymptotic agreement under Bayesian learning," *Theoretical Economics*, 11(1), 187–225.

Acemoglu, D., and J. A. Robinson (2021): "Culture, institutions and social equilibria: A framework," Discussion Paper 28832, National Bureau of Economic Research.

Adam, K., A. Marcet, and J. Beutel (2017): "Stock price booms and expected capital gains," *American Economic Review*, 107(8), 2352–2408.

Akbarpour, M., E. B. Budish, P. Dworczak, and S. D. Kominers (2021): "An economic framework for vaccine prioritization," *Available at SSRN 3846931*.

Allcott, H., M. Gentzkow, and L. Song (2022): "Digital addiction," *American Economic Review*, 112(7), 2424–63.

Allen, F., S. Morris, and H. S. Shin (2006): "Beauty contests and iterated expectations in asset markets," *The Review of Financial Studies*, 19(3), 719–752.

Allen, R., and J. Rehbeck (2021): "A Generalization of Quantal Response Equilibrium via Perturbed Utility," *Games*, 12(1), 20.

Altonji, J. G., T. E. Elder, and C. R. Taber (2005): "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of Political Economy*, 113(1), 151–184.

Alvarez, F., M. Beraja, M. Gonzalez-Rozada, and P. A. Neumeyer (2019): "From hyperinflation to stable prices: Argentina's evidence on menu cost models," *The Quarterly Journal of Economics*, 134(1), 451–505.

ALVAREZ, F. E., F. LIPPI, AND L. PACIELLO (2011): "Optimal price setting with observation and menu costs," *The Quarterly journal of economics*, 126(4), 1909–1960.

ALVES, F., G. KAPLAN, B. MOLL, AND G. L. VIOLANTE (2020): "A further look at the propagation of monetary policy shocks in HANK," *Journal of Money, Credit and Banking*, 52(S2), 521–559.

AMELIO, A., AND B. JULLIEN (2012): "Tying and freebies in two-sided markets," *International Journal of Industrial Organization*, 30(5), 436–446.

ANDRE, P., I. HAALAND, C. ROTH, AND J. WOHLFART (2022): "Narratives about the macroeconomy," Discussion Paper DP17305, CEPR.

ANGELETOS, G.-M., F. COLLARD, AND H. DELLAS (2018): "Quantifying confidence," *Econometrica*, 86(5), 1689–1726.

ANGELETOS, G.-M., AND J. LA'O (2010): "Noisy business cycles," *NBER Macroeconomics Annual*, 24(1), 319–378.

——— (2013): "Sentiments," *Econometrica*, 81(2), 739–779.

ANGELETOS, G.-M., AND C. LIAN (2016): "Incomplete information in macroeconomics: Accommodating frictions in coordination," in *Handbook of Macroeconomics*, vol. 2, pp. 1065–1240. Elsevier.

ANGELETOS, G.-M., AND A. PAVAN (2007): "Efficient use of information and social value of information," *Econometrica*, 75(4), 1103–1142.

ANGELETOS, G.-M., AND K. A. SASTRY (2023): "Inattentive Economies," Working Paper w26413, Northwestern University.

ANGRIST, J. D., P. A. PATHAK, AND R. A. ZÁRATE (2019): "Choice and consequence: Assessing mismatch at chicago exam schools," Discussion paper, National Bureau of Economic Research.

ARNOSTI, N., AND P. SHI (2017): "Design of Lotteries and Waitlists for Affordable Housing Allocation," *Columbia Business School Research Paper*, 66(17-52), 2291–2307.

ASH, E., G. GAUTHIER, AND P. WIDMER (2021): "RELATIO: Text semantics capture political and economic narratives," Discussion Paper 2108.01720, arXiv.

ASHLAGI, I., AND P. SHI (2014): "Improving community cohesion in school choice via correlated-lottery implementation," *Operations Research*, 62(6), 1247–1264.

ATHEY, S., C. AVERY, AND P. ZEMSKY (2000): "Mentoring and diversity," *American Economic Review*, 90(4), 765–786.

AYGUN, O., AND I. BÓ (2021): "College admission with multidimensional privileges: The Brazilian affirmative action case," *American Economic Journal: Microeconomics*, 13(3), 1–28.

AYGÜN, O., AND B. TURHAN (2020): "Dynamic reserves in matching markets," *Journal of Economic Theory*, 188, 105069.

AZEVEDO, E. M., AND J. D. LESHNO (2016): "A supply and demand framework for two-sided matching markets," *Journal of Political Economy*, 124(5), 1235–1268.

BACHMANN, R., AND C. BAYER (2014): "Investment dispersion and the business cycle," *American Economic Review*, 104(4), 1392–1416.

BAKER, S. R., N. BLOOM, AND S. J. DAVIS (2016): "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131(4), 1593–1636.

Balinski, M., and T. Sönmez (1999): "A Tale of Two Mechanisms: Student Placement," *Journal of Economic Theory*, 84, 73–94.

Barnichon, R. (2010): "Productivity and unemployment over the business cycle," *Journal of Monetary Economics*, 57(8), 1013–1025.

Barrero, J. M. (2022): "The micro and macro of managerial beliefs," *Journal of Financial Economics*, 143(2), 640–667.

Basu, S., and J. G. Fernald (1997): "Returns to scale in US production: Estimates and implications," *Journal of Political Economy*, 105(2), 249–283.

Battaglini, M. (2021): "Chaos and Unpredictability in Dynamic Social Problems," Discussion Paper 28347, National Bureau of Economic Research.

Baxter, M., and R. G. King (1999): "Measuring business cycles: approximate band-pass filters for economic time series," *Review of Economics and Statistics*, 81(4), 575–593.

Beaudry, P., D. Galizia, and F. Portier (2020): "Putting the cycle back into business cycle analysis," *American Economic Review*, 110(1), 1–47.

Beaudry, P., and F. Portier (2006): "Stock prices, news, and economic fluctuations," *American Economic Review*, 96(4), 1293–1307.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012): "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, 80(6), 2369–2429.

