Massachusetts Institute of Technology
Department of Economics
Working Paper Series

OPTIMAL MONETARY POLICY WITH INFORMATIONAL FRICTIONS

George-Marios Angeletos
Jennifer La'O

Working Paper 11-22
October 4, 2011
Rev: October 31, 2011

Room E52-251
50 Memorial Drive
Cambridge, MA 02142

Optimal Monetary Policy with Informational Frictions*

George-Marios Angeletos
MIT and NBER

Jennifer La'O
University of Chicago
Booth School of Business

October 31, 2011

Abstract

We study optimal monetary policy in an environment in which firms' pricing and production decisions are subject to informational frictions. Our framework accommodates multiple formalizations of these frictions, including dispersed private information, sticky information, and certain forms of inattention. An appropriate notion of constrained efficiency is analyzed alongside the Ramsey policy problem. Similarly to the New-Keynesian paradigm, efficiency obtains with a subsidy that removes the monopoly distortion and a monetary policy that replicates flexible-price allocations. Nevertheless, "divine coincidence" breaks down and full price stability is no more optimal. Rather, the optimal policy is to "lean against the wind", that is, to target a negative correlation between the price level and real economic activity.

JEL codes: E32, E52, D61, D83.

Keywords: business cycles, dispersed information, sticky information, rational inattention, optimal policy, price stability.

*This paper extends, and subsumes, a 2008 working paper that circulated under the title "Dispersed Information over the Business Cycle: Optimal Fiscal and Monetary Policy." We benefited from comments by various colleagues and seminar participants at MIT, Chicago, Harvard, Indiana, John Hopkins, Ohio State, UC Irvine, UCLA, Universidad Carlos III, Wharton, the Federal Reserve Banks of Minneapolis and Philadelphia, the Bank of Spain, the 2008 Hydra Workshop on Dynamic Macroeconomics, the 2008 NBER Monetary Economics Meeting in Boston, and the 2011 SED Meeting in Ghent. We finally owe special thanks to Robert King and Philippe Bacchetta for discussing earlier versions of our paper. *Email:* `angelet@mit.edu`, `jenlao@chicagobooth.edu`.

1 Introduction

Informational frictions are paramount: people have incomplete information about the state of the economy and hold differing beliefs regarding its future prospects. This could be because information is dispersed (Lucas, 1972; Morris and Shin, 2002); because people update their information only infrequently (Mankiw and Reis, 2002); or because they pay limited attention to the available information (Sims, 2001, Woodford, 2002).

Recent research has pushed forward each of these complementary formalizations of informational frictions and has documented how they can have profound implications for the positive properties of the business cycle. Woodford (2002) and Mackowiak and Wiederholt (2009) argue that rational inattention can justify significant inertia in the response of firms' pricing decisions to nominal shocks. Lorenzoni (2009) uses informational frictions to develop a theory of demand shocks. Angeletos and La'O (2011a) show how such frictions can accommodate self-fulfilling beliefs, and forces akin to "animal spirits", within otherwise conventional unique-equilibrium macro models.

Despite these advances on the positive front, it remains unclear as to how the interaction of informational and nominal frictions affect the design of optimal monetary policy. For instance, Mankiw and Reis (2002), Woodford (2002), Mackowiak and Wiederholt (2009) and Lorenzoni (2009) alike study environments in which monetary policy is exogenous and, in fact, suboptimal. Our contribution in this paper is to fill the aforementioned gap.

Framework. The backbone of our framework is a monetary business-cycle economy akin to those in the core of the modern DSGE paradigm, with the exception that incomplete price adjustment is tied to informational frictions rather than Calvo-like price setting. More specifically, the economy is populated by a representative household and a continuum of monopolistic, price-setting firms. The key friction then is that the firm's pricing and production decisions are based on incomplete information about the state of the economy.

Our framework treats the information structure as exogenous, thus focusing on questions of allocative rather than informational efficiency. While somewhat limiting the scope of our exercise, this keeps our analysis directly comparable to the standard theory of optimal monetary policy, which also focuses exclusively on allocative efficiency. It also permits us to accommodate a much higher level of generality than the pertinent literature on informational frictions.

Thus, whereas each of the aforementioned works is tied to a particular specification of the informational friction, our framework accomodates multiple interpretations. The key observation that facilitates this generality is that *any* type of informational friction—whether this originates from dispersed information, sticky information, or some form of inattentiveness—ultimately reduces to a certain measurability constraint on the set of feasible allocations. Our framework allows for an arbitrary such constraint and, in this sense, for a generic informational friction.

Furthermore, our framework accommodates multiple formalizations of the origins of macroeconomic volatility. As in conventional DSGE models, our model features aggregate shocks to technology and preferences. But it also allows for noisy news about future productivity, as in Jaimovich and Rebelo (2009), Lorenzoni (2009), and Blanchard et al. (2011), as well as for extrinsic shocks in beliefs of economic activity, as in Angeletos and La'O (2011a). Fluctuations may thus obtain, not only from shocks to fundamentals, but also from forces akin to "animal spirits".

Finally, our framework allows for capital accumulation and accommodates more general preference and technology specifications than those featured in the pertinent work. It also allows informational frictions to affect not only the firm's nominal pricing decisions but also its real production choices. As we explain in due course, this permits us to move beyond degenerate cases where the optimal monetary policy achieves first-best, full-information outcomes.

Results. We start our analysis by defining and characterizing an appropriate notion of constrained efficiency. This notion is akin to conventional first-best efficiency, except that it embeds the informational friction as a measurability constraint on the set of feasible allocations. In essence, the planner cannot "cure" an agent's inattentiveness, nor can he communicate private information from one agent to another. This notion of efficiency therefore identifies the best allocation society could ever attain given the informational friction and, of course, the resource constraints.

We next shift focus to the set of Ramsey implementable allocations. This set is strictly smaller than the set of resource- and informationally-feasible allocations, for the planner's ability to fashion allocations is restricted by specific policy instruments. Furthermore, this set varies depending on whether prices are flexible or sticky, by which we mean whether nominal prices can adjust freely to realized market conditions (as when firms can post supply schedules) or whether nominal prices are restricted to be contingent on limited information. The question of interest is then whether the efficient allocation is contained within these sets.

Our first result is that the efficient allocation is always contained in the set of flexible-price allocations. This is because the bite of informational frictions on this set is similar to that of a technological constraint. Efficiency is thus attained merely by removing the monopoly distortion.

Our second result is that any flexible-price allocation can be replicated under sticky prices if and only if it satisfies a certain log-separability condition. We spell out this condition in due course. For now, we note that it is automatically satisfied for *all* flexible-price allocations if technology takes a conventional Cobb-Douglas specification, or if one focuses on log-linear approximations (as most of the literature does). For all practical purposes, this condition can thus be taken for granted.

Putting these results together, we conclude that, no matter the precise form of the informational friction, the constrained efficient allocation belongs to set of sticky-price allocations. Furthermore, the optimal policy has a familiar nature: it is given by a non-contingent subsidy that removes the monopoly distortion and a monetary policy that replicates flexible-price allocations.

The optimality of flexible-price allocations—and the consequent optimality of monetary policies that replicate them under sticky prices—is a cornerstone of modern macroeconomic theory. Versions of this lesson have been shown by Goodfriend and King (1997, 2001), Rotemberg and Woodford (1997), Woodford (2003), and Khan, King and Wolman (2003), among others. To the best of our knowledge, our paper is the first to establish the robustness of this lesson to informational frictions.

Typically, this lesson translates directly to optimality of price stability: in the standard New-Keynesian framework, targeting price stability is optimal precisely because it removes relative-price distortions and replicates the optimal flexible-price allocation. A sharp prediction then emerges: stabilizing the price level also guarantees full stabilization of the "output gap", leaving no room for a meaningful trade-off between output and price stabilization—a property that Blanchard and Gali (2007) have coined the "divine coincidence".

Our last result is that the aforementioned translation is *not* robust to informational frictions. Efficiency requires that firms' output levels respond to their beliefs about the state of the economy. For this to be implemented along a Ramsey equilibrium, relative prices must also respond to the firms' beliefs, and therefore the nominal price set by any given firm must vary with its own belief. Since the aggregate of these beliefs vary with the true state, the price level must also vary with it, and "divine coincidence" breaks down.[1]

We refine our characterization of the optimal policy under certain restrictions on preferences and technologies, which are empirically plausible and are trivially satisfied in previous work. We thereby show that the optimal policy is to "lean against the wind": the optimal policy induces a negative correlation between the price level and real economic activity, effectively trading off less price stabilization for more output stabilization. By contrast, a policy that targets price stability distorts relative prices, fails to replicate the efficient allocation, and exacerbates output gaps.

Related literature. The literature on optimal monetary policy is voluminous. A certain portion of this literature is concerned with implementation issues such as whether the Ramsey optimum can be implemented as the unique equilibrium with a simple Taylor rule; see, e.g., Atkeson, Chari, and Kehoe (2010) and Cochrane (2011) for recent critiques. We sidestep these issues by assuming that the monetary authority controls directly the level of aggregate nominal demand. Instead, we focus on the more fundamental question of the optimality of flexible-price allocations and its mapping to price stability. We thus complement a large body of work that has studied the same question in conventional New-Keynesian settings.[2]

To the best of our knowledge, our paper is actually the first to study the aforementioned question in the presence of informational frictions.[3] Closely related, however, are Lorenzoni (2010) and Paciello and Wiederholt (2011). These works are tied to more narrow specifications of the information structure and, most importantly, do not address either the optimality of flexible-price allocations or their relation to sticky-price allocations. Nevertheless, Lorenzoni (2010) complements our results by showing that the optimal policy may not eliminate the impact of noise on aggregate output, while Paciello and Wiederholt (2011) shift focus onto a question of informational efficiency, namely how monetary policy can control the incentives firms face in allocating attention to the underlying business-cycle shocks.

[1] It is well known that a trade off between price and output stabilization emerges in the New-Keynesian paradigm once one allows for "mark-up shocks", or for real-wage rigidities and other frictions that upset the optimality of flexible-price equilibria (e.g, Woodford, 2003; Blanchard and Gali, 2007). Our contribution here is to show that information frictions break divine coincidence *without* upsetting the optimality of flexible-price equilibria.

[2] See, inter alia, Goodfriend and King (1997, 2001), Rotemberg and Woodford (1997), Woodford (2003), Adao, Coreia and Teles (2003), Khan, King and Wolman (2003), Galí (2008).

[3] This is true at least for the earlier version of our paper (Angeletos and La'O, 2008); that paper contained the same key policy results, albeit within a less general framework.

Finally, our analysis builds on two methodological approaches. First, we follow the more abstract work of Angeletos and Pavan (2007, 2008) in studying an efficiency concept that bypasses the details of policy instruments and identifies the best allocation among those that are informationally feasible. Second, we follow the Ramsey literature in identifying the allocations that can be implemented with simple fiscal and monetary policy instruments (e.g., Lucas and Stokey, 1983; Chari and Kehoe, 1999; Adao, Correia and Teles, 2003; Benigno and Woodford, 2004), and in characterizing the optimal monetary policy in relation to the underlying flexible-price allocations (e.g., Woodford, 2003, Galí, 2008). This combination is key to the generality and sharpness of our results.

Layout. The rest of the paper is organized as follows. Section 2 introduces the model. Section 3 studies efficiency and establishes the optimality of flexible-price allocations. Section 4 characterizes the set of sticky-price allocations and identifies the Ramsey optimum. Section 5 documents the counter-cyclicality of the optimal price level. Section 6 illustrates our result in a setting where learning obtains slowly. Section 7 concludes.

2 The Model

Time is discrete and periods are indexed by $t \in \{0, 1, 2, ...\}$. There is a continuum of monopolistic firms, which produce differentiated goods and are indexed by $i \in I$. The "managers" of these firms are subject to informational frictions. Their products are used by a competitive retail sector as intermediate inputs into the production of a final good, which in turn can be either consumed or invested to capital. Finally, there is representative household, which consists of a price-taking consumer, a price-taking worker, and the continuum of managers that run the monopolistic firms.[4]

Retailers. The retail sector consists of a representative final-good firm. Its technology is given by a CES aggregator of all the intermediate goods produced by all the monopolistic firms:

$$Y_t = \left[\int_I y_{it}^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}}$$

where Y_t is the quantity of the final good, y_{it} is the quantity of the intermediate good produced by monopolistic firm i, and $\rho > 1$ is the elasticity of substitution across intermediate goods. The firm's objective is to maximize its profit, $P_t Y_t - \int_I p_{it} y_{it} di$, where P_t is the price level (the price of the final good) and p_{it} is the price of the intermediate good i.

Monopolists. Take the firm that produces intermediate good i. Its output is given by

$$y_{it} = A_{it} F\left(k_{it}, \ell_{it}, n_{it}\right),$$

where A_{it} is its productivity (which may contain an aggregate random component as well as an idiosyncratic one), k_{it} is its capital stock, ℓ_{it} is the effort level of its manager, n_{it} is its employment

[4]Our analysis allows firms to be informationally-constrained but maintains the assumption of a representative, fully-informed, consumer. This is consistent with much of the recent literature on informational frictions, which also maintains this assumption (e.g., Woodford, 2002, Mankiw and Reis, 2003, Wiederholt and Mackowiak, 2009). Introducing informational frictions on the consumer's side is a natural extension—but it can also be a challenging one, in so far as belief heterogeneity may then be conducive to uninsurable idiosyncratic consumption risk.

level, and F is a constant- or diminishing-returns-to- scale production function. Investment is made within the firm (from retained earnings). The firm's capital stock therefore evolves according to

$$k_{i,t+1} = (1-\delta)k_{it} + x_{it}$$

where x_{it} is gross investment and $\delta \in [0,1]$ is the depreciation rate. Finally, the firm's realized dividend, net of taxes, is given by

$$\Pi_{it} = (1-\tau_t)\, p_{it}y_{it} - w_t P_t n_{it} - P_t x_{it},$$

where τ_t is the marginal tax rate on firm revenue and w_t is the real wage.

Households. The utility of the representative household is given by

$$\sum_{t=0}^{\infty} \beta^t \left[U(C_t, \xi_t) - V(N_t, \xi_t) - \int_I H(\ell_{it}) di \right]$$

where C_t is consumption of the final good, N_t is labor supply, and ℓ_{it} is the effort of the manager that runs firm i, ξ_t is a preference shock, and $\beta \in (0,1)$ is the discount factor. The functions U, V, and H are continuously differentiable and strictly increasing in, respectively, C, N, and ℓ; U is strictly concave; V and H are strictly convex; and the usual Inada conditions are satisfied.[5]

The representative household holds two types of assets: nominal bonds and money. The monetary authority is assumed to pay interest on money holdings, making the household indifferent between money and bonds. The budget constraint is therefore given by

$$P_t C_t + B_{t+1} = \int_I \Pi_{it} di + w_t P_t N_t + T_t + R_t B_t,$$

where T_t are lump-sum taxes, B_t are nominal financial assets (bonds and money), and R_{t+1} is the nominal interest rate.

Markets. The markets for labor and for the final good are competitive. The wage adjusts so as to clear the labor market, while the interest rate adjusts so as to clear the final-good market. It follows that any equilibrium allocation satisfies

$$N_t = \int_I n_{it} di \qquad \text{and} \qquad C_t + \int x_{it} di = Y_t.$$

The intermediate-goods markets, on the other hand, are monopolistic: managers set prices and commit to supply as much output as needed in order to meet demand at the prices they have set.

Money and the government. As mentioned previously, we sidestep the micro-foundations of money and, instead, impose that following ad hoc cash-in-advance constraint on total expenditure:

$$P_t Y_t = M_t,$$

where M_t can be interpreted either as money supply or nominal aggregate demand. The latter is allowed to be contingent on the state of the economy and is assumed to be controlled by the monetary

[5]The fact that preferences are separable and that ξ_t does not enter H is only for expositional simplicity.

authority. As in Woodford (2003), this approach abstracts from the utility of money holdings and can be motivated by considering the limit of a "cashless economy" in which the monetary authority pays interest on money and appropriately adjusts the nominal interest rate so as to induce the desired level of nominal spending. Furthermore, by letting the monetary authority control directly the level of nominal spending, we sidestep the ongoing debate on whether interest-rate rules induce a unique equilibrium (Atkeson, Chari and Kehoe, 2010; Cochrane 2011).

While the monetary authority controls M_t, the fiscal authority controls the tax rate τ_t, the lump-sum transfer T_t, and the issue of new nominal assets B_{t+1}. Since there is no revenue from seigniorage (the government pays interest for money), the government budget is given by

$$T_t + R_t B_t = (1 - \tau_t) \int_I p_{it} y_{it} di + B_{t+1}$$

Furthermore, because Ricardian equivalence holds in our setting as in the vast majority of monetary business-cycle models, we can set $B_t = 0$ without any loss of generality. Finally, we allow the tax rate τ_t to be contingent on the realized state of the economy—although, as it will become clear, this contingency won't be needed for optimality.

