# Essays on Labor Market Dynamics

by

Christina Hyde Patterson

B.A., Columbia University (2011)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Economics

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2019



Signature redacted

Author .....
Department of Economics
April 25, 2019

Signature redacted

Certified by .....
Daron Acemoglu
Elizabeth and James Killian Professor of Economics
Thesis Supervisor

Signature redacted

Certified by .....
David Autor
Ford Professor of Economics
Thesis Supervisor

Signature redacted

Certified by .....
Heidi Williams
Associate Professor
Thesis Supervisor

Signature redacted

Accepted by .....
Ricardo Caballero
Ford International Professor of Economics
Chairman, Department Committee on Graduate Theses

**Essays on Labor Market Dynamics**

by

Christina Hyde Patterson

Submitted to the Department of Economics
on April 25, 2019, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Economics

## Abstract

This thesis consists of three chapters on labor market dynamics.

In the first chapter, I show empirically that the unequal incidence of recessions is a core channel through which aggregate shocks are amplified. I show that the aggregate marginal propensity to consume (MPC) is larger when income shocks disproportionately hit high-MPC individuals, and I define the *Matching Multiplier* as the increase in the output multiplier originating from the matching of workers to jobs with different income elasticities – a greater matching multiplier translates into more powerful amplification in a range of business cycle models. Using administrative data from the United States, I document that the earnings of individuals with a higher marginal propensity to consume are more exposed to recessions. I show that this covariance between worker MPCs and the elasticity of their earnings to GDP is large enough to increase shock amplification by 40 percent over a benchmark in which all workers are equally exposed. Using local labor market variation, I validate this amplification mechanism by showing that areas with higher matching multipliers experience larger employment fluctuations over the business cycle. Lastly, I derive a generalization of the matching multiplier in an incomplete markets model and show numerically that this mechanism is quantitatively similar within this structural framework.

In the second chapter, joint with David Autor, David Dorn, Lawrence Katz, and John Van Reenen, we explore the well-documented fall of labor's share of GDP in the United States and many other countries. Existing empirical assessments typically rely on industry or macro data, obscuring heterogeneity among firms. In this paper, we analyze micro panel data from the U.S. Economic Census since 1982 and document empirical patterns to assess a new interpretation of the fall in the labor share based on the rise of "superstar firms." If globalization or technological changes advantage the most productive firms in each industry, product market concentration will rise as industries become increasingly dominated by superstar firms. Since these firms have high markups and a low labor share of firm value-added and sales, this depresses the aggregate labor share. We empirically assess seven predictions of this hypothesis: (i) industry sales will increasingly concentrate

in a small number of firms; (ii) industries where concentration rises most will have the largest declines in the labor share; (iii) the fall in the labor share will be driven largely by reallocation rather than a fall in the unweighted mean labor share across all firms; (iv) the between-firm reallocation component of the fall in the labor share will be greatest in the sectors with the largest increases in market concentration; (v) the industries that are becoming more concentrated will exhibit faster growth of productivity and innovation; (vi) the aggregate markup will rise more than the unweighted firm markup; and (vii) these patterns should be observed not only in U.S. firms, but also internationally. We find support for all of these predictions.

In the third chapter, I explore how the distribution of tasks across industries affects labor market responses to shocks. I present a model in which task-level wages connect industries employing the same tasks, meaning that the distribution of tasks across industries insures some workers against shocks and alters their labor market experiences. Workers trained in more dispersed tasks (e.g. accountants) face less unemployment risk from industry-specific shocks than workers who do tasks that are concentrated in few industries (e.g. petroleum engineers). Using industry and regional data, I show empirical evidence that supports the model's predictions – industries that employ more specialized labor contract less in response to demand shocks than industries with less specialized labor.

**JEL Classifications:** E21, J23, D33

Thesis Supervisor: Daron Acemoglu
Title: Elizabeth and James Killian Professor of Economics

Thesis Supervisor: David Autor
Title: Ford Professor of Economics

Thesis Supervisor: Heidi Williams
Title: Associate Professor

## Acknowledgments

This thesis marks the end of a long road and it is not one that I traveled alone. In the six years that I have been at MIT, I was fortunate to have tremendous support both within the MIT community and beyond, for which I will be eternally grateful.

I am indebted to my primary advisors, David, Heidi, and Daron, who profoundly shaped this thesis and my understanding of economics. I learned a lot from the opportunity to RA and then co-author with David early on in my graduate school experience. He showed me how to write economics papers and he modeled thinking creatively about real-world questions not always in the realm of traditional economics. I particularly want to thank Heidi for sticking with me. As my thesis turned further from her primary fields, she could have easily advised that I find another advisor in a more closely related field. Instead, she continued to not only advise my research, but also provide invaluable professional and personal advice in critical times of uncertainty. I will be forever grateful to her for that guidance and mentorship. Many people have remarked on Daron's brilliance, but what I want to thank him for is the time he spent – I do not know of many advisors who would have churned through as many drafts as he did, each in record time, and each with a thoroughness that went from sweeping thematic comments to corrections of footnotes. I can only aspire to trying to hold myself to a similar standard of deep engagement with my own students.

In addition to my main academic advisors, I am grateful for the advice and guidance I received from John Van Reenen, Jonathan Parker, Ivan Werning, Lawrence Katz, David Dorn, and Martin Beraja. Jonathan Parker's class in second year demonstrated an expansive view of what constituted macro research and in many ways inspired the idea in the first chapter of my dissertation. I am also indebted to the support from the mentors I met before MIT and who have continued to support me over these past 6 years: Toshi Mukoyama, Marco Del Negro, Marc Giannoni, and, most of all, Aysegul Sahin. My interest in macro and labor was first formed in working with Aysegul while at the New York Fed, and she has been, and I hope will continue to be, a valuable mentor, co-author and role model as a woman in macro.

I would not have made it through graduate school without my classmates. From study groups in first year to practice job talks in sixth year, my office mates Ben Marx, Ray Kluender, Dan Waldinger, Jon Petkun, and Michael Stepner shared both the most difficult and most rewarding times in graduate school with me. I learned from them every day, and I thank them for asking and answering daily questions about both research and life. I am very grateful to have found a close group of female colleagues and friends in our Saturday spin group: Jane Choi, Maddie McKelway, Tamar Oostrom, and Carolyn Stein. Those spin classes and bagels were consistently one of the highlights of my week. Lastly, I am grateful for the other friendships I built in the department with both students in the years below me (Ryan Hill, Colin Gray, Joe Hazell, John Sturm, and Joel Flynn, among

others) and older cohorts (John Firth, Brendan Price, Alex Bartik, and Melanie Wasserman, as well as others).

My studies at MIT were made possible by the generous financial support of the MIT economics department, the MIT Initiative on the Digital Economy, the Alfred P. Sloan Foundation through the National Bureau of Economic Research, and the Washington Center for Equitable Growth Doctoral Grants and Junior Fellowship programs. My work has also been aided by the excellent research assistance from Drew Bent and Brendan Moore. This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by the following National Science Foundation Grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. All results using U.S. Census Data have been reviewed to ensure that no confidential information is disclosed. Any opinions and conclusions expressed herein mine and do not necessarily represent the views of the U.S. Census Bureau.

Life would be a lot less enjoyable without the company of good friends, and I am lucky enough to have had many who helped me through graduate school in various ways. I'm grateful to Lauren Goldstein for two decades of friendship, or, should I say, sisterhood. I want to thank my Columbia family, Helen Werbe, Josefina Aguila, and Ariana Orvell, for friendship, adventure, and support over the last 10 years. In particular, I am thankful to Ariana Orvell for almost weekly morning chats over the past six years. From venting to strategizing, these chats got me through difficult times and I feel lucky to have gotten to share our surprisingly similar graduate school experiences from afar.

Lastly, I am thankful for the endless support of my family. I grew up in Boston, and I was thrilled to be able to return to my home town for graduate school to be near my parents, Jane Manopoli and Robert Patterson. This proximity meant that my mom and I often went on morning walks around the esplanade before classes, and she listened to laughter, tears and complaints through those walks. My parents also welcomed my friends from Cambridge into their home for dinners and holidays, and in doing so, brought us all closer together. My older sister, Diana Patterson, has been a role model in many ways, and I particularly admire her strength in pursuing a grueling path to become an orthopedic surgeons. I have no doubt the grit and determination I learned from her helped me complete this thesis. I am also grateful to have married into a welcoming and supportive second family. Holly Russell and Donald Reid shared with me their much-needed perspective from a different academic life and encouraged me through the job market with kindness and excitement. Hadley Reid welcomed me into her family as a sister and I look forward to celebrating her successes over the coming years.

Most of all, I am grateful to MIT for bringing me to my husband, Otis Reid. I would not have made it through graduate school without his constant support and love. He was both my loudest cheerleader and a silent co-author, editing my papers, talking through empirical specifications, debugging code, and brainstorming big picture questions. He not only has given me pep-talks through each and every setback, but he has also shared

in and celebrated my successes. He often says to me "One team, one dream." Indeed, completing our theses at MIT and moving to Chicago for the next chapter in our lives was certainly a shared dream that we have realized together. I am constantly in awe of his intellectual curiosity and passion for rigorous thought, and I look forward to seeing what dreams we achieve together in the decades to come.

THIS PAGE INTENTIONALLY LEFT BLANK

# Contents

**1 The Matching Multiplier and the Amplification of Recessions** **21**

1.1 Introduction . . . . . 21

1.2 Defining the Matching Multiplier . . . . . 29

1.3 Data description . . . . . 32

1.3.1 Longitudinal Employer-Household Dynamics (LEHD) linked to American Community Survey . . . . . 32

1.3.2 Panel Study of Income Dynamics . . . . . 35

1.4 MPCs and the incidence of recessions . . . . . 37

1.4.1 Estimating marginal propensities to consume . . . . . 37

1.4.2 Heterogeneity in worker exposure . . . . . 43

1.4.3 Discussion . . . . . 47

1.5 Matching Multiplier estimates . . . . . 48

1.6 Commuting zone analysis . . . . . 52

1.6.1 Local recessions and the Matching Multiplier . . . . . 55

1.6.2 Tradable industries and the Matching Multiplier . . . . . 59

1.7 The Matching Multiplier in a calibrated model . . . . . 62

1.7.1 Environment . . . . . 63

1.7.2 Equilibrium and the output multiplier . . . . . 65

1.7.3 Special limiting case: The empirical Matching Multiplier . . . . . 69

1.7.4 Model calibration . . . . . 71
1.7.5 Model estimates and counterfactuals . . . . . 73
1.8 Conclusion . . . . . 77
References . . . . . 88
Figures and Tables . . . . . 89

**2 The Fall of the Labor Share and the Rise of Superstar Firms 101**
2.1 Introduction . . . . . 101
2.2 A Model of Superstar Firms . . . . . 106
2.3 Data . . . . . 111
2.3.1 Data Construction . . . . . 111
2.3.2 Initial Data Description . . . . . 114
2.4 Empirical Tests of the Predictions of Superstar Firm Model . . . . . 117
2.4.1 Rising Concentration Correlates with Falling Labor Shares . . . . . 117
2.4.2 Between-Firm Reallocation Drives Fall in Labor Share . . . . . 123
2.4.3 Between-Firm Reallocation is Strongest in Concentrating Industries . 128
2.4.4 Markup Analysis . . . . . 130
2.4.5 Concentrating Industries have Higher Innovation and Productivity Growth . . . . . 133
2.4.6 Superstar Firm Patterns are International . . . . . 136
2.4.7 Magnitudes . . . . . 141
2.5 Further Descriptive Evidence on Superstar Firms . . . . . 143
2.5.1 Import Exposure and Superstar Firms . . . . . 143
2.5.2 Compustat Analysis: Publicly Listed Superstar Firms . . . . . 144
2.5.3 Worker Power and the Rise in Concentration . . . . . 146
2.6 Conclusions . . . . . 147
References . . . . . 149
Figures and Tables . . . . . 156

**3 It's What You Know That Matters: Task-Based Networks and Industry-Specific Shocks** **177**

3.1 Introduction . . . . . 177

3.2 Theory . . . . . 181

3.2.1 Model Setup . . . . . 181

3.2.2 Simple Case: Shutting Down the Input-Output Network . . . . . 182

3.2.3 Full Model: Adding Back the Input-Output Matrix . . . . . 186

3.3 Data Description . . . . . 189

3.3.1 Task Composition of Industries . . . . . 189

3.3.2 Local Labor Market Data . . . . . 191

3.3.3 Industry Shocks . . . . . 193

3.4 Empirical Application . . . . . 194

3.5 Empirical Results . . . . . 199

3.5.1 National Changes . . . . . 199

3.5.2 Government Spending Shocks . . . . . 202

3.6 Conclusion . . . . . 203

References . . . . . 204

Figures and Tables . . . . . 206

A Appendix for Chapter 1 . . . . . 211

B Appendix for Chapter 3 . . . . . 284

THIS PAGE INTENTIONALLY LEFT BLANK

# List of Figures


1.1 Recession Exposure and MPC by Demographic Group . . . . . 89

1.2 Fraction of U.S. Employment in LEHD Sample . . . . . 91

1.3 Heterogeneity in Marginal Propensity to Consume Estimates . . . . . 92

1.4 The Distribution of Estimated Marginal Propensities to Consume . . . . . 93

1.5 Relative Consumption Response in Various Scenarios . . . . . 94

1.6 Demographic Group MPCs: Calibration vs. Experiment . . . . . 95

2.1 International Comparison: Labor Share by Country . . . . . 157

2.2 The Labor Share in Manufacturing . . . . . 158

2.3 Average Payroll-to-Sales Ratio . . . . . 159

2.4 Average Concentration Across Four Digit Industries by Major Sector . . . . 160

2.5 The Relationship Between Firm Size and Labor Share . . . . . 162

2.6 The Relationship Between the Change in Labor Share and the Change in Concentration Across Six Sectors . . . . . 164

2.7 Melitz-Polanec Decomposition of the Change in Labor Share in Manufacturing . . . . . 165

2.8 Melitz-Polanec Decomposition of the Change in Labor Share in all Six Sectors 166

2.9 Regressions of the Components of the Change in Labor Share on the Change in Concentration . . . . . 167

2.10 Markup changes . . . . . 168

3.1 Description of Task Dispersion Across Industries . . . . . 206
3.2 Cross Commuting Zone Variation in Industry Taks Concentration . . . . . 207
A1 LEHD Sample Selection . . . . . 247
A2 The Fraction of Food in Total Spending . . . . . 248
A3 Imputed Total Consumption Using CEX-Based Imputation . . . . . 249
A4 Imputed and Actual Consumption Using Attanasio and Pistaferri (2014) Imputation . . . . . 250
A5 Heterogeneity in the Marginal Propensity to Consume: First Stage and Reduced Form . . . . . 251
A6 Sensitivity of Baseline MPC Estimates to Specification of Demographic Variables . . . . . 252
A7 MPCs out of Individual and Household Incomes . . . . . 253
A8 Estimates of the Average Marginal Propensity to Consume in the Literature 254
A9 The Stability of Marginal Propensity to Consume Estimates . . . . . 255
A10 The Persistence of Unemployment Shocks on Income, by Demographics . . 256
A11 Marginal Propensities to Consume: Alternate Consumption Variables . . . . 257
A12 Marginal Propensities to Consume Estimated in the CEX . . . . . 258
A13 Heterogeneity in Tax-Rebate-Based MPCs . . . . . 259
A14 Earnings Sensitivity to GDP and MPCs: Intensive and Extensive Earnings Margins . . . . . 260
A15 Earnings Sensitivity to GDP and MPCs: Demographic Group Decomposition261
A16 Earnings Sensitivity to GDP, by Decile of the MPC and Income Distribution 262
A17 Employment in the Great Recession and the Local Matching Multiplier . . . 263
A18 Employment Changes and the Local Matching Multiplier Over Time . . . . 264
A19 Consumption Response to Aggregate Shock in Benchmark and Actual Scenarios . . . . . 265
A20 Consumption Response by the Length of the Shock . . . . . 265

A21 The Matching Multiplier and the Length of the Rationing Period . . . . . 266
B1 Cross Commuting Zone Variation in Average Taks Concentration . . . . . 284

THIS PAGE INTENTIONALLY LEFT BLANK

# List of Tables


1.1 Summary Statistics for LEHD Sample . . . . . 90

1.2 Earnings Elasticity to GDP by a Worker's Marginal Propensity to Consume 96

1.3 National Estimates of the Matching Multiplier . . . . . 97

1.4 Employment Sensitivity to GDP and the Local Matching Multiplier . . . . . 98

1.5 Tradable and Nontradable Employment and the Local Matching Multiplier 99

1.6 Model Calibration by Demographic Group . . . . . 100

2.1 Summary Statistics . . . . . 169

2.2 Industry Regressions of Change in Share of Labor on Change in Concentration, Manufacturing . . . . . 171

2.3 Industry Regressions of the Change in the Payroll-to-Sales Ratio on the Change in Concentration, Different Sectors . . . . . 172

2.4 Decompositions of the Change in the Payroll-to-Value-Added Ratio, Manufacturing . . . . . 173

2.5 Decompositions of the Change in the Payroll to Sales Ratio, All Sectors . . . 174

2.6 Characteristics of Concentrating Industries . . . . . 175

3.1 Examples of Concentrated Tasks and Occupations . . . . . 208

3.2 Relationship Between Local Industry Employment and Wages and National Output . . . . . 209

3.3 Relationship Between Local Industry Employment and Government Spending Shocks . . . 209
3.4 Relationship Between Local Industry Wages and Government Spending Shocks . . . 210
A1 Years in the Estimation Sample by State . . . 267
A2 Mobility Patterns: Cross-State . . . 268
A3 Mobility Patterns: Cross-Sector . . . 268
A4 Summary Statistics for the Panel Study of Income Dynamics . . . 269
A5 Summary Statistics for the Consumer Expenditure Survey . . . 270
A6 Coefficient Estimates for Individual Marginal Propensities to Consume . . . 271
A7 The Correlation of MPCs Using Different Identifying Income Shocks . . . 272
A8 Estimates of the Marginal Propensity to Consume: Adding Job-Level Characteristics . . . 273
A9 MPCs Estimated Using Job-Level and Geographic Variables . . . 274
A10 Robustness of Relationship Between MPCs and Earnings Elasticities: Alternate Specifications . . . 275
A11 Robustness of Relationship Between MPCs and Earnings Elasticities: Alternate MPC Imputations . . . 276
A12 Heterogeneity in Worker Exposure to Recessions: ACS Subsample With Household Characteristics . . . 277
A13 Decomposing the Relationship Between Worker MPCs and GDP Cyclicality 278
A14 Heterogeneity in Earnings Elasticity Among the Unemployed . . . 279
A15 National Estimates of the Matching Multiplier: Including the Unemployed . 279
A16 Employment Cyclicality and the Local Matching Multiplier: Robustness . . 280
A17 Tradable and Nontradable Employment and the Local Matching Multiplier: Robustness . . . 281
A18 Estimated Discount Rates ($\beta$) for Alternate Process Calibrations . . . 282

A19 Model-Based Estimates of the Matching Multiplier . . . . . 283
B1 Summary Statistics for Industry-by-Commuting Zone data in the QWI . . . 284
B2 Robustness: Relationship Between Local Industry Employment and National Industry Variation . . . . . 285
B3 Robustness: Relationship Between Local Industry Wages and National Industry Variation . . . . . 285
B4 Alternative Specifications: Relationship Between Local Industry Employment and National Output . . . . . 286
B5 Robustness: Relationship Between Local Industry Employment and Government Spending Shocks . . . . . 286
B6 Robustness: Relationship Between Local Industry Average Wages and Government Spending Shocks . . . . . 287

THIS PAGE INTENTIONALLY LEFT BLANK

# Chapter 1

# The Matching Multiplier and the Amplification of Recessions

## 1.1 Introduction

The postwar U.S. economy is characterized by periodic large recessions. In the 11 recessions since 1945, gross domestic product fell by an average of 2 percent, and the unemployment rate spiked by an average of 2.3 percentage points. Most recently, in the Great Recession – the most severe downturn in the post-war period – GDP contracted by more than 4 percent, consumption fell by almost 3 percent, and the economy shed 8.6 million jobs (Mian and Sufi (2014)). Recessions are also unequally distributed. In the labor market, the employment of small and young firms is particularly volatile, as are the earnings of both very low and very high earners (Fort et al. (2013), Guvenen et al. (2017)). This project proposes a link between the heterogeneous impact of cyclical shocks and the size of recessions. I show that the unequal incidence of recessions increases the aggregate marginal propensity to consume and via this channel significantly amplifies recessions. I summarize the amplification coming from the matching of workers to different job types by the matching multiplier, and further show that local economies with greater matching

multipliers experience more volatile business cycles.

The response of aggregate consumption to business cycle shocks has been a central focus of macroeconomic research (Mian, Rao, and Sufi (2013), Kaplan and Violante (2014)). In both traditional and modern models of the business cycle, the aggregate MPC plays a critical role. For example, in the simplest Keynesian models, the Keynesian cross captures how the aggregate MPC amplifies recessions. The Keynesian cross captures a simple, intuitive feedback loop: When there is a shock that decreases aggregate demand in the economy, some of that translates into lower incomes, which leads to depressed demand, which again feeds back into incomes, *ad infinitum*. Through this feedback mechanism, the initial demand shock is amplified, and the size of each of these feedback loops is determined by the aggregate marginal propensity to consume, or MPC, which measures what fraction of each unit of income is turned back into consumption. This cumulative effect is memorably summarized by the classic Keynesian output multiplier, $\frac{1}{1-MPC}$. The same economic forces are present in many micro-founded modern models, including Werning (2015), Auclert (2017), Flynn, Patterson, and Sturm (2018), Bilbiie (2018b) and Auclert, Rognlie, and Straub (2018).