Belloni, A., V. Chernozhukov, C. Hansen, and D. Kozbur (2016): "Inference in high-dimensional panel models with an application to gun control," *Journal of Business & Economic Statistics*, 34(4), 590–605.

Benhabib, J., X. Liu, and P. Wang (2016): "Endogenous information acquisition and countercyclical uncertainty," *Journal of Economic Theory*, 165, 601–642.

Benhabib, J., P. Wang, and Y. Wen (2015): "Sentiments and aggregate demand fluctuations," *Econometrica*, 83(2), 549–585.

Benhima, K. (2019): "Booms and busts with dispersed information," *Journal of Monetary Economics*, 107, 32–47.

Bergemann, D., A. Bonatti, and A. Smolin (2018): "The design and price of information," *American Economic Review*, 108(1), 1–48.

Bergemann, D., and S. Morris (2013): "Robust predictions in games with incomplete information," *Econometrica*, 81(4), 1251–1308.

Berger, D., and J. Vavra (2019): "Shocks versus Responsiveness: What Drives Time-Varying Dispersion?," *Journal of Political Economy*, 127(5), 2104–2142.

Berger, D. W., L. Bocola, and A. Dovis (2023): "Imperfect risk-sharing and the business cycle," *The Quarterly Journal of Economics*, forthcoming.

Berry, S. T., and P. A. Haile (2021): "Foundations of demand estimation," in *Handbook of Industrial Organization*, vol. 4, pp. 1–62. Elsevier.

Bhandari, A., J. Borovička, and P. Ho (2019): "Survey data and subjective beliefs in business cycle models," Working Paper 2763942, SSRN.

Blanchard, O., and J. Galí (2010): "Labor Markets and Monetary Policy: A New Keynesian Model with Unemployment," *American Economic Journal: Macroeconomics*, 2(2), 1–30.

Blanchard, O. J., and N. Kiyotaki (1987): "Monopolistic competition and the effects of aggregate demand," *The American Economic Review*, 77(4), 647–666.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003): "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, pp. 993–1022.

Bloch, F., and D. Cantala (2017): "Dynamic assignment of objects to queuing agents," *American Economic Journal: Microeconomics*, 9(1), 88–122.

Bloom, N. (2009): "The impact of uncertainty shocks," *Econometrica*, 77(3), 623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Saporta-Eksten, and S. J. Terry (2018): "Really Uncertain Business Cycles," *Econometrica*, 86(3), 1031–1065.

Bohren, J. A., and D. N. Hauser (2021): "Learning with heterogeneous misspecified models: Characterization and robustness," *Econometrica*, 89(6), 3025–3077.

Boldrin, M., and M. Woodford (1990): "Equilibrium models displaying endogenous fluctuations and chaos: a survey," *Journal of Monetary Economics*, 25(2), 189–222.

Bordalo, P., N. Gennaioli, S. Y. Kwon, and A. Shleifer (2021): "Diagnostic bubbles," *Journal of Financial Economics*, 141(3), 1060–1077.

Bordalo, P., N. Gennaioli, A. Shleifer, and S. J. Terry (2021): "Real credit cycles," Working Paper 28416, National Bureau of Economic Research.

Border, K. C. (1989): *Fixed point theorems with applications to economics and game theory*. Cambridge university press.

Boston Herald (2022): "New Boston point system dings students for good grades," *Boston Herald*, January 18, 2022.

Brock, W. A., and C. H. Hommes (1997): "A rational route to randomness," *Econometrica*, 65(5), 1059–1095.

Brown, G., and N. Kapadia (2007): "Firm-specific risk and equity market development," *Journal of Financial Economics*, 84(2), 358–388.

Budish, E., Y.-K. Che, F. Kojima, and P. Milgrom (2013): "Designing random allocation mechanisms: Theory and applications," *American economic review*, 103(2), 585–623.

Burnside, C., M. Eichenbaum, and S. Rebelo (2016): "Understanding booms and busts in housing markets," *Journal of Political Economy*, 124(4), 1088–1147.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu (2021a): "Business news and business cycles," Working Paper 29344, National Bureau of Economic Research.

——— (2021b): "Business News and Business Cycles," Working Paper 29344, National Bureau of Economic Research.

Caballero, R. J., and M. L. Hammour (1994): "The cleansing effect of recessions," *The American Economic Review*, 84(5), 1350–1368.

Caballero, R. J., and A. Simsek (2020): "A risk-centric model of demand recessions and speculation," *The Quarterly Journal of Economics*, 135(3), 1493–1566.

Caplin, A., and M. Dean (2013): "Behavioral implications of rational inattention with shannon entropy," Discussion paper, National Bureau of Economic Research.

——— (2015): "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105(7), 2183–2203.

Caplin, A., M. Dean, and J. Leahy (2019): "Rational inattention, optimal consideration sets, and stochastic choice," *The Review of Economic Studies*, 86(3), 1061–1094.

——— (2022): "Rationally inattentive behavior: Characterizing and generalizing Shannon entropy," *Journal of Political Economy*, 130(6), 1676–1715.

Caplin, A., and D. Martin (2015): "A testable theory of imperfect perception," *The Economic Journal*, 125(582), 184–202.

Carbajal, J. C., and J. C. Ely (2013): "Mechanism design without revenue equivalence," *Journal of Economic Theory*, 148(1), 104–133.

Carlsson, H., and E. Van Damme (1993): "Global games and equilibrium selection," *Econometrica*, pp. 989–1018.

Carroll, C. D. (2001): "The epidemiology of macroeconomic expectations," Working Paper 8695, National Bureau of Economic Research.

Carroll, C. D., and T. Wang (2022): "Epidemiological Expectations," in *Handbook of Economic Expectations*, vol. 1, p. Forthcoming. Elsevier.

Çelebi, O., and J. P. Flynn (2021): "Priorities *vs.* Quotas," *Available at SSRN 3562665*.