Stocks and information. Let $\tilde{s}_t$ be an exogenous random variable that is drawn in the beginning of period t and let $s^t = (\tilde{s}_0, ..., \tilde{s}_t)$ denote the history of this random variable up to period t. The variable s^t represents the exogenous state of the economy, or the "state of Nature", as of period t. As it will become clear, s^t pins down, not only the fundamentals of the economy (namely preference shock ξ_t and the cross-sectional profile of the firms' productivities A_{it}) but also the profile of the firms' information sets. The stochasticity in s^t is determined as follows. The initial state $s^0 = \tilde{s}_0$ is drawn from $\tilde{\mathcal{S}}_0$ according to some distribution $\mathcal{F}_0$. Take now an arbitrary period $t \geq 1$. Conditional on s^{t-1}, Nature draws $\tilde{s}_t$ from $\tilde{\mathcal{S}}_t$ according to a probability distribution $\mathcal{F}_t$, where $\mathcal{F}_t(\tilde{s}_t | s^{t-1})$ denotes the probability of $\tilde{s}_t$ conditional on s^{t-1}.

Informational frictions are introduced by restricting the information that firm managers have about s^t. The information set that manager i has in the begging of period t is represented by an exogenous random variable ω_i^t, which is itself given by the history of exogenous signals about the underlying state. More specifically, $\omega_i^t = (\tilde{\omega}_{i0}, ..., \tilde{\omega}_{it})$, where $\tilde{\omega}_{it}$ identifies the period-t innovation in the information of firm i (or, equivalently, the period-t signal). The latter is determined as follows. The initial signal $\omega_i^0 = \tilde{\omega}_{i0}$ is drawn from a probability distribution $\mathcal{G}_0$, where $\mathcal{G}_0(\tilde{\omega}_{i0} | s^0)$ denotes the probability of $\tilde{\omega}_{i0}$ conditional on s^0. Take now any $t \geq 1$ and any i. For any given s^t and ω_i^{t-1}, Nature draws $\tilde{\omega}_{it}$ from $\tilde{\Omega}_t$ according to a probability distribution $\mathcal{G}_t$, where $\mathcal{G}_t(\tilde{\omega}_{it} | s^t, \omega_i^{t-1})$ denotes the probability of $\tilde{\omega}_{it}$ conditional on s^t and ω_i^{t-1}.

Let $\mathcal{S}_t \equiv \tilde{\mathcal{S}}_0 \times ....\tilde{\mathcal{S}}_t$ and $\Omega^t \equiv \tilde{\Omega}_0 \times ....\tilde{\Omega}_t$ denote the supports of, respectively, s^t and ω_i^t. To simplify the exposition, these supports are assumed to be subsets of $\mathbb{R}^N$ ($N_t \geq 1 \forall t$) and a firm's own productivity is assumed to be contained in its manager's information set.[6] We can thus express a firm's productivity as $A_{it} = A(\omega_{it})$, and the preference shock as $\xi_t = \xi(s^t)$, for some functions

[6]The first assumption helps only avoid technical complications regarding measurability. The second assumption can readily be relaxed, only at the cost of letting the output y_{it} of a firm i depend on A_{it} separately from ω_{it}.

A and ξ. It follows that, as anticipated, s^t pins down both the underlying fundamentals and the distribution of information in the economy. Finally, we assume that neither s^t nor any other random variable in the economy is ever common knowledge among the managers: there always exists some idiosyncratic noise, although perhaps arbitrarily small, in the firms' beliefs of the underlying state.[7]

Interpretation of the information structure. Note that the distribution $\mathcal{G}_t$ from which $\tilde{\omega}_{it}$ is drawn depends on s^t. This means that $\tilde{\omega}_{it}$ serves, in effect, as a signal of s^t. Since the latter may contain, not only shocks to fundamentals, but also noise variables, our formulation allows for aggregate noise in the firms' beliefs of the underlying fundamentals, as well as in the entire hierarchy of beliefs. The business cycle may thus be driven, not only by conventional preference and technology shocks, but also by "news" and "noise shocks" as in, inter alia, Jaimovich and Rebelo (2009) and Lorenzoni (2009), or by "sentiment shocks" as in Angeletos and La'O (2011a). At the same time, the fact that $\mathcal{G}_t$ may depend on ω_i^{t-1} accommodates an arbitrary intertemporal correlation in the idiosyncratic noise of these signals, in which case the observation of $\tilde{\omega}_{it}$ may permit firms to "refine" the information contained in earlier signals. More generally, the only essential restriction in our analysis is that information sets are taken as exogenous to policies/allocations; the precise details of what enters these information sets or how they adjust over time are otherwise inconsequential.

Our formalization of the information structure is thus very flexible and can nest, or at least proxy, many competing micro-foundations of informational frictions. For example, consider models with "sticky information" as in Mankiw and Reis (2002). These models are nested in our framework by letting $\mathcal{G}_t$ assign probability λ to $\tilde{\omega}_{it} = s^t$ and probability $1 - \lambda$ to $\tilde{\omega}_{it} = \omega_i^t$, where $\lambda \in (0, 1)$ is the probability with which a firm gets to see the true state, while $1 - \lambda$ is the probability with which the firm fails to update its information set. Alternatively, consider models in which agents receive in each period noisy private and public signals of the underlying shocks, such as in Woodford (2002), Hellwig (2002), Nimark (2008), Angeletos and La'O (2009), and Lorenzoni (2009). To the extent that one abstracts from the potential endogeneity of these signals, these works are thus nested in our framework by letting ω_i^t be the collection of such signals observed by firm i up to the beginning of period t. Finally, consider models with rational inattention as in Sims (2003) and Mackowiak and Wiederholt (2009). In these works, the informational friction is motivated as a constraint in the agents' ability to allocate attention. It is then *as if* firms observe noisy signals of the underlying state, in which case ω_i^t is again the collection of current and past signals.

Finally, it is worth noting that our formalization allows for rich learning dynamics. For instance, one can interpret the arrival of new signals about the underlying state as the product of decentralized communication as in, inter alia, Amador and Weill (2011) and Angeletos and La'O (2011a). In Section 6 we illustrate our results with an example in which the learning dynamics are similar to those in Woodford (2002).

[7]This assumption can be interpreted as a finite bound on the managers' attention. In any event, as it will become clear in due course, this assumption serves only as a minor equilibrium refinement. It rules out situations where all firms make their prices proportionally contingent on a sunspot or another common-knowledge random variable—a possibility that would not affect the set of equilibrium allocations (real variables) but would enrich the set of possible price paths (nominal variables).

Quantity versus pricing choices: where do frictions bind? The preceding description has remained ambivalent about how firm choices adjust to realized demand. The standard practice in much of the literature is to embed a certain form of "schizophrenia" within the firm. To understand what we mean by this, take Woodford (2002), Mankiw and Reis (2003), or Mackowiak and Wiederholt (2008). In these papers, each firm is, in effect, split into two "divisions", or "personalities". One division sets a price and commits to deliver any quantity demanded at that price; think of this as the "pricing division". The other division is assigned the task of adjusting the firms' input and production choices so as to meet the realized demand, whatever the demand and the input prices might turn out to be; think of this as the "production division". The latter is fully attentive, observing and responding to the true state of the world, but cannot communicate that information to the pricing division, nor can it itself adjust prices.

Clearly, this form of "schizophrenia" is a convenient abstraction that helps capture how informational frictions impact pricing decisions while maintaining the usual equilibrium concepts: if firms fix prices and markets are to clear, quantities *must* adjust. Furthermore, since one of the margins of quantity adjustment is inventories, the assumption that certain real choices may adjust more easily than pricing choices seems plausible, even if not properly micro-founded. We are thus sympathetic to these modeling conventions and seek to accommodate them in our framework.

We nevertheless make them clear in order to highlight a certain limitation in the pertinent literature: the aforementioned works presume that *all* production choices are free to adjust to realized demand, so that the informational friction affects *only* the pricing decision. In this sense, the informational friction is nothing more than a nominal friction. By contrast, we seek to accommodate the more general, and more realistic, scenario in which the information friction affects *both* real and nominal decisions.

Thus take an arbitrary firm i. We assume that the production division is free to adjust one input—for concreteness, employment n_{it}—so as to meet demand. However, some other production choices—managerial effort ℓ_{it} and investment x_{it}—cannot adjust to realized demand. Rather, they are restricted to depend on the same information as the price p_{it}, namely on ω_i^t. Managerial effort and investment are thus proxies for all kinds of quantity choices that may be subject to informational frictions: managers are imperfectly informed, or "inattentive", not only when they make nominal pricing decisions, but also when they take a number of real production decisions.

Which particular production choices are restricted to be contingent on ω_i^t and which ones are allowed to adjust to s^t is immaterial for our results. The particular split we assume here, and the particular interpretation of the various production inputs, is only for concreteness. For example, one could re-interpret ℓ as the extensive margin of employment and n as the intensive margin, the idea being that firms choose hiring on the basis of incomplete information but adjust work effort in response to realized market conditions.[8]

[8] This re-interpretation may require that we relax the separability of preferences between ℓ and N, or that we introduce decentralized markets for ℓ_{it}; neither of this, however, would affect the results.

3 Efficient and flexible-price allocations

Before we proceed to the characterization of optimal monetary policy, we first analyze two useful normative benchmarks. The first embeds the informational friction in the set of feasible allocations and identifies the best allocation a benevolent planner could obtain if he can select any allocation in the aforementioned set. The second studies the Ramsey problem under flexible prices, in which case the planner's power in selecting allocations is restricted by specific policy instruments.

3.1 Feasibility and efficiency with informational frictions

The conventional approach to policy analysis in business-cycle models is based on first-best efficiency. This approach, however, does not seem appropriate for environments in which communication and/or cognitive frictions limit the ability of agents to acquire, digest, or respond to information. Surely enough, the planner could achieve first-best outcomes if he had enough instruments to fashion incentives and, in addition, could get rid of the informational friction. The question of interest for us, however, is what a government can achieve *given* the underlying informational friction.

With this in mind, we modify the conventional notions of feasibility and efficiency to allow for informational frictions to constrain the set of allocations that could be chosen by the planner.

Definition 1. *A feasible allocation is a collection of contingent plans for aggregate output, consumption, and labor, and firm-level contingent plans for effort, employment, investment, and production choices, that satisfy the following constraints:*

(i) resource feasibility:

$$C_t + \int x_{i,t} di = Y_t = Z_t \left[\int y_{it}^{\frac{\rho-1}{\rho}} \, di \right]^{\frac{\rho}{\rho-1}}$$

$$y_{it} = A_{it} F\left(k_{it}, \ell_{it}, n_{it}\right)$$

$$k_{i,t+1} = (1-\delta) k_{it} + x_{it}$$

(ii) n_{it} *and* y_{it} *are contingent on* (ω_i^t, s^t)*, while* ℓ_{it} *and* x_{it} *are contingent only on* ω_i^t*.*

Definition 2. *A (constrained) efficient allocation is a feasible allocation that maximizes the ex-ante utility of the representative household.*

The first constraint needs no justification. The second constraint embeds the informational friction as a measurability constraint in the planner's problem: an allocation is informationally-feasible only if the effort and investment choices it prescribes to any particular firm are contingent (at most) on its manager's information set, ω_i^t. If the informational friction represents a geographical segmentation as, e.g., in Lucas (1972) and Angeletos and La'O (2009), this constraint means that the planner cannot transfer information from one "island" to another. If the informational friction represents a cognitive limitation or a certain form of inattentiveness as, e.g., in Mankiw and Reis (2002) and Sims (2003), this constraint simply means that the planner cannot overcome people's inattentiveness. In short, the policy maker is neither a messenger nor a psychiatrist.

We henceforth express a feasible allocation with a collection of functions $(\ell, n, k, y, C, N, K, Y)$ such that $\ell_{it} = \ell(\omega_i^t)$, $x_{it} = x(\omega_i^t)$, $k_{it} = k(\omega_i^{t-1})$, $n_{it} = n(\omega_i^t, s^t)$, $y_{it} = y(\omega_i^t, s^t)$, $C_t = C(s^t)$, $N_t = N(s^t)$, $K_t = K(s^{t-1})$, and $Y_t = Y(s^t)$. We can then state the planner's problem as follows:

Planner's problem. *Choose the functions* $(\ell, n, k, y, C, N, K, Y)$ *so as to maximize*

$$\max \sum_{t=0}^{\infty} \beta^t \int \left[U\left(C(s^t), \xi(s^t)\right) - V\left(N(s^t, \xi(s^t))\right) - \int H\left(\ell(\omega_i^t)\right) d\mathcal{G}_t(\omega_i^t|s^t) \right] d\mathcal{F}_t(s^t)$$

subject to

$$\begin{aligned}
C\left(s^t\right) + K\left(s^t\right) &= Y\left(s^t\right) + (1-\delta)K\left(s^{t-1}\right) \\
Y(s^t) &= \left[\int y(\omega_i^t, s^t)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t) \right]^{\frac{\rho}{\rho-1}} \\
y(\omega_i^t, s^t) &= A(\omega_i^t) F\left(k\left(\omega_i^{t-1}\right), \ell\left(\omega_i^t\right), n\left(\omega_i^t, s^t\right)\right) \\
N\left(s^t\right) &= \int n\left(\omega_i^t, s^t\right) d\mathcal{G}_t(\omega_i^t|s^t) \\
K\left(s^t\right) &= \int k\left(\omega_i^t\right) d\mathcal{G}_t(\omega_i^t|s^t)
\end{aligned}$$

This is akin to the planner's problem in any conventional macroeconomic model. The only difference is that certain choices—here those associated with effort ℓ_{it} and investment x_{it}—are restricted to be contingent on noisy, individual-specific, signals of the underlying state. Notwithstanding this qualification, the characterization of the planner's problem can proceed in a similar fashion.

Because this problem is strictly concave, it has a unique solution, which is pinned down by first-order conditions. To economize on notation, we henceforth let, for any $z \in \{\ell, n, k\}$,

$$MP_z\left(\omega_i^t, s^t\right) \equiv \left(\frac{y(\omega_i^t, s^t)}{Y(s^t)} \right)^{-\frac{1}{\rho}} A(\omega_i^t) \tfrac{\partial}{\partial z} F\left(k\left(\omega_i^{t-1}\right), \ell\left(\omega_i^t\right), n\left(\omega_i^t, s^t\right)\right)$$

denote firm i's marginal product of input z in terms of the final good. We similarly let

$$U_c(s^t) \equiv \tfrac{\partial}{\partial C} U\left(C(s^t), \xi(s^t)\right), \quad V_n(s^t) \equiv \tfrac{\partial}{\partial N} V\left(N(s^t), \xi(s^t)\right), \text{ and } H_\ell(\omega_i^t) \equiv \tfrac{\partial}{\partial \ell} H\left(\ell(\omega_i^t)\right)$$

denote the relevant marginal utilities. We use these short-cuts throughout the paper, with the understanding that these marginal products and marginal utilities depend on the allocation under consideration. We also let $\mathcal{P}(.|\omega_i^t)$ denote the Bayesian posterior conditional on ω_i^t, and $\mathbb{E}[.|\omega_i^t]$ the corresponding expectation. We then reach the following characterization of the efficient allocation.

Proposition 1. *A feasible allocation is efficient if and only if it satisfies the following conditions:*

$$\begin{aligned}
V_n(s^t) - U_c(s^t) MP_n\left(\omega_i^t, s^t\right) &= 0 \quad \forall\, (\omega_i^t, s^t) && (1) \\
\mathbb{E}\left[\ H_\ell(\omega_i^t) - U_c(s^t) MP_\ell\left(\omega_i^t, s^t\right) \ \middle|\ \omega_i^t\ \right] &= 0 \quad \forall\, \omega_i^t && (2) \\
\mathbb{E}\left[\ U_c(s^t) - \beta U_c(s^{t+1}) \left\{1 - \delta + MP_k(\omega_i^{t+1}, s^{t+1})\right\} \ \middle|\ \omega_i^t\ \right] &= 0 \quad \forall\, \omega_i^t && (3)
\end{aligned}$$

The interpretation of this result is simple. Condition (1) equates the marginal product of labor in any given firm with the marginal rate of substitution between consumption and leisure. By implication, the marginal product of labor is equated across all firms. Condition (2) equates the marginal disutility of a manager's effort with the manager's expectation of the marginal valuation of the marginal product of his effort. Finally, condition (3) is an Euler condition that equates the expected marginal costs and benefits of investment conditional on the relevant information.

Clearly, these conditions are akin to those that characterize the first best. The only essential difference is that the informational friction manifests as a certain form of "noise", or "tremble", that moves the planner's solution away and around the frictionless, first-best benchmark. In particular, the marginal cost and benefits of effort and investment fail to be equated state by state, simply because these choices can be conditioned at most on a noisy indicator of the underlying state. Constrained efficiency then requires only that this noise is unpredictable in the eyes of the beholder (conditional on the information set of the corresponding manager). Finally, since managerial effort and investment affect the marginal product of labor, the aforementioned noise impacts also the allocation of labor, despite the fact that employment choices are free to adjust to the true state.

Needless to say, this noise would vanish and first-best outcomes would obtain if *all* real choices were free to adjust to the true state of the economy. By imposing that managerial effort and investment are subject to the informational friction, our framework allows us to move beyond this trivial case, which is actually hard-wired in most of the previous literature.

3.2 Flexible-price equilibria

We now turn our attention to the set of flexible-price allocations. By this we mean the set of feasible allocations that can be supported by an equilibrium in which firms are free to post state-contingent prices. One can think of this as the firms posting a supply schedule rather than a fixed price. Alternatively, one can think of the "inattentive" manager delegating the price choice to the "attentive" production/sales division: the latter observes market conditions and adjusts accordingly, not only the firm's employment and production levels, but also its price. One way or another, the key is that the relevant measurability constraint applies only on allocations, not on the prices that support them: p_{it} may depend, not only on ω_i^t, but also on s^t. The flexible-price concept we adopt therefore removes only the nominal rigidity: it relaxes the bite of informational frictions on nominal prices, while maintaining its hold on real decisions.