While the aggregate MPC is core to both Keynesian and New Keynesian models, it is not itself a fundamental parameter of the economy. Unlike time or risk preferences, which are inputs to a model, the aggregate MPC is a feature of the model that depends on the model's other assumptions. For example, in a representative agent New Keynesian model, the aggregate MPC in response to a transitory income shock is very small because the agent saves to smooth consumption, while in models featuring substantial heterogeneity (e.g., with constrained hand-to-mouth consumers), the aggregate MPC can be an order of magnitude larger (Bilbiie (2018b), Galí, López-Salido, and Vallés (2007), Kaplan and Violante (2014)). To elucidate the core mechanism through which the incidence of recessions of recessions in the labor market increases the aggregate MPC, consider the case where all income comes from labor earnings. In this case, the aggregate marginal

propensity to consume is simply given by

$$MPC = \sum_i \frac{dC_i}{dE_i}\frac{dE_i}{dY} = \underbrace{\sum_i \frac{dC_i}{dE_i}}_{\text{Avg. MPC}} \times \underbrace{\sum_i \frac{dE_i}{dY}}_{\substack{\text{Avg. earnings}\\ \text{change}}} + \underbrace{Cov(\frac{dC_i}{dE_i}, \frac{dE_i}{dY})}_{\text{Covariance}} \tag{1.1}$$

where $E_i$ is the income of individual $i$, $C_i$ is the consumption of individual $i$, and $Y$ is aggregate output.[1] The aggregate MPC is made up of three terms. First, it depends on the average level of individual MPCs, $\sum_i \left(\frac{dC_i}{dE_i}\right)$. When individual consumption responds more on average to changes in incomes, the MPC is higher. Second, it depends on the sensitivity of overall earnings to income, $\sum_i \left(\frac{dE_i}{dY}\right)$. When all output is earned by workers, this term is exactly 1. Third, and most importantly for this paper, it depends on the covariance between individual MPCs and the sensitivities of individual incomes to aggregate movements, $Cov(\frac{dC_i}{dE_i}, \frac{dE_i}{dY})$. This covariance term captures how the matching of workers with different MPCs to jobs with different sensitivities to aggregate fluctuations affects the magnitude of the aggregate MPC – when shocks disproportionately hit the incomes of individuals whose consumption is more sensitive, the aggregate MPC is larger. This covariance also implies a simple summary measure, motivated by the models of the business cycle mentioned above, of how this type of matching amplifies business cycles – a mechanism I term the *Matching Multiplier*. Formally, I define the matching multiplier as the component of the overall output multiplier that comes from this covariance between worker MPCs and the sensitivity of their income to shocks. This term captures the additional output response generated by the unequal exposure of worker earnings to recessions, relative to a benchmark in which all workers face the average earnings elasticity.

In the first part of this paper, I estimate this key covariance term and quantify the degree to which the earnings of high-MPC workers are more exposed to recessions. The

[1] While this is the aggregate MPC in models where labor is the only factor, the multiplier will not, in general, be $\frac{1}{1-MPC}$, as other model features, such as endogenous fiscal or monetary rules, may affect the multiplier. See Flynn, Patterson, and Sturm (2018) and Auclert, Rognlie, and Straub (2018) for examples.

key challenge in estimating this empirical moment has been that it requires detailed information on both consumption and income, and very few national data sets include both well-measured, detailed longitudinal earnings data and measures of consumption at the individual level.[2] I overcome this challenge by combining information from two data sets. I first use the Panel Study of Income Dynamics, or PSID, which is a longitudinal survey that includes measures of both consumption and income, to estimate the MPC for individuals based on their detailed demographic characteristics. Then, using those sample demographic characteristics, I impute marginal propensities to consume in the matched worker-firm earnings data recorded in the Longitudinal Employer-Household Dynamics data set, or LEHD. Using this imputation, I explore the degree to which workers with different MPCs are more exposed to movements in aggregate GDP.

In accordance with the large literature estimating individual marginal propensities to consume, I find that at the individual level, MPCs are sizable and heterogeneous. I identify individual MPCs using the heterogeneous consumption drop upon unemployment. Consistent with other patterns in the literature, I estimate an average MPC out of lost labor income of 0.5, with young, black, and poor workers having higher MPCs. Under several assumptions that I define and test, I am able to impute these MPC estimates in the LEHD.

I uncover a large, positive covariance between these estimates of a group's marginal propensity to consume and the sensitivity of the group's earnings to aggregate GDP. Figure 1.1 displays this key positive relationship. Each circle represents a detailed demo-

[2] One notable exception is the JPMorgan Chase data, analyzed in Ganong and Noel (2017), which includes data on both direct deposit earnings and detailed credit card spending. This data set, however, has two substantial weaknesses relative to the Longitudinal Employer-Household Dynamics. First, and most importantly, it covers a much shorter time series, beginning in 2012, making it impossible currently to explore any recession periods. Second, it does not include much information on the firms in which individuals work, limiting the scope to explore firm-level determinants of earnings heterogeneity. The PSID also includes data on both income and consumption. Specifically, while the PSID does cover multiple recessions, its earnings measures pale in comparison to the administrative earnings data in the Longitudinal Employer-Household Dynamics. The administrative earnings data are better measured, have a higher frequency, provide a much larger sample size, and include firm-level variables that enable me to explore the role that firms may play in determining earnings behavior.

graphic group, the y-axis plots the average earnings elasticity to GDP within that group, and the x-axis plots the average MPC of that demographic group. Across groups, there is a strong positive relationship, with the earnings of high-MPC workers being more exposed to aggregate fluctuations than those of low-MPC workers.[3] The magnitude of this overall covariance is large – the earnings of an individual with an MPC that is one standard deviation above the mean is 0.3, or 30 percent, more sensitive to aggregate movements in output than someone with the average MPC.

In the second part of this paper, I show that the magnitude of this covariance between worker MPCs and earnings elasticities is large enough to have a meaningful effect on the response of the aggregate economy to shocks. First, using the nationally derived estimates for individual MPCs, the relationship between those and individual earnings elasticities shown in Figure 1.1, and a partial equilibrium model, I estimate that the heterogeneous incidence of shocks increases the aggregate marginal propensity to consume out of labor income by 6 percentage points. Due to the non-linearity of the multiplier, this difference in MPCs leads to an estimate of the matching multiplier of 0.13, which implies that overall, the amplification of demand shocks (i.e. the multiplier minus 1) is almost 40 percent larger than it would be if all workers faced the earnings sensitivity of the average dollar in the economy. Since this multiplier captures the general equilibrium response to a broad class of demand shocks, this amplification mechanism has implications not only for the magnitude of recessions in general but also for the response of the economy to both fiscal and monetary policy.

While the model framework provides a measure of how the matching of workers to different types of jobs and income processes might impact the amplification of recessions, whether this measure has actually played an important role in business cycles is still an open question. To shed light on the empirical importance of this mechansim, I turn to the

[3]Additionally, while this empirical exercise does not focus on uncovering what factors are driving this empirical relationship or determining whether it is privately optimal for individuals to sort in this manner, I find that most of this relationship is driven by differences within the firm rather than by the sorting of workers across firms.

geographic variation in the administrative earnings data to test and validate the model's prediction that areas with a larger matching multiplier experience deeper recessions and larger booms. In the theoretical framework, the matching multiplier matters because, for a given shock, it leads to a larger consumption response, and thus amplifies aggregate fluctuations through the consumption multiplier. Therefore, we should see more pronounced fluctuations in economies with greater multipliers, and more specifically, greater matching multipliers. Under the assumption that a significant share of demand within a commuting zone is derived locally, output in a commuting zone, or CZ, with a higher matching multiplier should be more sensitive to shocks, and indeed, this is what I find. I estimate significant local variation in the magnitude of the matching multiplier, and show the CZs with higher matching multipliers suffer deeper employment losses during recessions. I also find that this difference is entirely concentrated in nontradable industries, where the local consumption response should be much more important. Together, these local estimates provide additional important evidence for the quantitative significance of this matching multiplier mechanism in explaining economic fluctuations, as they do not rely on the structure of any specific modeling assumptions.

While my empirical measures of the matching multiplier treat MPCs of different demographic groups as fixed, in practice, consumption behavior will be determined in equilibrium as a function of current and future income shocks and the expected variability of those shocks. I explore the importance of these dynamic responses by quantifying the matching multiplier within a more generalized heterogeneous agent model. Specifically, I augment a standard Bewley-Huggett-Aiyagari model of consumer demand to include endogenous labor supply and rich cross-demographic group heterogeneity. I pair it with a simple supply side featuring sticky wages in the short run and exogenous labor rationing. I first show that in a special case of the model, the matching multiplier I estimated in the empirical analysis is exactly the sufficient statistic for understanding the general equilibrium effect of demand shocks. I then calibrate the consumer side of the model to feature

heterogeneity in consumer demand and show that the model generates a matching multiplier similar in magnitude to the empirical results. Moreover, I show numerically that the sufficient statistic used throughout the empirical analysis well captures this amplification mechanism even as I relax the assumptions that define the special case in which it is exactly the right sufficient statistic. These findings both theoretically ground the main empirical moment and demonstrate the importance of this mechanism in a wide class of models.

The analysis in this paper adds to a large literature emphasizing that micro heterogeneity in the consumption responses to income changes is critical in determining aggregate dynamics. Important works in this area include empirical studies documenting substantial heterogeneity in marginal propensities to consume at the individual level (Johnson, Parker, and Souleles (2006), Fagereng, Holm, and Natvik (2018), Jappelli and Pistaferri (2014)) and quantitative models demonstrating the importance of agent heterogeneity in the determining the effectiveness of fiscal (Galí, López-Salido, and Vallés (2007), Kaplan and Violante (2014)) and monetary policy (Auclert (2017), Kaplan, Moll, and Violante (2018), McKay, Nakamura, and Steinsson (2016)). Additionally, several other papers have highlighted the key role that the general equilibrium redistribution of income plays in the effectiveness of automatic stabilizers (McKay and Reis (2016)) or fiscal redistribution over the business cycle (Oh and Reis (2012)). This analysis builds on these papers but differs from them in two critical respects. First, I focus on a particular amplification mechanism coming from the covariance of worker MPCs and the sensitivity of their incomes to the business cycle. Second, I undertake a detailed empirical analysis of this channel of amplification using microdata to both quantify this covariance and empirically evaluate its importance in the aggregate.

One important emphasis in the heterogeneous-agent New-Keynesian literature is that the introduction of constrained, high-MPC workers does little to increase the aggregate MPC, as when these models are calibrated to match the empirical distribution of wealth,

high-MPC workers comprise a small share of the economy (Kaplan, Violante, and Weidner (2014)). Recent works by Kaplan and Violante (2014) and Kaplan, Moll, and Violante (2018) generate a larger aggregate marginal propensity to consume by introducing two types of assets, liquid and illiquid, which introduces "wealthy hand-to-mouth" individuals, who have substantial assets but high MPCs out of transitory shocks. Alternatively, Krusell and Smith (1998) and Carroll et al. (2017) show that incorporating preference heterogeneity generates large aggregate MPCs while matching the wealth distribution. The covariance between worker MPC and their earnings elasticity, which is the focus of this paper, is an alternate amplification mechanism that increases the aggregate MPC through heterogeneity in the incidence of the aggregate shock in the labor market. While high MPC workers may constitute a small share of the economy, if their income is most affected by the aggregate shock, they will become disproportionately important in determining the aggregate response to the shock.[4]

The importance of heterogeneity in worker's earnings elasticity to the aggregate is also highlighted analytically in a series of papers by Bilbiie (2008, 2018a, and 2018b) and by Auclert (2017). In a two-agent New Keynesian model, Bilbiie (2008) derives the result that monetary policy shocks are amplified with agent heterogeneity only when the elasticity of income of the constrained agent is above 1. Extended to the case with multiple agents, Auclert (2017) similarly highlights the theoretical role that the covariance between worker MPCs and the sensitivity of their incomes to aggregate output plays in amplifying monetary policy shocks. This paper differs from those papers in its focus on the measurement of this key covariance and quantification of its aggregate consequences for the amplification of shocks. However, while I focus on the importance of this moment for amplification, Bilbiie (2018a) shows the positive estimated covariance also has important

[4] It is important to note that the mechanism emphasized in this paper also differs from the amplification that comes from countercyclical income risk, featured in various ways in Werning (2015), McKay (2017), Heathcote and Perri (2018), and Ravn and Sterk (2017), among others. The matching multiplier mechanism explored here focuses instead in the distribution of realized income, rather than cyclical changes in income risk. Indeed, recent work by Bilbiie (2018a) clearly disentangles these two channels and shows that these two forces – countercyclical income risk and heterogeneous incidence of shocks – reinforce each other.

implications for the determinacy of heterogeneous-agent New Keynesian models with interest rate rules and potentially exacerbates a wide class of New-Keynesian puzzles.

This project also closely connects to a growing empirical literature examining the incidence of recessions. Recessions can unequally distribute shocks through several channels, including the housing market (Mian, Rao, and Sufi (2013)) or the asset market (Glover et al. (2011)). This project specifically focuses on heterogeneity coming through the labor market. In that context, Hoynes, Miller, and Schaller (2012) show that the earnings of young, low-education men are more sensitive to business cycles, and, using high-quality administrative tax records, Guvenen et al. (2017) document that the earnings of both the very low and very high income workers are particularly exposed.[5] Similarly, Guvenen, Ozkan, and Song (2014) find that the fortunes of individuals during recessions are strongly predicted by an individual's past earnings history.[6] I expand on these findings and link them to the theoretical literature to understand their implications for macroeconomic stability.

The rest of the of the paper proceeds as follows. Section 1.2 defines the matching multiplier in a simplified two-period framework. Section 1.3 describes the two main data sets that I combine in my empirical analysis. Section 1.4 presents empirical estimates of the degree to which workers with different marginal propensities to consume are differentially exposed to aggregate shocks. Section 1.5 provides estimates of the matching multiplier nationally and Section 1.6 uses geographic variation to empirically test the importance of this amplification mechanism. Lastly, Section 1.7 derives the empirical matching multiplier as

[5] This evidence on the relationship between income sensitivities and lagged incomes does not have any immediate implications for the relationship between income sensitivities and MPCs. First, MPCs are not a direct function of income and vary with other characteristics such as time preferences (Parker (2017)) or liquid wealth (Kaplan and Violante (2014)). Second, since MPCs are generally linearly falling in wealth, there is not clear mapping to the nonlinear relationship between the elasticity of income to changes in aggregate output across the income distribution.

[6] Relatedly, Coibion et al. (2017) explore the degree to which the heterogeneous effects of monetary policy shocks on earnings affect the level of inequality in the economy. Using the Consumer Expenditure Survey, they find that contractionary monetary policy increases earnings inequality, a pattern that is consistent with my finding that the earnings of low-income workers are more exposed to shocks. See Mian and Sufi (2016) for an extensive survey of the literature on the cyclicality of incomes and consumption across individuals.

a special case of a more generalized heterogeneous agent model and shows numerically that it well approximates the strength of the mechanism in a more general setting. Section 1.8 concludes.

## 1.2 Defining the Matching Multiplier

The equilibrium assignment of workers to jobs affects the economy's response to aggregate shocks in a wide class of models (Auclert (2017), Carroll et al. (2017)). In this section, I illustrate the effect of the equilibrium assignment of workers to jobs in a simple two-period framework. I return to a more formal generalized case in Section 1.7, where I connect these expressions to those derived in a setting where consumption is allowed to respond to changes in interest rates and future income realizations. Indeed, I will show that the expression derived here for the matching multiplier is close to the sufficient statistic for the mechanism in the more generalized framework, and I clarify the conditions under which the generalized setting exactly collapses to the expressions derived in this section and used as the foundation of the empirical analysis.

To begin, consider the simple case in which worker $i$ has a consumption function given by

$$C_i = c(E_i(Y), \theta)$$

where $E_i$ are the earnings of individual $i$, which are given as a function of aggregate output $Y$, and $\theta$ are other parameters affecting the consumption of the individual, such as preferences, borrowing constraints, etc. I assume that all output is consumed by workers, meaning that the market clearing condition dictates $Y = C$. Since I am interested in understanding the importance of demand-side heterogeneity in propagating shocks, I assume that prices are fixed and that output is demand-determined. Therefore, the total derivative of the market-clearing condition yields

$$dY = \sum_i \frac{dc_i}{dE_i}\frac{dE_i}{dY}dY + \underbrace{\sum_i \frac{dc_i}{d\theta}d\theta}_{d\varepsilon}$$

where $d\varepsilon$ is the change in total demand in response to an exogenous shock before output is allowed to adjust. Define $\gamma_i = \frac{dE_i}{dY}\frac{Y}{E}$ as the elasticity of individual $i$'s earnings to the aggregate and $MPC_i = \frac{dc_i}{dE_i}$ as the marginal propensity to consume of individual $i$. Assuming that $\sum_i \frac{dc_i}{dE_i}\frac{dE_i}{dY} < 1$, the change in $Y$ can be expressed as

$$\frac{dY}{d\varepsilon} = \frac{1}{1 - \sum_i \frac{E_i}{Y} MPC_i \gamma_i} \tag{1.2}$$

where $\sum_i \frac{E_i}{Y} MPC_i \gamma_i = \sum_i \frac{dC_i}{dE_i}\frac{dE_i}{dY}$ is the *actual* aggregate marginal propensity to consume ($MPC^a$) in the economy – it captures how much of an additional unit of output is translated into an additional unit of consumption demand, taking into account the distribution of the aggregate shock. The multiplier, which is given in this setting by $\frac{1}{1-MPC^a}$, is critically important, as it determines the economy's response to any demand shock $d\epsilon$. Equation 1.2 can be rewritten to highlight the role of earnings heterogeneity as

$$\frac{dY}{d\varepsilon} = \frac{1}{1 - \left(\overline{\gamma}\,\overline{MPC} + Cov(MPC_i, \gamma_i)\right)} \tag{1.3}$$

where $\overline{MPC}$ is the earnings-weighted average MPC in the economy, $\overline{\gamma}$ is the elasticity of the average dollar in the economy to aggregate output, and $Cov(MPC_i, \gamma_i)$ is the earnings-weighted covariance between MPCs and earnings elasticities.[7] In the benchmark case in which every worker has an earnings elasticity equal to the average, $Cov(MPC_i, \gamma_i) = 0$ and the aggregate MPC is $MPC^b = \bar{\gamma}\,\overline{MPC}$. However, when the labor earnings of high-MPC workers are more exposed to aggregate movements in output, $Cov(MPC_i, \gamma_i) > 0$, and the aggregate MPC is larger. To explicitly capture the contribution of the covariance

[7] In this simple model, the earnings-weighted average elasticity in the economy is 1, but in the data, this number will be different, as labor earnings may not move one for one with output due to various frictions.

term to the multiplier, I define the *matching multiplier, or* $MM$, as difference between the multiplier when workers face their actual earnings elasticity and a benchmark multiplier where the covariance between worker MPCs and earnings elasticities is 0. The essential role of the covariance between worker MPCs and earnings elasticities can be seen clearly in the first-order approximation of this difference:[8]

$$MM = \frac{1}{1 - MPC^a} - \frac{1}{1 - MPC^b} \approx \frac{Cov(MPC_i, \gamma_i)}{(1 - MPC^a)^2} \tag{1.4}$$

The key driver of the matching multiplier is the covariance between worker MPCs and the elasticity of worker earnings to aggregate shocks, which exactly captures the difference in the aggregate marginal propensities to consume between the actual and benchmark scenarios. The denominator rescales this key difference to capture the nonlinearity in the multiplier – an increase in the MPC affects the multiplier more when the MPC is closer to 1 than when the MPC is closer to 0. This first-order approximation in Equation 1.4 will guide my empirical focus on estimating the covariance between worker MPCs and earnings elasticities to aggregate output.

[8] A first-order Taylor expansion of this benchmark multiplier around the actual MPC is given by

$$\text{Multiplier} = \frac{1}{1 - MPC^a} + \frac{MPC^a - MPC^b}{(1 - MPC^a)^2}$$

The matching multiplier is the multiplier when $MPC^a \neq MPC^b$ minus the multiplier when $MPC^a = MPC^b$, thus yielding Equation 1.4. I linearly approximate the change in the multiplier to abstract from the potentially large outliers that are induced by very low or high aggregate MPCs. While this is not an issue nationally, it would affect the CZ-level estimates in Section 1.6, which include both more true variation and estimation noise.

## 1.3 Data description

### 1.3.1 Longitudinal Employer-Household Dynamics (LEHD) linked to American Community Survey

The main data set for this analysis is the Longitudinal Employed-Household Dynamics (LEHD) data set from the U.S. Census Bureau. The LEHD is a longitudinal data set that provides quarterly earnings for all workers covered by the state-level Unemployment Insurance, or UI, Program. The data set in this project includes a subset of 23 states in an unbalanced panel from 1995 to 2011, a period that covers two recessions. Figure 1.2 plots the fraction of total U.S. private sector employment that is covered in this subset of states. By the late 1990s, this subset contained almost 50 percent of total U.S. private employment.[9] In addition to detailed information on the quarterly earnings of these workers, the data set also includes information on the establishment (location, industry, firm size, and age), as well as demographic information on the workers (age, race, gender, and education).[10] Reported quarterly earnings in the LEHD include gross wages and salaries, bonuses, stock options, tips, and other gratuities, as well as the value of meals and lodging (Spletzer (2014)).

In implementing the analysis, I make several important restrictions to the raw LEHD data. First, I exclude the first two years that an individual appears in any state within my sample. Those observations mechanically have zero earnings for at least part of the

[9] In a given state, this data set covers about 95 percent of private sector employment. See Appendix Table A1 for the list of included states, as well as the years for which each state is in the sample. I include all states to which I was given access. Appendix Table A4 shows, within the PSID, that the demographic and labor market characteristics of workers in the LEHD states are very similar to those nationally over the same sample period.

[10] Data on worker demographics in the LEHD come from the internal Census Personal Characteristic file, which covers 95 percent of the sample and is used to identify individuals across all census transactions and the long- and short-form censuses, which cover 61 and 12 percent of the sample, respectively (Vilhuber and McKinney (2014)). Specifically, worker age and gender are recorded in the person characteristic file, race is sourced from the long-form census, and education is sourced from the short-form census. In all cases, missing demographic data are imputed by the Census Bureau. Because education is imputed for almost 90 percent of the sample, I use this variable only in robustness exercises but not in the baseline results reported throughout this analysis. See Abowd et al. (2009) for further details on the data construction.

previous two years, and as lagged individual earning will be a critical input to the imputation of individual marginal propensities to consume, I exclude these observations. Second, in order to abstract from education or retirement decisions, I exclude workers who are younger than age 25 or older than age 62. Appendix Figure A1 shows that together, these restrictions exclude about 28 percent of all individual-year observations in the sample. Third, for computational reasons, I restrict my attention to an annual fourth-quarter snapshot of the data.[11] Table 1.1 shows simple summary statistics for this sample. Over the full sample, I observe an average of 38 million workers each year, about half of whom are male and who earn an average of $44,000 dollars per year.