CDC (2018): "Interim updated planning guidance on allocating and targeting pandemic influenza vaccine during an influenza pandemic," https://www.cdc.gov/flu/pandemic-resources/pdf/2018-Influenza-Guidance.pdf.

Çelebi, O. (2022): "Diversity Preferences, Affirmative Action and Choice Rules," *Working Paper*.

Çelebi, O., and J. P. Flynn (2022): "Priority design in centralized matching markets," *The Review of Economic Studies*, 89(3), 1245–1277.

Cerreia-Vioglio, S., R. Corrao, and G. Lanzani (2020): "Robust opinion aggregation and its dynamics," Working paper, IGIER, Università Bocconi.

Cesarini, D., E. Lindqvist, M. J. Notowidigdo, and R. Östling (2017): "The effect of wealth on individual and household labor supply: evidence from Swedish lotteries," *American Economic Review*, 107(12), 3917–46.

Chahrour, R., and R. Ulbricht (2023): "Robust predictions for DSGE models with incomplete information," *American Economic Journal: Macroeconomics*, 15(1), 173–208.

Chan, J., and E. Eyster (2003): "Does banning affirmative action lower college student quality?," *American Economic Review*, 93(3), 858–872.

Charoenwong, B., Y. Kimura, A. Kwan, and E. Tan (2021): "Investment Plans, Uncertainty, and Misallocation," Working Paper 3662347, SSRN.

Chiang, Y.-T. (2021): "Strategic Uncertainty over Business Cycles," Mimeo, University of Chicago.

Chicago Public Schools (2022): "Selective Enrollment Admissions Process: Community Feedback," *Chicago Public Schools*, March 9, 2022.

Choi, J. P., and D.-S. Jeon (2021): "A Leverage Theory of Tying in Two-Sided Markets with Nonnegative Price Constraints," *American Economic Journal: Microeconomics*, 13(1), 283–337.

Christiano, L. J., C. L. Ilut, R. Motto, and M. Rostagno (2008): "Monetary policy and stock market boom-bust cycles," *ECB working paper*.

Coibion, O., Y. Gorodnichenko, and S. Kumar (2018a): "How Do Firms Form Their Expectations? New Survey Evidence," *American Economic Review*, 108(9), 2671–2713.

——— (2018b): "Replication data for: How Do Firms Form Their Expectations? New Survey Evidence," American Economic Association (Publisher) and Inter-university Consortium for Political and Social Research (Distributor). Available at: https://doi.org/10.3886/E113095V1.

Collinson, R., I. G. Ellen, and J. Ludwig (2015): "Low-income housing policy," in *Economics of Means-Tested Transfer Programs in the United States, Volume 2*, pp. 59–126. University of Chicago Press.

Corrao, R., J. P. Flynn, and K. A. Sastry (2023): "Nonlinear Pricing with Underutilization: A Theory of Multi-part Tariffs," *American Economic Review*, 113(3), 836–860.

Costain, J., and A. Nakov (2015): "Precautionary price stickiness," *Journal of Economic Dynamics and Control*, 58, 218–234.

——— (2019): "Logit price dynamics," *Journal of Money, Credit and Banking*, 51(1), 43–78.

Daley, B. (1999): "Plan Drops Race Role in Enrollment, Compromise Misses Point, Critics Say," *Boston Globe B*, 1.

David, J. M., H. A. Hopenhayn, and V. Venkateswaran (2016): " Information, Misallocation, and Aggregate Productivity," *The Quarterly Journal of Economics*, 131(2), 943–1005.

David, J. M., and V. Venkateswaran (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109(7), 2531–67.

Davis, S. J., J. Haltiwanger, R. Jarmin, J. Miranda, C. Foote, and E. Nagypal (2006): "Volatility and dispersion in business growth rates: Publicly traded versus privately held firms," *NBER Macroeconomics Annual*, 21, 107–179.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The rise of market power and the macroeconomic implications," *The Quarterly Journal of Economics*, 135(2), 561–644.

Dean, M., and N. L. Neligh (2022): "Experimental Tests of Rational Inattention," Working paper, Columbia University.

Decker, R. A., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2020): "Changing business dynamism and productivity: Shocks versus responsiveness," *American Economic Review*, 110(12), 3952–90.

Demirer, M. (2020): "Production function estimation with factor-augmenting technology: An application to markups," Mimeo, MIT.

Deng, L., M. A. Khan, and T. Mitra (2022): "Continuous unimodal maps in economic dynamics: On easily verifiable conditions for topological chaos," *Journal of Economic Theory*, 201, 105446.

Denti, T. (2018): "Posterior-separable cost of information," Discussion Paper 10, working paper.

Denti, T. (2023): "Unrestricted information acquisition," *Theoretical Economics*, forthcoming.

Dew-Becker, I., and S. Giglio (2020): "Cross-sectional uncertainty and the business cycle: evidence from 40 years of options data," Discussion Paper w27864, National Bureau of Economic Research.

Doğan, B. (2016): "Responsive affirmative action in school choice," *Journal of Economic Theory*, 165, 69–105.

Dur, U., S. D. Kominers, P. A. Pathak, and T. Sönmez (2018): "Reserve design: Unintended consequences and the demise of Boston's walk zones," *Journal of Political Economy*, 126(6), 2457–2479.

Dur, U., P. A. Pathak, and T. Sönmez (2020): "Explicit vs. statistical targeting in affirmative action: theory and evidence from Chicago's exam schools," *Journal of Economic Theory*, 187, 104996.

Dur, U., M. Thayer, and W. Phan (2022): "Allocating Medical Resources During a Pandemic," *Working Paper*.

Echenique, F., and M. B. Yenmez (2015): "How to control controlled school choice," *American Economic Review*, 105(8), 2679–94.

Edmond, C., V. Midrigan, and D. Y. Xu (2018): "How Costly Are Markups?," Working Paper 24800, National Bureau of Economic Research.