Definition 3. *Prices are "flexible" or "state-contingent" if and only if p_{it} can be contingent on both ω_i^t and s^t.*

Definition 4. *A flexible-price Ramsey equilibrium is a feasible allocation along with a collection of state-contingent prices and fiscal and monetary policies such that*

(i) the household and all firms are at their respective optima;

(ii) the government budget constraint and the cash-in-advance constraint are satisfied;

(iii) markets clear.

Definition 5. *The flexible-price Ramsey optimum is the feasible allocation that maximizes welfare (i.e., the ex ante utility of the representative household) among the ones that are part of a flexible-price Ramsey equilibrium.*

The above definitions are standard, except that feasibility already embeds the informational friction: ℓ_{it} and x_{it} are restricted to be measurable in ω_i^t rather than s^t.

To characterize the set of flexible-price allocations, consider first the behavior of the representative household and the final-good firm. The household's labor supply satisfies

$$V_n(s^t) = U_c(s^t)w(s^t),$$

while the final-good sector's input demand satisfies

$$y\left(\omega_i^t, s^t\right) = \left(\frac{p(\omega_i^t, s^t)}{P(s^t)}\right)^{-\rho} Y(s^t). \tag{4}$$

As usual, the former condition pins down the real wage, while the latter condition gives the demand function faced by the typical intermediate-good monopolist.

Consider now the behavior of this monopolist. Its objective is to maximize its expectation of the present value of $\mathcal{M}_t\Pi_{it} - H(\ell_{it})$, where Π_{it} is its nominal profit, $\mathcal{M}_t \equiv U_c(C_t, \xi_t)/P_t$ is the marginal value of nominal income for the representative household, and $H(\ell_{it})$ is the cost (disutility) of effort. The firm internalizes the fact that the demand for its product is given by (4), and hence that its real profit, net of taxes and investment costs, is given by

$$\frac{\Pi_{it}}{P_t} = (1-\tau_t)\frac{p_{it}y_{it}}{P_t} - w_t n_{it} - x_{it} = (1-\tau_t)Y_t^{1/\rho}y_{it}^{1-1/\rho} - w_t n_{it} - x_{it}.$$

Recall then that the "production unit" adjusts employment after observing the state of the economy, while the "manager" chooses effort and investment conditional only on ω_i^t. With this qualification in mind, we conclude that we can write the firm's problem as follows:

$$\max_{\ell,n,x,k,y} \mathbb{E}\left[\sum_t^\infty \left\{U_c(s^t)\left[\left(1-\tau(s^t)\right)Y(s^t)^{1/\rho}y(\omega_i^t, s^t)^{1-1/\rho} - w(s^t)n(s^t) - x(\omega_i^t)\right] - H\left(\ell(\omega_i^t)\right)\right\}\right]$$

subject to

$$y(\omega_i^t, s^t) = A(\omega_i^t)F\left(k(\omega_i^{t-1}), \ell(\omega_i^t), n(\omega_i^t, s^t)\right)$$
$$k(\omega_i^t) = (1-\delta)k(\omega_i^{t-1}) + x(\omega_i^t)$$

Since ρ is greater than 1, F is concave, and H is strictly convex, the above problem is strictly concave, guaranteeing that the following first-order conditions pin down the firm's optimal plan:

$$\begin{aligned}
\left(\tfrac{\rho-1}{\rho}\right)\left(1-\tau(s^t)\right)MP_n(\omega_i^t, s^t) - w\left(s^t\right) &= 0\\
\mathbb{E}\left[U_c(s^t)\left(\tfrac{\rho-1}{\rho}\right)\left(1-\tau(s^t)\right)MP_\ell(\omega_i^t, s^t)\middle|\omega_i^t\right] - H_\ell(\omega_i^t) &= 0\\
\mathbb{E}\left[\ U_c(s^t) - \beta U_c(s^{t+1})\left\{1-\delta+\left(\tfrac{\rho-1}{\rho}\right)\left(1-\tau(s^{t+1})\right)MP_k(\omega_i^{t+1}, s^{t+1})\right\}\ \middle|\ \omega_i^t\ \right] &= 0
\end{aligned}$$

where MP_z is, once again, a short-cut for the marginal product of input z (where $z \in \{\ell, n, k\}$) in terms of the final good. The principal difference between these conditions is that investment and managerial effort are chosen under uncertainty about the current state s^t, while employment is free to adjust to s^t. As with the planner's problem, this explains why the optimality conditions for ℓ_{it} and x_{it} (equivalently, k_{it+1}) involve taking expectations over s^t conditional on ω_i^t.

Building on these observations, we reach the following result, which characterizes the set of allocations that can be part of an equilibrium when prices are flexible.

Proposition 2. *A feasible allocation is part of a flexible-price Ramsey equilibrium if and only if there exists a function $\phi : \mathcal{S}^t \to \mathbb{R}_+$ such that the following hold:*

$$V_n(s^t) - U_c(s^t)\phi(s^t)MP_n\left(\omega_i^t, s^t\right) \;=\; 0 \quad \forall\, (\omega_i^t, s^t) \tag{5}$$

$$\mathbb{E}\left[\; H_\ell(\omega_i^t) - U_c(s^t)\phi(s^t)MP_\ell\left(\omega_i^t, s^t\right) \;\middle|\; \omega_i^t \;\right] \;=\; 0 \quad \forall\, \omega_i^t \tag{6}$$

$$\mathbb{E}\left[\; U_c(s^t) - \beta U_c(s^{t+1})\left\{1 - \delta + \phi(s^{t+1})MP_k(\omega_i^{t+1}, s^{t+1})\right\} \;\middle|\; \omega_i^t \;\right] \;=\; 0 \quad \forall\, \omega_i^t \tag{7}$$

The interpretation of this result is similar to Proposition 1. The only difference between the efficient allocation and any flexible-price allocation is the wedge $\phi(s^t)$ showing up in the above conditions. This wedge is simply the product of the monopoly markup and the tax rate:

$$\phi(s^t) = \left(\frac{\rho - 1}{\rho}\right)\left(1 - \tau(s^t)\right).$$

By appropriately choosing the tax, the Ramsey planner can induce an arbitrary such wedge, and in so doing implement any of the allocations identified by Proposition 2. This explains why ϕ is a "free variable" in this proposition: the set of flexible-price allocations is spanned by varying ϕ.

The following is then an immediate implication.

Theorem 1. *With flexible prices, the Ramsey optimum coincides with the constrained efficient allocation. The latter is implemented with a non-contingent subsidy that merely offsets the monopoly distortion, namely $\tau(s^t) = \tau^* \equiv -\frac{1}{\rho-1} \quad \forall\, s^t$.*

This result provides an important benchmark for the normative properties of business cycles under informational frictions. Once the monopoly distortion is removed, there is no way to improve upon the flexible-price equilibrium but by relaxing the informational friction: the economy's response to the underlying technology shocks, as well as to any noise in information, is efficient given the informational friction. In short, although informational frictions help disconnect equilibrium outcomes from fundamentals, they do not by themselves open the door to stabilization policy.

Translating this result in terms of the familiar notions of "output gaps" and "price dispersion", we should iterate that, as long as real production choices are subject to the informational friction (here, effort and investment), the constrained efficient allocation generally does not coincide with the first-best allocation. The informational friction thus perturbs the Ramsey optimum away the from first-best outcomes. At the aggregate level, this perturbation may manifest as noise-driven volatility in aggregate output gaps; at the cross-sectional level, it may manifest as noise-driven dispersion relative prices.

Note, however, that this perturbation away from the first best reflects merely the bite of the informational friction on the set of feasible allocations. This explains why the associated levels of output-gap volatility and relative-price dispersion are optimal in the eyes of the Ramsey planner: although there exist state-contingent policies that can attain either lower volatility in output gaps or lower dispersion in relative prices, these policies are dominated by the non-contingent one that merely offsets the monopoly power. In turn, this anticipates a related property of the optimal monetary policy: the optimal policy may not eliminate the impact of noise on either the aggregate or the cross-section of the economy, even if this might be feasible once prices are sticky.

It is also worth noting that this result does not rely on any specific assumptions about either the information structure or preferences and technologies. The key insight is that the informational friction is a measurability constraint that constrains both efficient and flexible-price allocations: the inability of certain agents to observe or process certain information is akin to a technological constraint. This is a simple but important observation.

4 Sticky-price allocations and optimal monetary policy

We now move on to characterize the set of equilibria and the Ramsey optimum when prices are sticky. As anticipated, by "sticky prices" we mean that firms can no more post contingent price plans (or supply functions). Rather, managers have to fix a price conditional on their information sets and then the "production unit" of each firm is left to adjust employment and output so as to meet the realized demand at the price set by the manager.

Definition 6. *Prices are "sticky" or "non-contingent" if and only if p_{it} can be contingent at most on ω_i^t, but not on s^t.*

Definition 7. *A sticky-price equilibrium is an allocation along with a collection of non-contingent prices and fiscal and monetary policies such that*

(i) the household and all firms are at their respective optima;

(ii) the government budget constraint and the cash-in-advance constraint are satisfied;

(iii) markets clear.

Definition 8. *The sticky-price Ramsey optimum is the feasible allocation that maximizes welfare among the ones that can be implemented as part of a sticky-price Ramsey equilibrium.*

These definitions are similar to the ones we introduced under flexible prices. The only difference is that the price p_{it} must now be measurable in ω_i^t, underscoring how the informational friction is now tied to a nominal friction: nominal prices cannot adjust to information that is revealed at the time trading takes place. These definitions are consistent with the notions of equilibrium that are implicit in all the recent monetary business-cycle models with informational frictions, including Mankiw and Reis (2002), Woodford (2002), Nimark (2008), Mackowiak and Wiederholt (2009), Lorenzoni (2009, 2010), and Paciello and Wiederholt (2011). We are merely making these notions explicit, highlighting how the nominal friction is superimposed on the informational one, and enriching the class of economies under consideration.

4.1 Sticky-price equilibria

To characterize the set of sticky-price equilibria, note first that the behavior of the representative household and the representative final-good firm remains the same as with flexible prices. The following conditions thus old for the real wage and the relative prices:

$$V_n(s^t) = U_c(s^t)w(s^t), \tag{8}$$

$$y\left(\omega_i^t, s^t\right) = \left(\frac{p(\omega_i^t)}{P(s^t)}\right)^{-\rho} Y(s^t). \tag{9}$$

It is thus only on the behavior of the monopolistic, intermediate-good firms that the informational friction now has an extra bite, as prices can no more be contingent on s^t.

Using condition (9), we have that firm i's real profit is given by

$$\frac{\Pi_{it}}{P_t} = (1-\tau_t)\frac{p_{it}y_{it}}{P_t} - w_t n_{it} - x_{it} = (1-\tau_t)\left(\frac{p_{it}}{P_t}\right)^{1-\rho} Y_t - w_t n_{it} - x_{it}.$$

We can thus state the monopolistic firm's optimal pricing and production problem as follows:

$$\max_{p,\ell,n,x,k} \mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t\left\{U_c(s^t)\left[\left(1-\tau(s^t)\right)\left(\frac{p\left(\omega_i^t\right)}{P(s^t)}\right)^{1-\rho} Y(s^t) - w(s^t)n\left(\omega_i^t, s^t\right) - x\left(\omega_i^t\right)\right] - H\left(\ell(\omega_i^t)\right)\right\}\middle|\,\omega_i^t\right]$$

subject to

$$k\left(\omega_i^t\right) = (1-\delta)\,k\left(\omega_i^{t-1}\right) + x\left(\omega_i^t\right)$$

$$A(\omega_i^t)F\left(k\left(\omega_i^{t-1}\right), \ell\left(\omega_i^t\right), n\left(\omega_i^t, s^t\right)\right) = \left(\frac{p\left(\omega_i^t\right)}{P(s^t)}\right)^{-\rho} Y\left(s^t\right)$$

The first constraint is simply the law of motion for capital. The second constraint follows from combining condition (9) with the production function and dictates how employment and output adjust so as to meet the realized demand.

Let $\beta^t U_c(s^t)\lambda\left(\omega_i^t, s^t\right)$ be the Lagrange multiplier on the second constraint. The first order conditions with respect to labor, effort, and capital are given by the following:

$$w(s^t) - \lambda\left(\omega_i^t, s^t\right) A(\omega_i^t) f_n\left(\omega_i^t, s^t\right) = 0 \tag{10}$$

$$\mathbb{E}\left[U_c(s^t)\lambda\left(\omega_i^t, s^t\right) A(\omega_i^t) f_\ell\left(\omega_i^t, s^t\right) - H_\ell(\omega_i^t)\middle|\,\omega_i^t\right] = 0 \tag{11}$$

$$\mathbb{E}\left[U_c(s^t) - \beta U_c(s^{t+t})\left\{1-\delta+\lambda\left(\omega_i^{t+1}, s^{t+1}\right) A(\omega_i^{t+1}) f_k\left(\omega_i^{t+1}, s^{t+1}\right)\right\}\middle|\,\omega_i^t\right] = 0, \tag{12}$$

where, for any $z \in \{\ell, n, k\}$, $f_z(\omega_i^t, s^t)$ is a short-cut for $\frac{\partial}{\partial z} F(k(\omega_i^{t-1}), \ell(\omega_i^t), n(\omega_i^t))$, that is, the physical marginal product of input z in terms of the intermediate good.[9] The first-order condition with respect to the price, on the other hand, can be stated as follows:

$$\mathbb{E}\left[U_c(s^t) y\left(\omega_i^t, s^t\right) \left\{ \left(1 - \tau(s^t)\right) \left(\tfrac{\rho - 1}{\rho}\right) \frac{p\left(\omega_i^t\right)}{P(s^t)} - \lambda\left(\omega_i^t, s^t\right) \right\} \middle| \, \omega_i^t \right] = 0 \tag{13}$$

To interpret these conditions, note that $\lambda_{it} = \lambda(\omega_i^t, s^t)$ identifies the real marginal cost of serving demand, that is, the real marginal cost the firm has to incur so as to produce that amount of output that is required in order to meet the realized demand at the pre-set price. Condition (10) then means that this marginal cost is equal to the ratio of the real wage rate to the marginal product of n_{it}. More generally, this should be interpreted as saying that the marginal cost is equated to the real price of any input that the firm can adjust in order to meet demand, discounted by the marginal product of that input.

Turning attention to the inputs that are contingent on incomplete information—here effort ℓ_{it} and investment x_{it}—note that the benefit of increasing any of them consists of the expected savings in terms of lower ex-post costs of meeting the demand. Conditions (11) and (12) then say that these marginal savings must be equated to the costs of these inputs.

Finally, to interpret condition (13), note that $(1 - \tau_t) \frac{\rho-1}{\rho} \frac{p_{it}}{P_t}$ is marginal revenue (in real terms) net of taxes. If prices were flexible, marginal revenue and marginal costs would have to be equated state-by-state: $(1 - \tau_t) \frac{\rho-1}{\rho} \frac{p_{it}}{P_t}$ would equal λ_{it} for all (ω_i^t, s^t). Now that prices are sticky, the realized marginal revenue and marginal costs will generally fail to coincide as demand varies with information upon which the price cannot be contingent. This is the essence of how informational frictions interact with nominal frictions both in our framework and in all of the related literature: the key is that prices are contingent only on a noisy indicator of the true state. At the same time, optimality in the firm's price-setting behavior requires that the realized pricing "error" (i.e., the difference between marginal revenue and marginal cost) be orthogonal to the variation in output under the appropriate, risk-adjusted, expectation operator of the manager. This is the meaning of condition (13).

To recap, any sticky-price equilibrium is pinned down by the optimality conditions (8)-(13) along with the budgets constraints and the cash-in-advance constraint. We now move on to characterize these equilibria by a set of necessary and sufficient conditions that solve out for the equilibrium prices and the associated policies and, instead, put restrictions only on the real allocations.

To this goal, we first introduce the following definition:

Definition 9. *An allocation is log-separable if and only if there exist positive-valued functions Ψ^ω and Ψ^s such that*

$$\log y\left(\omega_i^t, s^t\right) = \log \Psi^\omega(\omega_i^t) + \log \Psi^s(s^t)$$

After some manipulation of conditions (10)-(13), which we leave for the appendix, we can then reach the following result.

[9] This is not to be confused with MP_z, the marginal product in terms of the final good, which we defined earlier.

Proposition 3. *A feasible allocation can be part of a sticky-price equilibrium if and only if*

(i) there exist functions $\phi : \mathcal{S}^t \to \mathbb{R}_+$ and $\chi : \Omega_t \times \mathcal{S}^t \to \mathbb{R}_+$ such that the following hold:

$$\begin{aligned} V_n(s^t) - U_c(s^t)\phi(s^t)\chi(\omega_i^t, s^t)MP_n\left(\omega_i^t, s^t\right) &= 0 \quad \forall\ (\omega_i^t, s^t) \quad (14)\\ \mathbb{E}\left[H_\ell(\omega_i^t) - U_c(s^t)\phi(s^t)\chi(\omega_i^t, s^t)MP_\ell\left(\omega_i^t, s^t\right)\middle|\,\omega_i^t\right] &= 0 \quad \forall\ \omega_i^t \quad (15)\\ \mathbb{E}\left[U_c(s^t) - \beta U_c(s^{t+1})\left\{1-\delta+\phi(s^{t+1})\chi(\omega_i^{t+1}, s^t)MP_k\left(\omega_i^{t+1}, s^{t+1}\right)\right\}\middle|\,\omega_i^t\right] &= 0 \quad \forall\ \omega_i^t \quad (16)\\ \mathbb{E}\left[U_c(s^t)Y\left(s^t\right)^{1/\rho} y\left(\omega_i^t, s^t\right)^{1-1/\rho}\phi(s^t)\left\{\chi(\omega_i^t, s^t)-1\right\}\middle|\,\omega_i^t\right] &= 0 \quad \forall\ \omega_i^t \quad (17) \end{aligned}$$

(ii) the allocation is log-separable (in the sense of Definition 7).