In order to refine elements of my empirical analysis, I take advantage of a linkage between the administrative earnings data and the rich survey data contained in the American Community Survey, or ACS. Beginning in 2000, the ACS has interviewed approximately 1 percent of the population every year on an array of topics including labor market activities, educational decisions, and housing situations. While the ACS has includes detailed information at the individual level, it does not include longitudinal information at the worker level. Therefore, I link these two data sources at the individual level; the ACS is designed to be a 1 percent sample each year, and I am able to match almost exactly 1 percent of the workers in my LEHD sample to an observation in the ACS in a given year.

I use the information in the American Community Survey for two primary purposes. First, I expand the set of worker observables for a subset of workers, allowing for a richer analysis of worker heterogeneity. Second, I explore and validate my assumptions on the labor market activities of individuals as they transition in and out of the LEHD sample. While the LEHD provides a comprehensive snapshot of employment, it does not provide information on the labor market activity of those who are not currently employed in that sample. Throughout the analysis, I assume that prime-age workers who leave

[11] It is common in the literature to restrict analysis of the LEHD to an annual snapshot. See Sorkin (2018) and Abowd, Lengermann, and McKinney (2003). I explore the robustness of the results to using annual earnings data rather than fourth-quarter earnings and find that results are similar. See the discussion in Appendix A.

employment in my sample make no labor market earnings in those quarters. This assumption will affect both my measures of a worker's past earnings history (and thus of their marginal propensities to consume) and their earnings profiles over the business cycle. This assumption of zero earnings in periods of nonemployment would be violated if individuals move to a job either outside the LEHD coverage (i.e., to the military, federal employment, or self-employment) or to a state that is not in my sample. I find that among workers who leave my LEHD sample between $t-1$ and $t$, and are in the ACS in year $t$, 24 percent report being employed elsewhere in the sample.

While these patterns in the ACS do suggest that there is some measurement error, this error would be a problem for my analysis if workers of different characteristics are differentially likely to move outside the LEHD sample over the business cycle. For example, if younger workers are disproportionately likely to move to states outside my sample during recessions, I will overstate the unemployment of the young workers in recessions and thus will erroneously conclude that the earnings of young workers are more sensitive to recessions when, in fact, they are not.[12] However, I show in Appendix A that there is no evidence that higher-MPC individuals are more likely to either move to employment not covered by unemployment insurance or out of state during recessions, suggesting that this potential bias is likely small.

### 1.3.2 Panel Study of Income Dynamics

I supplement the detailed administrative data on earnings with marginal propensities to consume estimated using the Panel Study of Income Dynamics. Each household in the

[12] In addition to overstating the sensitivity of these workers' earnings to GDP, I may also overstate the difference in MPCs between workers of different ages. As I discuss in Section 1.4.1, I use a worker's past earnings to impute a worker's MPC. Differences in cross-state mobility by worker characteristic would bias the lagged earnings of those workers, and this in turn would bias my estimate of that worker's MPC. Since I mostly focus on the cross section of employment and I remove the initial employment periods, this bias in the measurement of lagged earnings would only appear for workers who moved out of my sample to non-covered employment and then moved back into my sample in a future period. Additionally, the average MPC imputed in the LEHD sample (0.405) is very similar to the average MPC in the PSID sample (0.399), suggesting that cross-state mobility is not biasing the estimate of the average MPC.

PSID is interviewed every year from 1968 to 1997 and every other year after that. Among other things, households are asked detailed questions on their demographics and labor market experiences of the household head and spouse. In order to include secondary earners in my analysis and estimate a marginal propensity to consume at the individual level for both household heads and spouses, I transform the data to have up to two observations per household, one for the head and one for the spouse. This transformation means that each individual within the household has the same consumption, but the head and spouse differ in their demographics and labor market variables. Appendix Table A4 shows summary statistics for the PSID sample that I include in the estimation.[13] Because I include all samples within the PSID to estimate MPCs (i.e., the nationally representative sample and the oversample of low-income groups), the sample has a higher share of black and low-income workers than the total population.

For most of the PSID sample, the main expenditure variable is food consumption. There are several reasons to suspect that the response of food expenditure to income changes is not representative of overall expenditure responses. First, as food is a necessity, its share of total consumption varies across the income distribution (see Appendix Figure A2 and Bonaparte and Fabozzi (2011)) Second, the provision of food stamps potentially distorts food consumption decisions on the margin and likely dampens fluctuations in food expenditure relative to overall expenditures (Hastings and Shapiro (2017)). In order to address these issues, the main measure of expenditure I use throughout the analysis is total expenditure, which I impute using overlapping information in the Panel Study of Income Dynamics and the Consumer Expenditure Survey, or CEX, following the methodology introduced in Blundell, Pistaferri, and Preston (2008) and expanded in Guvenen

[13] I start with the source and SEO samples of the PSID from 1968 to 2015. I drop any observations that do not have two lags, used to define previous income, and one lead, used to define current income. See Appendix A for a richer discussion of the timing of the PSID survey. As in the LEHD analysis, I also drop those individuals younger than age 25 or older than age 62. I also drop observations with missing food consumption or missing income. Lastly, I drop observations where the two-year change in either food consumption or income is more than fourfold. These restrictions on outliers are similar to those in Hendren (2017), who excludes individuals with more than a threefold change in food consumption, and Gruber (1997), who excludes observations with a greater than 1.1 log change in food consumption.

and Smith (2014). Using the CEX, I estimate a demand for food expenditure as a function of durable consumption, nondurable consumption, demographic variables, and relative prices. Under the assumption that food demands are monotonic, this demand function can be inverted to get predicted total consumption based on the food expenditures and demographics of the household in the PSID.[14] Appendix Figure A3 shows that this imputation not only matches the time series and levels of consumption in the CEX but also captures important cross-sectional patterns in the CEX such as the food share of consumption and age profile of consumption. Section 1.4.1 further describes the details of how I utilize this imputed measure of consumption and the unique panel structure of the PSID to estimate worker marginal propensities to consume.

## 1.4 MPCs and the incidence of recessions

As I discussed in Section 1.2, the matching multiplier is a direct function of the covariance between two individual-level statistics – an individual's marginal propensity to consume and the sensitivity of individual earnings to aggregate shocks. In the next section, I describe how I estimate each of these parameters. In Section 1.4.1, I present estimates of MPCs by worker characteristic in the PSID, where there is both consumption and earnings data. In Section 1.4.2, I impute these MPCs in the LEHD using overlapping information on worker demographics and estimate the degree to which workers of different MPCs are differentially exposed to aggregate shocks.

[14] An alternative method for imputing more comprehensive consumption measures in the PSID, proposed by Attanasio and Pistaferri (2014), is to use the expanded consumption categories in the later years of the PSID. Specifically, they impute relationship between food consumption and overall consumption in the 1999-2013 survey years to impute the total consumption in the previous years of the sample. Appendix A discusses this imputation measure in more detail and shows that estimates of marginal propensities to consume using this methodology show similar patterns to the CEX-based imputation.

### 1.4.1 Estimating marginal propensities to consume

Since I do not observe the consumption behavior of individuals in the LEHD, I estimate the marginal propensity to consume for workers with different characteristics using the panel structure of the PSID and impute these values in the LEHD. While the strategy of imputing MPCs within the LEHD based on individual characteristics is novel, my estimation of MPCs borrows from a long line of literature that explores the response of individual consumption to income changes. A consistent finding within this large and heterogeneous literature is that households exhibit high marginal propensities to consume out of transitory income shocks and that the magnitude of these responses differs across the population.[15]

I build most directly on a line of research beginning with Gruber (1997), who examines the consumption drop upon unemployment. Using the panel structure of the PSID, I estimate

$$\Delta C_{t,i} = \sum_x (\beta_x \Delta E_{t,i} \times X_{t,i} + \alpha_x X_{t-1,i}) + \delta_{t,s} + \epsilon_{t,i} \tag{1.5}$$

where $C_{t,i}$ is total household consumption of individual $i$ at time $t$, imputed from the CEX as explained in Section 1.3.2;[16] $E_{t,i}$ is labor earnings of individual $i$; $\delta_{t,s}$ are state-by-year fixed effects, which capture any time variation that is common to all individuals within a state and year; and $X_{t,i}$ is a characteristic of the individual.[17] In order to include data

[15] In particular, a series of recent papers estimate that upon receiving tax rebates, workers, on average, spend ore than half of the windfall within two quarters but that individuals with few financial resources spend more than those with more cash on hand (Johnson, Parker, and Souleles (2006), Parker et al. (2013)). See also Gelman (2016), Gross, Notowidigdo, and Wang (2017), Jappelli and Pistaferri (2014), Kaplan and Violante (2014), and Jappelli and Pistaferri (2010) for a comprehensive survey.

[16] Figure A11 explores the sensitivity of these estimates to using alternative consumption measures. While the levels of the MPCs differ, the cross-sectional patterns are very similar. Appendix A discusses an alternate estimation of MPCs that uses only the short panel structure of the CEX. While noisier, the patterns that result from this separate estimation are similar to the baseline estimates in Figure 1.3.

[17] In baseline specifications, $x$ includes: five lagged income bins, a quadratic in age, female and black dummies, black interacted with age, and female interacted with black. See Appendix A for additional details on sample restrictions.

from 1997 to 2015, when the PSID becomes every other year, I consider two-year changes in both income and consumption. The coefficients of interest are $\beta_x$, which capture the change in consumption per dollar change in income associated with an individual of characteristic $x$. If $X$ only includes characteristics that are common to both the PSID and the LEHD, then using the estimated $\beta_x$, I can impute the MPC in the LEHD as

$$\widehat{MPC}_{i,t} = \sum_x \widehat{\beta}_x X_{i,t}$$

Since there are many factors that could simultaneously move income and consumption at the individual level, I identify the causal relationship between income and consumption in Equation 1.5 using a shock to an individual's income as an instrument for $\Delta E_{t,i}$. In a general class of models, the marginal propensity to consume of the individual is a function of the type of income shock.[18] My baseline estimates are identified using unemployment as the shock to income (Gruber (1997), Hendren (2017)), and later, I explore the sensitivity of my estimates to the type of shock. In addition to providing a large shock to income, the unemployment shock likely captures income variation that is most relevant for understanding recessions. Much of the income risk associated with recessions comes through the extensive margin of employment, and thus capturing the MPC out of lost labor income upon unemployment captures the relevant variation. Note that the particular income shock of unemployment likely is partially anticipated and is somewhat permanent (Hendren (2017), Jacobson, LaLonde, and Sullivan (1993)), and thus, the estimates of $\gamma_x$ in Equation 1.5 reflect the average response to a shock with those properties. In esti-

[18] For example, as in Jappelli and Pistaferri (2010), consider the standard problem of a consumer who can borrow and lend freely and thus maximizes her expected utility subject to an intertemporal budget constraint and future stream of uncertain labor income. If we assume that interest rates are constant and preferences are quadratic, the consumer's optimality condition implies that the consumer will smooth expected consumption over time $c_{it} = E_t[c_{it+1}]$, and thus, changes in consumption reflect only revisions to expectations of the future stream of labor income. The magnitude of the revision of an individual's expectations of her discounted future income stream will thus depend on the nature of the income process and the type of income shock. See Appendix A for an expanded discussion of the relationship between the individual income process and marginal propensities to consume.

mating Equation 1.5 with unemployment as an instrument, I limit the sample to the set of individuals who are employed in the previous period and identify $\beta_x$ by comparing the change in consumption between $t-2$ and $t$ of those who remain employed to the change among those who lose their jobs.

### Description of MPC estimates

Before exploring the full distribution of MPCs that result from Equation 1.5, Figure 1.3 shows the patterns in MPC heterogeneity with bivariate regressions for select subgroups for the set of covariates that I include in the full estimation. The farthest-left estimate shows that, using unemployment as the income shock, the average marginal propensity to consume in the estimation sample is just more than 0.5. This average estimate is similar to other estimates in the literature that use comparable identifying variation.[19] Importantly, the coefficients to the right show that there is substantial variation in MPCs around this average – younger, black, and poorer workers, on average, have a larger consumption drop per dollar of lost income. The lagged income measure along which I allow MPCs to vary is the average earnings of workers in years $t-2$ and $t-3$, which is meant to capture differences in permanent income across workers stemming from different educations or other skills.[20] Women and men in Figure 1.3 have similar MPCs, on average, but women also earn less than men, and Appendix Table A6 shows that once you control for differences in earnings levels, women have lower MPCs than men. These cross-demographic

[19] Using income and consumption data from JPMorgan Chase and similar identifying variation, Ganong and Noel (2017) estimate an average MPC of 0.4. They also find demographic patterns similar to those in the PSID – they find higher MPCs among individuals with lower incomes and assets, as well as among those who are younger. Similarly, using the Nielsen Consumer Panel, McKee and Verner (2015) estimate an MPC out of unemployment insurance benefits of between 0.6 and 0.9, and Jappelli and Pistaferri (2014) use survey data in Italy and find an average MPC out of transitory income of 0.48. See Appendix Figure A8 for a graphical comparison to other similar estimates in the literature and Jappelli and Pistaferri (2010) for a comprehensive review.

[20] I use the average earnings over the previous two years to balance capturing a more permanent measure of earnings capacity against a loss of sample that comes with the more stringent within-individual panel. However, Appendix Figure A6 shows that the patterns are similar when using income either lagged further, averaged over longer intervals, or fixed at a given age.

patterns are broadly consistent both quantitatively and qualitatively with other estimates of MPC heterogeneity by demographic group in the literature (Ganong and Noel (2017), Parker et al. (2013), Parker (2017), and McKee and Verner (2015)). Putting this all together, Figure 1.4 shows the full distribution of marginal propensities to consume in the PSID that result from Equation 1.5.[21] There is a substantial amount of variation, with the large mass between 0 and 1, and a small number of estimates above 1.

While I allow for heterogeneity in worker MPCs along only four demographic dimensions (race, age, gender, and earnings history), it is not necessary for the MPCs to be a direct function of those specific characteristics. Rather, it is likely the case that these characteristics are correlated with other underlying economic circumstances that directly affect MPCs. Specifically, in models of precautionary savings or credit constraints, a key source of heterogeneity in MPCs is the individual's cash on hand, and indeed, several studies find heterogeneity in MPCs along this margin (Jappelli and Pistaferri (2014), Kaplan, Violante, and Weidner (2014), Parker et al. (2013)). The demographic patterns in Figure 1.3 are consistent with that source of heterogeneity, as the literature has shown that young, black, and low-income workers also have less liquid wealth (Dettling et al. (2017)). Indeed, in Appendix Figure A6, I replicate the finding that households with assets below median have much larger MPCs than households with assets above median. These differences in the amount of cash on hand across demographic groups may reflect different income processes across groups, which when combined with borrowing constraints or precautionary savings lead to different liquidity positions (Deaton (1991), Carroll (2001)). Alternatively, it may reflect differences in risk preferences or discount rates across demographic groups, or differences in access to credit (Gelman (2016), Parker (2017)).[22] It is

[21] Specifically, I include the following parametrization of the variables described in Figure 1.3: five approximately equally-sized lagged earnings bins ($< \$22,000$, $\$22,000 - \$35,000$, $\$35,000 - \$48,000$, $\$48,000 - \$65,000$ and $> \$65,000$); a quadratice in age; female and black indicators; an interaction between black and age; and an interaction between female and black. See Appendix Table A6 for the regression coefficients underlying the distribution in Figure 1.4.

[22] Using novel data from a financial app, Gelman (2016) explores the nature of the observed heterogeneity in MPCs. He finds that about half of the variation in cash on hand is within individual over time, while about

also possible that heterogeneity in $\widehat{\beta}_x$ reflects differences in the persistence of the unemployment shock across individuals, although empirically I find that the persistence of the unemployment shock on labor market earnings is relatively similar across demographic groups.[23]

### Assumptions for MPC imputation

While the above method for estimating MPCs closely follows existing methods in the literature, my subsequent imputation of these MPCs in the LEHD necessitates several important additional assumptions that warrant further discussion. First, in imposing that MPCs only vary by worker demographics, I assume that individual MPCs are invariant to the sign and magnitude of the income shock. As I discussed above, in a standard model in which agents maximize their expected utility subject to an intertemporal budget constraint, an individual's MPC depends on the persistence of the shock, not on the magnitude or sign. However, with liquidity constraints, the marginal propensity to consume out of small shocks may depend on the size of the shock as well as the sign (Kaplan and Violante (2014)). I explore the importance of the identifying shock by comparing MPC estimates that result from labor earnings shocks of differing size and sign.[24] Appendix Figure A9 shows that the average marginal propensity to consume is similar across these various identifying shocks. Specifically, the MPC is similar when using smaller income shock, identified from movements in industry unemployment rates or changes in state gross domestic product, or when using the positive income shock of finding a job. Additionally, Appendix Table A7 shows that not only are the averages similar but also that

half is across individual, suggesting that temporary income shocks and persistent worker characteristics are almost equally important. Parker (2017) also finds that behavioral differences across individuals play an important role in explaining variation in the consumption response to tax rebates.

[23] Appendix Figure A10 compares the persistence of the income shocks associated with unemployment across demographic groups and shows that in all cases, earnings growth recovers after one to two years.

[24] While all of the shocks in Figure A9 are shocks to labor income, in Appendix A, I also explore the demographic heterogeneity in MPCs out of tax rebates, as in Parker et al. (2013). While the estimates are noisier, I again find that the young, less educated and poorer have higher MPCs.

alternate MPC estimates are highly correlated at the individual level. This does not mean that the MPC out of all different types of shocks is the same, but rather that the MPC out of labor market earnings is similar for changes in earnings of different signs and sizes.

A second key stability assumption embedded in the imputation of MPCs in the LEHD is that for an individual income shock of a given magnitude, the consumption response is constant over the business cycle. This assumption could be violated for several reasons. It may be that liquidity constraints are more likely to bind in recessions, in which case the average MPC of a demographic group is likely to be higher. Conversely, unemployment insurance is higher in recessions, leading to lower realized MPCs out of lost labor income in recessions. Existing empirical evidence on this cyclicality of MPCs is scarce – Gross, Notowidigdo, and Wang (2017) find that the marginal propensity to consume out of liquidity is higher in recessions, but a calibration by Carroll et al. (2017) finds that MPCs are roughly constant over time. I explore this in my setting by adding an interaction of changes in income with the state unemployment rate, thereby allowing the MPC by demographic to vary over the business cycle, and find that differences, both on average and for each demographic group, are statistically and economically small.[25]

Third, I impose that at the individual level, the marginal propensity to consume is a function only of the characteristics that I include in $x$ (i.e., age, earnings history, gender, and race). While this is obviously an approximation, this assumption would be a problem for my analysis if within each demographic bin there is sorting across jobs such that it was precisely the higher-MPC workers within the group who are at cyclically insensitive jobs. If this were the case, I would be inaccurately capturing the heterogeneity in exposure of workers to business cycles by their MPC.[26] While my data do not allow me to fully address this, I explore the sensitivity of my MPC estimates in the PSID to including job-

[25] See Appendix Figure A9 for a graph showing the estimates of MPCs at different points in the business cycle.

[26] The sorting pattern wherein high-MPC workers sort into cyclically insensitive firms would be implied by a model of risk sharing between the worker and the firm. However, an alternate model of sorting is one in which high-MPC workers are credit-constrained, unable to finance long job searches, and thus take less stable and more cyclically sensitive jobs (Herkenhoff, Phillips, and Cohen-Cole (2016)).

level characteristics. If sorting across jobs of different characteristics were important in explaining MPC heterogeneity within demographic group, then these terms should have additional explanatory power. Appendix Table A9 shows that no additional job-level variables or regional controls meaningfully change the MPC estimates. While none of these definitively rule out the sorting of workers within demographic groups to jobs based on their MPC, they do suggest that this cross-job sorting within demographic group is small and unlikely to be driving the patterns.

### 1.4.2 Heterogeneity in worker exposure

Using these estimates of an individual's marginal propensity to consume, I move to the individual-level earnings data in the LEHD and estimate the degree to which workers of different marginal propensities to consume are differentially sensitive to aggregate shocks. To do so, I estimate the following equation:

$$\Delta E_{i,t} = \alpha_1 MPC_{i,t-1} + \alpha_2 MPC_{i,t-1} \times \Delta \log G_t + \delta_t + \epsilon_{i,t} \tag{1.6}$$

where $E_{i,t}$ is a measure of the individual's fourth-quarter earnings; $MPC_{i,t-1}$ is the imputed MPC of individual $i$; $\delta_t$ are year-fixed effects, soaking up any variation in earnings that is common across individuals in a given year; and $\Delta \log G_t$ is the annual change in the log of national GDP.[27] The sample includes the set of all workers employed in year $t-1$.[28] The coefficient of interest is $\alpha_2$, which captures the degree to which the earnings

[27] For computational reasons, I estimate Equation 1.6 on a 5 percent random subsample of the data. Appendix Table A10 shows an alternate estimation of Equation 1.6, where MPCs are discretized into bins of 0.01 and the data are aggregated such that each observation is the earnings change of all individuals in the earnings bin $b$ in year $t-1$. This estimation method preserves much of the variation in MPCs while utilizing the entire sample. Results are very similar to the baseline estimates in Table 1.2.

[28] I restrict attention to the set of employed workers for several reasons. First, the LEHD is a data set of employment, not the labor force, and thus does not have complete coverage of the unemployed. I also earnings weight the regressions; thus, including the unemployed would necessitate an alternate weighting strategy. Finally, a large fraction of earnings is earned by the employed rather than the new hires. However, in Appendix Section A, I use the link between the ACS and the LEHD to provide estimates for the heterogeneity in the incidence of hiring from unemployment over the business cycle. I find that the differential sensitivities among the unemployed are small and contribute very little to the aggregate. See Appendix Table A14 for

of workers of different MPCs are differentially sensitive to movements in aggregate GDP.