Ehlers, L., I. E. Hafalir, M. B. Yenmez, and M. A. Yildirim (2014): "School Choice with Controlled Choice Constraints: Hard Bounds versus Soft Bounds," *Journal of Economic Theory*, 153, 648–683.

Eisfeldt, A. L., and A. A. Rampini (2006): "Capital reallocation and liquidity," *Journal of Monetary Economics*, 53(3), 369–399.

Eliaz, K., and R. Spiegler (2020): "A model of competing narratives," *American Economic Review*, 110(12), 3786–3816.

Ellison, G., and P. A. Pathak (2016): "The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools," *American Economic Review*, 111(3), 943–975.

——— (2021): "The Efficiency of Race-Neutral Alternatives to Race-Based Affirmative Action: Evidence from Chicago's Exam Schools," *American Economic Review*, 111(3), 943–75.

eMarketer Insider Intelligence (2020): "Google's US Ad Revenues to Drop for the First Time," *Newsroom*, June 22. Retrieved from: https://www.emarketer.com/newsroom/index.php/google-ad-revenues-to-drop-for-the-first-time/.

Epple, D., R. Romano, and H. Sieg (2008): "Diversity and affirmative action in higher education," *Journal of Public Economic Theory*, 10(4), 475–501.

Erdil, A., and H. Ergin (2008): "What's the matter with tie-breaking? Improving efficiency in school choice," *American Economic Review*, 98, 669–689.

——— (2017): "Two-sided matching with indifferences," *Journal of Economic Theory*, 171, 268–292.

Erdil, A., and T. Kumano (2019a): "Efficiency and stability under substitutable priorities with ties," *Journal of Economic Theory*, 184, 104950.

——— (2019b): "Efficiency and stability under substitutable priorities with ties," *Journal of Economic Theory*, 184, 104950.

Esponda, I., and D. Pouzo (2016): "Berk–Nash equilibrium: A framework for modeling agents with misspecified models," *Econometrica*, 84(3), 1093–1130.

Fernald, J. (2014): "A quarterly, utilization-adjusted series on total factor productivity," Accessed from `https://www.johnfernald.net/`.

Flynn, J. P., C. Patterson, and J. Sturm (2021): "Fiscal policy in a networked economy," Working Paper 29619, National Bureau of Economic Research.

Flynn, J. P., and K. A. Sastry (2021): "Strategic Mistakes," Mimeo, MIT.

Flynn, J. P., and K. A. Sastry (2022a): "Attention Cycles," Working paper, MIT.

Flynn, J. P., and K. A. Sastry (2022b): "Attention Cycles," Working Paper 3592107, SSRN.

Flynn, Z., J. Traina, and A. Gandhi (2019): "Measuring markups with production data," Working Paper 3358472, SSRN.

Fosgerau, M., E. Melo, A. De Palma, and M. Shum (2020): "Discrete choice and rational inattention: A general equivalence result," *International Economic Review*, 61(4), 1569–1589.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, firm turnover, and efficiency: Selection on productivity or profitability?," *American Economic Review*, 98(1), 394–425.

Foster, L., J. C. Haltiwanger, and C. J. Krizan (2001): *Aggregate Productivity Growth: Lessons from Microeconomic Evidence*pp. 303–372. University of Chicago Press.

Fragiadakis, D., and P. Troyan (2017): "Improving matching under hard distributional constraints," *Theoretical Economics*, 12(2), 863–908.

Frankel, D. M., S. Morris, and A. Pauzner (2003): "Equilibrium selection in global games with strategic complementarities," *Journal of Economic Theory*, 108(1), 1–44.

Fudenberg, D., R. Iijima, and T. Strzalecki (2015): "Stochastic choice and revealed perturbed utility," *Econometrica*, 83(6), 2371–2409.

Gabaix, X. (2014): "A sparsity-based model of bounded rationality," *The Quarterly Journal of Economics*, 129(4), 1661–1710.

——— (2016): "A behavioral New Keynesian model," *American Economic Review*, 110(8), 2271–2327.

Gabaix, X., and R. S. Koijen (2020): "Granular instrumental variables," Working Paper 28204, National Bureau of Economic Research.

Gale, D., and L. S. Shapley (1962): "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–15.

Galí, J. (2008): *Monetary policy, Inflation, and the Business Cycle: an Introduction to the New Keynesian Framework and its Applications*. Princeton University Press.

Galí, J., and L. Gambetti (2009): "On the sources of the Great Moderation," *American Economic Journal: Macroeconomics*, 1(1), 26–57.

——— (2019): "Has the US wage Phillips curve flattened? A semi-structural exploration," Working Paper 25476, National Bureau of Economic Research.

Galí, J., and T. Van Rens (2021): "The vanishing procyclicality of labour productivity," *The Economic Journal*, 131(633), 302–326.

Gennaioli, N., Y. Ma, and A. Shleifer (2016): "Expectations and investment," *NBER Macroeconomics Annual*, 30(1), 379–431.

Gentry, R. J., J. S. Harrison, T. J. Quigley, and S. Boivie (2021): "A database of CEO turnover and dismissal in S&P 1500 firms, 2000–2018," *Strategic Management Journal*, 42(5), 968–991.

Gershkov, A., B. Moldovanu, P. Strack, and M. Zhang (2021): "A theory of auctions with endogenous valuations," *Journal of Political Economy*, 129(4), 1011–1051.

Geyer, J., and H. Sieg (2013): "Estimating a model of excess demand for public housing," *Quantitative Economics*, 4(3), 483–513.

Goeree, J., T. R. Palfrey, and C. A. Holt (2016): *Quantal Response Equilibrium: A Stochastic Theory of Games.* Princeton University Press.

Goetzmann, W. N., D. Kim, and R. J. Shiller (2022): "Crash Narratives," Working Paper 30195, National Bureau of Economic Research.

Goldstein, D. (2012): "Bostonians Committed to School Diversity Haven't Given Up on Busing," *The Atlantic, October*, 10.