Part (i) revisits, in effect, the implementability constraints that the Ramsey planner faced when prices were flexible. If we compare this result to Proposition 2, we see that the only difference between the set of sticky-price allocations and the set of flexible-price allocations is captured by the additional wedge $\chi_{it} = \chi(\omega_i^t, s^t)$ showing up in conditions (14), (15) and (16) relative to the corresponding conditions for flexible-price allocations, namely conditions (5), (6) and (7).

This wedge captures the extra power that sticky prices give the planner relative to flexible prices. Just as the planner can induce the wedge ϕ_t by appropriately choosing the tax rate, the planner can also induce the wedge χ_{it} by appropriately choosing monetary policy. However, while the planner is entirely free to choose any ϕ_t he wishes, his choice over χ_{it} are restricted by condition (17). In essence, this condition says that this wedge must be unforecastable under an appropriate risk-adjusted expectation operator.[10] This reflects the fact that, as always, the power of monetary policy rests exclusively on incomplete price adjustment: monetary policy can impact real allocations only in so far as it reacts to contingencies that are not included in those upon which prices are set and, in this sense, are unforecastable at the time prices are set.

Any ex post variation in χ_{it} then captures variation in the realized monopoly markup of firm i that obtains as monetary policy and aggregate demand react to contingencies that were unobserved (or not attended to) at the moment firms set their prices. The realized markup for firm i coincides with the ex ante optimal one if and only if $\chi_{it} = 1$, which in turn is the case if and only if there is no surprise in realized demand. By contrast, the realized markup is lower (respectively, higher) than the optimal one if demand turns out to be higher (respectively, lower) than what expected.

To recap the above discussion, part (i) of Proposition 3 revisits essentially known properties of the set of sticky-price allocations, albeit now within the context of environments with informational frictions. Moving to part (ii), we see that informational frictions introduce an additional and less familiar implementability constraint: for an allocation to be part of a sticky-price Ramsey equilibrium, it must also be log-separable in the sense of Definition 7.

This constraint has to do with the mapping between nominal and relative prices. Fix a period t and a state s^t, and take an arbitrary pair of firms i and j that have incomplete information (or imperfect attention) about the current state s^t. For an allocation to be implementable as part of

[10] To see this, note that (17) can be restated as $\hat{\mathbb{E}}_{it}[\chi_{it}] = 1$, where $\hat{\mathbb{E}}_{it}[X] \equiv \mathbb{E}_{it}[U_c(C_t)\phi_t p_{it} y_{it} \cdot X]$ for any random variable X.

a sticky-price Ramsey equilibrium, it must be that the *nominal* price set by firm i is contingent at most on ω_i^t, and similarly the nominal price set by firm j must be contingent at most on ω_{jt}. At the same time, the *relative* price of two firms i and j is pinned down, from the consumer's side, by the ratio of their output levels. The latter are themselves sensitive, in general, to the true state s^t. Putting the two properties together, it must be that

$$\log p(\omega_i^t) - \log p(\omega_{jt}) = -\rho \left[\log y(\omega_i^t, s^t) - \log y(\omega_i^t, s^t)\right]$$

Clearly, this can be true only if the dependence of y_{it} on s^t cancels with that of y_{jt} on s^t, so as to make the right-hand-side of the above condition invariant to s^t—which is precisely what the log-separability assumption guarantees. Indeed, unless this condition is satisfied by a candidate allocation, relative prices along this allocation have to vary with s^t, which in turn would require that at least one of the two corresponding nominal prices be contingent on s^t.

We conclude that part (ii) highlights the interaction of informational and nominal frictions: log-separability is not an issue if nominal prices can adjust to the true state s^t, but becomes relevant as soon as nominal prices must be pre-set on the basis of imprecise information about s^t. The following is then a direct implication of this observation.

Proposition 4. *A flexible-price equilibrium allocation can be replicated as a sticky-price equilibrium allocation if and only if it is log-separable.*

To see this, take any flexible-price allocation. By Proposition 1, this allocation necessarily satisfies conditions (5)-(7). Conditions (14)-(17) are then trivially satisfied once we let $\chi(\omega_i^t, s^t) = 1$ for all (ω_i^t, s^t). That is, any flexible-price allocation necessarily satisfies part (i) of Proposition 2. It follows that a flexible-price allocation can be replicated under sticky prices if and only if it satisfies part (ii) of that proposition, that is, if and only if it is log-separable.

4.2 Optimal monetary policy

We are now ready to state our main result about the nature of the optimal allocation.

Theorem 2. *The Ramsey optimum with sticky prices attains the constrained efficient allocation if and only if the latter is log-separable. When this is the case, the optimal tax offsets the monopoly distortion and the optimal monetary policy replicates the corresponding flexible-price allocation.*

This result follows from the combination of Theorem 1 and Proposition 4. By Theorem 1, we know that the constrained efficient allocation always belongs to the set of flexible-price allocations and is indeed attained with a non-contingent tax that removes the monopoly distortion. Next, by Proposition 4 together with the assumption that this allocation is log-separable, we know that this allocation can be replicated under sticky prices. It follows that the optimal monetary policy exists and is indeed one that replicates the optimal flexible-price allocation.

As with Theorem 1 and Proposition 4, this result does not hinge on the details of either the information structure or the preference and technology specification of the economy. It nevertheless requires the efficient allocation to be log-separable: unless this is the case, there is *no* monetary policy that can attain constrained efficiency.

Notwithstanding this qualification, it is important to note that Theorem 2 revisits a key policy lesson of the last three decades. In the standard New-Keynesian framework, the underlying flexible-price allocations are first-best efficient as soon as the monopoly distortion is removed. It follows that the best monetary policy can do is to replicate the underlying flexible-price allocations. Versions of this lesson have been established for the canonical Keynesian framework by, inter alia, Goodfriend and King (1997, 2001), Rotemberg and Woodford (1997), Woodford (2002), and Khan, King and Wolman (2003). To the best of our knowledge, our paper is the first to establish the robustness of this important lesson to settings where firms face informational frictions.[11]

We now revisit the log-separability restriction, upon which the ability of monetary policy to implement the constrained efficient allocation rests.

Proposition 5. *Suppose that the technology exhibits constant elasticity with respect to the input that can adjust to realized demand:*

$$\frac{\partial \log F(k,\ell,n)}{\partial \log n} = \alpha \tag{18}$$

for some constant $\alpha \in (0,1)$. Then, every flexible-price equilibrium allocation—including the constrained efficient one—is log-separable, and can thus be replicated under sticky prices.

Note that this result nests the familiar Cobb-Douglas production function. Also, the aforementioned specification of the technology suffices for *all* flexible-price allocations to be log-separable. If we are only interested in the Ramsey optimum, this property can be relaxed by assuming that a log-linear relation holds between y_{it} and n_{it} along the constrained efficient allocation. Finally, to the extent that one focuses on log-linear approximations (which is what most of the applied policy literature does), this condition is automatically satisfied. Combined, these observations imply that, for all practical purposes, the log-separability restriction can be taken for granted.

4.3 Remark

Before proceeding to address the (sub)optimality of price stability, it is worth contrasting our results to policy prescriptions that fail to properly incorporate how informational frictions restrict the set of allocations which the planner can induce with different fiscal and monetary policies.

To appreciate what we mean, consider the following example, which can be solved in closed form. First, abstract from capital accumulation and taste shocks; let preferences be given by $U(C,\xi) = C$, $V(N,\xi) = N^2$, and $H(\ell) = \ell^2$; and let production be given by

$$y_{it} = A_{it}\ell_{it}^{1-\alpha}n_{it}^{\alpha},$$

for some $\alpha \in (0,1)$. Second, let the log of the productivity of a firm be given by

$$a_{it} \equiv \log A_{it} = \bar{a}_t + \zeta_{it},$$

[11] This property was first noted in the 2008 version of our paper, albeit in a much less general framework.

where $\bar{a}_t \sim \mathcal{N}(0, \sigma_a^2)$ represents an aggregate productivity shock (i.i.d. across t) and $\zeta_{it} \sim \mathcal{N}(0, \sigma_\zeta^2)$ represents a purely idiosyncratic shock (i.i.d. across both i and t, and orthogonal to $\bar{a}_t$). Third, let the information that manager i receives in period t be given by $\tilde{\omega}_{it} \equiv (a_{it}, z_{it})$, where a_{it} is the firm's own productivity and z_{it} is a noisy signal of aggregate productivity; the latter is given by

$$z_{it} = \bar{a}_t + \varepsilon_t + \eta_{it},$$

where $\eta_{it} \sim \mathcal{N}(0, \sigma_\eta^2)$ is idiosyncratic noise (i.i.d. across both i and t) and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is common noise (i.i.d across t but perfectly correlated across i). Finally, consider the class of fiscal and monetary policies such that $(1 - \tau_t)$ and M_t can be expressed as log-linear functions of $\tilde{s}_t \equiv (\bar{a}_t, \varepsilon_t)$.

For any such pair of policies, the equilibrium ℓ_{it} and p_{it} can be expressed as a log-linear functions of $\tilde{\omega}_{it}$, while the equilibrium n_{it} and y_{it} can be expressed as log-linear functions of $\tilde{\omega}_{it}$ and $\tilde{s}_t$.[12] By implication, aggregate output Y_t can be expressed as a log-linear function of the aggregate productivity shock $\bar{a}_t$ and the aggregate noise shock ε_t. What is more, by appropriately choosing the response of τ_t and/or M_t to these shocks, the policy maker is free to induce *any* sensitivity he wishes in the response of aggregate output to either of these shocks.[13]

One may thus be tempted to argue that the optimal policy should is on that eliminates the impact of the noise and replicates the first-best level of aggregate output. While this policy is feasible in the above example, it is actually suboptimal, for it fails to take into account the distortions it induces in the cross-section of the economy once firms are informationally constrained. When information is complete, the first best has each firm's output depend positively, not only on its own productivity, but also on the aggregate productivity. The response of aggregate output to aggregate productivity shocks then reflects the joint response of individual output to own and aggregate productivity. Once information is incomplete, eliminating the impact of the noise on aggregate output requires that each firm's output be insensitive to z_{it}. But then the first-best level of aggregate output can be obtained only if each firm's reaction to its own productivity is raised enough to compensate for lack of direct dependence on aggregate productivity. It follows that the aforementioned policies come at the cost of excessive cross-sectional dispersion in relative quantities and relative prices, as well as excessive firm-level volatility, which explains why they are not optimal.

Moving beyond specific examples like the above, it might be impossible to characterize in closed form the allocations that are induced by different policies, and it might be hard to comprehend in their totality the multiple ways through which noise may manifest. Different allocations may

[12]This follows directly from the power-form specification of the preferences and technologies and the Gaussian specification of the various shocks and noises.

[13]To understand this, consider the flexible-price allocations and let $\log(1 - \tau_t) = \varphi_a \bar{a}_t + \varphi_\varepsilon \varepsilon_t$, where φ_a and φ_ε are scalars that parameterize the response of the policy to the underlying shocks. It is then easy to show that aggregate output is given, up to a constant, by $\log Y_t = \Phi_a \bar{a}_t + \Phi_\varepsilon \varepsilon_t$, where the vector $(\Phi_a, \Phi_\varepsilon)$ is an invertible linear function of the vector $(\varphi_a, \varphi_\varepsilon)$. By appropriately choosing the latter two coefficients, the policy maker is thus free to choose any sensitivities he wishes for the response of aggregate output to either the productivity shock $\bar{a}_t$ or the noise shock ε_t. Intuitively, a negative φ_ε induces a firm's production to respond less to the noisy signal z_{it}, while a positive φ_a induces a firm's production to respond more to productivity a_{it}. Similar arguments apply to monetary policy once prices are sticky.

involve different responses of the economy, not only to aggregate and idiosyncratic fundamentals, but also to the entire hierarchy of the firms' beliefs regarding the underlying state. The "beauty" of our efficiency concept is that it encapsulates, in a concise yet complete way, all the various trade-offs that the planner may be facing when choosing across different allocations. Of course, if the planner could get rid of the noise in the first place and attain full-information outcomes, this would be wonderful. To the extent, however, that the planner cannot alter the information structure itself, our analysis provides the right welfare-based benchmark for optimal policy.[14]

5 On the Suboptimality of Price Stability

The analysis so far has established the optimality of monetary policies that replicate flexible-price allocations. Within the New-Keynesian paradigm, replicating flexible-price allocations is usually synonymous to targeting price stability. We now show that this is not the case in an environment with informational frictions.[15]

Towards this goal, we henceforth assume that the technology satisfies (18). As already discussed, this is without serious loss of generality for all practical purposes, and is instrumental for the efficient allocation to be implementable. We then obtain the following result, which holds irrespective of the details of the stochastic structure of the economy and the behavior of the efficient allocation.

Proposition 6. *Suppose the technology satisfies (18) and consider any sticky-price equilibrium that implements the efficient allocation. There exists a deterministic sequence J_t such that the following property holds:*

$$\log P(s^t) = J_t - \frac{1}{\rho(1-\alpha)+\alpha}\left\{(1-\alpha)\log Y\left(s^t\right) + \alpha \log \frac{Y\left(s^t\right)}{N\left(s^t\right)}\right\} \tag{19}$$

This result establishes that, along an optimal path, the price level can be decomposed into two components: a deterministic trend, which is given by J_t; and a stochastic component, which is given by the remaining part of the left-hand-side of condition (19) and is a negative transformation of aggregate output and labor productivity.

The deterministic trend captures a certain form of nominal indeterminacy: because nominal prices are free to be contingent on calendar date, the optimal monetary policy and the associated

[14]Starting from this benchmark, it might be natural to extend the analysis in settings where the information structure is endogenous. Alternatively, one may find it useful to restrict the planner further, for example by limiting the contingencies which policy can react to, or requiring that the allocations be implemented with "realistic" Taylor rules. These issues are clearly relevant, but are beyond the scope of this paper.

[15]The optimality of price stability in New-Keynesian settings has to be qualified by a definition of the appropriate price index. For example, if commodity prices are flexible but wages are sticky, then the appropriate price index is a wage index. The arguments we make in this section have nothing to do with this issue: we establish sub-optimality of stabilizing the "right" price index. More specifically, since the prices that are sticky in our framework are those of intermediate-good producers, the appropriate price index is the CES aggregator P_t of the prices of these producers. This keeps our analysis directly comparable to, inter alia, Goodfriend and King (1997, 2001), Rotemberg and Woodford (1997), Woodford (2003), Adao, Coreia and Teles (2003), Khan, King and Wolman (2003), Galí (2008).

price level are pinned down only up to an arbitrary deterministic time path. This kind of indeterminacy is generic to *any* model in which incomplete nominal adjustment rests merely on informational frictions. We expect it to vanish in models that combine informational frictions with Calvo-like nominal rigidity, but this is beyond the scope of this paper—and it is largely orthogonal to the point we wish to make, which regards the *stochastic* properties of the price level. Without serious loss of generality, we henceforth select the optimal policy for which $J_t = 0$.[16]

The stochastic component in condition (19) then clarifies the precise nature of the suboptimality of price stability: to replicate flexible price-allocations and maximize welfare, the monetary authority must target a certain negative correlation between the price level and the level of real economic activity, as measured by aggregate output Y_t and aggregate labor productivity Y_t/N_t.

Let us explain this result. Since the technology satisfies (18), there must be an increasing function $g : \mathbb{R}^2_+ \to \mathbb{R}_+$ such that, for all (k, ℓ, n),

$$F(k, \ell, n) = g(k, \ell)n^{\alpha}.$$

Firm output can thus be expressed as

$$y_{it} = q_{it}n_{it}^{\alpha}$$

where $q_{it} \equiv A_{it}g(k_{it}, \ell_{it})$ is a composite of the firm's TFP, capital, and managerial effort. Note then that α parameterizes the fraction of (log) output that can adjust to the true state, while q_{it} identifies the effective productivity of labor, that is, the input that is free to adjust.

Different firms have different productivities, q_{it}, either because they are exogenously endowed with different TFP levels, A_{it}, and/or because they endogenously choose different levels of capital, k_{it}, and managerial effort, ℓ_{it}. Difference in capital and effort in turn emerge, not only because firms have different TFP levels, but also because they hold different beliefs about the state of the economy: optimistic managers may work harder and/or invest more, leading to an increase in the ex post productivity of labor in their firms.

Wherever the differences in q_{it} come from, efficiency requires that these differences are reflected in output levels: more productive firms should produce more. Equilibrium, in turn, requires that differences in output levels are reflected in relative prices: more productive firms should sell at lower relative prices. It follows that the relative price of a firm must be decreasing in its own q_{it}, no matter whether the latter is driven by TFP or the firm's belief about the state of the economy.