I explore the different dimensions of earnings cyclicality with various specifications of the outcome variable $E_{i,t}$. I capture the intensive margin of earnings elasticity by using the log of earnings, thus restricting the sample to the set of individuals who remain employed across years, and I capture the extensive margin of employment using an indicator for whether the individual is employed in time $t$. Lastly, I combine the intensive and extensive margin into one estimate using $\Delta E_{i,t} = \frac{E_{i,t}-E_{i,t-1}}{.5\times E_{i,t}+.5\times E_{i,t-1}}$. This transformation defines and bounds the earnings losses of those who lose their job between periods, thus providing an estimate for overall earnings elasticities.[29] Importantly, each individual in the regression is weighted by their share of overall earnings. What matters for the economy as a whole is not the differential elasticity of individuals but the differential elasticity of *dollars* earned in the economy.

Before going directly to estimates of $\alpha_2$ from Equation 1.6, recall Figure 1.1 from the introduction to demonstrate the variation underlying this key relationship. Each point represents a demographic group aggregated to the level of heterogeneity in the MPC imputation (for example, white men, ages 25 to 35 with lagged earnings between \$22,000 and \$35,000) and shows the relationship between group MPCs and earnings elasticities to GDP.[30] This figure clearly shows that there is a tight, almost linear relationship between these two variables – higher-MPC demographic groups are much more exposed to recessions. For example, the consumption of black men from ages 25 to 35 who earned less than \$22,000 in the previous year is very sensitive to changes in their earnings, with an

details.

[29] In Appendix Table A10, I explore alternate transformations for estimating the overall earnings elasticity. I show that the estimates using either ($\log(E_{i,t}+100)$) or normalizing the level of the earnings change by the average earnings in the individual's group produce similar patterns but in general produce larger variation in elasticities across workers.

[30] Specifically, the points on the y-axis of Figure 1.1 refer to the $\widehat{\gamma}_g$ from

$$\Delta \log E_{gt} = \gamma_1 + \gamma_g \Delta \log G_t + \epsilon_{gt}$$

estimated separately on each subgroup $g$ in the data from 1995 to 2011. The subgroup is defined in year $t-1$ and earnings are unconditional on the subgroup into which individuals move in period $t$.

average MPC of above 1, yet it is precisely these workers whose earnings are most sensitive to aggregate shocks.[31] The elasticity of their earnings is just above 2, meaning that when GDP falls by 1 percent, their earnings, on average, fall by 2 percent. In contrast, the consumption of nonblack women over age 55 who earned \$48,000-\$65,000 in the previous year is, on average, much less sensitive to changes in their income, yet the earnings of these workers are almost completely insensitive to movements in aggregate GDP.

The slope of the fitted-value line in Figure 1.1 is precisely the estimate of $\alpha_2$ from Equation 1.6.[32] The first column of Table 1.2 reports the estimate corresponding to the slope of Figure 1.1.[33] Columns 2 and 3 show this overall estimate decomposed into the intensive and extensive margin of earnings. Demographic groups that are cyclically sensitive on the intensive margin also tend to be cyclically sensitive on the extensive margin of employment, and thus, there is a strong positive relationship between the sensitivity of a worker's income to GDP and a worker's MPC along both the intensive and extensive margins. Both margins contribute similarly to the overall heterogeneity in exposure.

The subsequent columns of Table 1.2 show the estimated relationship controlling individually for various worker characteristics ($X_i$) along which MPCs vary. I include the

[31] In Appendix A, I replicate the findings in Guvenen et al. (2017) that the earnings of the very high and very low earnings are most exposed to GDP, as well as discuss how these patterns affect the relationship between MPCs and sensitivity to GDP.

[32] Note that the slope of the line in Figure 1.1 does not exactly equal the individual-based estimates in Table 1.2 because of the slightly different functional form of earnings imposed in the estimation of Equation 1.6 – since demographic groups as a whole never exit the labor market, the y-axis of Figure 1.1 uses the change in log earnings, while the individual-level regression uses a transformation to include entering and exiting workers.

[33] The standard errors reported in Table 1.2 are clustered at the individual level, which accounts for correlation in the error over time within individuals. However, they do not take into account the additional noise stemming from the MPC imputation. Indeed, the MPC estimates in the LEHD rely on two imputations – an imputation of total consumption in the PSID from the CEX, and an imputation of the MPC in the LEHD from the PSID. Column 2 of Appendix Table A10 shows standard errors adjusted to incorporate the noise in these imputations. I obtain these estimates by taking 500 random draws from both the PSID and CEX data to calculate 500 imputations of a worker's MPC in the LEHD. I then estimate Equation 1.6 for each imputation and combine the point estimates and standard errors following multiple imputation methods in Rubin (1987). As expected, the adjusted standard errors are larger, but the point estimates are still highly statistically significant. Due to the large computational burden imposed by this bootstrapping procedure and the small qualitative difference in results, I proceed with the clustered standard errors for the rest of the analysis. See Appendix A for more details on this bootstrapping procedure.

variable both individually and interacted with GDP. If all of the variations in the relationship between MPCs and earning elasticities were driven by heterogeneity along one dimension, the coefficient on the MPC and GDP interaction would drop to 0 when that characteristic was included. However, the coefficients on the main interaction term between MPCs and GDP are relatively stable across Columns 4 through 8 as each worker characteristic is added individually. Even in Column 9, when all individual controls are included, there is still a substantial coefficient on MPCs. These results show that there is no one dimension of heterogeneity that is driving the relationship but instead that the joint distribution across these characteristics together drives the relationship.[34]

This positive relationship between the cyclical sensitivity of earnings and the MPC of the worker is also very robust across empirical specifications. Appendix Table A10 shows that both the direction and magnitude of the relationship between a worker's MPC and their earnings elasticity is robust to the functional form of the earnings outcome variable as well as decisions about the level of individual earnings aggregation. Specifically, while earnings in Table 1.2 are measured at the worker level using aggregate fourth-quarter earnings, Column 5 of Appendix Table A10 shows that the patterns are robust to using annual rather than fourth-quarter earnings. Column 7 in Appendix Table A10 shows that patterns are similar when using state-level GDP as a cyclical indicator rather than aggregate GDP, suggesting that these patterns are present both across and within states. Appendix Table A11 shows that the relationship is also robust across the various methods for imputing MPCs. For all imputations, a 1 standard deviation increase in the MPC of the individual is associated with an increase in the elasticity of earnings with respect to GDP of between 0.33 and 0.39.

Lastly, in Appendix Table A12, I use the expanded information in the ACS subsample

[34] To see these relationships visually, Appendix Figure A14 shows versions of Figure 1.1 separately for the intensive and extensive margins of earnings. Appendix Figure A15 reproduces Figure 1.1 but highlights the points associated with different worker characteristics. As with the estimates in Table 1.2, all characteristics contribute to the overall relationship, and even within the various subgroups, there is a positive relationship between earnings sensitivity to GDP and worker MPCs.

to explore the possibility of additional sorting patterns within demographic groups that could bias the estimated relationship between worker MPCs and earnings elasticities. One possible concern would be that workers of the same demographic group but with different household structures sort into jobs with different sensitivity to aggregate movements. However, I find that the slope of the relationship between MPCs and earnings sensitivity to GDP is similar when I use an alternate MPC that is allowed to vary by the number of children in the household and an individual's marital status or when I aggregate fully to the household level.[35]

### 1.4.3 Discussion

Although the source of this MPC-earnings elasticity is not immediately relevant for understanding its contribution to macroeconomic stability, it is an interesting object in its own right. The finding that the earnings of high-MPC workers are more exposed to aggregate conditions is at odds with a simple model of firm-worker risk-sharing, which would predict that workers for whom fluctuations in income are costly would sacrifice part of their expected earnings to enter into contracts in which their wages are less sensitive to aggregate demand fluctuations (Bailey (1974)). These patterns are not, at face value, inconsistent with the idea that workers sort across jobs according to their risk preferences (Schulhofer-Wohl (2011)). However, for this allocation to be privately optimal, there would need to be a very strong negative relationship between risk aversion and worker MPCs.

One probable explanation for the observed covariance between worker MPCs and aggregate earnings risk is that differences with either observed or unobserved skills drive the sorting across firms and are correlated with worker MPCs. However, as I discuss in more detail in Appendix Section A, I find that the large majority of heterogeneity in busi-

[35] Appendix Figure A7 shows that the MPCs out of individual income are highly correlated with MPCs out of household income, suggesting that the importance of different demographic groups in household income does not drive the estimated differences in MPC heterogeneity.

ness cycle exposure by MPC occurs within the firm rather than between firms, suggesting that the sorting of workers across firms based on their geographic location or industry-specific skills is only a small part of the story. Within the firm, I find substantial heterogeneity in exposure by MPC even within an occupation, much of which is explained by differences in the earnings history of the workers, with low-income workers' earnings being more sensitive to changes in GDP than the earnings of higher-income workers.

The within-occupation correlation between worker MPCs and aggregate earnings risk may still reflect sorting across jobs based on the unobserved skills of the individual. However, the relative unimportance of firms and the importance of earnings history suggest that this pattern is also consistent with models where workers' exposure to aggregate shocks and their MPCs are explicitly linked. For example, these patterns are consistent empirically with a large literature showing that one unemployment spell increases the risk of future unemployment spells (Stevens (1997)). Theoretically, they could be consistent with a job ladder model as in Jarosch (2014), wherein workers search for both more productive and more secure jobs, and thus, as they climb the job ladder, they sort into higher-paying and more-secure jobs. This model of worker flows would result in a pattern where within the firm, the worker with higher lagged earnings, both because she is in a higher-paying job further up the ladder and because she has not recently experienced an unemployment spell, is less exposed to shocks than the lower-earning recently hired worker.

## 1.5 Matching Multiplier estimates

Together with the framework specified in Section 1.2, the derived estimates of individual MPCs and heterogeneity in earnings exposure by a worker's MPC allow me to calculate the matching multiplier and thus quantify the macroeconomic consequences of this inequality in worker exposure to business cycles. Recall from Section 1.2 that the matching multiplier is defined as the first-order difference in the Keynesian multiplier, with the em-

pirical incidence of aggregate shocks and the multiplier in a benchmark case in which all workers faced the same earnings elasticity:

$$MM = \frac{MPC^a - MPC^b}{(1 - MPC^a)^2}$$

In order to bring this to the data and account for the empirical fact that not all output is earned by workers in the form of wages, I make two small adjustments and calculate

$$MM = \frac{\frac{E}{Y}\sum_i \frac{E_i}{E} MPC_i(\widehat{\gamma}_i - \overline{\gamma})}{\left(1 - \left(\frac{E}{Y}\sum_i \frac{E_i}{E} MPC_i\widehat{\gamma}_i + \frac{Y-E\gamma}{Y} MPC_{Y-E\gamma}\right)\right)^2} \tag{1.7}$$

where $\widehat{\gamma}_i$ comes from the coefficient from Equation 1.6,[36] $E$ is the total earnings of the employed, $Y$ is GDP, and $\overline{\gamma}$ is the benchmark earnings elasticity, defined as the elasticity of the average dollar in the economy.

There are two additional terms in Equation 1.7 relative to Equation 1.4 that are needed to quantify this channel in the data. In the framework used to derive Equation 1.4, all output is earned by workers in the form of wages; thus, $Y = E$. However, this is not true empirically, as workers earn income from other sources (e.g., capital and profits). Therefore, in the empirical estimates of the matching multiplier, I need to make adjustments for (1 the MPC out of non-labor income ($MPC_{Y-E}$); and (2 the share of overall output going to workers in the form of wages ($\frac{E}{Y}$). In the baseline estimates, I assume that the labor share is two-thirds and I set the MPC out of nonlabor income such that the total benchmark MPC ($MPC^b$) is exactly $\overline{MPC}\overline{\gamma}$, which means that I assume that the MPC out of nonlabor income is around 0.23.[37]

[36] Estimates of Equation 1.6 in Table 1.2 do not directly provide estimates of $\gamma_i$, but rather $\gamma_i = \widehat{\alpha}_1 + \widehat{\alpha}_2 MPC_i$. In terms of the $\alpha$ estimates, $\overline{\gamma} = \overline{\alpha} + \sum_i \widehat{\alpha}_2 MPC_i \frac{E_i}{E}$ and thus, $\gamma_i - \overline{\gamma} = \widehat{\alpha}_2 - \sum_i \frac{E_i}{E} MPC_i \widehat{\alpha}_2$.

[37] While it is difficult to validate this assumption explicitly, recent empirical evidence from Sweden in Di Maggio, Kermani, and Majlesi (2018) suggests that the MPC out of non-labor income is meaningful, with an MPC out of dividends of 0.35 and the aggregate MPC out of capital gains of 0.06, suggesting that an overall MPC of 0.23 is not unreasonable. This average estimate for the MPC out of capital gains comes from a back-of-the-envelope calculation based on Table 3 in Di Maggio, Kermani, and Majlesi (2018). Individuals in the bottom half of the wealth distribution have an average MPC of 0.13, those in the top half have an MPC of 0.05, and the bottom half of the wealth distribution owns 7 percent of overall stocks. See Appendix A for

Panel A in Table 1.3 shows calculations of Equation 1.7 using the various empirical estimates of earnings heterogeneity derived above in Section 1.4.2. The top row shows the overall estimates of the matching multiplier. Column 1 shows the estimate of the benchmark MPC, defined as the earnings-weighted individual MPC (equal to 0.42) times the earnings elasticity of the average dollar in the economy (equal to 0.6). Column 2 shows that the actual MPC is 0.06 percentage points higher than this benchmark, meaning that heterogeneity in worker exposure to recessions increases the aggregate MPC by 28 percent.

Across the remaining columns of Table 1.3, I translate this difference in MPCs into a difference in the multiplier. Column 3 shows that these estimates imply a benchmark multiplier of 1.3, meaning that if all workers had the same exposure of their earnings to aggregate shocks, for each dollar of an initial shock, output changes by an additional 30 cents on that dollar. Column 6 shows that this multiplier increases by 0.13, or 10 percent, with the empirical incidence of shocks, bringing the overall multiplier to 1.42.

The subsequent rows of Panel A show similar estimates for $MM$ using the alternate methods for estimating heterogeneity in elasticities by worker MPCs. Row 2 shows $MM$ estimated using the demographic-group-specific earnings elasticities from each point in Figure 1.1. This does not impose the strict linear functional form, yet the estimates are very similar. Similarly, row 3 shows the estimates using the estimates for heterogeneity for each decile of the MPC distribution, and again, shows that the estimates are very similar.[38]

Panel B of Table 1.3 shows estimates of $MM$ for different benchmarks. The baseline estimates consider reallocating dollars across the entire economy to equalize earnings heterogeneity, but Panel B explores the benchmarks in the case when the reallocation of earnings is more restricted. The first row shows $MM$ in a benchmark case in which each worker is given the earnings elasticity of the average dollar in their commuting zone.

a derivation of Equation 1.7.

[38] Appendix Figure A16 shows the estimates of earnings elasticities by MPC deciles.

For example, this equates the earnings elasticities of everyone in Detroit, allowing for the earnings of workers in Detroit to be more sensitive to recessions than those for workers in Austin. The estimate of the matching multiplier is very similar, suggesting substantial within-region heterogeneity in worker exposure. Row 2 shows the slightly more restrictive benchmark in which workers are given the average earnings elasticity of their industry and commuting zone. For example, this scenario equates the earnings elasticities of all workers in car manufacturing in Detroit but allows for different elasticities across industries within Detroit and across cities within car manufacturing. The matching multiplier falls only slightly, a finding that further illustrates the small role that the sorting of workers across industries plays in driving this pattern.[39]

The last panel of Table 1.3 explores the importance of the various assumptions on the importance of nonlabor income. In the first row, I re-estimate the matching multiplier under an alternative extreme assumption that the MPC out of nonlabor income is 0. In this case, the matching multiplier drops only slightly to 0.10, but this represents a larger percentage increase in the amplification because the baseline multiplier also falls. Rows 2 and 3 explore the magnitudes as I vary the labor share. Intuitively, the higher the labor share, the larger the role of this mechanism in the overall economy. Indeed, the effect of worker heterogeneity on shock amplification rises to 70 percent in the simple case where all output is earned by workers.

## 1.6 Commuting zone analysis

The national estimates discussed above quantify the importance of this worker heterogeneity using the structure of the framework in Section 1.2, and show that the magnitude

[39] Appendix Table A15 presents modified estimates of the matching multiplier that include the heterogeneity in earnings elasticities from the unemployed and the employed. This overall matching multiplier is a weighted average of the amplification coming from the employed, presented in Table 1.3, and the unemployed, where the weight is the fraction of overall earnings in each group. Because the new hires earn a small fraction of overall income, the addition of the unemployed has only a small effect on the overall estimate of the matching multiplier.

of the covariance between worker MPCs and earnings elasticities is large enough to imply sizable effects on aggregate fluctuations. In this section, I move beyond the model structure and empirically test the importance of the matching multiplier mechanism in determining the economy's response to shocks. Under the assumption that demand in commuting zones is largely locally determined, the model in Section 1.2 would predict that output in areas with a higher measured matching multiplier would be more sensitive to shocks than output in areas with a smaller matching multiplier. I test this prediction by utilizing the fine geographic detail in the LEHD to provide empirical estimates for how this unequal incidence of recessions affects the economy's response to shocks. If the mechanism were not important, whether because (a) consumption is not an amplification channel (b) this covariance term doesn't matter or (c) the assumptions underlying my empirical analysis do not hold, we should see no relationship between the local matching multiplier and the size of local recessions. However, if, as I hypothesize, the matching multiplier is a meaningful amplification channel, I should find that areas with a higher matching multiplier have more severe recessions and that this effect is concentrated in nontradable employment, where demand is locally determined.

These local empirical tests of the strength of this mechanism are important because they do not rely on the simplifying assumptions embedded in the framework. Indeed, the ways in which a higher matching multiplier translates into aggregate output and employment outcomes in a local economy may depend on general equilibrium factors at the local level that are beyond the scope of the model in Section 1.2. In particular, the multiplier is not a deep structural parameter of a region, and thus, the ways in which the covariance between worker MPCs and income sensitivities affects output responses is potentially a function of other features of the economy. For example, while monetary policy and other national policies are held constant across regions, local prices may adjust in response to changes in local demand, partially offsetting the matching mechanism, or local

fiscal authorities could respond to offset the amplification induced by this mechanism.[40] These local estimates will therefore include both the first-order effect of the matching multiplier on the local economy and any additional local general equilibrium responses to this mechanism.[41]

I explore the strength of this mechanism at the local labor market level by separately estimating the matching multiplier in each commuting zone.[42] My LEHD sample includes 270 commuting zones, and in each one of those areas, I separately estimate Equation 1.6 to get a CZ-specific estimate of the degree to which workers of different MPCs are differentially exposed to recessions (i.e., $\widehat{\alpha}_{2,c}$), which I use to calculate $\widehat{MM}_c$ in each commuting zone using a slight modification of Equation 1.4:

$$\widehat{MM}_c = \frac{MPC_c^a - MPC_c^b}{(1 - MPC_c^a)^2} = \frac{\sum_{i \in c} \frac{E_i}{E_c} MPC_i(\widehat{\alpha}_{2,c} MPC_i - \overline{\alpha})}{(1 - \sum_{i \in c} \frac{E_i}{E_c} MPC_i(\widehat{\alpha}_{2,c} MPC_i - \overline{\alpha}) - \overline{\gamma}\, \overline{MPC}_c)^2}$$

where $\overline{MPC}_c$ is the earnings-weighted average MPC of the area, $E_c$ is total earnings in the commuting zone, $\overline{\alpha} = \sum_{i \in c} \widehat{\alpha}_{2,c} MPC_i \frac{E_i}{E_c}$, and $\overline{\gamma}$ is the national average elasticity of earnings with respect to GDP. The true local matching multiplier would include $\overline{\gamma}_c$, the average elasticity of earnings in the area, in place of $\overline{\gamma}$, the national average elasticity of earnings. However, as this exercise explores the relationship between $\widehat{MM}_c$ and local

[40] While both of these mechanisms would downward bias the true effect of the matching multiplier and therefore make the local estimate a lower bound for the full effect of the multiplier, factor mobility across local labor markets could upward bias the result (Chodorow-Reich (2017)).

[41] The commuting zone estimates also mitigate the possibility that the assumptions embedded in the construction of the MPC estimates are driving the findings. For example, since I only allow MPCs to vary by demographic group, if the unobservably high-MPC workers within that demographic group all sort to the low cyclicality jobs while the unobservably low-MPC workers sort to the high cyclicality jobs, there could be no relationship at the individual level between income sensitivities and worker MPCs even though I detect one at the demographic-group level. While I do not find evidence for this within the PSID (see Table A9), a finding that local labor markets with larger estimated matching multipliers experience deeper recessions further dispels these concerns – if the mechanism were purely driven by mismeasurement, it would not be related to observed recession intensity.

[42] I restrict attention to commuting zones where the 23 states in the subsample cover at least 90 percent of employment in that commuting zone. Since many commuting zones cross state lines and include states not in my sample, this excludes an average 12 percent of workers and earnings in each sample year.

cyclicality, $\widehat{MM}_c$ should not be a direct function of the average cyclicality of earnings in the area. Therefore, I rescale all estimates of the local covariance not with their local $\overline{\gamma}_c$ but with the average $\overline{\gamma}$. There is substantial variation in this measure across local labor markets – the cross-commuting zone average covariance between worker MPCs and earnings elasticities is 0.076, but the standard deviation is also 0.075, implying a coefficient of variation of 1.[43] I also calculate the benchmark multiplier in each commuting zone (which I notate with $\widehat{B}_c$), which is the multiplier in each area in the case that the covariance between worker MPCs and earnings elasticities was 0. The total multiplier in a commuting zone would be given, to first order, by $\widehat{B}_c + \widehat{MM}_c$.