Google AdSense (2020): "Online Terms of Service," Accessed November 22, 2020, from https://www.google.com/adsense/localized-terms.

Gorodnichenko, Y. (2008): "Endogenous information, menu costs and inflation persistence," Discussion paper, National Bureau of Economic Research.

Goto, M., F. Kojima, R. Kurata, A. Tamura, and M. Yokoo (2017): "Designing matching mechanisms under general distributional constraints," *American Economic Journal: Microeconomics*, 9(2), 226–62.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, capacity utilization, and the real business cycle," *The American Economic Review*, pp. 402–417.

Grigoryan, A. (2021): "Effective, fair and equitable pandemic rationing," *Available at SSRN 3646539*.

Grigsby, J., E. Hurst, and A. Yildirmaz (2019): "Aggregate nominal wage adjustments: New evidence from administrative payroll data," *American Economic Review*, 111(2), 428–71.

Grubb, M. D. (2009): "Selling to overconfident consumers," *American Economic Review*, 99(5), 1770–1807.

Grubb, M. D., and M. Osborne (2015): "Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock," *American Economic Review*, 105(1), 234–71.

Guernsey, L. (1999): "Can it Pay to Surf the Web?," *The New York Times*, November 22.

Guynn, J., and K. McCoy (2018): "Apple co-founder Steve Wozniak says he's left Facebook over data collection," *USA Today*, April 8.

Hafalir, I. E., M. B. Yenmez, and M. A. Yildirim (2013): "Effective Affirmative Action in School Choice," *Theoretical Economics*, 8, 325–363.

Hall, R. E. (1978): "Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence," *Journal of Political Economy*, 86(6), 971–987.

Handley, K., and J. F. Li (2020): "Measuring the effects of firm uncertainty on economic activity: New evidence from one million documents," Working Paper 27896, National Bureau of Economic Research.

Hansen, L. P., and R. Jagannathan (1991): "Implications of security market data for models of dynamic economies," *Journal of political economy*, 99(2), 225–262.

Harsanyi, J. C. (1973): "Oddness of the number of equilibrium points: a new proof," *International Journal of Game Theory*, 2(1), 235–250.

Hassan, T. A., S. Hollander, L. van Lent, M. Schwedeler, and A. Tahoun (2020): "Firm-Level Exposure to Epidemic Diseases: COVID-19, SARS, and H1N1," Working Paper 26971, National Bureau of Economic Research.

Hassan, T. A., S. Hollander, L. Van Lent, and A. Tahoun (2019): "Firm-level political risk: Measurement and effects," *The Quarterly Journal of Economics*, 134(4), 2135–2202.

Hassan, T. A., J. Schreger, M. Schwedeler, and A. Tahoun (2021): "Sources and transmission of country risk," Discussion paper, National Bureau of Economic Research.

Hauser, J. R., and B. Wernerfelt (1990): "An evaluation cost model of consideration sets," *Journal of Consumer Research*, 16(4), 393–408.

Hébert, B. M., and J. La'O (2022): "Information Acquisition, Efficiency, and Non-Fundamental Volatility," Discussion paper, National Bureau of Economic Research.

Hébert, B. M., and M. Woodford (2020): "Neighborhood-Based Information Costs," Discussion paper, National Bureau of Economic Research.

Hellwig, C., and L. Veldkamp (2009): "Knowing what others know: Coordination motives in information acquisition," *The Review of Economic Studies*, 76(1), 223–251.

Hellwig, C., and V. Venkateswaran (2009): "Setting the right prices for the wrong reasons," *Journal of Monetary Economics*, 56, S57–S77.

Hoffman, M., F. Bach, and D. Blei (2010): "Online learning for Latent Dirichlet Allocation," *Advances in Neural Information Processing Systems*, 23.

Hopenhayn, H., and R. Rogerson (1993): "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of political Economy*, 101(5), 915–938.

Hsieh, C.-T., and P. J. Klenow (2009): "Misallocation and manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403–1448.

Huo, Z., and M. Pedroni (2020): "A single-judge solution to beauty contests," *American Economic Review*, 110(2), 526–68.

Huo, Z., and N. Takayama (2022): "Higher-Order Beliefs, Confidence, and Business Cycles," Working Paper 4173060, SSRN.

IHS Markit (2017): "The Economic Value of Behavioural Targeting in Digital Advertising," Research report, Retrieved from: https://iabeurope.eu/knowledgehub/policy/the-economic-value-of-data-driven-advertising/.

Ilut, C., M. Kehrig, and M. Schneider (2018): "Slow to hire, quick to fire: Employment dynamics with asymmetric responses to news," *Journal of Political Economy*, 126(5), 2011–2071.

Ilut, C. L., and R. Valchev (2020): "Economic Agents as Imperfect Problem Solvers," Working Paper 27820, National Bureau of Economic Research.

ISABELLA, L. A. (1990): “Evolving interpretations as a change unfolds: How managers construe key organizational events,” *Academy of Management Journal*, 33(1), 7–41.

JENSEN, M. K. (2018): “Aggregative games,” in *Handbook of Game Theory and Industrial Organization, Volume I*. Edward Elgar Publishing.

JIANG, F., J. LEE, X. MARTIN, AND G. ZHOU (2019): “Manager sentiment and stock returns,” *Journal of Financial Economics*, 132(1), 126–149.

JUNG, J., J. H. J. KIM, F. MATĚJKA, AND C. A. SIMS (2019): “Discrete Actions in Information-Constrained Decision Problems,” *The Review of Economic Studies*, 86(6), 2643–2667.

JURADO, K., S. C. LUDVIGSON, AND S. NG (2015): “Measuring uncertainty,” *American Economic Review*, 105(3), 1177–1216.

KACPERCZYK, M., S. VAN NIEUWERBURGH, AND L. VELDKAMP (2016): “A Rational Theory of Mutual Funds’ Attention Allocation,” *Econometrica*, 84(2), 571–626.