But now note that q_{it} is, in effect, private information to the firm: this object is pinned down by the firm's own TFP and the firm's own belief about the state of the economy, neither of which is common knowledge to other firms. Hence, for the *relative* price of a given firm i to fall with its own q_{it}, it must be that the *nominal* price p_{it} set by the firm is a decreasing function of its q_{it}. But

[16] To be more precise, indeterminacy pertains with regard to the contingency of the price level to *any* variable that is common knowledge. This explains the role of our assumption that the realization of *no* random variable is common knowledge to the firms: if some random variables were common knowledge (say, the weather), the price level could be an arbitrary function of those variables too. If we were to relax the aforementioned assumption, condition (19) would apply only once one partials out from (P_t, Y_t, N_t) the effect of any common-knowledge shocks.

then it follows that the price level must be a decreasing function of the aggregate of the q's: when the majority of firms happens to have higher labor productivity either because of an aggregate TFP shock or because of a correlated shock in their beliefs about the state of the economy, the aggregate price level must fall, otherwise the right relative prices would not be induced.

This explains, in effect, that the price level must be inversely related to aggregate labor productivity, as measured by the aggregate of q_{it}. What then remains is to understand the mapping from this object to aggregate output. Normally, one expects aggregate output and aggregate productivity to be positively correlated. This, however, might fail to be the case if taste shocks happen to move the marginal rate of substitution between consumption in the opposite direction, so that booms are periods of low disutility of work, high employment, and low marginal product of labor. This explains why the co-variation between the price level and aggregate output ultimately hinges on the cyclical properties of labor productivity, as indicated by condition (19).

Building on these observations, we reach the following key result regarding the cyclical properties of the price level along the Ramsey optimum.

Theorem 3. *Suppose that either of the following three conditions holds:*

(i) $\alpha = 0$*; or*

(ii) Y_t/N_t *is procyclical along the constrained efficient allocation; or*

(iii) capital is fixed and there are no taste shocks.

Then, the optimal policy induces a negative co-variation between the price level and real output.

The sufficiency of the first two conditions follow directly from Proposition 6: when α is zero (or close enough to zero), or when Y_t/N_t is procyclical (or not too countercyclical), condition (19) guarantees that P_t is negatively correlated with Y_t. The first condition imposes, in effect, that most production is subject to informational frictions. The second condition means that, if some inputs are free to adjust to realized demand, the marginal product of these inputs is procylical. While somewhat restricted, both conditions seem quite plausible empirically.

The third condition then identifies exogenous properties that suffice for the second one to hold: if capital is fixed and there are no shocks to the marginal rate of substitution between consumption and leisure, then labor productivity Y_t/N_t is necessarily procyclical along the efficient allocation. In fact, in this case that actually nests most of the pertinent literature on informational frictions, there exists an increasing function $\Gamma : \mathbb{R} \to \mathbb{R}$ such that, along the optimal path,

$$\log P_t = -\Gamma(\log Y_t).$$

The optimal policy can thus be represented as a simple target for the price level in terms of a negative function of real output.[17]

We should emphasize that the aforementioned conditions are sufficient but not necessary. To overturn the countercyclicality of the optimal price level, one needs *both* of the following: a sufficiently high α, so that a significant part of production is free to adjust to the true state; and the

[17] Perhaps interestingly, there actually exist special cases in which this reduces to a stabilization target for nominal GDP. These cases correspond to $\Delta = 1$ in the example of the next section.

bulk of the business cycle to be driven by taste shocks, so that booms are associated with sufficiently low labor productivity. While theoretically possible, this scenario seems highly implausible from an empirical perspective; we thus ignore it for the remainder of this discussion.

Our findings thus offer an interesting contrast to previous policy lessons. In the standard New-Keynesian framework, the optimality of flexible-price allocations is synonymous to the optimality of targeting price stability. Here, by contrast, we have shown that price stability is suboptimal even though flexible-price allocations remain optimal (Theorem 2). Furthermore, to replicate flexible-price allocations and maximize welfare, the monetary authority must actually "lean against the wind": during booms, the optimal policy contracts, causing prices to fall and moderating the increase in output (relative to the one that would have obtained under price stability); during recessions, the converse is true. A policy that targets price stability would therefore not only distort relative prices but also induce real output to overreact to any underlying business cycle disturbances.

At this point, it is worth emphasizing that our results do not hinge on the precise nature of the shocks that drive the business cycle. Whether aggregate fluctuations are triggered by innovations in true productivity as in standard DSGE, or by noise shocks as in Lorenzoni (2009) and Angeletos and La'O (2009), or by higher-order beliefs and forces akin to "animal spirits" as in Angeletos and La'O (2011a), the monetary policy must act counter-cyclically in order to induce firms to make the right pricing and production choices. This is because our result depends only on the fact that relative prices must reflect differences in firm beliefs and output levels, not on the precise nature of the underlying structural shocks. We will illustrate this point further in the next section.

Furthermore, we should clarify that the suboptimality of price stability pertains as long as the informational friction has a non-trivial role in the planner's problem. By this we mean to rule out degenerate cases in which either firms have no private information about the state of the economy or this information plays no essential allocative role along the optimum.

Suppose, in particular, that there are no idiosyncratic productivity risk and that all real production choices can adjust to the true state of the economy. In this case, the (full-information) first best has that all firms choose the same inputs and produce the same amount of output. But then the planner can trivially implement this allocation by targeting price stability. This is because all relative prices are entirely constant along the Ramsey optimum and, by implication, there is no need to induce the nominal pricing decisions of the firms to react to any kind of information.

But now suppose that at least some of the firms' real production choices have to be contingent on incomplete information about the state of the economy. That immediately renders price stability suboptimal, precisely because relative prices must move with firms' beliefs of the state.

Moreover, even if all production choices can adjust to the realized aggregate state, price stability will remain suboptimal to the extent that firms cannot disentangle idiosyncratic productivity shocks from aggregate ones when they set prices. To see this more clearly, suppose that when setting its nominal price, the only information a firm has about the aggregate productivity shock is its own productivity. Clearly, efficiency requires that the relative price of a firm falls with the firm's relative productivity. For this to be the case in equilibrium, the nominal price set by any firm must be decreasing in its own productivity level for any given aggregate productivity level. Because the firm has no other information about the aggregate state, it follows that there must exist a decreasing function g such that $p_{it} = g(A_{it})$. But then the aggregate price level has to be a decreasing function of the aggregate productivity shock, verifying our claim.[18]

Last but not least, we should highlight that the optimal cylicality of the price level does not hinge on the details of the informational friction. To see this, note that condition (19) identifies a structural condition between the price level and real economic activity that is invariant to either the specification of the underlying productivity shocks or the specification of the information structure. As anticipated in the preceding discussion, this is because our result is driven merely by the need to make relative prices reflect differences in the firms' beliefs and their associated output levels, and does not depend on the details of the shocks and noises that drive these beliefs. We illustrate this point further in the next section.

These observations underscore the robustness and generality of our insights. Any reasonable perturbation away from the convenient but unrealistic assumption of common knowledge appears to suffice for "divine coincidence" to break down. As long as information is incomplete, efficiency dictates that a firm's production choices, and hence its relative price, must respond to its beliefs about the state of the economy. For this to be implemented along a sticky-price equilibrium, the nominal price set by that firm must co-move with its own belief. But this implies that the aggregate price level must co-move with the average belief, and hence with the true state of the economy. Price stability is thus suboptimal, despite the fact that flexible-price allocations remain optimal.

6 An illustration

To further illustrate the nature of the optimal allocation, we revisit the example that we introduced in subsection 4.3. With aggregate productivity now been subject to persistent shocks, firms continuously update their beliefs about the underlying state of the economy. This example is thus closely related to those studied in Woodford (2002), Lorenzoni (2009), Angeletos and La'O (2009), and Mackowiak and Wiederholt (2009).

As in subsection 4.3, we abstract from capital and let a firm's output be given by

$$y_{it} = A_{it}\ell_{it}^{1-\alpha}n_{it}^{\alpha}, \tag{20}$$

where $\alpha \in (0,1)$. The log-productivity of the firm is once again given by $a_{it} \equiv \log A_{it} = \bar{a}_t + \zeta_{it}$, where $\zeta_{it} \sim \mathcal{N}(0, \sigma_\zeta^2)$ is the idiosyncratic productivity shock and $\bar{a}_t$ is the aggregate one. The latter is now assumed to follow a log-normal AR(1) process:

$$\bar{a}_t = \varrho\bar{a}_{t-1} + v_t,$$

[18] Incidentally, this qualifies one of the results in Paciello and Wiederholt (2011). The version of their model with productivity shocks and exogenous information is a special case of the framework we have considered here. Price stability is found to be optimal in that setting only because the informational friction is inconsequential for the planner's problem. Allowing firms to have private information about both idiosyncratic and aggregate productivity shocks would render price stability suboptimal in that setting just as it does in our paper.

where $\upsilon_t \sim \mathcal{N}(0, \sigma_\upsilon^2)$ is the period-t innovation. σ_υ parameterizes the volatility of the productivity process and $\varrho \in (0,1)$ its persistence. The information that manager i receives in period t consists of the firms's own productivity and of the noisy signal

$$z_{it} = \bar{a}_t + \varepsilon_t + \eta_{it}, \tag{21}$$

where $\eta_{it} \sim \mathcal{N}(0, \sigma_\eta^2)$ is the idiosyncratic noise component and $\varepsilon_t \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ is the common noise component. Finally, we abstract from taste shocks and let

$$U(C,\xi) = \frac{C^{1-\gamma}}{1-\gamma}, \quad V(N,\xi) = \frac{1}{1+\epsilon} N^{1+\epsilon} \quad \text{and} \quad H(\ell) = \frac{1}{1+\epsilon}\ell^{1+\epsilon}, \tag{22}$$

where $\gamma, \epsilon > 0$. The aggregate state is thus given by $s^t = (\tilde{s}_0, ..., \tilde{s}_t)$ where $\tilde{s}_t = (\bar{a}_t, \varepsilon_t)$, while the information set of a firm is given by $\omega_i^t = (\tilde{\omega}_{i0}, ..., \tilde{\omega}_{it})$ where $\tilde{\omega}_{it} = (a_{it}, z_{it})$.

As in Woodford (2002), we can interpret the idiosyncratic noise in the private signal z_{it} interchangeably as either the product of dispersed information or the product of inattention. At the same time, we introduce correlation in this noise—through the common shock ε_t—in order to accommodate "news" and "noise shocks" as in Jaimovich and Rebelo (2009), Lorenzoni (2009), and Blanchard et al. (2011). Note, in particular, that the case of public (or nearly public) signals can be nested by taking the limit as $\sigma_\eta \to 0$, while the case of purely private information can be nested by taking the limit as $\sigma_\varepsilon \to 0$. Finally, the common noise ε_t in the aforementioned signal can also be thought of as a tractable proxy for richer forms of correlated errors in higher-order beliefs, which may manifest as forces akin to "animal spirits" (Angeletos and La'O, 2011a).

At the technical level, the procedure we follow to solve for the efficient allocation is based on a Kalman filter and is similar to that used in Woodford (2002); see the Appendix for details. At the conceptual level, however, there is an important difference. In Woodford (2002), productivity is constant and the business cycle is driven by exogenous shocks to money supply. Here, instead, the business cycle is driven by productivity and noise shocks, and monetary policy is optimally set in response to these shocks. A similar qualification applies to Lorenzoni (2009), which focuses on productivity and noise shocks but, unlike us, imposes a monetary policy rule that fails to implement the efficient allocation.

To illustrate the dynamics of the Ramsey optimum, we henceforth focus on a particular parameterization. We interpret a period as a quarter. Accordingly, we set $\sigma_\nu = 0.02$ for the standard deviation of the productivity innovation and $\varrho = 0.98$ for its persistence, which is roughly consistent with US data about labor productivity. Following Woodford (2003), we set $\gamma = .2$ and $\epsilon = .3$; these values help capture the empirical regularity that output and employment move closely together over the business cycle, while real wages are almost acyclical. Next, we set $\alpha = 0$, so that all quantity choices are subject to informational frictions, and $\rho = 2$, which helps generate a non-trivial complementarity in real production choices as in Angeletos and La'O (2009).[19] Finally, we set $\sigma_\zeta = 10\sigma_\nu$,

[19] The new-Keynesian literature often associates the elasticity parameter ρ with the size of the markup. However, while observed monopoly markup provide direct information about the level of profits (and, perhaps, about the level

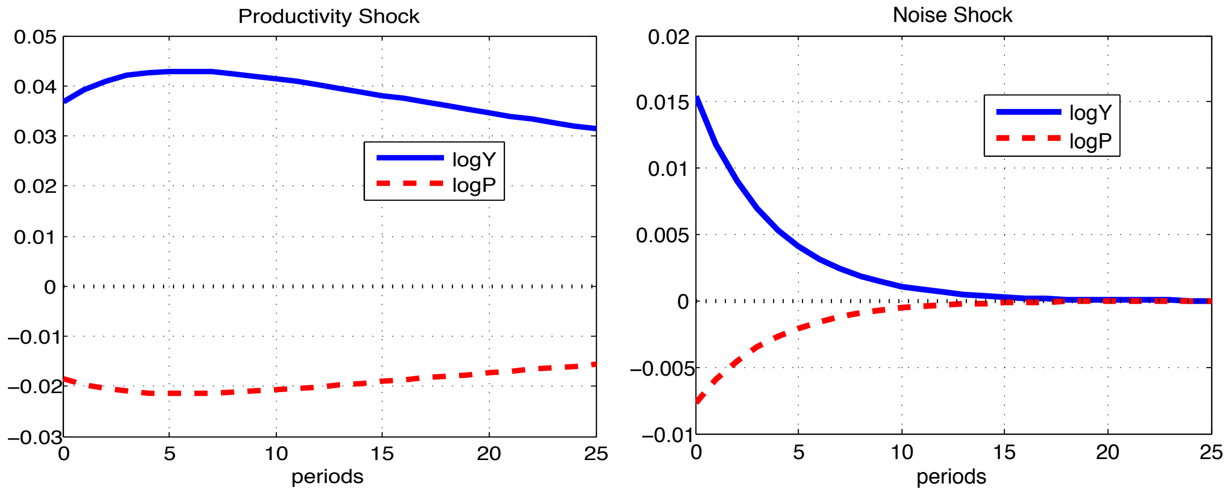

Figure 1: Impulse responses of aggregate output and the price level to an aggregate productivity shock (left panel) and an aggregate noise shock (right panel).

$\sigma_\varepsilon = 4\sigma_\nu$ and $\sigma_\eta = 0$; these levels of noise help slow down learning and are comparable to those featured in Woodford (2002), Lorenzoni (2009), and Mackowiak and Wiederholt (2009).[20]

With this parameterization at hand, Figure 1 depicts the impulse responses of real output and the price level to a productivity shock (a positive innovation in $\bar{a}_t$) and a noise shock (a positive innovation in ε_t). In each case, the size of the shock is equal to one standard deviation and the impulse responses are represented in terms of log-deviations from steady-state values.

As evident in the left panel of the figure, real output responds positively to the innovation in productivity, pretty much as in any standard RBC model. The informational friction, however, induces some inertia in this response: as aggregate productivity is not common knowledge, it takes time for the constrained-efficient level of output to build up towards the first-best level. Furthermore, because the productivity shock eventually dies away, the boom also dies away. The combination of these two forces explain why the impulse response of real output has a hump-shape. The impulse response of the price level is then, in effect, the mirror image of that of output. As output expands, the price level falls; when output reverts, the price level increases. Inflation is thus negatively correlated with real output growth.

The right panel of Figure 1 shifts the focus from productivity shocks to noise shocks. The origin and the dynamics of the business cycle are now quite different. The boom now originates

of distortion in the economy), they do not by themselves identify either the complementarity of real decisions or the elasticity of relative prices to relative quantities. Indeed, it is straightforward to disconnect the monopoly markup in the theory from the relevant elasticity and complementarity parameters, either by introducing limit pricing (in which case the monopoly markup is pinned down by an outside competitive fringe) or by adding a nested CES structure as in Angeletos and La'O (2009).

[20]Specifically, we have that the ratio of the standard deviation of the signal noises to the standard deviation of the innovation to the fundamental (productivity) is between 4 and 10. For comparison, the corresponding ratio is 4 to 6.3 in Woodford (2002), and 10 to 30 in Lorenzoni (2009).

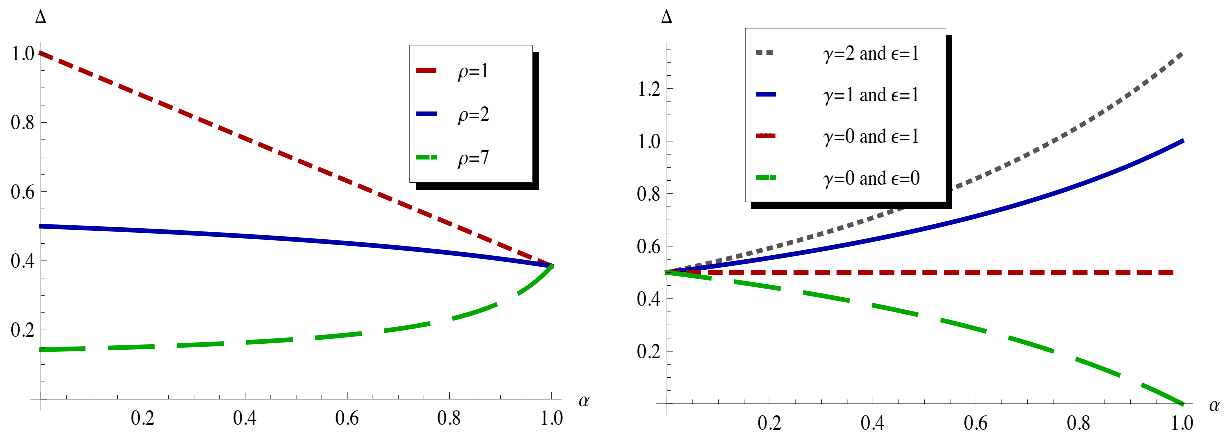

Figure 2: Comparative statics of Δ, the ratio of the standard deviation of inflation to that of output growth along the Ramsey optimum.

from optimistic beliefs that are driven by noise rather than the true productivity shocks. Initially firms are (rationally) confused, they expect the economy to expand, they hire more workers, and they help sustain a boom that would have been unwarranted if the true fundamentals (aggregate productivity) were common knowledge. But as time passes and firms keep receiving more and more signals about the state of the economy, their optimistic beliefs are corrected, and the boom dies away. Notwithstanding these intriguing differences in the positive properties of the business cycle, the qualitative picture that emerges for the co-movement of output and prices along the Ramsey optimum is the same: prices are below trend whenever output is above trend, and inflation is once again negatively correlated with output growth.