I explore the degree to which local markets with a higher matching multiplier are more sensitive to aggregate shocks by estimating variants on the following equation:

$$\Delta \log L_{c,t} = \phi_1 \widehat{MM}_c \times \Delta \log G_t + \phi_2 \widehat{B}_c \times \Delta \log G_t + X'\Phi + \delta_c + \delta_t + \epsilon_{c,t} \tag{1.8}$$

were $L_{c,t}$ is total employment in commuting zone $c$ in year $t$, $\widehat{MM}_c$ is the matching multiplier in commuting zone $c$, $B_C$ is the benchmark multiplier in the area, $\delta_c$ are commuting-zone fixed effects, $G_t$ is national GDP, and $X$ is a series of additional CZ-level controls. The year fixed effects $\delta_t$ absorb the average cyclicality of employment across commuting zones, and thus the coefficient of interest is $\phi_1$. The theory predicts that $\phi_1$ is positive – all else equal, areas with a higher matching multiplier should be more sensitive to aggregate shocks. The matching multiplier is an amplification of a baseline local consumption multiplier, and thus the theory also predicts that $\phi_2$ is positive, meaning that all else equal, areas with a benchmark multiplier should also more cyclical.

Causally identifying the relationship between $\widehat{MM}_c$ and local cyclicality is challenging, as the size of the matching multiplier locally is likely correlated with many other

[43] For computational efficiency, I estimate commuting-zone level estimates using aggregated binned regressions. Specifically, I round the MPC of each individual to the nearest 0.01 and aggregate all workers in the commuting zone to MPC bins. I then estimate Equation 1.6 on the binned data, weighting each bin by the share of earnings in that bin in year $t-1$.

features of the local labor market that also affect the cyclicality of the area. There are potentially many factors contributing to differences in the matching multiplier across commuting zones. The local industry mix could drive differences in the matching multiplier, since industries differ both in their sensitivities to recessions and in the average MPCs of their workers. Additionally, local labor markets may differ in the degree to which workers of different demographics are sorted across firms or in the within-firm incidence of shocks across workers.[44] The estimates of the local matching multipliers will encompass all of these factors. My first empirical strategy is to estimate Equation 1.8 using an array of controls for other local characteristics that would affect the cyclicality of the area and may be correlated with the local matching multiplier. I then build on this with a second approach that explores additional interactions that further isolate the particular local demand channel through which the matching multiplier should operate. In particular, I test the prediction that the differential effects on employment across regions should be concentrated in nontradable industries, which are subject to local demand, rather than tradable industries, which are subject to a more national demand.

### 1.6.1 Local recessions and the Matching Multiplier

The results of estimating Equation 1.8 are presented in Table 1.4. Before moving to the full estimation of Equation 1.8, Column 1 first reports a specification that includes only the total multiplier ($B_c + MM_c$) and does not separately estimate the contribution of the matching multiplier. I find that, as predicted, areas with a higher overall multiplier are more sensitive to business cycles. Column 2 then breaks apart the overall multiplier into the two components that are the focus of this paper – $MM_c$ and $B_c$. The positive coefficients on each of these variables show that both components of the multiplier contribute to the overall descriptive relationship between local cyclicality and the multiplier; across CZ,

[44] For example, some areas may have more cross-firm racial segregation or more assortative matching of high-wage workers to high-wage firms. Differences within the firm in the incidence of recessions could stem from differences in the power or scope of labor unions or from differences in fairness norms or managerial practices across regions (Bloom et al. (2017)).

having both a higher baseline multiplier and a higher matching multiplier is correlated with being more cyclical. Across the subsequent columns, I add a series of increasingly demanding CZ-level controls. The matching multiplier is a highly nonlinear function of the joint distribution of the firm and individual characteristics in an area, but the composition of firms and demographics could also affect the cyclicality of an area through several other mechanisms. For example, the industry distribution of an area obviously affects the response of employment to aggregate shocks, as could the age or racial composition of an area (Jaimovich and Siu (2009)). Therefore, in Column 3, I include controls for these local demographic variables and the industry structure of the area, each on their own and interacted with aggregate GDP.[45] When these controls are added, the coefficient on $\widehat{MM}_c$ drops by 70 percent, but remains large and statistically significant. Note that the estimate of $\phi_2$, which is the coefficient on the benchmark multiplier interacted with GDP, becomes small and statistically insignificant. Given the methodology for constructing MPCs, the average MPC of an area is strictly a function of the distribution of local demographics, and were I to be able to control flexibly enough for these demographics, this term would drop out of the regression entirely. Therefore, it is not surprising that after controlling for the demographics of the area, there is not enough variation left to identify the effect of the average MPC on local cyclicality. The matching multiplier, however, is a function not of the *level* of the controls but of the degree to which income changes are distributed across those groups, and therefore, these controls do not have the same effect on the matching

[45] Specifically, the demographic controls included are the share of CZ-level employment in each two-digit industry, the average worker age, percentage of workers who are black, average lagged worker incomes, the fraction of workers who are female, and the fraction of the area that is employed. I also include a control for the average cyclicality of the area ($\overline{\gamma}_c$), which I include alone and interacted with GDP. See Appendix Table A16 for a specification that includes each demographic characteristic of the area interacted with time dummies to completely nonparametrically control for heterogeneity along these dimensions. The results are similar, suggesting that allowing for a differential cycle and trend is sufficient to capture heterogeneity across CZ along this dimension. Appendix Figure A18 shows the year-by-year estimates of the relationship between $\widehat{MM}_c$ and employment changes, which underlie the coefficient on the interaction between GDP and $MM_c$ in Equation 1.8. These annual estimates show that areas with a higher matching multiplier experienced both larger employment losses during the Great Recession and greater employment gains during the boom years of the mid-2000s.

multiplier. In Column 4, I include additional controls for the size and age composition of the firms in the area, again both alone and interacted with GDP, and find that the coefficient is insensitive to these additional controls.

A growing body of literature has documented the important role that household debt plays in explaining geographic variation in the dynamics of consumption and employment over the business cycle. In particular, Mian, Rao, and Sufi (2013) show that there is a large elasticity of consumption to housing net worth, and Mian and Sufi (2014) show that the deterioration of household balance sheets played a significant role in explaining the cross-sectional patterns of employment declines during the Great Recession. In Column 5 of Table 1.4, I explore the possibility that the local matching multiplier is correlated with local household financial positions by including various controls for changes in the financial wealth of a commuting zone, both independently and interacted with aggregate GDP.[46] I find that coefficients on $\widehat{MM}_c$ are largely insensitive to these additional controls, showing that the relationship is not driven by a spurious correlation with local household wealth or debt. Lastly, in Column 6, I add a full set of state-by-year fixed effects, thus controlling for any differences in cyclicality that are common across states due to state-level economic policies and identifying the effect of $\widehat{MM}_c$ by comparing the experiences of commuting zones within the same state. I find that while the estimates are somewhat noisier, the coefficient on $\widehat{MM}_c$ remains stable.[47]

Across all columns, the magnitudes of the coefficients on the matching multiplier in Table 1.4 are substantial. The estimate in Column 3 implies that areas with the average matching multiplier have an elasticity to aggregate GDP that is 0.17 percentage points, or

[46] I follow Kaplan, Mitman, and Violante (2016) in defining housing and financial net worth in each commuting zone. Specifically, I include separately changes in housing and financial wealth per capita, changes in house prices, and the initial level of household debt per capita. I include each of these variables independently, allowing for different employment trends along each dimension, as well as interacted with aggregate GDP, allowing for different cyclicalities along each dimension as well.

[47] In this specification, I exclude commuting zones that include multiple states. Appendix Table A16 shows the robustness of the estimates in Table 1.4 to alternate functional forms of the multiplier. Specifically, I show that the patterns and magnitudes are similar when using only MPC differences rather than differences in the multiplier.

27 percent, higher than, an area with a matching multiplier of 0.[48] This estimate is similar to, although somewhat smaller than, the estimates of the national matching multiplier in Table 1.3. Additionally, a CZ that has a matching multiplier in the 90th percentile of the cross-CZ distribution has employment that is 0.33 percentage points more sensitive to GDP than an area with a matching multiplier that is in the 10th percentile, meaning that it is 1.5 times as sensitive to aggregate fluctuations.

All of the estimates in Table 1.4 explore the differential sensitivity of the commuting zones to the same aggregate shock. These estimates therefore include two dimensions of geographic heterogeneity – a first-order difference in the degree to which the local area is exposed to the aggregate shock and differences in the higher-order effect of the shock coming through the consumption response of households to the initial shock. An alternative approach is to use a standardized local labor demand shock, which eliminates potential differences in this first-order effect, and thus isolates the local amplification channel. Column 2 in Appendix Table A16 shows that the patterns are similar when using a local Bartik-style shock rather than aggregate GDP as the demand shock.[49] The magnitude of this estimate is also similarly large – a commuting zone with the average matching multiplier has an employment response that is 40 percent larger than an area with a matching multiplier of 0.

One feature of the above analysis is that both $\widehat{MM}_c$ itself and the relationship between $\widehat{MM}_c$ and employment cyclicality are estimated over the same sample period. Even though I purge $\widehat{MM}_c$ explicitly of the differences in the average cyclicality across areas, it may be that the nature of the particular set of shocks in the sample period jointly

[48] The average elasticity of employment in an area to GDP is 0.799. The coefficient in Column 1 of Table 1.4 implies that the elasticity of an area with $\widehat{MM}_c = 0$ is $0.799 - 0.199 * 0.853$ and that an area with the average $\widehat{MM}_c$ is 0.799. This implies that the average area has an elasticity that is $100 * \frac{0.199*0.853}{0.799-0.199*0.853} = 27$ percent higher.

[49] When I use this Bartik shock, I *both* re-estimate $\widehat{MM}_c$ using this shock and then re-estimate Equation 1.8 replacing aggregate GDP with this shock. This ensures that both the matching multiplier and the shock have the same scale. See Appendix A for further details on the construction and estimation of these local shocks.

generates a big recession in the area and a high matching multiplier.[50] I explore this possibility by re-estimating $\widehat{MM}_c$ in the pre-Great Recession period (i.e., 1995 to 2006) and look at the relationship between the change in employment in the Great Recession and the matching multiplier using the pre-recession $\widehat{MM}_c$ as an instrument for the full-sample $\widehat{MM}_c$ estimate.[51] The left panel of Appendix Figure A17 shows the relationship between the percentage change in employment from 2007 to 2010 and the matching multiplier estimated on the full sample, and the right panel shows the relationship with the full sample $\widehat{MM}_c$ instrumented with $\widehat{MM}_c$ estimated in the pre-recession sample. In both cases, there is an economically and statistically significant negative relationship between employment falls in the Great Recession and $\widehat{MM}_c$, but the slope is actually steeper when using the pre-recession estimates of $\widehat{MM}_c$ as an instrument. The finding that the relationship is robust to the separate estimation periods suggests that bias from the simultaneous estimation of $MM$ and employment responses is not driving the patterns.

### 1.6.2 Tradable industries and the Matching Multiplier

Since the matching multiplier affects the cyclicality of an area through its effect on local consumption, the relative employment effect of the local matching multiplier across regions should appear for nontradable industries, which are subject to local demand, rather than for tradable industries, which are subject to a more national demand. Note that this does not mean that the matching multiplier does not affect tradable employment overall – to the extent that the matching multiplier affects national demand for tradable goods, there should be an employment effect on tradable employment. The prediction is rather

[50] To fix ideas, consider two commuting zones that differ in the share of their employment in the construction industry, which has, on average, high-MPC workers. The Great Recession, which featured a particularly large shock to construction, would simultaneously generate a higher $\widehat{MM}_c$ in the area with more construction and cause that area to be more cyclical, resulting in a spurious relationship between $\widehat{MM}_c$ and local cyclicality. I control explicitly for the industry structure of the area, so industry specific shocks would not affect this, but other less obvious particularities of the Great Recession, such as local financial relationships, could cause similar problems in the joint estimation.

[51] While the estimates of pre-recession $\widehat{MM}_c$ are noisier, the correlation with full-sample $\widehat{MM}_c$ is high at 0.55 and the first stage relationship is strong.

that differences in the matching multiplier across regions should not affect the *relative* employment in tradable industries.

I explore the relative effect for tradable and nontradable industries by moving to the three-digit industry by CZ level and estimating a modified version of Equation 1.8

$$\Delta \log L_{i,c,t} = \phi_1 \widehat{MM}_c \times \Delta \log G_t \times T_i + \phi_2 B_c \times \Delta \log G_t \times T_i + X'\Phi + \delta_{ci} + \delta_{it} + \epsilon_{i,c,t} \quad (1.9)$$

where $L_{i,c,t}$ is total employment in industry $i$ in commuting zone $c$ at time $t$, and $T_i$ is an indicator for whether the industry is a tradable industry.[52] $\delta_{ci}$ are CZ-by-year fixed effects, which allow for each industry to have a different employment trend across commuting zones. $\delta_{it}$ are industry-by-year fixed effects, which flexibly control for any business cycle variation that is common to the industry. The sample only includes the set of industries that are classified as tradable or nontradable, and the parameter of interest is $\phi_1$, which captures the differential effect of the matching multiplier on the cyclicality of tradable industries relative to nontradable industries. Theory predicts that $\phi_1$ is negative, meaning that the matching multiplier affects the employment of tradable industries less than the employment of nontradable industries. Theory also predicts that $\phi_2$ is negative, meaning that the baseline multiplier also has a larger relative effect on employment in nontradable industries than it does in tradable industries. $X_{i,c,t-1}$ includes all terms supporting the key triple interaction, as well as various CZ-level controls, all of which have been suppressed for notational brevity.[53]

Table 1.5 reports the finding that the relationship between the matching multiplier and local employment cyclicality is concentrated in nontradable industries . Column 1 first re-

[52] Tradable industries are defined as in Mian and Sufi (2014). I use their "simple" definition in which retail and restaurants are nontradable and industries in global trade data are tradable. Each regression only includes the set of industries classified as either tradable or nontradable, and thus excludes those industries in the middle.

[53] The suppressed terms supporting the triple interaction are $\widehat{MM}_c \times \Delta \log G_t$ and $MPC_{t-1,c} \times \Delta \log G_t$ and $MPC_{t-1,c} \times T_i$. The other two-way interactions are absorbed by the fixed effects.

ports the overall relationship between the matching multiplier and local cyclicality using the industry-level specification, and shows a larger coefficient than in the aggregated CZ-level specification. Column 2 restricts only to the subsample of tradable and nontradable industries and shows that, on average, the overall relationship between local cyclicality and $\widehat{MM}_c$ is smaller for this subsample of industries. More importantly, Columns 3 and 4 estimate the regression separately for tradable and nontradable industries. Column 3 reports the relationship for employment in tradable industries, showing that the relationship is small, negative, and statistically insignificant. Conversely, Column 4 shows the relationship for employment in nontradable industries, where the relationship is positive, large, and statistically significant. A comparison of Columns 3 and 4 clearly shows that the overall effect of the matching multiplier on local cyclicality is concentrated in nontradable industries. For completeness, Column 5 reports the full joint estimation in Equation 1.9 and shows that the coefficients are very stable, and the coefficient on the triple interaction is almost exactly the difference between the main coefficient in Columns 3 and 4. Column 6 shows that magnitude of the coefficient is robust to a specification including a full set of CZ-by-year fixed effects, thus controlling for any employment movements common to industries within a commuting zone in a given year and identifying only the degree to which the relative movement of tradable and nontradable employment covaries with the matching multiplier.

While the fixed effects in Column 6 control for any factor in the area that affects all industries at a given time within a commuting zone, in Column 7, I also control for local financial positions, which also affect local demand and thus the relative employment of tradable and nontradable industries (Mian and Sufi (2014)). Importantly, I include these housing and wealth variables alone, interacted with GDP, and triple-interacted with a tradable indicator, allowing for the relative cyclicality of tradable and nontradable employment to vary with the financial wealth of an area. I find that the coefficient identifying the relative effect of $\widehat{MM}_c$ on tradable industries is unaffected by these additional

controls, ruling out the possibility that these patterns are driven by confounding differences in house price movements or debt positions.

Appendix Table A17 shows that these patterns are also robust to the redefinition of several of the key variables in Equation 1.9. Specifically, the pattern that the matching multiplier affects the employment cyclicality of nontradable industries more than tradable industries is robust to disaggregating to four-digit NAICS codes, which allows for an even more rigorous control for differences in industry experiences across commuting zones. It is also robust to using alternate functional forms for the multiplier or using different MPC imputations at the individual level. Together, these specifications show that the relative patterns across tradable and nontradable industries are very stable.

## 1.7 The Matching Multiplier in a calibrated model

The above national and local estimates demonstrate that the covariance between worker MPCs and earnings elasticities is large enough to have a meaningful effect on the aggregate response of the economy to shocks. While the empirical measure of the matching multiplier treats the MPCs of different demographic groups as fixed, in practice, consumption behavior will be determined as a function of current and future income shocks, and thus, will change as a function of the shock and its incidence. In this section, I turn to the structure of a model to relax this assumption and numerically show that the empirical statistic is a good approximation for the mechanism even in a more generalized setting.

Specifically, I explore this mechanism within a standard Bewley-Huggett-Aiyagari model augmented along three dimensions. First, I introduce endogenous labor supply and rich consumer heterogeneity. Second, I pair this model of aggregate demand with fixed wages in the short run, which capture the role that this mechanism plays in demand-driven amplifications. Third, I introduce an exogenous labor rationing process that generates labor income fluctuations in the presence of fixed wages. Within the context of this model, I clarify the conditions under which the matching multiplier that I use through-

out the empirical exercise is exactly the right sufficient statistic for capturing the general equilibrium effect of a demand shock. I then calibrate the demand side of the model and show that this statistic is a good approximation for the mechanism even as I relax those conditions.

### 1.7.1 Environment

The following setting is similar to the generalized setting in Auclert, Rognlie, and Straub (2018) and is a simplified version of the multisector model in Flynn, Patterson, and Sturm (2018). Since the focus of this exercise is to understand the role of heterogeneity in the demand side, I allow for rich heterogeneity among consumers and keep the supply side intentionally simple. Consider an economy in discrete time with $T$ periods. The economy is populated by a continuum of agents of $I$ types, where each type $I$ has a mass $\nu_i$ of individuals such that $\sum_i \nu_i = 1$. Households within each group are ex ante homogenous, but households face idiosyncratic risk in their productivity or labor supply $e(s)$, a process that may vary across demographic groups $I$. Households across groups may differ both in the income process they face and in their discount rates ($\beta_i$). Households have preferences over both consumption and leisure, and agents have the ability to borrow and save into a real asset ($a_{i,t}$) to smooth consumption but are subject to a borrowing constraint that $a_{i,t} \geq b$. The household's problem therefore is to choose paths for their consumption $c_{i,t}$ and labor supply $l_{i,t}$ to maximize their utility taking wages, prices, and the real interest rate as given:

$$\max_{c_{i,t}, l_{i,t}} \sum_{t=0}^{T} \beta_i^t E[u(c_{i,t}) - v(l_{i,t})] \tag{1.10}$$

subject to

$$w_t l_{i,t} e(s) + r_t a_{i,t-1} - \tau_{i,t} = p_t c_{i,t} + a_{i,t} \tag{1.11}$$

$$a_{i,t} \geq b \tag{1.12}$$

where Equation 1.11 is the agent's flow budget constraint, $\tau_{i,t}$ are lump sum taxes, and $p_t$ is the price of the final good at time $t$. I assume that $u(c) = \frac{c^{1-\omega^{-1}}}{1-\omega^{-1}}$ and $v(l) = \frac{l^{1+\psi^{-1}}}{1+\psi^{-1}}$, where $\omega$ is the intertemporal elasticity of substitution and $\psi$ is the Frisch labor supply elasticity.[54]

While the demand side of the model features rich heterogeneity, the supply side of the model is simple.[55] All workers are employed by a representative competitive firm, which produces with constant returns to scale technology and takes labor as the only input:

$$Y_t = L_t \ \forall \ t \tag{1.13}$$

Firm profit maximization implies that

$$w_t = p_t \ \ \forall \ t \tag{1.14}$$

The government sets potentially individual-specific lump sum taxes $\tau_{i,t}$ to finance government spending $G_t$. Rather than balance its budget strictly between periods, the government can issue bonds $B_t$ to smooth fluctuations across periods and therefore is subject to an intertemporal budget constraint:

[54] These functional forms are used for the quantitative exercise in this section. However, the results in this section apply to a broader set of preferences. See Auclert (2017) or Flynn, Patterson, and Sturm (2018) for a discussion using more general preferences. These preferences have the advantage that they guarantee that the resulting labor supply and Marshallian demand functions are continuous and differentiable in $r_t$. The assumed CRRA utility function also exhibits sufficient diminishing marginal utility of consumption to guarantee the existence of an equilibrium. See Appendix A for a discussion of the assumptions required to guarantee the existence of equilibrium.

[55] In a similar framework, Auclert, Rognlie, and Straub (2018) explore the importance of worker MPCs in a model with an enriched supply side that includes capital, sticky prices, and a Taylor rule for monetary policy. They show that while these modifications reduce the overall size of the multiplier, worker MPCs still remain crucial in determining the output response to fiscal policy. Specifically, they show that in a model that matches the empirical estimates of intertemporal MPCs and with deficit-financed spending, impact multipliers can be above 1, even with active monetary policy, distortionary taxation, and investment crowd out.

$$\sum_t \frac{\tau_{i,t}}{\prod_{i \leq t}(1+r_i)} = \sum_t \frac{p_t G_t}{\prod_{i \leq t}(1+r_i)} \tag{1.15}$$

I assume that government spending preferences are given exogenously by $\theta_G$, such that $G_t = G(r_t, \tau_t, \theta_G)$. Even though this fiscal rule is specified exogenously, government spending still responds to interest rate changes in order to maintain the budget constraint in Equation 1.15.

### 1.7.2 Equilibrium and the output multiplier

Consider first the case where all prices are fully flexible. The household problem in Equation 1.10 results in a demand for consumption and a labor supply function given by

$$c_{i,t} = c_i(\{\lambda_t\}_{t \in T}, \{\tau_{i,t}\}_{t \in T}, \beta_i, b) \tag{1.16}$$

$$l_{i,t} = l_i(\{\lambda_t\}_{t \in T}, \{\tau_{i,t}\}_{t \in T}, \beta_i, b) \tag{1.17}$$

where $\lambda_t = \{r_t, w_t, p_t\}$ is the vector of prices. Note that these are Marshallian demands, and thus these functions only depend directly on exogenous parameters $(\beta_i, b, \tau_{i,t})$ and prices. The goods market clearing condition is given by

$$Y_t = C_t + G_t = \sum_i \nu_i c_{i,t} + G_t \quad \forall\, t \tag{1.18}$$

and the labor market clearing condition is given by[56]

$$L_t = \sum_i \nu_i l_{i,t} e(s) \quad \forall\, t \tag{1.19}$$

An allocation of $\{c_{i,t}\ ,\ l_{i,t},\ \tau_{i,t}\ ,\ r_t\ ,\ p_t\ ,\ w_t\ ,\ G_t\ ,\ Y_t\}$ that satisfies Equations 1.14, 1.15, 1.16, 1.17, 1.18, and 1.19 characterizes the flexible price equilibrium.