KAMADA, Y., AND F. KOJIMA (2017): “Stability concepts in matching under distributional constraints,” *Journal of Economic Theory*, 168, 107–142.

——— (2018): “Stability and strategy-proofness for matching with constraints: A necessary and sufficient condition,” *Theoretical Economics*, 13(2), 761–793.

KANG, Y. P. (2000): “It Pays to Cheat, Not Surf,” *Wired*, July 10.

KAUSTIA, M., AND V. RANTALA (2021): “Common analysts: method for defining peer firms,” *Journal of Financial and Quantitative Analysis*, 56(5), 1505–1536.

KEHRIG, M. (2015): “The cyclical nature of the productivity distribution,” Working Paper 1854401, SSRN.

KELLER, W., AND S. R. YEAPLE (2009): “Multinational enterprises, international trade, and productivity growth: firm-level evidence from the United States,” *The Review of Economics and Statistics*, 91(4), 821–831.

KESTEN, O., M. KURINO, AND A. S. NESTEROV (2017): “Efficient lottery design,” *Social Choice and Welfare*, 48(1), 31–57.

KESTEN, O., AND M. U. ÜNVER (2015): “A theory of school-choice lotteries,” *Theoretical Economics*, 10(2), 543–595.

KEYNES, J. M. (1936): *The General Theory of Employment, Interest, and Money*. Macmillan, London.

KOJIMA, F. (2012): “School Choice: Impossibilities for Affirmative Action,” *Games and Economic Behavior*, 75, 685–693.

KOMINERS, S. D., AND T. SÖNMEZ (2016): “Matching with slot-specific priorities: Theory,” *Theoretical Economics*, 11(2), 683–710.

KOZLOWSKI, J., L. VELDKAMP, AND V. VENKATESWARAN (2020): “The tail that wags the economy: Beliefs and persistent stagnation,” *Journal of Political Economy*, 128(8), 2839–2879.

KROENCKE, T. A. (2017): “Asset pricing without garbage,” *The Journal of Finance*, 72(1), 47–98.

LAFFONT, J.-J., AND J. TIROLE (1986): “Using cost observation to regulate firms,” *Journal of Political Economy*, 94(3, Part 1), 614–641.

LANDSMARK, T. (2009): “It’s time to end busing in Boston,” *Boston Globe, Opinion-Editorial*.

Leshno, J. (2017): "Dynamic matching in overloaded waiting lists," *SSRN 2967011*.

Levinson, M. L., J. M. Noonan, J. Z. Fay, A. M. Mantil, C. J. Buttimer, and J. D. Mehta (2012): "First-round analysis of BPS proposed 6-zone, 9-zone, 11-zone, and 23-zone school assignment plans," .

Li, T.-Y., and J. A. Yorke (1975): "Period Three Implies Chaos," *The American Mathematical Monthly*, 82(10), 985–992.

Liberti, J. M., and M. A. Petersen (2019): "Information: Hard and soft," *Review of Corporate Finance Studies*, 8(1), 1–41.

Loewenstein, J., W. Ocasio, and C. Jones (2012): "Vocabularies and vocabulary structure: A new approach linking categories, practices, and institutions," *Academy of Management Annals*, 6(1), 41–86.

Lorenzoni, G. (2009): "A theory of demand shocks," *American Economic Review*, 99(5), 2050–84.

Loughran, T., and B. McDonald (2011a): "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," *The Journal of Finance*, 66(1), 35–65.

——— (2011b): "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *The Journal of Finance*, 66(1), 35–65.

Ma, Y., T. Ropele, D. Sraer, and D. Thesmar (2020): "A Quantitative Analysis of Distortions in Managerial Forecasts," Working Paper 26830, National Bureau of Economic Research.

Macaulay, A. (2020): "Cyclical attention to saving," Working Paper 140, Oxford University.

Macaulay, A., and W. Song (2022): "Narrative-Driven Fluctuations in Sentiment: Evidence Linking Traditional and Social Media," Working paper, University of Oxford Department of Economics.

Maćkowiak, B., F. Matejka, and M. Wiederholt (2020): "Rational Inattention: A Review," Discussion Paper 15408, CEPR.

Maćkowiak, B., and M. Wiederholt (2009): "Optimal sticky prices under rational inattention," *American Economic Review*, 99(3), 769–803.

——— (2018): "Lack of preparation for rare events," *Journal of Monetary Economics*, 100, 35–47.

Maitlis, S. (2005): "The social processes of organizational sensemaking," *Academy of Management Journal*, 48(1), 21–49.

Mäkinen, T., and B. Ohl (2015): "Information acquisition and learning from prices over the business cycle," *Journal of Economic Theory*, 158, 585–633.

Mani, A., S. Mullainathan, E. Shafir, and J. Zhao (2013): "Poverty impedes cognitive function," *Science*, 341(6149), 976–980.

Manski, C. F. (1993): "Identification of endogenous social effects: The reflection problem," *The Review of Economic Studies*, 60(3), 531–542.

Marcet, A., and T. J. Sargent (1989a): "Convergence of least-squares learning in environments with hidden state variables and private information," *Journal of Political Economy*, 97(6), 1306–1322.

——— (1989b): "Convergence of least squares learning mechanisms in self-referential linear stochastic models," *Journal of Economic Theory*, 48(2), 337–368.

Maskin, E., and J. Riley (1984): "Monopoly with incomplete information," *The RAND Journal of Economics*, 15(2), 171–196.

Matějka, F. (2015): "Rationally Inattentive Seller: Sales and Discrete Pricing," *The Review of Economic Studies*, 83(3), 1125–1155.

Matějka, F., and A. McKay (2015): "Rational inattention to discrete choices: A new foundation for the multinomial logit model," *American Economic Review*, 105(1), 272–98.

Mathevet, L. (2010): "A contraction principle for finite global games," *Economic Theory*, 42(3), 539.