We conclude this section by illustrating how the optimal counter-cyclicality of the price level depends on the exogenous parameters of the economy. To this goal, we first note that, for the class of economies considered in this section, condition (19) reduces to

$$\log P_t = -\Delta \log Y_t,$$

where $\Delta \equiv \frac{(\epsilon+1)-\alpha(1-\gamma)}{(\epsilon+1)(\rho-\alpha(\rho-1))} > 0$. When $\Delta = 1$, the above reduces to a stabilization target for nominal GDP. More generally, Δ pins down the volatility of the inflation rate relative to that of output growth:

$$\frac{Var(\log P_t - \log P_{t-1})}{Var(\log Y_t - \log Y_{t-1})} = \Delta^2.$$

Echoing a more general observation we made in the previous section, we see that the aforementioned ratio is invariant to either the productivity process or the information structure: these primitives are crucial for the precise dynamics of the business cycle, but are irrelevant for determining the optimal counter-cyclicality of the price level. Instead, the latter depends only on the preference parameters (ϵ, γ, ρ), which govern the cyclical behavior of labor supply and the sensitivity of relative prices to relative quantities, and the technology parameter α, which governs the fraction of output that can adjust to the realized state (and hence the bite of the informational friction).

With these observations in mind, Figure 2 illustrates the sensitivity of Δ to α and ρ (left panel) and to ϵ and γ (right panel). The following patterns are evident. First, Δ is decreasing in ρ, highlighting how the counter-cylicality of the optimal price level hinges precisely on the need for relative prices to respond to cross-firm differences in beliefs and output choices: the higher the elasticity of demand to prices, the smaller the variation in equilibrium prices that is needed in order to induce the optimal variation in equilibrium quantities. Second, the impact of α is generally non-monotonic, while higher values for ϵ and γ tend to strengthen the optimal counter-cylicality of the price level. Finally, the value of Δ is about 0.5 for our baseline parameterization (meaning that inflation is half as volatile as output growth), it lies above 0.3 for a wide range of plausible parameter values, and it is often above 0.5. Notwithstanding the fact that the exercise conducted here—despite its proximity to previous work—is too stylized to permit a serious quantitative assessment, this last finding indicates that the optimal counter-cylicality of the price level can be non-trivial in magnitude.

7 Concluding remarks

The modern literature on optimal monetary policy starts with two key lessons: (i) the efficiency of flexible-price allocations in the baseline New-Keynesian model, and (ii) and the consequent optimality of policies that target price stability. In central-bank jargon, stabilizing the price level also guarantees perfect stabilization of the appropriate, welfare-based measure of the "output gap". Researchers have sought to move away from this kind of "divine coincidence", and hence open the door for active output-stabilization policies, either by perturbing the efficiency of flexible-price allocations (e.g., by adding real mark-up shocks) or by introducing additional nominal frictions (e.g., by adding sticky wages along with sticky prices). A distinct strand of the literature has studied various macroeconomic implications of informational frictions. However, this literature has largely shied away from examining how informational frictions impact the normative properties of the business cycle and the design of optimal monetary policy. This paper builds a bridge between these two strands of the literature and shows that informational frictions break the aforementioned "divine coincidence" without disturbing the efficiency of flexible-price allocations.

More specifically, our first key result is that informational frictions do not by themselves break the optimality of flexible-price allocations: the (constrained) efficient allocation is contained in the set of flexible-price Ramsey allocations. As a result, the optimal monetary policy remains the one that replicates flexible-price allocations, pretty much as in the New-Keynesian paradigm.

Our second key result is then that price stability is no more consistent with replicating flexible-price allocations. Rather, the optimal monetary policy induces a negative relation between the nominal price level and the level of real output. In this sense, the monetary authority must lean against the wind: it must be less accommodative than the one that would have stabilized the price level. If the monetary authority fails to do so, inefficiency will manifest, not only in cross-sectional price dispersion, but also in the response of aggregate output to the productivity shocks; stabilizing the appropriate measure of the output gap thereby requires moving away from price stability.

These results were derived in a framework that allowed for a generic formulation of the informational friction: we bypassed any specific "micro-foundation" of the informational friction and instead captured the latter by an arbitrary measurability constraint on the set of feasible allocations and prices. In this regard, our results are quite general and provide a clean benchmark for understanding how informational frictions impact the efficiency of flexible-price allocations and the nature of optimal monetary policy.

As with any other benchmark, however, our analysis is subject to certain limitations. The one that seems most intriguing from a conceptual perspective regards the potential endogeneity of the informational friction. In our setting, information sets were assumed to be exogenous to the allocation under consideration. Clearly, this ceases to be the case as soon as one recognizes that agents may optimally choose the amount of information they collect about (or the attention they pay to) the state of the economy, or that the informativeness of price and other macroeconomic signals is likely to be affected by how agents respond in the first place to their available information. While these forms of endogeneity in the information structure need not upset the essence of our insights regarding allocative efficiency, exploring their implications for optimal policy is certainly worthwhile. Complementary in this respect is the ongoing work by Paciello and Wiederholt (2011), which focuses on optimal collection of information (or optimal attention), and that by Angeletos and La'O (2011b), which focuses on learning through prices and macro statistics.

From a practical perspective, on the other hand, it seems natural to extend the analysis to settings that combine informational frictions with conventional, Calvo-like, price stickiness. Clearly, this will not interfere with our result regarding the optimality of flexible-price allocations: the nature of the nominal rigidity is irrelevant for the characterization of either the constrained efficient allocation or the set of flexible-price allocations. Furthermore, to the extent that the efficient allocation remains implementable, we expect our result regarding the sub-optimality of price stability to survive: the *state-variation* in the price level we have documented here is a necessary condition for the replication of the right relative prices. At the same time, such an extension will enrich our understanding of the conditions under which the constrained efficient allocation may, or may not, be implementable. It will also permit one to study implementation details that might not affect our key insights but remain relevant for practical purposes—such as the precise shape of the Taylor rule that may implement the efficient allocation.

References

[1] Adao, Bernardino, Isabel Correia, and Pedro Teles (2003), "Gaps and Triangles," *Review of Economic Studies* 70(4), 699-713.

[2] Amador, Manuel, and Pierre-Olivier Weill (2011), "Learning from Private and Public Observations of Others Actions,"*Journal of Economic Theory*, forthcoming.

[3] Angeletos, George-Marios, and Jennifer La'O (2008), "Dispersed Information over the Business Cycle: Optimal Fiscal and Monetary Policy," MIT mimeo.

[4] Angeletos, George-Marios, and Jennifer La'O (2009), "Noisy Business Cycles," *NBER Macroeconomics Annual 2009*, 24, 319-378.

[5] Angeletos, George-Marios, and Jennifer La'O (2011a), "Decentralization, Communication, and the Origins of Fluctuations," NBER Working Paper 17060.

[6] Angeletos, George-Marios, and Jennifer La'O (2011b), "Learning through Prices and Optimal Monetary Policy," work in progress.

[7] Angeletos, George-Marios, and Alessandro Pavan (2007), "Efficient Use of Information and Social Value of Information," *Econometrica* 75(4), 1103-1142.

[8] Angeletos, George-Marios, and Alessandro Pavan (2008), "Policy with Dispersed Information," *Journal of the European Economic Association* 7, 11-60.

[9] Atkeson, Andrew, Varadarajan V. Chari, and Patrick J. Kehoe (2010), "Sophisticated Monetary Policies,"*Quarterly Journal of Economics* 125, 47-89.

[10] Benigno, Pierpaolo, and Michael Woodford (2004), "Optimal Monetary and Fiscal Policy: A Linear-Quadratic Approach," in M. Gertler and K. Rogoff, eds., *NBER Macroeconomics Annual 2003*, vol. 18, Cambridge: MIT Press, 271-333.

[11] Blanchard, Olivier, and Jordi Gali (2007), "Real Wage Rigidities and the New Keynesian Model," *Journal of Money, Credit and Banking* 39(1), 35-65.

[12] Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni (2010), "News, Noise, and Fluctuations: an empirical exploration." MIT mimeo.

[13] Cochrane, John (2011), "Determinacy and Identification with Taylor Rules, " *Journal of Political Economy* 119, 565-615.

[14] Chari, Varadarajan V., and Patrick J. Kehoe (1999), "Optimal fiscal and monetary policy," in J. B. Taylor & M. Woodford, eds., *Handbook of Macroeconomics*, vol. 1, chapter 26, 1671-1745, Elsevier.

[15] Galí, Jordi (2008), *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework,* Princeton: Princeton University Press.

[16] Goodfriend, Marvin, and Robert King (1997), "The New Neoclassical Synthesis and the Role of Monetary Policy," *NBER Macroeconomics Annual 1997*, Cambridge: MIT Press.

[17] Goodfriend, Marvin, and Robert King (2001), "The Case for Price Stability," NBER Working Paper 8423.

[18] Hellwig, Christian (2002), "Public Announcements, Adjustment Delays and the Business Cycle," UCLA mimeo.

[19] Khan, Aubhik, Robert King, and Alexander Wolman (2003), "Optimal Monetary Policy," *Review of Economic Studies* 70, 825-860.

[20] Jaimovich, Nir, and Sergio Rebelo (2009), "Can News About the Future Drive the Business Cycle?," *American Economic Review* 99, 1097-1118.

[21] Lorenzoni, Guido (2009), "A Theory of Demand Shocks," *American Economic Review* 99(5), 2050-2084.

[22] Lorenzoni, Guido (2010), "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information."*Review of Economic Studies* 77, 305-38.

[23] Lucas, Robert E., Jr. (1972), "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4, 103-124.

[24] Lucas, Robert E., Jr., and Nancy Stokey (1983), "Optimal fiscal and monetary policy in an economy without capital," *Journal of Monetary Economics* 12(1), 55-93.

[25] Mackowiak, Bartosz, and Mirko Wiederholt (2009), "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99, 769-803.

[26] Mankiw, N. Gregory and Ricardo Reis (2002), "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics* 117, 1295-1328.

[27] Morris, Stephen, and Hyun Song Shin (2002), "The Social Value of Public Information," *American Economic Review* 92, 1521-1534.

[28] Nimark, Kristoffer (2008), "Dynamic Pricing and Imperfect Common Knowledge," *Journal of Monetary Economics* 55, 365-382.

[29] Paciello, Luigi, and Mirko Wiederholt (2011), "Exogenous Information, Endogenous Information, and Optimal Monetary Policy," Northwestern/EIEF Mimeo.

[30] Rotemberg, Julio, and Michael Woodford (1999), "The Cyclical Behavior of Prices and Costs.," in J. B. Taylor & M. Woodford (eds.), *Handbook of Macroeconomics*, Elsevier.

[31] Sims, Christopher (2003), "Implications of Rational Inattention," *Journal of Monetary Economics* 50, 665-690.

[32] Woodford, Michael (2002), "Imperfect Common Knowledge and the Effects of Monetary Policy," in P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton: Princeton University Press.

[33] Woodford, Michael (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton: Princeton University Press.

Appendix

Proof of Proposition 1. This follows from taking the first-order conditions of the planner's problem defined in page 3.1 and noting that is problem is strictly convex, guaranteeing that these conditions are both necessary and sufficient for efficiency. **QED**

Proof of Proposition 2. To prove necessity, note first that the orthogonality conditions (5)-(7) follow directly from the analysis in the main text. All that remains is therefore to show that, as usual, resource feasibility follows from the combination of budgets and market clearing. Thus consider the nominal budget constraint of the household, which is given by

$$P\left(s^t\right)C\left(s^t\right) = \int \Pi\left(\omega_i^t, s^t\right) d\mathcal{G}_t(\omega_i^t|s^t) + w\left(s^t\right)P\left(s^t\right)N\left(s^t\right) + T\left(s^t\right).$$

Firm profits, in turn, are given by

$$\Pi\left(\omega_i^t, s^t\right) = (1-\tau_t)\, p\left(\omega_i^t\right) y\left(\omega_i^t, s^t\right) - w\left(s^t\right)P\left(s^t\right) n\left(\omega_i^t, s^t\right) - P\left(s^t\right) x\left(\omega_i^t\right)$$

Aggregating over profits, using the budget constraint of the government, and capital accumulation equation, we get that

$$P\left(s^t\right)C\left(s^t\right) = \int p\left(\omega_i^t\right) y\left(\omega_i^t, s^t\right) d\mathcal{G}_t(\omega_i^t|s^t) - P\left(s^t\right)\left(K\left(s^t\right) - (1-\delta)\,K\left(s^{t-1}\right)\right)$$

Using then the fact that $\int p\left(\omega_i^t\right) y\left(\omega_i^t, s^t\right) d\mathcal{G}_t(\omega_i^t|s^t) = P_t Y_t$, we arrive at the resource constraint, as expected.

To prove sufficiency, we need to find prices and policies that sustain the candidate allocation as an equilibrium. Thus, take any allocation that satisfies conditions (5)-(7) and let us propose the following contingent nominal prices:

$$p(\omega_i^t, s^t) = \left(\frac{y(\omega_i^t, s^t)}{Y(s^t)}\right)^{-\frac{1}{\rho}}$$

It follows that price level is constant:

$$P(s^t) = 1$$

Clearly, these nominal prices implement the right relative prices. Next, let the real wage, the tax rate and the monetary policy be given by

$$w(s^t) = \frac{V_n(s^t)}{U_c(s^t)}, \qquad \tfrac{\rho-1}{\rho}(1-\tau_t(s^t)) = \phi(s^t), \qquad \text{and} \qquad M(s^t) = Y(s^t),$$

Finally, let the nominal interest rate—also the real one since the price level is constant—be

$$R_t = \frac{U_c(s^t)}{\mathbb{E}\left[U_c(s^{t+1})|\, s^t\right]}$$

and set $B(s^t) = 0$ and $T(s^t) = \tau(s^t)\int p(\omega_i t) y(\omega_i^t, s^t) d\mathcal{F}(\omega_i^t|s^t)$.

With the prices and the policies defined as above, the following are true. First, the optimality conditions of the household and the final-good sector, the cash-in-advance constraint, and the government budget are satisfied automatically. Second, the budget constraint of the household and market clearing then follow from the resource constraints. Finally, the optimality conditions of the monopolistic firms follow from conditions (5)-(7). We conclude that the aforementioned prices and policies sustain that candidate allocation as part of a flexible-price equilibrium, which completes the proof. **QED**

Proof of Theorem 1. Follows from the main text.

Proof of Proposition 3. To prove necessity, note that feasibility follows again from the combination of budgets and market clearing. Next, from the analysis in the main text, we have that any sticky-price allocation satisfied the following conditions:

$$\begin{aligned}
\mathbb{E}\left[U_c(s^t) y\left(\omega_i^t, s^t\right)\left\{\phi(s^t)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}} - \lambda\left(\omega_i^t, s^t\right)\right\}\middle|\, \omega_i^t\right] &= 0 \qquad \forall\, \omega_i^t \\
\mathbb{E}\left[U_c(s^t) - \beta U_c(s^t)\left\{1-\delta+\lambda\left(\omega_i^{t+1}, s^{t+1}\right) A(\omega_i^{t+1}) f_k\left(\omega_i^{t+1}, s^{t+1}\right)\right\}\middle|\, \omega_i^t\right] &= 0 \qquad \forall\, \omega_i^t \\
\mathbb{E}\left[H_\ell(\omega_i^t) - U_c(s^t)\lambda\left(\omega_i^t, s^t\right) A(\omega_i^t) f_\ell\left(\omega_i^t, s^t\right)\middle|\, \omega_i^t\right] &= 0 \qquad \forall\, \omega_i^t \\
V_n(s^t) - U_c(s^t)\lambda\left(\omega_i^t, s^t\right) A(\omega_i^t) f_n\left(\omega_i^t, s^t\right) &= 0 \qquad \forall\, \omega_i^t, s^t
\end{aligned}$$

where, for any $z \in \{\ell, n, k\}$, $f_z\left(\omega_i^t, s^t\right) \equiv \frac{\partial}{\partial z} F\left(k\left(\omega_i^{t-1}\right), \ell\left(\omega_i^t\right), n\left(\omega_i^t, s^t\right)\right)$. Conditions (10)-(13) in part (i) follow directly from the above once we let

$$\chi(\omega_i^t, s^t) = \frac{\lambda(\omega_i^t, s^t)}{\phi(s^t)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}}}$$

Finally, part (ii), namely the fact that any equilibrium allocation is log-separable, follows directly from condition (9) if we let $\Psi^\omega\left(\omega_i^t\right) \equiv p\left(\omega_i^t\right)^{-\rho}$ and $\Psi^s\left(s^t\right) \equiv Y(s^t) P(s^t)^\rho$.