[56] Asset market clearing ($\sum \nu_i a_{i,t} = B_t$) is implied by Walras' law.

The exercise in this paper will be to consider the response of an economy, initially at this flexible price equilibrium, to an unanticipated demand shock when wages are fixed for $k+1$ periods.[57] This assumption of fixed wages will isolate the importance of this mechanism in explaining demand-driven fluctuations. Equation 1.14 immediately implies that prices are also fixed over this period.[58] In this case, the interest rate does not adjust to clear the labor market, and thus, workers are off their labor supply curves for the first $k$ periods (i.e., in response to a negative shock, there are workers who would like to work more but cannot because a firm is not willing to hire them).[59] Rather, in those periods, labor supply is rationed, and a worker's labor supply is imposed exogenously as

$$l_{i,t} = n_{i,t}(Y_t) \tag{1.20}$$

such that $\sum_i \nu_i n_{i,t} = L_t$ .[60] This rationing function takes as inputs the full distribution across workers of their labor supply in the last period and aggregate change in output and returns labor supply $n_{i,t}$ for each individual. This function is what determines the change in the worker's earnings in response to an aggregate demand shock. This reduced form specification, similar to Werning (2015), captures the notation that, for example, in response to a negative demand shock, workers are not able to work as much as they would like. In order to capture the relationship between the exposure of worker earnings to aggregate shocks and worker MPCs documented in Section 1.4.2, I parametrize the rationing function $N(Y_t, l_{i,t-1}) = \{n_{i,t}\}$ as

[57] See Auclert (2017) for the case where wages are sticky indefinitely.

[58] This formulation of wage rigidity and corresponding price rigidity builds on that in Werning (2015).

[59] The only relevant price in this model is the real interest rate. When wages and prices are fixed for $k+1$ periods, the real interest rate is only fixed for $k$ periods, and thus, labor is only rationed in $k$ periods.

[60] A more complete alternative to this rationing function is to explicitly model heterogeneity in the labor market through search frictions. See Ravn and Sterk (2016) for a recent example. Additionally, while the resulting mechansim is similar, the endogenous redistribution mechansim here differs from that in Bilbiie (2008) and ?. In his setting, cyclical inequality comes from the redistribution of firm profits – the government taxes firm profits (held by the unconstrained agents) and rebates them lump-sum to the constrained agents. I allow for a rationing function that is reduced form but disciplined by labor market data.

$$n_{it} = \frac{Y_t}{L_t^*}\left(1 - \gamma\overline{MPC} + \gamma MPC_i\right) l_{it}^* \tag{1.21}$$

where $l_{i,t}^*$ is the amount that individual $i$ would like to work at the given wage, $L_t^* = \sum_i \nu_i l_{i,t}^*$, $\overline{MPC}$ is the earnings-weighted average MPC in the economy and $\gamma$ is the slope of this incidence function with respect to GDP.[61] This formulation of the rationing function implies that the elasticity of worker $i$'s earnings to the aggregate is linear in the worker's MPC and is given by $\gamma_i = 1 - \gamma\overline{MPC} + \gamma MPC_i$.

Since the interest rate is not pinned down by the labor market clearing condition, it must be set by monetary policy.[62] Assume for simplicity that the central bank targets fixed real interest rate[63]

$$r_t = \overline{r} \tag{1.22}$$

In this rationing equilibrium, consumer demand becomes

$$c_{i,t} = c_i(\{y_{i,t}\}_{t\leq k}, \{\lambda_t\}_{t\in T}, \{\tau_{i,t}\}_{t\in T}, \beta_i, b) \tag{1.23}$$

These are similar to the flexible price conditions (Equations 1.16 and 1.17), except that now they are a function of incomes in periods 1 through $k$, as these are now *exogenously* given by the rationing function.

[61] Specifically, in order for the rationing function to clear the market,

$$\overline{MPC} = \sum_i \frac{\nu_i e_{i,t} l_{i,t}^*}{L_t^*} MPC_i$$

[62] Note that the real interest rate, through the Fisher equation, is given by $r \approx i - \pi = i - \frac{p_t}{p_{t-1}}$. When wages are flexible, the real interest rate adjusts with the change in the real wage to clear the labor market. However, when wages and thus prices are fixed, inflation is 0 and the real interest rate can only move with the nominal interest rate.

[63] See Appendix A for a discussion of the multiplier with a more generalized monetary policy rule in which $r_t = r(Y_t)$. Since my focus is on quantifying the importance of heterogeneity in the labor market, I abstract from potential offsetting effects coming from countervailing monetary policy.

**Definition** *The rationing equilibrium is defined as the set of* $\{c_{i,t}\ ,\ l_{i,t},\ \tau_{i,t}\ ,\ r_t\ ,\ p_t\ ,\ w_t\ ,\ G_t\ ,\ Y_t\}$ *such that firms optimize as in Equation 1.14, consumers optimize consumption according to Equation 1.23 and supply labor according to* $N(Y_t, l_{i,t-1})$ *for periods* 1 *through* $k$ *and according to Equation 1.17 for* $t \geq k$*, and goods and labor markets clear in each period as in Equations 1.18 and 1.19.*

Using bold variables to represent vectors of aggregate variables (i.e. $\mathbf{Y} = \{Y_t\}$), I derive the response of the economy to shocks in Proposition 1. I define the partial equilibrium effect of the shock on output as the response of the economy to a shock to any of the parameters of the model, before accounting for any of the general equilibrium responses of the economy.[64] In this economy, this amounts to all effects on the economy *before* I allow for incomes or interest rates to change in response to the shock. Let $\partial\mathbf{Y}$ be the vector of the partial equilibrium change in output in each period. Define $C_Y$ to be a matrix where the $k, j$ entry is given by $\frac{dc_k}{dY_j} = \sum_i \frac{dc_i}{dy_j}\gamma_i \frac{l_{i,j}}{L_j}$, which is the aggregate response of consumption at time $k$ to income in time $j$.

**Proposition 1.** *Under the assumption that wages are sticky for* $k+1$ *periods, for any shock to parameters* $(\theta, \tau, \theta_G)$*, the total change in output from an initial flexible price allocation is given to first order by:*

$$d\mathbf{Y} = \left(I - \mathbf{C_Y} J_k - (\mathbf{C_r} + \mathbf{G_r}) J_{T-k} (\mathbf{L_r})^{-1}\right)^{-1} \partial\mathbf{Y} \tag{1.24}$$

*where subscripts denote partial derivatives (i.e.* $\mathbf{C_r}$ *is the partial derivative of consumption with respect to* $r$*), and* $J_k$ *and* $J_{T-k}$ *are diagonal matrices with 1s in the first* $k$ *or the last* $T-k$ *entries, respectively.*

The proof for the proposition can be found in Appendix A. The first term (i.e. $C_y J_k$) captures the heterogeneous agent intertemporal version of the traditional Keynesian mul-

[64] See the proof of Proposition 1 in Appendix A for a more detailed delineation of partial and general equilibrium effects.

tiplier and embeds the mechanism that is the key focus of this paper. This matrix of intertemporal MPCs features prominently in Auclert, Rognlie, and Straub (2018), who argue in a similar setting that these moments are essential for determining general equilibrium effects in heterogeneous agent models. Due to the forward-looking nature of the consumer's problem, what matters for the total consumption response today is not just the change in today's income but also the change in future period incomes, as those will further alter the consumer's consumption patterns today. The matrix $J_k$ simply captures the fact that wages are only fixed for some period; thus, the consumer only responds directly to income changes in those periods. The heterogeneous incidence of labor shocks directly affects the magnitude of $C_Y$ – when high-MPC workers have higher $\gamma_i$, the components of the $C_Y$ matrix are larger.

The second term (i.e. $\mathbf{C_r} + \mathbf{G_r})J_{T-k}(\mathbf{L_r})^{-1}$) captures the movements in the interest rate in periods after $k$. In those periods, the interest rate will adjust to bring workers back onto their labor supply curves and clear the labor market. Workers today anticipate the future market adjustment, and their consumption today will depend on the change in the interest rate in the future. Proposition 1 shows that this matrix is a sufficient statistic for characterizing the first-order general equilibrium effect on output of a demand shock.

### 1.7.3 Special limiting case: The empirical Matching Multiplier

Corollary 1 shows that the empirical matching multiplier derived in Section 1.2 is exactly the sufficient statistic in Proposition 1 in a special case of the model defined by three additional assumptions.

**Corollary 1.** *In the case where (1) wages are sticky for exactly two periods and (2) $C_r + G_r = 0$, for any small unanticipated and temporary shock to parameters $(\theta, \tau, \theta_G)$, the total change in output from an initial flexible price allocation is given in the first period to first order by:*

$$dY_1 = \frac{\partial Y_1}{1 - \sum_i \frac{y_{i,1}}{Y_1} MPC_i \gamma_i} \tag{1.25}$$

*Additionally, the difference in the output response between the actual case and the case where $\gamma_i = 1$ is given by:*

$$MM = \frac{1}{1 - \sum_i \frac{y_{i,1}}{Y_1} MPC_i \gamma_i} - \frac{1}{1 - \sum_i \frac{y_{i,1}}{Y_1} MPC_i} \tag{1.26}$$

The first assumption about the length of the fixed wage period governs the degree to which consumption in period 1 responds to future income changes. The assumption of price stickiness for at least two periods means that individual MPCs will directly affect aggregate demand as in Proposition 1.[65] The assumption that prices are sticky for *no more* than two periods breaks the intertemporal dependence of MPCs, as income is only exogenously assigned in the first period; thus, there is no direct response of consumption to income in future periods.[66] The second assumption governs the importance of the second term in the multiplier in Proposition 1, which captures the response of consumption today to changes in the interest rate in future nonrationed periods. While this movement is necessary to clear the market in those periods, Corollary 1 considers the case where this response limits to 0, meaning that consumption today responds negligibly to expected changes in future interest rates. Lastly, to get from the multiplier to the *matching multiplier* used in the empirics, worker MPCs must not be sensitive to the incidence of the shock, and thus the demand shock must be unanticipated and last for only one period. In this case, the period 1 response of consumption is given by the distribution of steady state marginal propensities to consume (i.e. $\frac{dc_{i,1}}{dy_{i,1}} = MPC_i$), and thus, the incidence of the shock will not affect the period-1 MPC of the individual. Certainly, the MPCs of individuals will change in future periods in response to the shock, but the period 1 MPC, which is the focus of the empirical exercise, will be unaffected.

[65] See Appendix A for a derivation of the multiplier in the case where prices and wages are fully flexible. In this case, the multiplier is not a direct function of covariance between worker MPCs and the incidence of recessions.

[66] In other words, this simplification means that the only non-0 column of the matrix $C_Y$ is the first one; thus, the first period can be solved in isolation.

This special case clarifies the relationship between the empirical formulation of the matching multiplier and the multiplier in a more generalized model. However, it also illustrates that this case may be very special. Wages and prices may, in reality, be fixed for multiple periods, reintroducing the importance of intertemporal MPCs. While it is unlikely that agents can anticipate the timing of recessions, they may be able to anticipate the incidence of recessions. In a dynamic setting, MPCs are themselves a function of a worker's earnings volatility; thus, if workers know they are in the benchmark scenario in which shocks will be distributed in proportion to earnings, they may respond by changing their asset accumulation in advance of the shock.

In the following, I calibrate the model to explore quantitatively how the strength of the matching multiplier mechanism depends on the simplifying assumptions that went into deriving the empirical formula. First, I explore the magnitude of this mechanism in a reasonably calibrated model and show that the first-order approximation used to derive Equation 1.24 well captures the overall effect. I then explore the sensitivity of the mechanism to two of the simplifying assumptions that map the general case to the empirical moment, and in doing so, I find that special case closely approximates, or even understates, this mechanism in the more general setting.

### 1.7.4 Model calibration

In order to match the empirical exercise, where I consider heterogeneity in MPCs and earnings sensitivities across both demographic and income groups, I calibrate an economy with eight demographic groups characterized by the combination of two genders, two education bins, and two race bins.[67] Within each demographic group, agents are ex ante homogenous, but across demographic groups, agents differ along several ex ante dimensions. First, I allow each demographic group to have a different income process. Using the Panel Study of Income Dynamics and following the methodology in Heathcote, Perri,

[67] I solve the consumer problem using the method of endogenous gridpoints from Carroll (2006).

and Violante (2010), I estimate the parameters of an AR(1) process for the log of income for each demographic group, which are the persistence parameter ($\rho$), the variance of the persistent shock to earnings ($\sigma^2_{persistent}$), the variance of the transitory shock ($\sigma^2_{transitory}$), and an average earnings level.[68] The resulting estimates are reported in Columns 2 through 5 of Table 1.6.[69] I pick up the well-documented differences in the average earnings by gender, race, and education, but I also find that black workers and women tend to have more volatile earnings than other demographic groups. All groups have earnings that are similarly (and highly) persistent.

For each group, I set the intertemporal elasticity of substitution to be 1.5, assume that agents cannot borrow (i.e., $b = 0$), and set the supply of assets such that the equilibrium interest rate is 1.02.[70] Taking the earnings process for each demographic group as given, I choose the discount rate that matches the earnings-weighted average MPC in the data for each demographic group.[71] Column 6 of Table 1.6 shows the resulting annual discount rates, which vary from 0.68 to 0.92. Generally, groups with higher MPCs and more volatile

[68] See Appendix A for details on the estimation of the income process as well as an exploration of the robustness of the estimates to alternate methods. In the baseline estimates, I restrict attention to the labor incomes of those individuals who report being unemployed or employed at the time of the PSID survey, are between the ages of 25 and 62, and for whom I can impute an MPC (i.e., observations with at least two lags of earnings). In order to include periods of 0 earnings, I use the following transformation

$$\log(y_i t) = \log(y_{it} + \overline{UI})$$

where $\overline{UI}$ is the average annual unemployment insurance payment reported in the PSID.

[69] These parameter estimates are similar to, although slightly higher than, comparable estimates in the literature. For example, Heathcote, Perri, and Violante (2010) estimate an average permanent component of around 0.015 and a transitory component of around 0.1. Carroll, Slacalek, and Tokuoka (2015) review the literature and show that estimates in the literature for the variance of the permanent component range from 0.01 to 0.054 and that estimates of the transitory variance range from 0.01 to 0.2.

[70] I follow Kaplan and Violante (2014) in choosing a value for the EIS above 1. While empirical studies using aggregate consumption data typically find very low values for the elasticity of intertemporal substitution, or EIS, (Hall (1988)), as is discussed in Bansal, Kiku, and Yaron (2012) and Kaplan and Violante (2014), this traditional approach may be downward-biased because of attenuation or endogeneity bias. Empirical studies that deal with this bias tend to find much larger estimates. For example, Gruber (2013) leverages exogenous variation in the after-tax interest rate using shifts in the tax rate and estimates an EIS of 2.

[71] The average MPC that I match in the model is the current-period consumption response to a current-period income shock. This object most closely maps to the empirical MPC estimates, which measure the contemporaneous consumption response to a realized income change. An alternate calibration is to match the average MPC rather than the weighted average MPC. That calibration produces similar results.

incomes have lower discount rates. Since agents are risk-averse, agents facing more income volatility want to accumulate assets, which pushes them away from their borrowing constraint and brings down the average MPC of that group. However, since it is precisely those groups that have higher measured MPCs, a higher degree of impatience is needed to match the MPC of the group in the model. While these discount rates are lower than is typically assumed in these calibrations, the estimates are actually in line with empirical estimates, both on average and in their patterns across demographics. For example, using experimental evidence in Denmark, Harrison, Lau, and Williams (2002) find annual discount factors of around 0.78, with higher estimates for the rich, skilled, and educated. Note also that the discount rates in the model need not reflect pure differences in time preferences across individuals and may also capture unmodeled differences in access to the banking sector of costs of borrowing.

Lastly, the key parameter to calibrate for quantifying the importance of this mechanism is $\gamma$ in the rationing function. Note these the earnings process for each demographic group will affect the steady state of the model, but the earnings process is distinct from the shape of the rationing function, which determines the degree to which the earnings of higher-MPC workers are more exposed to a demand shock. I take this parameter directly from Table 1.2 and set $\gamma = 1.3$. With this specification and calibration, Equation 1.21 means that the elasticity of each individual $i$ (i.e., $\gamma_i$) to the shock is given as a linear function of their MPC.

### 1.7.5 Model estimates and counterfactuals

In this section, I summarize the model-based estimates of the matching multiplier mechanism by comparing the change in consumption across two different incidences of the aggregate shock. In the "actual case," I consider an aggregate shock that is distributed such that the earnings elasticity of each worker is exactly equal to $\gamma_i$, which I calibrated to match the empirical incidence of recessions across the MPC distribution. In the "bench-

mark case," I consider an aggregate shock of the same size that is distributed such that the earnings elasticity of all workers is equal to 1. I measure the strength of this mechanism in the model by comparing the aggregate consumption response in period 1 in the actual case to the aggregate consumption response in period 1 in the benchmark case. In order to normalize this metric across settings, I divide by the change in consumption in the baseline case, and thus I report the percentage increase in the response of consumption to the shock that comes from the unequal incidence of the shock.[72]

I begin by calculating Equation 1.26 within the model, which is the model-based measure that most closely matches the statistic used in the empirical analysis. It is the sufficient statistic for the matching multiplier in the special case of the model, and it is simply the marginal propensity to consume of individuals in the steady state of the model averaged across different distributions of the shock. Column 1 of Figure 1.5 shows the estimate – according to the sufficient statistic, the consumption response to the aggregate shock is 13 percent higher with the empirical distribution of the shock than it would have been if the shock were evenly distributed. This sufficient statistic is most directly comparable to the empirical statistic used throughout the analysis.

In Column 2, I move away from the sufficient statistic approach and instead calculate the *actual* aggregate consumption response in the model to a one-time unexpected income shock. As before, I compare the aggregate consumption response in the actual and benchmark scenarios.[73] Since the sufficient statistic is derived to closely capture the overall response of the shock, the relative consumption response in Column 2 is reassuringly close to that implied by the sufficient statistic in Column 1.[74] Not only are the

[72] An alternative representation of this would be to plot the difference in the MPC across the two scenarios. See Table A19 for those statistics, as well as other metrics summarizing this mechanism.

[73] Specifically, beginning from the initial steady state of the model, I shock the income of each individual in period 1 such that aggregate income falls by 1 percent. In one case, the size of the income shock for each individual is calculated so that it induces a $\gamma_i$ percent change in the earnings of those workers. In the other case, the size of the individual income shock is instead calibrated so that it induces a uniform 1 percent change in the earnings of each worker.

[74] While I focus on the magnitude of this mechanism in the initial period, in Appendix Figure A19, I show the full dynamic path for the consumption differential in the simulated case. I find that the consumption boost

estimates close on average, but the left panel of Figure 1.6 shows that the sufficient statistic also captures the patterns across demographic groups. The horizontal axis shows the earnings-weighted average MPC for each demographic group in the steady state of the model. As is explained in Section 1.7.4, these are calibrated to match the MPCs for each demographic group in the PSID. On the vertical axis, the red squares show the aggregate MPC for each demographic group (i.e. $\frac{dC}{dY}$) that results from the uniform income shock. As the sufficient statistic formula suggests, the two measure are highly correlated.[75]

In the subsequent columns of Figure 1.5, I explore the sensitivity of the differential consumption response to two of the assumptions that defined the empirical simplification of the model described above. First, in Columns 3 and 4, I explore the importance of the assumption that the shock is completely unanticipated, meaning that as the incidence of the shock changes, workers MPCs in the initial period remain the same. However, if workers know they are in the benchmark scenario in which aggregate shocks are more evenly distributed, they are likely *also* facing a different overall income process, and thus when the aggregate shock eventually hits, they will have accumulated different assets and may have different MPCs. Therefore, the assumption that I can change the incidence of the shock without changing the distribution of MPCs may be incorrect. There is no formal relationship between the shape of the rationing function and income volatilities in the model, so I explore the importance of this assumption in a reduced form way. Specifically, I re-estimate the steady state of the model in a case where all demographic groups face the same transitory *and* persistent variance of income, which is given by the earnings-weighted average in the population. I then define the benchmark consumption change as the percent drop in consumption in response to the evenly distributed shock but begin-

coming from the empirical distribution of the shock is temporary, and in fact, in future periods, consumption in the benchmark case is higher. Since overall consumption is lower, the aggregate savings in response to the shock in the benchmark case are higher than in the empirical scenario, meaning that more assets are brought to future periods.

[75] The aggregate MPC and the calibrated MPC may differ from each other for two reasons. First, the individual MPC refers to marginal changes while the income shock that I consider is 1 percent of output, which is not marginal. Second, the sufficient statistic is derived from a first-order approximation and thus important higher order effects could cause these two measures to deviate.

ning from this new steady state. I call this benchmark consumption response the "endogenous benchmark" to refer to the endogenous response of initial MPCs to the incidence of the aggregate shock. The actual case stays the same and continues to be calculated beginning at the original steady state. Column 3 in Figure 1.5 shows the sufficient statistic using this endogenous benchmark scenario in the which workers stochastic income processes are equalized. This is directly comparable to the sufficient statistic in Column 1. Column 4 instead shows the relative response to an income-shock in the actual and the endogenous benchmark case, which is directly comparable to the estimate in Column 2. The results in both Column 3 and 4 show that this endogenous MPC response makes little difference for the strength of this mechanism. Furthermore, since equating the sensitivity of earnings to the aggregate across individuals does not fully equate the overall volatility of the income process across groups, this exercise likely provides an upper bound for the response of MPCs to this change.