Mattsson, L.-G., and J. W. Weibull (2002): "Probabilistic choice and procedurally bounded rationality," *Games and Economic Behavior*, 41(1), 61–78.

Maxted, P. (2020): "A macro-finance model with sentiment," Working paper, Harvard University.

McDonald, B. (2021): "Software Repository for Accounting and Finance," Accessed June 2020. Available at: https://sraf.nd.edu/.

McFadden, D. L. (1973): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, pp. 105–142. Academic Press, New York.

McKelvey, R. D., and T. R. Palfrey (1995): "Quantal response equilibria for normal form games," *Games and Economic Behavior*, 10(1), 6–38.

Melo, E. (2022): "On the uniqueness of quantal response equilibria and its application to network games," *Economic Theory*, 74(3), 681–725.

Milgrom, P., and J. Roberts (1990): "Rationalizability, learning, and equilibrium in games with strategic complementarities," *Econometrica: Journal of the Econometric Society*, pp. 1255–1277.

Milgrom, P., and I. Segal (2002): "Envelope theorems for arbitrary choice sets," *Econometrica*, 70(2), 583–601.

Milgrom, P., and C. Shannon (1994): "Monotone comparative statics," *Econometrica*, pp. 157–180.

Molavi, P. (2019): "Macroeconomics with learning and misspecification: A general theory and applications," Working paper, Northwestern University.

Morris, S., R. Rob, and H. S. Shin (1995): "p-Dominance and belief potential," *Econometrica*, pp. 145–157.

Morris, S., and H. S. Shin (1998): "Unique equilibrium in a model of self-fulfilling currency attacks," *American Economic Review*, pp. 587–597.

——— (2002): "Social value of public information," *American Economic Review*, 92(5), 1521–1534.

Morris, S., and M. Yang (2019): "Coordination and Continuous Stochastic Choice," .

Mullainathan, S., and E. Shafir (2013): *Scarcity: Why Having too Little Means so Much*. Macmillan.

Mussa, M., and S. Rosen (1978): "Monopoly and product quality," *Journal of Economic Theory*, 18(2), 301–317.

Nakamura, E., J. Steinsson, P. Sun, and D. Villar (2018): "The elusive costs of inflation: Price dispersion during the US great inflation," *The Quarterly Journal of Economics*, 133(4), 1933–1980.

Nickell, S. (1981): "Biases in dynamic models with fixed effects," *Econometrica*, 49(6), 1417–1426.

Nöldeke, G., and L. Samuelson (2007): "Optimal bunching without optimal control," *Journal of Economic Theory*, 134(1), 405–420.

Olley, G. S., and A. Pakes (1996): "The dynamics of productivity in the telecommunications equipment industry," *Econometrica*, 64(6), 1263–1297.

Oster, E. (2019): "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 37(2), 187–204.

Ottonello, P., and T. Winberry (2020): "Financial heterogeneity and the investment channel of monetary policy," *Econometrica*, 88(6), 2473–2502.

Parker, J. A., and C. Julliard (2005): "Consumption risk and the cross section of expected returns," *Journal of Political Economy*, 113(1), 185–222.

Pathak, P. A., A. Rees-Jones, and T. Sönmez (2020): "Immigration Lottery Design: Engineered and Coincidental Consequences of H-1B Reforms," Discussion paper, National Bureau of Economic Research.

Pathak, P. A., and T. Sönmez (2008): "Leveling the Playing Field: Sincere and Sophisticated Players in the Boston Mechanism," *American Economic Review*, 98, 1636–1652.

Pathak, P. A., T. Sönmez, M. U. Unver, and M. B. Yenmez (2021): "Leaving No Ethical Value Behind: Triage Protocol Design for Pandemic Rationing," *NBER Working Paper*, (26951).

Peterman, W. B. (2016): "Reconciling micro and macro estimates of the Frisch labor supply elasticity," *Economic inquiry*, 54(1), 100–120.

Piscitello, G. M., E. M. Kapania, W. D. Miller, J. C. Rojas, M. Siegler, and W. F. Parker (2020): "Variation in Ventilator Allocation Guidelines by US State During the Coronavirus Disease 2019 Pandemic: A Systematic Review," *JAMA network open*, 3(6).

Pomatto, L., P. Strack, and O. Tamuz (2023): "The Cost of Information: The Case of Constant Marginal Costs," Working paper, arXiv:1812.04211.

Posner, E. A., and E. G. Weyl (2018): *Radical Markets: Uprooting Capitalism and Democracy for a Just Society*. Princeton University Press.

Restuccia, D., and R. Rogerson (2008): "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 11(4), 707–720.

Rochet, J.-C. (1987): "A necessary and sufficient condition for rationalizability in a quasi-linear context," *Journal of Mathematical Economics*, 16(2), 191–200.

Rosenthal, R. W. (1989): "A bounded-rationality approach to the study of noncooperative games," *International Journal of Game Theory*, 18(3), 273–292.

Roth, A. E. (2002): "The Economist As Engineer: Game Theory, Experimentation, and Computation As Tools for Design Economics," *Econometrica*, 70, 1341–1378.

Roth, A. E., and E. Peranson (1999): "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89, 748–780.

Roth, A. E., T. Sönmez, and M. U. Ünver (2004): "Kidney exchange," *The Quarterly journal of economics*, 119(2), 457–488.

Sartori, E. (2021): "Competitive provision of digital goods," Working paper, Center for Studies in Economics and Finance.

Savov, A. (2011): "Asset pricing with garbage," *The Journal of Finance*, 66(1), 177–201.

Schwartzstein, J., and A. Sunderam (2021): "Using models to persuade," *American Economic Review*, 111(1), 276–323.

SEC (2011): "Investor Bulletin: How to Read a 10-K," Article, Securities and Exchange Commission Office of Investor Education and Advocacy, Available at: https://www.sec.gov/files/reada10k.pdf.

Sergeyev, D., C. Lian, and Y. Gorodnichenko (2022): "The Economics of Financial Stress," Working paper, UC Berkeley.