Consider now sufficiency. Take any allocation that satisfies properties (i) and (ii) in the proposition; we need to find prices and policies that sustain the candidate allocation as an equilibrium. Because the allocation is separable, we have that

$$y\left(\omega_i^t, s^t\right) = \Psi^\omega(\omega_i^t)\Psi^s(s^t)$$

for some functions Ψ^ω and Ψ^s. Let us then propose the following nominal prices:

$$p(\omega_i^t) = \Psi^\omega \left(\omega_i^t\right)^{-\frac{1}{\rho}},$$

which are clearly measurable in ω_i^t. It follows that the price level satisfies

$$P(s^t) = \left[\int p\left(\omega_i^t\right)^{1-\rho} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}} = \left[\int \Psi^\omega \left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}},$$

while aggregate output satisfies

$$Y\left(s^t\right) = \Psi^s\left(s^t\right)\left[\int \Psi^\omega \left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{\rho}{\rho-1}},$$

and therefore relative prices satisfy

$$\frac{p(\omega_i^t)}{P(s^t)} = \frac{\Psi^\omega \left(\omega_i^t\right)^{-\frac{1}{\rho}}}{\left[\int \Psi^\omega \left(\omega_i^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}}} = \left(\frac{y(\omega_i^t, s^t)}{Y(s^t)}\right)^{-\frac{1}{\rho}}$$

That is, we can find nominal prices that implement the right relative prices while being measurable in ω_i^t. Next, let the real wage, the tax rate and the monetary policy be given by

$$w(s^t) = \frac{V_n(s^t)}{U_c(s^t)}, \qquad \tfrac{\rho-1}{\rho}(1-\tau_t(s^t)) = \phi(s^t), \qquad \text{and} \qquad M(s^t) = P(s^t)Y(s^t).$$

Finally, let the nominal interest rate be

$$R_t = \frac{U_c(s^t)/P(s^t)}{\mathbb{E}\left[U_c(s^{t+1})/P(s^{t+1})|\, s^t\right]}$$

and set $B(s^t) = 0$ and $T(s^t) = \tau(s^t)\int p(\omega_i^t)y(\omega_i^t, s^t)d\mathcal{F}(\omega_i^t|s^t)$. With the prices and policies defined as above, the optimality conditions of the household and the final-good sector, the cash-in-advance constraint, and the government budget are satisfied automatically; the budget constraint of the household and market clearing then follow from the resource constraints; and the optimality conditions of the monopolistic firms follow from conditions (10)-(13). We conclude that the aforementioned prices and policies sustain the candidate allocation as part of a sticky-price equilibrium, which completes the proof. **QED**

Proof of Proposition 4. Follows from the main text.

Proof of Theorem 2. Follows from the main text.

Proof of Proposition 5. Take any flexible-price equilibrium. From Proposition 2 we know that $n(\omega_i^t, s^t)$ is pinned down by condition (5). Since the technology satisfies (18), it can be expressed as $F(k, \ell, n) = g(k, \ell)n^\alpha$ for some function g. Let us define

$$q(\omega_i^t) \equiv A_{it}g\left(\ell(\omega_i^t), k(\omega_i^t)\right) \tag{23}$$

This implies that

$$y(\omega_i^t, s^t) = q(\omega_i^t) n\left(\omega_i^t, s^t\right)^{\alpha}, \tag{24}$$

where $q\left(\omega_i^t\right)$ is determined on the basis of ω_i^t alone, whereas $n\left(\omega_i^t, s^t\right)$ adjusts to the realized state. Thus, $q(\omega_i^t)$ pins down the marginal productivity of employment or, equivalently, the reciprocal of the marginal cost of serving demand. Next, condition (5) reduces to the following:

$$\frac{V_n(s^t)}{U_c(s^t)} = \phi(s^t)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}} \alpha \frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t, s^t\right)} \tag{25}$$

Combining (24) and (25), and solving for output and employment, we obtain

$$y\left(\omega_i^t, s^t\right) = \Psi^{\omega}(\omega_i^t)\Psi^s(s^t) \tag{26}$$

$$n\left(\omega_i^t, s^t\right) = \Psi^{\omega}(\omega_i^t)^{\frac{\rho-1}{\rho}}\Psi^s(s^t)^{\frac{1}{\alpha}} \tag{27}$$

with

$$\Psi^{\omega}(\omega_i^t) \equiv q(\omega_i^t)^{\frac{1}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}} \quad \text{and} \quad \Psi^s(s^t) \equiv \left[\frac{\phi(s^t)U'\left(C(s^t)\right)Y(s^t)^{\frac{1}{\rho}}}{V'(N\left(s^t\right))}\right]^{\frac{\alpha}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}} \tag{28}$$

This confirms that, with the assumed specification for the production function F, *every* flexible-price equilibrium allocation is log-separable. **QED**

Proof of Proposition 6. From the proof of Proposition 5, we have that the efficient allocation satisfies

$$y\left(\omega_i^t, s^t\right) = \Psi^{\omega}(\omega_i^t)\Psi^s(s^t) \tag{29}$$

$$n\left(\omega_i^t, s^t\right) = \Psi^{\omega}(\omega_i^t)^{\frac{\rho-1}{\rho}}\Psi^s(s^t)^{\frac{1}{\alpha}} \tag{30}$$

with $\Psi^{\omega}(\omega_i^t)$ and $\Psi^s(s^t)$ defined as in (28), except that now $\phi(s^t) = 1$ because we are at the efficient allocation. Let us define

$$Q\left(s^t\right) \equiv \left[\int q(\omega_i^t)^{\frac{\frac{\rho-1}{\rho}}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1-\alpha\left(\frac{\rho-1}{\rho}\right)}{\frac{\rho-1}{\rho}}} \tag{31}$$

with $q(\omega_i^t) \equiv A_{it} g\left(\ell(\omega_i^t), k(\omega_i^t)\right)$. We may then write aggregate employment and output as follows:

$$Y\left(s^t\right) = \left[\int y\left(\omega_i^t, s^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{\rho}{\rho-1}} = \Psi^s(s^t)Q\left(s^t\right)^{\frac{1}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}} \tag{32}$$

$$N\left(s^t\right) = \int n\left(\omega_i^t, s^t\right) d\mathcal{G}_t(\omega_i^t|s^t) = \Psi^s(s^t)^{\frac{1}{\alpha}} Q\left(s^t\right)^{\frac{\frac{\rho-1}{\rho}}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}} \tag{33}$$

Next, the optimality condition for labor is given by

$$V'(N\left(s^t\right)) = U'\left(C\left(s^t\right)\right)\left(\frac{y\left(\omega_i^t, s^t\right)}{Y(s^t)}\right)^{-\frac{1}{\rho}} \alpha \frac{y\left(\omega_i^t, s^t\right)}{n\left(\omega_i^t, s^t\right)} \tag{34}$$

Aggregating the above across firms, using (32) and (28), and solving for Y, we obtain the following representation of the efficient level of aggregate output:

$$Y\left(s^t\right) = \left(\frac{V'\left(N\left(s^t\right)\right)}{U'\left(C\left(s^t\right)\right)}\right)^{-\frac{\alpha}{1-\alpha}} Q\left(s^t\right)^{\frac{1}{1-\alpha}} \tag{35}$$

We now consider prices. Following the proof of Proposition 3, for any arbitrary common-knowledge process J_t', nominal prices are given by

$$p(\omega_i^t) = e^{J_t'}\Psi^\omega\left(\omega_i^t\right)^{-\frac{1}{\rho}}$$

implement the efficient allocation and are measurable in ω_i^t. It follows that the aggregate price level is given by

$$P(s^t) = \left[\int p\left(\omega_i^t\right)^{1-\rho} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}} = e^{J_t'}\left[\int q(\omega_i^t)^{\frac{\frac{\rho-1}{\rho}}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}} d\mathcal{G}_t(\omega_i^t|s^t)\right]^{\frac{1}{1-\rho}}$$

We may thus express the aggregate price level in terms of $Q\left(s^t\right)$ as follows

$$P(s^t) = e^{J_t'}Q\left(s^t\right)^{-\frac{1}{\rho-\alpha(\rho-1)}} \tag{36}$$

Next, combining (35) with (36) to eliminate Q, we obtain the following necessary condition for the efficient price level:

$$P(s^t) = e^{J_t'}\left[Y\left(s^t\right)^{1-\alpha}\left(\frac{V'\left(N\left(s^t\right)\right)}{U'\left(C\left(s^t\right)\right)}\right)^{\alpha}\right]^{-\frac{1}{\rho-\alpha(\rho-1)}}$$

Furthermore, we may rewrite the efficiency condition for employment (34) as follows

$$\frac{V'(N\left(s^t\right))}{U'\left(C\left(s^t\right)\right)}n\left(\omega_i^t, s^t\right) = Y(s^t)^{\frac{1}{\rho}}\alpha y\left(\omega_i^t, s^t\right)^{\frac{\rho-1}{\rho}}$$

Aggregating this condition across all firms,

$$\frac{V'(N\left(s^t\right))}{U'\left(C\left(s^t\right)\right)}\int n\left(\omega_i^t, s^t\right) d\mathcal{G}_t(\omega_i^t|s^t) = Y(s^t)^{\frac{1}{\rho}}\alpha\int y\left(\omega_i^t, s^t\right)^{\frac{\rho-1}{\rho}} d\mathcal{G}_t(\omega_i^t|s^t)$$

we obtain the following expression:

$$\frac{V'(N\left(s^t\right))}{U'\left(C\left(s^t\right)\right)} = \alpha\frac{Y\left(s^t\right)}{N\left(s^t\right)} \tag{37}$$

The economy therefore admits an aggregate intratemporal efficiency condition which relates the aggregate marginal rate of substitution to the aggregate marginal product of labor. Combining this with (36) and taking logs, we find that the aggregate price level may be expressed as a function of aggregate output and aggregate labor productivity as follows

$$\log P(s^t) = J_t - \frac{1-\alpha}{\rho(1-\alpha)+\alpha}\log Y\left(s^t\right) - \frac{\alpha}{\rho(1-\alpha)+\alpha}\log\frac{Y\left(s^t\right)}{N\left(s^t\right)}$$

where $J_t = J_t' - \frac{\alpha}{\rho(1-\alpha)+\alpha}\log\alpha$. **QED**

Proof of Theorem 3. From condition (19), it is evident that either $\alpha = 0$ or procylicality of Y_t/N_t suffice for P_t to be countercyclical. To prove the theorem, we thus only need to show that condition (iii) implies condition (ii).

Thus suppose capital is fixed ($\delta = 0$ and $k_{it} = 1$ for all i, t) and there are no taste shocks (U and V do not depend on ξ_t). Combining conditions (32) and (33), we get that

$$Y\left(s^t\right) = Q\left(s^t\right) N\left(s^t\right)^{\alpha} \tag{38}$$

Furthermore, the optimality condition for labor is given by

$$\frac{V'(N\left(s^t\right))}{U'\left(Y\left(s^t\right)\right)} = \alpha \frac{Y\left(s^t\right)}{N\left(s^t\right)} \tag{39}$$

Finally, since capital is fixed and there is no depreciation, the resource constraint reduces to

$$C\left(s^t\right) = Y\left(s^t\right).$$

These observations imply that the economy is isomorphic to a representative-agent economy in which the production function is given by (38) and productivity is given by Q_t.

Using this last observation, it is now straightforward to show that both output and labor productivity are strictly increasing functions of Q_t, which in turn implies that output and labor productivity are themselves increasing transformation of one another along the efficient allocation.

Consider first the claim that output is increasing in Q_t. We prove this by contradiction. Suppose aggregate output Y is weakly decreasing in Q. From the aggregate production function (38), this implies that N must be strictly decreasing in Q. However, if N is strictly decreasing in Q and Y is weakly decreasing in Q, then by convexity of V and concavity of U the marginal rate of substitution $V'(N\left(s^t\right))/U'\left(Y\left(s^t\right)\right)$ must be strictly decreasing in Q. Substituting for N from the production function into the marginal product of labor, condition (39) implies

$$\frac{V'(N\left(s^t\right))}{U'\left(Y\left(s^t\right)\right)} = \alpha Q\left(s^t\right)^{\frac{1}{\alpha}} Y\left(s^t\right)^{1-\frac{1}{\alpha}}$$

But if $V'(N\left(s^t\right))/U'\left(Y\left(s^t\right)\right)$ is strictly decreasing in Q, the above condition implies that $Y\left(s^t\right)$ must be strictly increasing in Q. But this contradicts our original statement. Therefore, Y must be strictly increasing in Q.

Consider now the claim that labor productivity is increasing in Q. We again prove this by contradiction. Suppose the labor productivity $Y\left(s^t\right)/N\left(s^t\right)$ is weakly decreasing in Q. Substituting for Y from the production function, we have that

$$\frac{Y\left(s^t\right)}{N\left(s^t\right)} = Q\left(s^t\right) N\left(s^t\right)^{\alpha-1}$$

This, combined with the fact that $Y\left(s^t\right)/N\left(s^t\right)$ is weakly decreasing in Q, implies that N must be strictly increasing in Q. We know from the proof of part (i) that Y is strictly increasing in Q. However, if N and Y are both strictly increasing in Q, then by convexity of V and concavity of U

the marginal rate of substitution $V'(N(s^t))/U'(Y(s^t))$ must be strictly increasing in Q. But the intratemporal condition (39) then implies that labor productivity $Y(s^t)/N(s^t)$ must be strictly increasing in Q. This contradicts our original statement. Therefore, the term $Y(s^t)/N(s^t)$ is strictly increasing in $Q(s^t)$.

Since both Y and Y/N are increasing functions of Q, we infer labor productivity is necessarily procyclical, which completes the proof of the theorem. Furthermore, note that in this case condition (19) can be expressed as $\log P_t = J_t - \Gamma(\log Y_t)$ for some increasing function Γ, proving that there is actually an one-to-one negative relation between the price level and aggregate output. **QED**

Analysis of example in Section 6. Consider the example studied in Section 6 and let

$$X_t \equiv [\bar{a}_t, \log Q_t]$$

where $\bar{a}$ is the aggregate productivity shock and where Q_t is defined in (31). We next show that X_t is a sufficient state variable for the aggregate dynamics of the economy along the Ramsey optimum.

Claim A. *There exists a matrix Λ, and vectors $\Delta_v, \Delta_\varepsilon, \Gamma$ such that, along any equilibrium that implements the efficient allocation, the following are true:*

$$\begin{aligned} X_t &= \Lambda X_{t-1} + \Delta_v v_t + \Delta_\varepsilon \varepsilon_t \\ [\log Y_t, \log P_t] &= [0, J_t] + \Gamma \log Q_t \end{aligned}$$

where J_t is a deterministic sequence.

As in the proof of Theorem 3, the specification of the production function (18) and the absence of capital imply that we can express aggregate output $Y(s^t)$ and aggregate employment $N(s^t)$ as in (32) and (33), where $\Psi^s(s^t)$ is as defined in (28). Using the assumed specification for U and V in (22), we guess and verify that the equilibrium functions $\Psi^s(s^t), N(s^t), Y(s^t)$ can be expressed in terms of $Q(s^t)$ alone, as follows

$$\begin{aligned} \Psi^s(s^t) &= Q(s^t)^{\chi_\Psi} && (40) \\ N(s^t) &= Q(s^t)^{\chi_N} && (41) \\ Y(s^t) &= Q(s^t)^{\chi_Y} && (42) \end{aligned}$$

for some coefficients $(\chi_\Psi, \chi_N, \chi_Y)$. Conditions (28), (32), and (33) then imply that these coefficients $(\chi_\Psi, \chi_N, \chi_Y)$ must satisfy the following linear system of equations

$$\begin{aligned} \chi_\Psi &= \frac{\alpha}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}\left(\chi_Y\left(\frac{1}{\rho}-\gamma\right)-\epsilon\chi_N\right) \\ \chi_N &= \frac{\chi_\Psi}{\alpha} + \frac{\frac{\rho-1}{\rho}}{1-\alpha\left(\frac{\rho-1}{\rho}\right)} \\ \chi_Y &= \chi_\Psi + \frac{1}{1-\alpha\left(\frac{\rho-1}{\rho}\right)} \end{aligned}$$

This system has a unique solution given by

$$\chi_\Psi = \frac{(\epsilon+1)}{(\epsilon+1)-\alpha(1-\gamma)} - \frac{1}{1-\alpha\left(\frac{\rho-1}{\rho}\right)} \tag{43}$$

$$\chi_N = \frac{(1-\gamma)}{(\epsilon+1)-\alpha(1-\gamma)} \tag{44}$$

$$\chi_Y = \frac{(\epsilon+1)}{(\epsilon+1)-\alpha(1-\gamma)} \tag{45}$$

Furthermore, following the proof of Theorem 3, we may also express the aggregate price level in terms of $Q\left(s^t\right)$ as follows

$$P(s^t) = e^{J_t} Q\left(s^t\right)^{\chi_P} \tag{46}$$

where

$$\chi_P \equiv -\frac{1}{\rho-\alpha(\rho-1)} \tag{47}$$

Therefore, given functions $q\left(\omega_i^t\right)$ and $Q\left(s^t\right)$, the equilibrium behavior of $Y\left(s^t\right)$ and $P\left(s^t\right)$ along the Ramsey optimum are pinned down by equations (42), and (46). What then remains to be characterized is the behavior of $q\left(\omega_i^t\right)$ and $Q\left(s^t\right)$.