The patterns displayed in the right panel of Figure 1.6 illuminate why the endogenous response of worker MPCs to the change in their income processes does not matter for the magnitude of the matching multiplier. The blue circles in Figure 1.6, on the y-axis, show the endogenous benchmark MPCs of each group, which characterize the steady state when the stochastic income process is equalized across groups. On the x-axis are the calibrated MPCs. Even though everyone faces the same income risk, the dispersion across demographic groups remains in the equal income process case because of remaining differences in discount factors across groups. While these baseline and endogenous MPCs are highly correlated, there are substantial within-group responses of consumption behavior to the income process. For example, the average MPC of black women with at least some college increases by 13 percent, as their asset accumulation falls in response to the drop in their income volatility in the endogenous scenario. However, in bringing the income volatilities to the average, other groups experience an increase in their income volatility, meaning that the increase in the MPC of black women is offset by a fall in in-

come in other demographic groups. Since all workers in the benchmark scenario face the same earnings cyclicality, the benchmark MPC reduces to the average elasticity of earnings to the shock times the earnings-weighted average MPC. Therefore, the endogenous response of MPCs to the changing income process only matters insofar as they affect aggregate earnings-weighted MPC. The green "x" in Figure 1.6 shows that the new average MPC lies very close to the 45-degree line – it is virtually unchanged across the scenarios.

In the next columns of Figure 1.5, I explore the importance of the length of the rationing period by reporting the strength of the mechanism as I increase the length of the shock. Specifically, in Columns 5 and 6, I continue to fix the interest rate but instead consider the response of the economy in time 1 when the shock lasts for two and three periods, respectively. Importantly, at time 1, agents know the length of the shock, and thus, their consumption response in the first period will depend on the expected length of the shock as they anticipate a future income change and may respond to that today. I find that as the length of the rationing period increases, so does the strength of the mechanism – when the shock is more persistent, the distribution of the shock is even more important. The intuition is that when the shock is persistent and unevenly distributed in both periods, the same workers who face a large shock today also anticipate facing a larger shock in the next period, and thus the drop in their consumption today is even larger. The heterogeneous exposure of the future shock amplifies the heterogeneity today, and thus, the overall strength of this mechanism increases.[76]

[76] The quantitative patterns here echo the analytical findings in Bilbiie (2018a), where precautionary savings motives interact with the cyclical inequality channel to further increase the amplification of shocks. Additionally, Appendix Figure A20 shows that with the extended shock, the amplification coming from the matching multiplier is also more persistent, although the bulk of the consumption amplification is still in the first period. Appendix Figure A21 shows a similar analysis using the sufficient statistics approach. Specifically, I estimate the full intertemporal MPC matrix in Proposition 1, and I plot the matching multiplier as I extend the rationing period. As in Figure 1.5, as the length of the rationing period increases, so too does the matching multiplier. The matching multiplier is 8 percent of the total multiplier when rationing lasts for one period, but it goes up above 30 percent when the length of the rationing period increases to 10 periods.

## 1.8 Conclusion

This paper explores the link between inequality in the labor market and macroeconomic stability. I demonstrate that the aggregate marginal propensity to consume is higher when there is a positive covariance between worker MPCs and the elasticity of their earnings to aggregate output, a mechanism I call the matching multiplier. Empirically, it is precisely the high-MPC workers whose earnings are most exposed to recessions, and this relationship is large enough to have meaningful effects on the response of output to shocks. This mechanism matters for the amplification of shocks – employment in areas with a higher matching multiplier are more sensitive to shocks, particularly in nontradable industries, where local demand channels should be especially strong. Lastly, I explore the importance of this mechanism in amplifying demand shocks in a more generalized theoretical setting and find that the empirical matching multiplier well captures the strength of the channel in a more generalized setting.

Uncovering the linkages between labor market inequality and the consumption multiplier has potentially important implications for macroeconomic stabilization policy. Indeed, policies can be made more effective in part by explicitly targeting this covariance between earnings heterogeneity and worker MPCs. For example, the government could consider this covariance when deciding how to target fiscal stimulus across industries or firms. While much of the covariance between MPCs and earnings sensitivities to GDP occurs within the firm, there is still some scope for targeting particular industries or firm types (e.g., young and small firms), where higher-MPC workers are more likely to be employed. Additionally, unemployment insurance that is targeted toward high-MPC workers could provide greater aggregate consumption stabilization benefits. Several countries, such as Germany, make unemployment insurance more generous for older workers, but the results in this paper suggest a rationale for making unemployment insurance benefits more generous for young workers, who have higher MPCs and more volatile earnings.

Lastly, these results may suggest another reason for policymakers to be alarmed by

rising inequality in the economy. As wealth becomes more unequally distributed, MPCs in the population may become more dispersed, with a wide swath of consumption being greatly affected by aggregate shocks. A concurrent economic phenomenon of the past decade is that workers have become increasingly sorted across firms, occupations, and even types of employment contracts. These two economic forces could combine to further strengthen this mechanism and contribute to larger recessions in the future.

# Bibliography


Abowd, John M., Paul A. Lengermann, and Kevin L. McKinney "The Measurement of Human Capital in the U.S. Economy," Technical Paper 2002-09, U.S. Census Bureau, 2003.

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon Woodcock *The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators*, 149–230: University of Chicago Press, 2009.

Altonji, Joseph G. and Lewis M. Segal "Small-Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business & Economic Statistics*, 1996, 14 (3) , 353–366.

Attanasio, Orazio and Luigi Pistaferri "Consumption Inequality over the Last Half Century: Some Evidence Using the New PSID Consumption Measure," *American Economic Review*, 2014, 104 (5) , 122–26.

Auclert, Adrien "Monetary Policy and the Redistribution Channel," Working Paper 23451, National Bureau of Economic Research, 2017.

Auclert, Adrien, Matthew Rognlie, and Ludwig Straub "The Intertemporal Keynesian Cross," Working Paper 25020, National Bureau of Economic Research, 2018.

Bailey, Elizabeth E. "Innovation and regulation," *Journal of Public Economics*, 1974, 3 (3) , 285–295.

Bansal, Ravi, Dana Kiku, and Amir Yaron "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 2012, 1 (1) , 183–221.

Barth, Erling, Alex Bryson, James C. Davis, and Richard Freeman "It's Where You Work: Increases in the Dispersion of Earnings across Establishments and Individuals in the United States," *Journal of Labor Economics*, 2016, 34 (S2) , S67–S97.

Bilbiie, Florin "Limited Asset Markets Participation, Monetary Policy, and (Inverted) Aggregate Demand Logic," *Journal of Economic Theory*, 2008, 140 , 162–196.

——— "Hetergeneity and Determinacy: Amplification without Puzzles," *Working paper*, 2018a.

——— "The New Keynesian Cross," *Working Paper*, 2018b.

Bloom, Nicholas, Erik Brynjolfsson, Lucia Foster, Ron S Jarmin, Megha Patnaik, Itay Saporta-Eksten, and John Van Reenen "What Drives Differences in Management?" Working Paper 23300, National Bureau of Economic Research, 2017.

Blundell, Richard, Luigi Pistaferri, and Ian Preston "Consumption Inequality and Partial Insurance," *American Economic Review*, 2008, 98 (5) , 1887–1921.

Bonaparte, Yosef and Frank J. Fabozzi "Is Food Consumption a Good Proxy for Nondurable Consumption?" *Economics Letters*, 2011, 111 .

Carroll, Christopher D. "A Theory of the Consumption Function, with and without Liquidity Constraints," *Journal of Economic Perspectives*, 2001, 15 (3) , 23–45.

——— "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," *Economics Letters*, 2006, 91 (3) , 312 – 320.

Carroll, Christopher D., Jiri Slacalek, and Kiichi Tokuoka "Buffer-stock saving in a Krusell–Smith world," *Economics Letters*, 2015, 132 , 97 – 100.

Carroll, Christopher, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White "The Distribution of Wealth and the Marginal Propensity to Consume," *Quantitative Economics*, 2017, 8 (3) , 977–1020.

Chodorow-Reich, Gabriel "Geographic Cross-Sectional Fiscal Spending Multipliers: What Have We Learned?" Working Paper 23577, National Bureau of Economic Research, 2017.

Coibion, Olivier, Yuriy Gorodnichenko, Lorenz Kueng, and John Silvia "Innocent Bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, 2017, 88 (C) , 70–89.

Crouzet, Nicolas and Neil R. Mehrotra "Small and Large Firms over the Business Cycle," Working Paper 741, Federal Reserve Bank of Minneapolis, 2017.

Deaton, Angus "Saving and Liquidity Constraints," *Econometrica*, 1991, 59 (5) , 1221–1248.

Dettling, Lisa J., Joanne W. Hsu, Lindsay Jacobs, Kevin B. Moore, and Jeffrey P. Thompson "Recent Trends in Wealth-Holding by Race and Ethnicity: Evidence from the Survey of Consumer Finances,"Technical report, 2017.

Di Maggio, Marco, Amir Kermani, and Kaveh Majlesi "Stock Market Returns and Consumption," Working Paper 24262, National Bureau of Economic Research, 2018.

Fagereng, Andreas, Martin B. Holm, and Gisle J. Natvik "MPC Heterogeneity and Household Balance Sheets," Working Paper 7134, 2018.

Farber, Hank "Mobility and stability: The dynamics of job change in labor markets," *Handbook of Labor Economics*, 1999, 3 (B) , 2439–2483.

Flynn, Joel, Christina Patterson, and John Sturm "Shock Propagation and Optimal Policy in a Heterogenous-Agent Multi-Good Economy," *Mimeo*, 2018.

Fort, Teresa C., John Haltiwanger, Ron S. Jarmin, and Javier Miranda "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size," *IMF Economic Review*, 2013, 61 (3) , 520–559.

Galí, Jordi, J. David López-Salido, and Javier Vallés "Understanding the Effects of Government Spending on Consumption," *Journal of the European Economic Association*, 2007, 5 (1) , 227–270.

Ganong, Peter and Pascal Noel "How Does Unemployment Affect Consumer Spending?"Technical report, 2017.

Gelman, Michael "What Drives Heterogeneity in the Marginal Propensity to Consume? Temporary Shocks vs Persistent Characteristics," *Mimeo*, 2016.

Glover, Andrew, Jonathan Heathcote, Dirk Krueger, and José-Víctor Ríos-Rull "Intergenerational Redistribution in the Great Recession," Working Paper 16924, National Bureau of Economic Research, 2011.

Gross, Tal, Matthew J. Notowidigdo, and Jialan Wang "The Marginal Propensity to Consume Over the Business Cycle," Working Paper 22518, National Bureau of Economic Research, 2017.

Gruber, Jonathan "The Consumption Smoothing Benefits of Unemployment Insurance," *American Economic Review*, 1997, 87 (1) , 192–205.

——— "Unemployment Insurance, Consumption Smoothing, and Private Insurance: Evidence from the PSID and CEX," *Research in Employment Policy*, 1998, 1 , 3–32.

——— "A Tax-Based Estimate of the Elasticity of Intertemporal Substitution," *Quarterly Journal of Finance*, 2013, 3 (1) .

Guvenen, Fatih, Serdar Ozkan, and Jae Song "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 2014, 122 (3) , 621–660.

Guvenen, Fatih, Sam Schulhofer-Wohl, Jae Song, and Motohiro Yogo "Worker Betas: Five Facts about Systemic Earnings Risk," *American Econommic Review: Papers and Proceedings*, 2017, 107 (5) , 398–403.

Guvenen, Fatih and Anthony A. Smith "Inferring Labor Income Risk and Partial Insurance from Economic Choices," *Econometrica*, 2014, 82 (6) , 2085–2129.

Hall, Robert E. "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 1988, 96 (2) , 339–357.

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams "Estimating Individual Discount Rates in Denmark: A Field Experiment," *American Economic Review*, 2002, 92 (5) , 1606–1617.

Hastings, Justine S. and Jesse M. Shapiro "How Are SNAP Benefits Spent? Evidence from a Retail Panel," Working Paper 23112, National Bureau of Economic Research, 2017.

Heathcote, Jonathan and Fabrizio Perri "Wealth and Volatility," *The Review of Economic Studies*, 2018, 85 (4) , 2173–2213.

Heathcote, Jonathan, Fabrizio Perri, and Giovanni L. Violante "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006," *Review of Economic Dynamics*, 2010, 13 (1) , 15–51.

Hendren, Nathaniel "Knowledge of Future Job Loss and Implications for Unemployment Insurance," *American Economic Review*, 2017, 107 (7) , 1778–1823.

Herkenhoff, Kyle, Gordon Phillips, and Ethan Cohen-Cole "How Credit Constraints Impact Job Finding Rates, Sorting & Aggregate Output," Working Paper 22274, National Bureau of Economic Research, 2016.

Hoynes, Hilary, Douglas L. Miller, and Jessamyn Schaller "Who Suffers During Recessions?" *The Journal of Economic Perspectives*, 2012, 26 (3) , 27–48.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan "Earnings Losses of Displaced Workers," *The American Economic Review*, 1993, 83 (4) , 685–709.

Jaimovich, Nir and Henry E. Siu "The Young, the Old, and the Restless: Demographics and Business Cycle Volatility," *American Economic Review*, 2009, 99 (3) , 804–26.

Jappelli, Tullio and Luigi Pistaferri "The Consumption Response to Income Changes," *Annual Review of Economics*, 2010, 2 (1) , 479–506.

——— "Fiscal Policy and MPC Heterogeneity," *American Economic Journal: Macroeconomics*, 2014, 6 (4) , 107–36.

Jarosch, Gregor "Searching for Job Security and the Consequences of Job Loss," 2014.

Johnson, David S., Jonathan A. Parker, and Nicholas S. Souleles "Household Expenditure and the Income Tax Rebates of 2001," *American Economic Review*, 2006, 96 (5) , 1589–1610.

Kaplan, Greg, Kurt Mitman, and Giovanni L Violante "Non-durable Consumption and Housing Net Worth in the Great Recession: Evidence from Easily Accessible Data," Working Paper 22232, National Bureau of Economic Research, 2016.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante "Monetary Policy According to HANK," *American Economic Review*, 2018, 108 (3) , 697–743.

Kaplan, Greg and Giovanni L. Violante "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 2014, 82 (4) , 1199–1239.

Kaplan, Greg, Giovanni L. Violante, and Justin. Weidner "The Wealthy Hand-to-Mouth," *Brookings Papers on Economic Activity*, 2014, 2014 (1) , 77–138.

Krusell, Per and Anthony A. Smith "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 1998, 106 (5) , 867–896.

McKay, Alisdair "Time-varying idiosyncratic risk and aggregate consumption dynamics," *Journal of Monetary Economics*, 2017, 88 , 1 – 14.

McKay, Alisdair, Emi Nakamura, and Jón Steinsson "The Power of Forward Guidance Revisited," *American Economic Review*, 2016, 106 (10) , 3133–58.

McKay, Alisdair and Ricardo Reis "The Role of Automatic Stabilizers in the U.S. Business Cycle," *Econometrica*, 2016, 84 , 141–194.

McKee, Graham and Emil Verner "The Consumption Response to Extended Unemployment Benefits in the Great Recession," Working Paper 1-056, Kilts Booth Marketing, 2015.

Mian, Atif, Kamalesh Rao, and Amir Sufi "Household Balance Sheets, Consumption, and the Economic Slump," *Quarterly Journal of Economics*, 2013, 128 (4) , 1687–1726.

Mian, Atif and Amir Sufi "What Explains the 2007–2009 Drop in Employment?" *Econometrica*, 2014, 82 (6) , 2197–2223.

——— "Who Bears The Cost Of Recessions? The Role Of House Prices And Household Data," *Handbook of Macroeconomics*, 2016.

Oh, Hyunseung and Ricardo Reis "Targeted transfers and the fiscal response to the great recession," *Journal of Monetary Economics*, 2012, 59 (S) , 50–64.

Ouimet, Paige and Rebecca Zarutskie "Who works for startups? The relation between firm age, employee age, and growth," *Journal of Financial Economics*, 2014, 112 (3) , 386 – 407.

Parker, Jonathan A. "Why Don't Households Smooth Consumption? Evidence from a $25 Million Experiment," *American Economic Journal: Macroeconomics*, 2017, 9 (4) , 153–83.

Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson, and Robert McClelland "Consumer Spending and the Economic Stimulus Payments of 2008," *American Economic Review*, 2013, 103 (6) , 2530–53.

Ravn, Morten O. and Vincent Sterk "Macroeconomic Fluctuations with HANK & SAM: An Analytical Approach," discussion papers, Centre for Macroeconomics (CFM), 2016.

——— "Job uncertainty and deep recessions," *Journal of Monetary Economics*, 2017, 90 , 125 – 141.

Rubin, D. B. *Multiple Imputation for Nonresponse in Surveys*: J. Wiley & Sons, 1987.

Schulhofer-Wohl, Sam "Heterogeneity and Tests of Risk Sharing," *Journal of Political Economy*, 2011, 119 (5) , 925–958.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter "Firming Up Inequality," Working Paper 21199, National Bureau of Economic Research, 2015.

Sorkin, Isaac "Ranking Firms Using Revealed Preference," *The Quarterly Journal of Economics*, 2018, 133 (3) , 1331–1393.

Spletzer, James R. "Inequality Statistics from the LEHD," *CES Working Paper Series*, 2014.

Stevens, Ann Huff "Persistent Effects of Job Displacement: The Importance of Multiple Job Losses," *Journal of Labor Economics*, 1997, 15 (1) , 165–188.

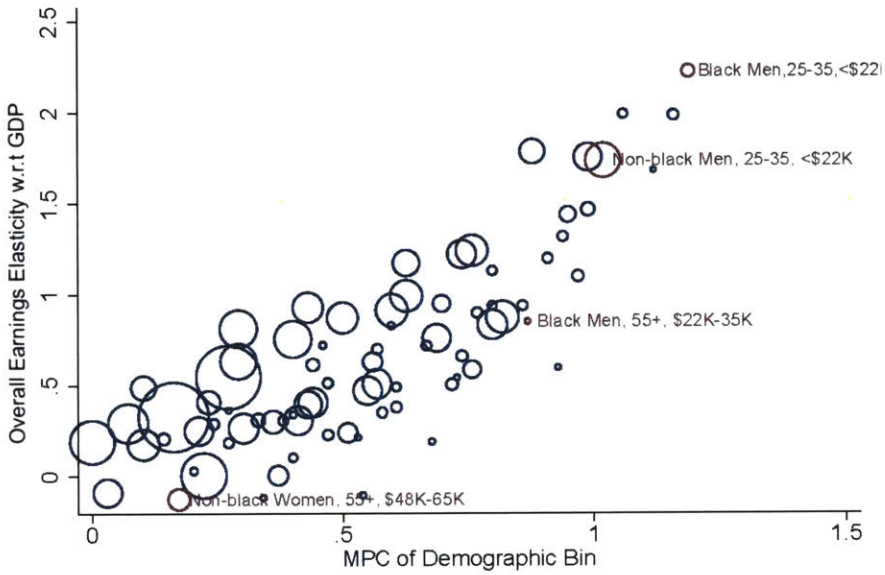
Vilhuber, Lars and Kevin McKinney "LEHD Infrastructure Files in the Census RDC," working papers, Centre for Economic Studies (CES), 2014.

Werning, Iván "Incomplete Markets and Aggregate Demand," Working Paper 21448, National Bureau of Economic Research, 2015.

Wiswall, Matthew and Basit Zafar "Preference for the Workplace, Investment in Human Capital, and Gender," *The Quarterly Journal of Economics*, 2018, 133 (1) , 457–507.

# Figures and Tables

Figure 1.1: Recession Exposure and MPC by Demographic Group



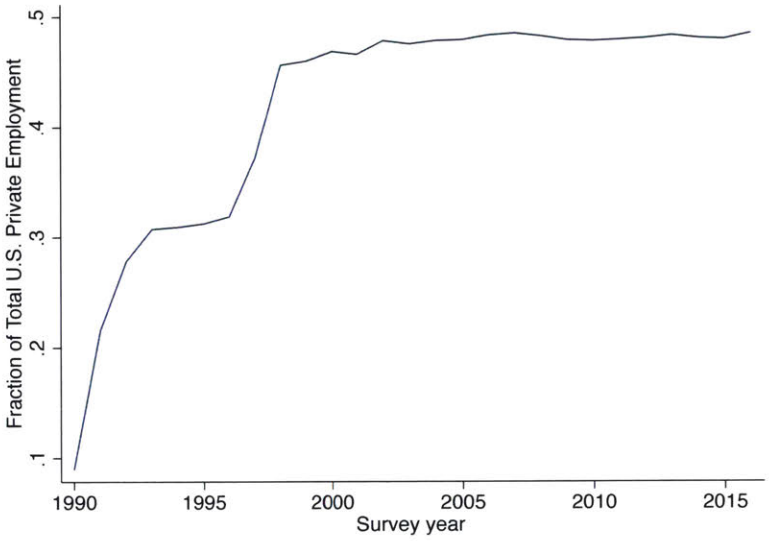

Notes: Sample includes the set of all workers employed in a sample state in year $t-1$ from 1995 to 2011. The dependent variable in the regression producing the y-axis estimates is the total change in log earnings for the demographic group. The size of each bubble represents the earnings share of that demographic group. The coefficient on the fitted line for this plot is 1.33. Appendix Figure A14 shows the corresponding figure separately for the intensive and extensive margin of earnings.