Shaked, M., and J. G. Shanthikumar (2007): *Stochastic Orders*. Springer.

Shiller, R. J. (2017): "Narrative economics," *American Economic Review*, 107(4), 967–1004.

——— (2020): *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Sieg, H., and C. Yoon (2020): "Waiting for affordable housing in New York City," *Quantitative Economics*, 11(1), 277–313.

Simon, H. A. (1947): *Administrative Behavior*. Macmillan, New York.

Sims, C. A. (1998): "Stickiness," in *Carnegie-Rochester Conference Series on Public Policy*, vol. 49, pp. 317–356. Elsevier.

——— (2003): "Implications of rational inattention," *Journal of monetary Economics*, 50(3), 665–690.

——— (2006): "Rational inattention: Beyond the linear-quadratic case," *American Economic Review*, 96(2), 158–163.

Solon, G., R. Barsky, and J. A. Parker (1994): "Measuring the cyclicality of real wages: how important is composition bias?," *The quarterly journal of economics*, 109(1), 1–25.

Song, W., and S. Stern (2021a): "Firm inattention and the transmission of monetary policy: A text-based approach," Working Paper 3900641, SSRN.

Song, W., and S. Stern (2021b): "Firm Inattention and the Transmission of Monetary Policy: A Text-Based Approach," Mimeo, University of Michigan.

Sönmez, T. (2013): "Bidding for army career specialties: Improving the ROTC branching mechanism," *Journal of Political Economy*, 121(1), 186–219.

Sönmez, T., and M. B. Yenmez (2020a): "Affirmative Action with Overlapping Reserves," Discussion paper, Boston College.

——— (2020b): "Constitutional implementation of Reservation Policies in India," Discussion paper, Boston College Department of Economics.

——— (2021): "Can Economic Theory Be Informative for the Judiciary? Affirmative Action in India Via Vertical and Horizontal Reservations," *Unpublished mimeo*.

——— (2022a): "Affirmative action in India via vertical, horizontal, and overlapping reservations," *Econometrica*, 90(3), 1143–1176.

——— (2022b): "Constitutional Implementation of Affirmative Action Policies in India," *arXiv preprint arXiv:2203.01483*.

Spiegler, R. (2016): "Bayesian networks and boundedly rational expectations," *The Quarterly Journal of Economics*, 131(3), 1243–1290.

Stahl, D. O. (1990): "Entropy control costs and entropic equilibria," *International Journal of Game Theory*, 19(2), 129–138.

Stevens, L. (2019): "Coarse Pricing Policies," *The Review of Economic Studies*, 87(1), 420–453.

Stewart, N., J. C. Mays, and M. Haag (2019): "Facing Homeless Crisis, New York Aims for 1,000 New Apartments a Year," *New York Times*, 1.

Stokey, N. L., R. E. Lucas, and E. C. Prescott (1989): *Recursive Methods in Economic Dynamics*. Harvard University Press.

Strausz, R. (2017): "A theory of crowdfunding: A mechanism design approach with demand uncertainty and moral hazard," *American Economic Review*, 107(6), 1430–76.

Sundararajan, A. (2004): "Nonlinear pricing of information goods," *Management Science*, 50(12), 1660–1673.

Temnyalov, E. (2021): "An information theory of efficient differential treatment," *Available at SSRN 3302150*.

Tuoi Tre News (2017): "Vietnamese college hopefuls miss dream school despite acing national exam," *Tuoi Tre News*, August 4, 2017.

Tversky, A., and D. Kahneman (1973): "Availability: A heuristic for judging frequency and probability," *Cognitive Psychology*, 5(2), 207–232.

Vaara, E., S. Sonenshein, and D. Boje (2016): "Narratives as sources of stability and change in organizations: Approaches and directions for future research," *Academy of Management Annals*, 10(1), 495–560.

Van Damme, E. (1991): *Stability and perfection of Nash equilibria*, vol. 339. Springer.

Van Zandt, T., and X. Vives (2007): "Monotone equilibria in Bayesian games of strategic complementarities," *Journal of Economic Theory*, 134(1), 339–360.

Waldinger, D. (2018): "Targeting in-kind transfers through market design: A revealed preference analysis of public housing allocation," *American Economic Review, Forthcoming*.

Weinstein, J., and M. Yildiz (2007): "Impact of higher-order uncertainty," *Games and Economic Behavior*, 60(1), 200–212.

Weisstein, E. W. (2001): "Logistic map," *Wolfram MathWorld*.

Weitzman, M. L. (1974): "Prices vs. quantities," *The Review of Economic Studies*, 41(4), 477–491.

Werning, I. (2017): "Comment on: Is the Macroeconomy Locally Unstable and Why Should We Care?," in *NBER Macroeconomics Annual 2016*, vol. 31, pp. 540–552. University of Chicago Press.

Westkamp, A. (2013): "An analysis of the German university admissions system," *Economic Theory*, 53(3), 561–589.

Wilson, R. B. (1993): *Nonlinear Pricing*. Oxford University Press, New York.

Woodford, M. (2003a): "Imperfect Common Knowledge and the Effects of Monetary Policy," *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps.*

——— (2003b): *Interest and Prices: Foundations of a Theory of Monetary Policy.* Princeton University Press.

——— (2012): "Inattentive valuation and reference-dependent choice," Working paper, Columbia University.

——— (2020): "Modeling imprecision in perception, valuation, and choice," *Annual Review of Economics*, 12, 579–601.

Yang, K. H. (2022): "Selling consumer data for profit: Optimal market-segmentation design and its consequences," *American Economic Review*, 112(4), 1364–93.

Yang, M. (2015): "Coordination with flexible information acquisition," *Journal of Economic Theory*, 158, 721–738.

Yasar, M., R. Raciborski, and B. Poi (2008): "Production function estimation in Stata using the Olley and Pakes method," *The Stata Journal*, 8(2), 221–231.