We now show that there exists a fixed point in $q\left(\omega_i^t\right)$ and $Q\left(s^t\right)$ which pins down their joint solution. First, using the specification for H in (22) and the Cobb-Douglas technology in (20), we rewrite the optimality condition for managerial effort (2) as follows:

$$\left(\frac{q(\omega_i^t)}{A\left(\omega_i^t\right)}\right)^{\frac{\epsilon+1}{1-\alpha}} = (1-\alpha)\,\mathbb{E}\left[Y\left(s^t\right)^{\frac{1}{\rho}-\gamma} y\left(\omega_i^t, s^t\right)^{\frac{\rho-1}{\rho}} \middle| \omega_i^t \right]$$

Substituting (20) for individual output into this condition, and using our expressions for $\Psi^s\left(s^t\right)$ and $Y\left(s^t\right)$ from (40) and (42), we find that the equilibrium value of $q(\omega_i^t)$ is given by the fixed point to the following functional equation:

$$q(\omega_i^t)^{\frac{\epsilon+1}{1-\alpha} - \frac{\frac{\rho-1}{\rho}}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}} = (1-\alpha)\, A_{it}^{\frac{\epsilon+1}{1-\alpha}}\,\mathbb{E}\left[Q\left(s^t\right)^{\chi_Y\left(\frac{1}{\rho}-\gamma\right)+\chi_\Psi\left(\frac{\rho-1}{\rho}\right)} \middle| \omega_i^t \right] \tag{48}$$

This fixed-point representation pins down the equilibrium $q(\omega_i^t)$ for any information structure. However, given the Gaussian information structure we have imposed, we propose that the equilibrium $q(\omega_i^t)$ and $Q(s^t)$ are jointly log-normal. This implies that the aforementioned functional equation can be restated in the following log-linear form:

$$\log q(\omega_i^t) = (1-\chi)\,\theta \log A_{it} + \chi \mathbb{E}\left[\log Q\left(s^t\right) \middle| \omega_i^t\right] \tag{49}$$

with

$$\chi \equiv \left[\frac{\epsilon+1}{\frac{\epsilon+1}{1-\gamma}-\alpha} - \frac{\frac{\rho-1}{\rho}}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}\right]\left[\frac{\epsilon+1}{1-\alpha} - \frac{\frac{\rho-1}{\rho}}{1-\alpha\left(\frac{\rho-1}{\rho}\right)}\right]^{-1} \qquad \text{and} \qquad \theta \equiv \frac{(\epsilon+1)-\alpha(1-\gamma)}{(\epsilon+1)-(1-\gamma)}.$$

We now proceed to solve for the fixed point to the above condition. The information that a firm obtains in period t is summarized by the vector

$$\tilde{\omega}_{it} = \begin{bmatrix} a_{it} \\ z_{it} \end{bmatrix} = \begin{bmatrix} \bar{a}_t + \zeta_{it} \\ \bar{a}_t + \varepsilon_t + \eta_{it} \end{bmatrix}$$

The information set of manager i at time t is thus given by $\omega_i^t \equiv \{\tilde{\omega}_{i,\tau}\}_{\tau=-\infty}^t$, the sequence of past and current signals. The method we thus follow to solve for the Ramsey optimum is based on the Kalman filter, and is similar to that used in Woodford (2002).

State Vector and Law of Motion. We guess and verify that the state vector X_t defined by

$$X_t \equiv \begin{bmatrix} \bar{a}_t \\ \log Q_t \end{bmatrix}$$

is a sufficient statistic for the entire aggregate dynamics of the economy. Managers use a Kalman filter to update their information in each period about X_t, and the equilibrium is determined by the fixed point condition (49).

Claim B. *The dynamics of X_t are given by the following law of motion*

$$X_t = \Lambda X_{t-1} + \Delta_v v_t + \Delta_\varepsilon \varepsilon_t \tag{50}$$

with

$$\Lambda \equiv \begin{bmatrix} \varrho & 0 \\ \Lambda_{21} & \Lambda_{22} \end{bmatrix}, \Delta_v \equiv \begin{bmatrix} 1 \\ \Delta_{v2} \end{bmatrix}, \Delta_\varepsilon \equiv \begin{bmatrix} 0 \\ \Delta_{\varepsilon 2} \end{bmatrix} \tag{51}$$

The coefficients $(\Lambda_{21}, \Lambda_{22}, \Delta_{21})$ *are given by*

$$\Lambda_{21} = \varrho \left(K_{21} + K_{22}\right) \tag{52}$$

$$\Lambda_{22} = \varrho \left(1 - \frac{1}{\theta}\left(K_{21} + K_{22}\right)\right) \tag{53}$$

$$\Delta_{v2} = \left(1 - \chi\right)\theta + \chi\left(K_{21} + K_{22}\right) \tag{54}$$

$$\Delta_{\varepsilon 2} = \chi K_{22} \tag{55}$$

and

$$K \equiv \begin{bmatrix} K_{11} & K_{21} \\ K_{21} & K_{22} \end{bmatrix}$$

is the matrix of kalman gains, defined by

$$K \equiv \mathbb{E}\left[\left(X_t - \mathbb{E}_{i,t-1}\left[X_t\right]\right)\left(\tilde{\omega}_{it} - \mathbb{E}_{i,t-1}\left[\tilde{\omega}_{it}\right]\right)'\right]\mathbb{E}\left[\left(\tilde{\omega}_{it} - \mathbb{E}_{i,t-1}\left[\tilde{\omega}_{it}\right]\right)\left(\tilde{\omega}_{i,t} - \mathbb{E}_{i,t-1}\left[\tilde{\omega}_{it}\right]\right)'\right]^{-1} \tag{56}$$

We verify this claim in the following analysis and describe the procedure for finding the fixed point.

Observation Equation. In each period t, manager i observes private signal $\tilde{\omega}_{it}$,. In terms of the aggregate state and error terms, manager i's observation equation takes the form

$$\tilde{\omega}_{it} \equiv \begin{bmatrix} e_1' \\ e_1' \end{bmatrix} X_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \varepsilon_t + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \zeta_{it} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \eta_{it} \tag{57}$$

where e_j is defined as a column vector of length two where the j-th entry is 1 and all other entries are 0.

Forecasting and Inference. Manager i's $t-1$ forecast of $\tilde{\omega}_{it}$ is given by

$$\mathbb{E}_{i,t-1}\left[\tilde{\omega}_{it}\right] = \begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \mathbb{E}_{i,t-1}\left[X_t\right]$$

where $\mathbb{E}_{i,t-1}\left[X_t\right]$ is manager's i's $t-1$ forecast of X_t. Combining this with the law of motion (50), it follows that $\mathbb{E}_{i,t-1}\left[X_t\right] = \Lambda \mathbb{E}_{i,t-1}\left[X_{t-1}\right]$.

To form minimum mean-squared-error estimates of the current state, managers use the kalman filter to update their forecasts. Updating is done via

$$\mathbb{E}_{i,t}\left[X_t\right] = \mathbb{E}_{i,t-1}\left[X_t\right] + K\left(\tilde{\omega}_{it} - \mathbb{E}_{i,t-1}\left[\tilde{\omega}_{it}\right]\right), \tag{58}$$

where K is the 2×2 matrix of Kalman gains, defined in (56). Substitution of manager i's $t-1$ forecast of $\tilde{\omega}_{it}$ into (58) gives us

$$\mathbb{E}_{i,t}\left[X_t\right] = \left(I - K \begin{bmatrix} e_1' \\ e_1' \end{bmatrix}\right) \Lambda \mathbb{E}_{i,t-1}\left[X_{t-1}\right] + K\tilde{\omega}_{i,t} \tag{59}$$

Let $\bar{\mathbb{E}}_t\left[X_t\right] \equiv \int_I \mathbb{E}_{i,t}\left[X_t\right] di$ be the time t average expectation of the current state. Aggregation over (59) implies

$$\bar{\mathbb{E}}_t\left[X_t\right] = \left(I - K \begin{bmatrix} e_1' \\ e_1' \end{bmatrix}\right) \Lambda \bar{\mathbb{E}}_{t-1}\left[X_{t-1}\right] + K \int \tilde{\omega}_{it} di$$

Finally, using the fact that aggregration over signals yields $\int \tilde{\omega}_{it} di = \begin{bmatrix} e_1' \\ e_1' \end{bmatrix} X_t + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \varepsilon_t$, it follows that the average expectation evolves according to

$$\begin{aligned} \bar{\mathbb{E}}_t\left[X_t\right] \quad = \quad & K \begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \Lambda X_{t-1} + \left(I - K \begin{bmatrix} e_1' \\ e_1' \end{bmatrix}\right) \Lambda \bar{\mathbb{E}}_{t-1}\left[X_{t-1}\right] \\ & + K \begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \Delta_v v_t + K\left(\begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \Delta_\varepsilon + \begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \varepsilon_t \end{aligned} \tag{60}$$

where Λ,Δ_v, Δ_ε are given by (51).

Characterizing Aggregate Output. Local output in each island is determined by the best-response-like condition in (49), which may be rewritten as $\log q_{i,t} = (1-\chi)\,\theta a_{i,t} + \chi e_2' \mathbb{E}_{i,t}\left[X_t\right]$. Aggregating over this condition, we find that aggregate output must satisfy

$$\log Q_t = (1-\chi)\,\theta \bar{a}_t + \chi e_2' \bar{\mathbb{E}}_t\left[X_t\right] \tag{61}$$

Substituting our expression for $\bar{\mathbb{E}}_t\left[X_t\right]$ from (60) into (61), gives us

$$\begin{aligned} \log Q_t \quad = \quad & \left[(1-\chi)\,\theta\varrho + \chi\varrho\left(K_{21} + K_{22}\right)\right] \bar{a}_{t-1} + \left[\chi\Lambda_{21} - \chi\varrho\left(K_{21} + K_{22}\right)\right] \bar{\mathbb{E}}_{t-1}\left[\bar{a}_{t-1}\right] \\ & + \chi\Lambda_{22} \bar{\mathbb{E}}_{t-1}\left[\log Q_{t-1}\right] + \left[(1-\chi)\,\theta + \chi\left(K_{21} + K_{22}\right)\right] v_t + \chi K_{22} \varepsilon_t \end{aligned}$$

Moreover, rearranging condition (61), we find that $\bar{\mathbb{E}}_t [\log Q_t] = \frac{1}{\chi} (\log Q_t - (1-\chi) \theta \bar{a}_t)$. Finally, using this condition in the above equation gives us

$$\begin{aligned} \log Q_t &= [(1-\chi)\theta\varrho + \chi\varrho(K_{21}+K_{22}) - \Lambda_{22}(1-\chi)\theta]\bar{a}_{t-1} + \Lambda_{22}\log Q_{t-1} \\ &\quad + [\chi\Lambda_{21} - \chi\varrho(K_{21}+K_{22})]\bar{\mathbb{E}}_{t-1}[\bar{a}_{t-1}] + [(1-\chi)\theta + \chi(K_{21}+K_{22})]v_t + \chi K_{22}\varepsilon_t \end{aligned}$$

For this to coincide with the law of motion conjectured in (50) and (51) for every $(\bar{f}_{t-1}, \log Q_{t-1}, v_t, \varepsilon_t)$, it is necessary and sufficient that the coefficients $(\Lambda_{21}, \Lambda_{22}, \Delta_{v2}, \Delta_{\varepsilon 2})$ solve the following system:

$$\begin{aligned} \Lambda_{21} &= (1-\chi)\theta\varrho + \chi\varrho(K_{21}+K_{22}) - \Lambda_{22}(1-\chi)\theta \\ \Delta_{v2} &= (1-\chi)\theta + \chi(K_{21}+K_{22}) \\ \Delta_{\varepsilon 2} &= \chi K_{22} \\ 0 &= \chi\Lambda_{21} - \chi\varrho(K_{21}+K_{22}) \end{aligned}$$

The unique solution to this system for $(\Lambda_{21}, \Lambda_{22}, \Delta_{v2}, \Delta_{\varepsilon 2})$ is the one given in the proposition. Therefore, given the kalman gains matrix K, we can uniquely identify the coefficients of the law of motion of X_t.

Kalman Filtering. Let us define the variance-covariance matrices of forecast errors as

$$\begin{aligned} \Sigma &\equiv \mathbb{E}\left[(X_t - \mathbb{E}_{i,t-1}[X_t])(X_t - \mathbb{E}_{i,t-1}[X_t])'\right] \\ V &\equiv \mathbb{E}\left[(X_t - \mathbb{E}_{i,t}[X_t])(X_t - \mathbb{E}_{i,t}[X_t])'\right] \end{aligned}$$

These matrices will be the same for all islands i, since their observation errors are assumed to have the same stochastic properties. Using these matrices, we may write K as the product of two components:

$$\mathbb{E}_i\left[(X_t - \mathbb{E}_{i,t-1}[X_t])(\tilde{\omega}_{it} - \mathbb{E}_{i,t-1}[\tilde{\omega}_{it}])'\right] = \Sigma \begin{bmatrix} e_1 & e_1 \end{bmatrix} + \sigma_\varepsilon^2 \Delta_\varepsilon \begin{bmatrix} 0 & 1 \end{bmatrix}$$

and

$$\begin{aligned} \mathbb{E}_i\left[(\tilde{\omega}_{it} - \mathbb{E}_{i,t-1}[\tilde{\omega}_{it}])(\tilde{\omega}_{it} - \mathbb{E}_{i,t-1}[\tilde{\omega}_{it}])'\right] &= \begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \Sigma \begin{bmatrix} e_1 & e_1 \end{bmatrix} + \sigma_\zeta^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \sigma_\eta^2 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ &\quad + \sigma_\varepsilon^2 \left(\begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \Delta_\varepsilon \begin{bmatrix} 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Delta_\varepsilon' \begin{bmatrix} e_1 & e_1 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) \end{aligned} \tag{62}$$

Therefore, K is given by

$$K = \left(\Sigma \begin{bmatrix} e_1 & e_1 \end{bmatrix} + \sigma_\varepsilon^2 \Delta_\varepsilon \begin{bmatrix} 0 & 1 \end{bmatrix} \right) \left(\sigma_z^2\right)^{-1} \tag{63}$$

where $\sigma_z^2 \equiv \mathbb{E}_i\left[(\tilde{\omega}_{it} - \mathbb{E}_{i,t-1}[\tilde{\omega}_{it}])(\tilde{\omega}_{i,t} - \mathbb{E}_{i,t-1}[\tilde{\omega}_{it}])'\right]$ is given by (62).

Finally, what remains to determine is the matrix Σ. The law of motion implies that matrices Σ and V satisfy

$$\Sigma = \Lambda V \Lambda' + \sigma_v^2 \Delta_v \Delta_v' + \sigma_\varepsilon^2 \Delta_\varepsilon \Delta_\varepsilon'$$

In addition, the forecasting equation (59) imply these matrices must further satisfy

$$V = \Sigma - \left(\Sigma \begin{bmatrix} e_1 & e_1 \end{bmatrix} + \sigma_\varepsilon^2 \Delta_\varepsilon \begin{bmatrix} 0 & 1 \end{bmatrix}\right) \left(\sigma_z^2\right)^{-1} \left(\begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \Sigma + \sigma_\varepsilon^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Delta_\varepsilon'\right)$$

Combining the above two equations, we obtain the stationary *Ricatti Equation* for Σ:

$$\begin{aligned} \Sigma &= \Lambda \Sigma \Lambda' - \Lambda \left(\Sigma \begin{bmatrix} e_1 & e_1 \end{bmatrix} + \sigma_\varepsilon^2 \Delta_\varepsilon \begin{bmatrix} 0 & 1 \end{bmatrix}\right) \left(\sigma_z^2\right)^{-1} \left(\begin{bmatrix} e_1' \\ e_1' \end{bmatrix} \Sigma + \sigma_\varepsilon^2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Delta_\varepsilon'\right) \Lambda' \\ &\quad + \sigma_v^2 \Delta_v \Delta_v' + \sigma_\varepsilon^2 \Delta_\varepsilon \Delta_\varepsilon' \end{aligned} \tag{64}$$

where Λ, Δ_v, Δ_ε are functions of the kalman gains matrix K, and K is itself a function of Σ and Δ_ε. The variance-covariance matrix Σ, the kalman gains matrix K, and the law of motion matrices Λ, Δ_v, Δ_ε are thus obtained by solving the large non-linear system of equations described by (52)-(55), (63), and (64). This system is too complicated to allow further analytical results; we thus solve for the fixed point numerically.

Characterizing Aggregate Output and Prices. The numerical solution described above pins down the values for Λ, Δ_v, Δ_ε for the law of motion (50) of X_t. Aggregate output and the price level are then determined by equations (42), and (46):

$$\begin{aligned} \log Y_t &= \chi_Y \log Q_t \\ \log P_t &= \chi_P \log Q_t + J_t \end{aligned}$$

where the coefficients (χ_Y, χ_P) are as in (45), and (47). This verifies Claim B, with

$$\Gamma = \begin{bmatrix} \chi_Y \\ \chi_P \end{bmatrix}.$$

The results reported in Section 6 are then based on a numerical implementation of the above characterization, and with J_t normalized to zero.