Table 1.1: Summary Statistics for LEHD Sample

| | 1995 Snapshot | 2011 Snapshot | Full Sample |
|---|---|---|---|
| Number of workers | 22,680,000 | 45,380,000 | 38,078,235 |
| Number of establishments | 1,015,000 | 3,464,000 | 2,772,529 |
| ***Worker Characteristics*** | | | |
| Fraction Male | 0.53 | 0.51 | 0.52 |
| Average Worker Age | 40.61 | 43.05 | 42.10 |
| Average 2-year Lagged Income | 36,600 | 42,210 | 40,373 |
| Fraction Black | 0.09 | 0.11 | 0.10 |
| Fraction College Educated | 0.30 | 0.28 | 0.29 |
| ***Job Characteristics*** | | | |
| Average Total Quarterly Earnings | 10,650 | 11,920 | 11,794 |
| Average Annual Earnings | 40,140 | 45,430 | 44,119 |
| Average No. Jobs per quarter per worker | 1.16 | 1.14 | 1.15 |
| Fraction with multiple jobs | 0.13 | 0.11 | 0.12 |

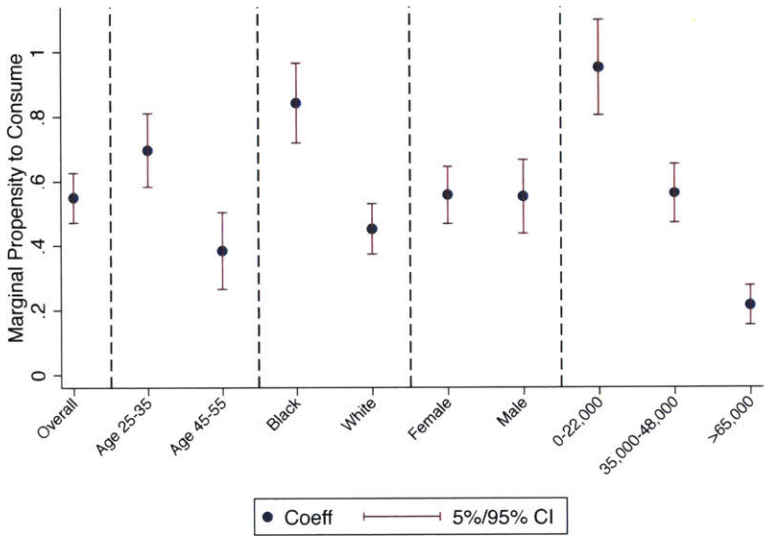
Notes: Sample includes all individuals in the baseline sample. Averages are unweighted, and nominal values are expressed in 2010 dollars. Column 1 shows the data in 1995, Column 2 shows the data in 2011, and Column 3 shows the sample averaged from 1995 to 2011. Counts for the number of workers and the number of establishments are rounded to comply with U.S. Census disclosure requirements.

Figure 1.2: Fraction of U.S. Employment in LEHD Sample



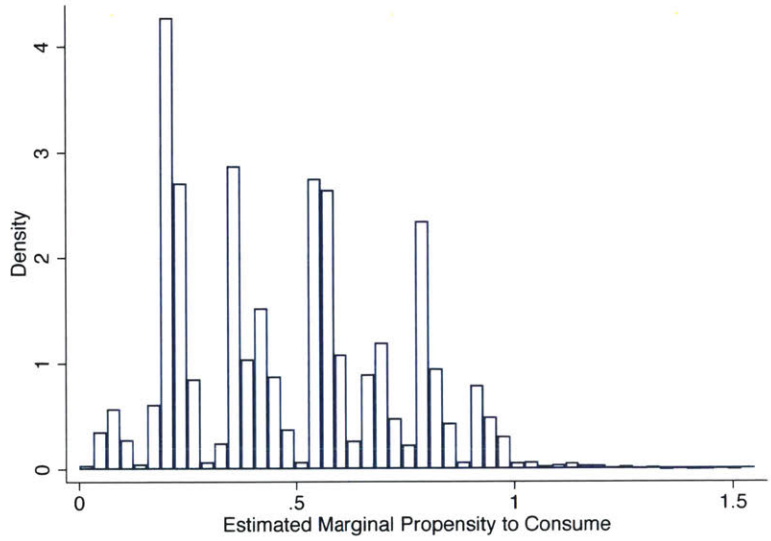

Notes: Figure is constructed using data from the Current Population Survey Annual Social and Economic Supplement. States included in the LEHD sample are Arkansas; Arizona; California; Colorado; Washington D.C.; Delaware; Florida; Iowa; Illinois; Indiana; Kansas; Maryland; Maine; Montana; New Mexico; Nevada; Oklahoma; Oregon; Pennsylvania; Tennessee; Washington; and West Virginia. See Table A1 for the sample years for each state.

Figure 1.3: Heterogeneity in Marginal Propensity to Consume Estimates



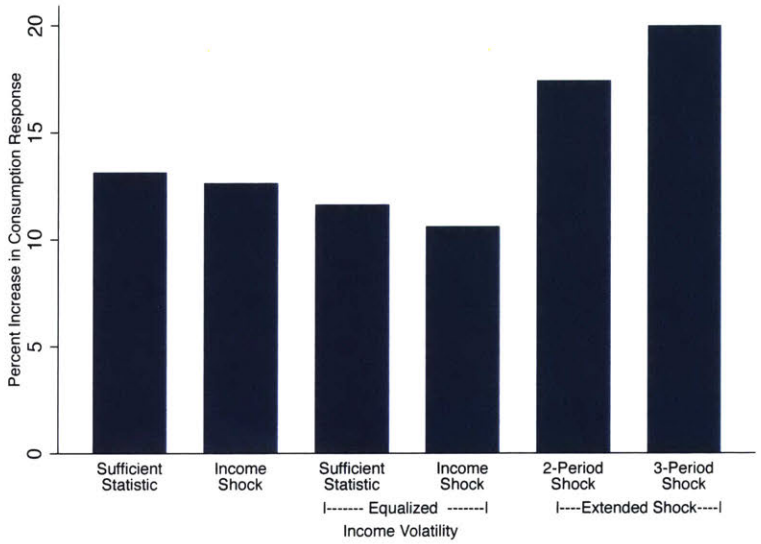

Notes: Each estimate represents a separate regression including only the stated demographic group. In all cases, consumption is measured using total consumption, imputed using the method in Blundell, Pistaferri, and Preston (2008). Income is measured using individual labor income. The instrument for income changes is unemployment. The sample includes the set of workers who were employed two years before the current year. The sample in the PSID excludes observations with more than a 400 percent change in food consumption or income over a given two-year period. Lagged income is measured as the average labor marker earnings of the individual in $t-2$ and $t-3$. All regressions include year-by-state fixed effects and observations from 1992 to 2013. Appendix Figure A5 shows the reduced form and first stage for these regressions.

Figure 1.4: The Distribution of Estimated Marginal Propensities to Consume



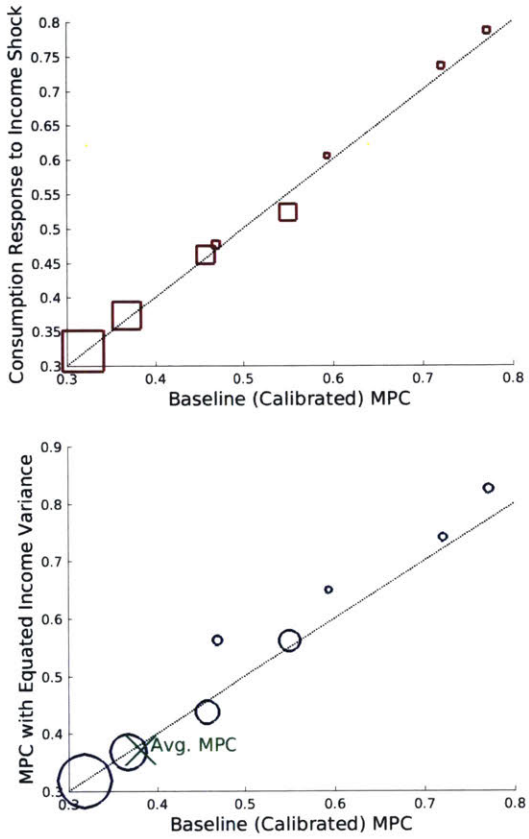

Notes: See Appendix Table A6 for the coefficients that underlie this imputation. Negative imputed marginal propensities to consume are set to 0. Consumption is measured using total consumption, imputed using the method in Blundell, Pistaferri, and Preston (2008). Income is measured using individual labor income. The instrument for income changes is unemployment. The sample includes the set of workers who were employed two years before the current year. The sample in the PSID excludes observations with more than a 400 percent change in food consumption or income over a given two-year period. Lagged income is measured as the average labor marker earnings of the individual in $t-2$ and $t-3$. Regression includes year-by-state fixed effects and observations from 1992 to 2013.

Figure 1.5: Relative Consumption Response in Various Scenarios




Notes: The y-axis plots the difference between the change in consumption in period 1 in the actual scenario, in which the aggregate shock is distributed according to the empirical distribution, and the change in consumption in period 1 in the benchmark scenario, in which the aggregate shock is distributed evenly, normalized by the change in consumption in the benchmark scenario. Column 1 uses the sufficient statistic formula in Equation 1.25, and Column 2 shows the result from the one-period income shock. Columns 3 and 4 shows the results using an alternate benchmark in which workers anticipate the benchmark income volatility. Columns 5 and 6 show the result for a persistent two- and three-period shock, respectively. Lastly, Column 6 allows the interest rate to adjust in future periods to clear the market. See Appendix Table A19 for the level changes in consumption and the corresponding MPC levels underlying these estimates and Appendix Figure A19 for an extended time series path of the experiment in Column 2.

Figure 1.6: Demographic Group MPCs: Calibration vs. Experiment




Notes: In the left panel, the y-axis plots the total change in consumption in the demographic group when each group is hit by a uniform negative 1 percent shock, divided by the total change in earnings of that group. In the right panel, the y-axis plots the average group-level MPC in the case where the variance of earnings is equated across groups. In both panels, the x-axis plots the steady state MPC of each group, which is calibrated to match the empirical estimates from the PSID. The 45-degree line is marked by the black dashed line. Each bubble represents one of the eight demographic groups, and the size of the bubble reflects that group's share of total earnings.

Table 1.2: Earnings Elasticity to GDP by a Worker's Marginal Propensity to Consume

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | | Earnings Margin | | Partialled Out MPC Components ($X_{i,t-1}$) | | | | |
| | Overall | Intensive | Extensive | Gender | Race | Income | Age | All Together |
| $MPC_{i,t-1}$ | -0.245 | -0.061 | -0.110 | -0.244 | -0.245 | -0.136 | -0.290 | -0.103 |
| | (0.001) | (0.001) | (0.000) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) |
| $MPC_{i,t-1} * \Delta GDP_t$ | 1.300 | 0.586 | 0.482 | 1.257 | 1.324 | 1.662 | 1.310 | 1.140 |
| | (0.028) | (0.025) | (0.012) | (0.028) | (0.028) | (0.047) | (0.028) | (0.053) |
| $X_{i,t-1}$ | | | | 0.008 | 0.008 | 0.033 | -0.006 | |
| | | | | (0.000) | (0.000) | (0.000) | (0.000) | |
| $X_{i,t-1} * \Delta GDP_t$ | | | | -0.378 | -0.131 | 0.163 | 0.004 | |
| | | | | (0.015) | (0.027) | (0.016) | (0.001) | |
| No. Observations (Million) | 29.20 | 25.50 | 29.20 | 29.20 | 29.20 | 28.46 | 29.20 | 28.46 |
| R-Squared | 0.009 | 0.002 | 0.010 | 0.009 | 0.009 | 0.009 | 0.017 | 0.018 |
| Avg. MPC | .431 | .423 | .431 | .431 | .431 | .431 | .431 | .431 |

Notes: Each regression is estimated on a 5 percent random subsample. All observations are weighed by the individual's earnings in $t-1$. Earnings in each regression are defined as total quarterly earnings for each individual in the fourth-quarter of the year. The number of observations is rounded to the nearest 100 to comply with U.S. Census disclosure requirements. Standard errors are clustered at the individual level. The outcome variable in Columns 1 and 4 through 9 is $\Delta E_{i,t} = \frac{E_{i,t}-E_{i,t-1}}{.5\times E_{i,t}+.5\times E_{i,t-1}}$. The outcome variable in Column 2 is the change in the log of fourth-quarter earnings, and the outcome variable in Column 3 is an indicator for being employed in time $t$. Columns 4 through 8 add the noted demographic characteristic $x_{i,t-1}$, both independently and interacted with GDP. Gender is a dummy for whether the worker is female, race is a dummy for whether the worker is black, and income is the log of a worker's earnings in the time $t-2$ and $t-3$. See Appendix Tables A10 and A11 for alternate specifications of these estimates.

Table 1.3: National Estimates of the Matching Multiplier

| | $MPC^b$ | $MPC^a$ | Pct. Increase in MPC | Actual Multiplier | Benchmark Multiplier | Matching Multiplier | Pct. Increase in Amplification |
|---|---|---|---|---|---|---|---|
| *Panel A: Alternate Specification* | | | | | | | |
| Linear | 0.23 | 0.29 | 28% | 1.42 | 1.30 | 0.13 | 42.62% |
| Demographic Group | 0.25 | 0.31 | 25% | 1.46 | 1.33 | 0.13 | 40.05% |
| Decile | 0.28 | 0.35 | 25% | 1.54 | 1.39 | 0.17 | 42.64% |
| *Panel B: Alternate Benchmarks* | | | | | | | |
| Within Commuting Zone | 0.23 | 0.29 | 26% | 1.42 | 1.30 | 0.12 | 40.34% |
| Within Industry and Commuting Zone | 0.24 | 0.29 | 22% | 1.42 | 1.32 | 0.10 | 33.04% |
| *Panel C: Alternate Benchmarks* | | | | | | | |
| Non-Labor MPC of zero | 0.15 | 0.22 | 41% | 1.28 | 1.18 | 0.10 | 57.20% |
| Labor Share = 0.6 | 0.23 | 0.29 | 25% | 1.40 | 1.30 | 0.11 | 37.67% |
| Labor Share = 1 | 0.23 | 0.33 | 41% | 1.48 | 1.30 | 0.21 | 70.09% |

Notes: $MPC^b$ is the earnings-weighted average MPC times the average elasticity of earnings. Linear estimates of earnings heterogeneity are taken from Table 1.2, demographic-group-specific estimates are taken from Figure 1.1, and decile-based estimates are presented in Appendix Figure A16. Panels B and C use the linear specification. The benchmark in Panel A and Panel C is that each worker faces the earnings elasticity of the average dollar in the economy.

Table 1.4: Employment Sensitivity to GDP and the Local Matching Multiplier

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| $(MM_c + B_c) * \Delta \log GDP_t$ | 2.505 | | | | | |
| | 0.486 | | | | | |
| $MM_c * \Delta \log GDP_t$ | | 2.751 | 0.853 | 0.882 | 0.740 | 0.855 |
| | | (0.473) | (0.303) | (0.319) | (0.344) | (0.370) |
| $B_c * \Delta \log GDP_t$ | | 0.220 | -1.748 | -1.069 | -3.562 | -3.845 |
| | | (0.911) | (1.223) | (1.368) | (2.208) | (2.983) |
| Year FE | X | X | X | X | X | X |
| Demographic Controls | | | X | X | X | X |
| Firm Controls | | | | X | | |
| Financial Controls | | | | | X | X |
| Year by State. FE | | | | | | X |
| No. Observations | 2245 | 2245 | 2245 | 2245 | 1658 | 1441 |
| R-Squared | 0.513 | 0.514 | 0.751 | 0.754 | 0.788 | 0.848 |

Notes: In all columns, the dependent variable is the change on log employment at the commuting zone level. All regressions include CZ-fixed effects, and observations are weighted by the share of employment in the commuting zone in $t-1$. $\widehat{MM}_c$ is winsorized at the 5th and 95th percentile to account for large outliers ($\widehat{MM}_c$ is unbounded). Demographic controls include the share of employment in the two-digit industry; the average age and lagged earnings of the area; and the fraction of the CZ that is female, black, and in the labor force in $t-1$, each included separately and interacted with $\Delta \log G_t$. Firm controls include average firm size and age in $t-1$, included independently and interacted with $\Delta \log G_t$. Financial controls include the change in per-capita housing and financial wealth, the change in house prices, and lagged levels of debt per capita, each included individually and interacted with $\Delta \log G_t$. Standard errors are clustered at the commuting zone. See Appendix Table A16 for additional specifications.

Table 1.5: Tradable and Nontradable Employment and the Local Matching Multiplier

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| $MM_c * GDP_t$ | 1.503 | 0.425 | -0.232 | 1.027 | 1.113 | | |
| | (0.364) | (0.300) | (0.637) | (0.246) | (0.294) | | |
| $MM_c * GDP_t * Tradable_i$ | | | | | -1.475 | -1.427 | -1.261 |
| | | | | | (0.658) | (0.686) | (0.716) |
| $B_c * GDP_t$ | -0.326 | -0.228 | -0.085 | -0.526 | -0.824 | | |
| | (0.664) | (0.489) | (1.212) | (0.300) | (0.479) | | |
| $B_c * GDP_t * Tradable_i$ | | | | | 1.327 | 1.243 | 1.445 |
| | | | | | (0.452) | (0.466) | (0.498) |
| Industry*Year FE | X | X | X | X | X | X | X |
| CZ*Industry FE | X | X | X | X | X | X | X |
| CZ*Year FE | | | | | | X | X |
| Demographic Controls | X | X | X | X | X | X | X |
| Financial Controls | | | | | | | X |
| Included Industries | All | T + N | T | N | T | T +N | T +N |
| No. Observations | 183789 | 50269 | 32979 | 17290 | 50269 | 50269 | 41173 |
| R-Squared | 0.408 | 0.392 | 0.317 | 0.470 | 0.394 | 0.434 | 0.439 |

Notes: The row labeled "included industries" specifies whether the regression includes all industries (All), tradable industries (T), nontradable industries (N), or both (T+N). The dependent variable in all regressions is the change in log employment in a three-digit NAICS within a commuting zone. The earnings change is winsorized at the 5th and 95th percentiles, as are the estimates of $\widehat{MM}_c$. Demographic controls include the average age and lagged earnings of the area, as well as the fraction of the CZ that is female, black, and in the labor force, each separately, interacted with $G_t$, and further interacted with $T_i$. Financial controls include the change in per-capita housing and financial wealth and lagged levels of debt per capita, each included individually, interacted with $G_t$ and further interacted with $T_i$. All observations are weighted by the share of employment in $t-1$. Standard errors are clustered at the commuting zone. See Appendix Table A17 for alternate specifications.

Table 1.6: Model Calibration by Demographic Group

| | MPC | Income Processes | | | | $\beta$ |
|---|---|---|---|---|---|---|
| | | $\rho$ | $\sigma^2_{persistent}$ | $\sigma^2_{transitory}$ | $\overline{\log y_{it}}$ | |
| *High School or Less* | | | | | | |
| Non-Black Men | 0.46 | 0.94 | 0.05 | 0.29 | 10.74 | 0.88 |
| Black Men | 0.78 | 0.99 | 0.01 | 0.45 | 10.41 | 0.68 |
| Non-Black Women | 0.55 | 0.98 | 0.02 | 0.37 | 10.12 | 0.85 |
| Black Women | 0.59 | 0.97 | 0.02 | 0.45 | 9.97 | 0.82 |
| *Some College or More* | | | | | | |
| Non-Black Men | 0.32 | 0.97 | 0.03 | 0.36 | 11.23 | 0.92 |
| Black Men | 0.72 | 0.95 | 0.04 | 0.42 | 10.76 | 0.75 |
| Non-Black Women | 0.37 | 0.96 | 0.05 | 0.35 | 10.60 | 0.90 |
| Black Women | 0.47 | 0.97 | 0.02 | 0.44 | 10.41 | 0.86 |

Notes: The MPC reported in Column 1 is the earnings-weighted average MPC estimated in the PSID as described in Section 1.4.1. Columns 2 through 5 report the parameters of the income process, which are estimated separately by demographic group using data from the PSID from 1969 to 2013. See Appendix A for details. Column 6 reports the discount factor that is needed to match the earnings-weighted average MPC given the other parameters.

# Chapter 2

# The Fall of the Labor Share and the Rise of Superstar Firms

Joint with David Autor, David Dorn, Lawrence F. Katz and John Van Reenen.

## 2.1 Introduction

Much research has documented a decline in the share of GDP going to labor in many nations over recent decades (e.g., Blanchard, 1997; Elsby, Hobjin and Sahin, 2013; Karabarbounis and Neiman, 2013; Piketty 2014). Dao et al. (2017) point to a decline in the labor share between 1991 and 2014 in 29 large countries that account for about two-thirds of world GDP in 2014. Figure 2.1 illustrates this general decline in labor's share in twelve OECD countries with the fall in the United States particularly evident since 2000. The erstwhile stability of the labor share of GDP throughout much of the twentieth century was one of the famous Kaldor (1961) "stylized facts" of growth. The macro-level stability of labor's share was always, as Keynes remarked, "something of a miracle," and indeed disguised a lot of instability at the industry level (Elsby, Hobijn and Sahin, 2013; Jones, 2005). Although there is controversy over the degree to which the fall in the labor share of GDP is due to measurement issues such as the treatment of capital depreciation (Bridgman,

2014), housing (Rognlie, 2015), self-employment (Elsby, Hobjin, and Sahin, 2013; Gollin, 2002), intangible capital (Koh, Santaeulalia-Lopis and Zheng, 2016) and business owners taking capital instead of labor income (Smith, Yagan, Zidar and Zwick, 2017), there is a general consensus that the fall is real and significant.[1]

There is less consensus, however, on what are the *causes* of the recent decline in the labor share. Karabarbounis and Neiman (2013) hypothesize that the cost of capital relative to labor has fallen, driven by rapid declines in quality-adjusted equipment prices especially of Information and Communication Technologies (ICT), which could lower the labor share if the capital-labor elasticity of substitution is greater than one.[2] Elsby, Hobjin and Sahin (2013) argue for the importance of trade and international outsourcing especially with China. Like them, we will explore the role of trade, but we do not find that manufacturing industries with greater exposure to exogenous trade shocks differentially lose labor share relative to other manufacturing industries (although they do decline in terms of employment). Additionally, we observe a decline in labor's share in largely non-traded sectors such as wholesale trade, retail trade, and utilities, where trade exposure is more limited. Piketty (2014) stresses the role of social norms and labor market institutions, such as unions and the real value of the minimum wage. As we will show, the broadly common experience of a decline in labor shares across countries with different levels and evolution of unionization and other labor market institutions somewhat vitiates this argument.[3]

[1]The main issue in terms of housing is the calculation of the contribution of owner-occupied housing to GDP which is affected by property price fluctuations. We sidestep this by focusing on the Economic Census which includes firms (the "corporate sector" of the NIPA), not households. Similarly, the Census enumerates only employer firms, so does not have the self-employed. There remains an issue of how business owners allocate income, but Smith Smith, Yagan, Zidar and Zwick (2017) show that this can account for only an eighth of the labor share decline. We discuss some of the other factors, such as intangible capital below.

[2]Karabarbounis and Neiman (2013) argue for this, but the bulk of the empirical literature suggests an elasticity of below one (e.g., Lawrence, 2015; Oberfield and Raval, 2014; Antras, 2004; Hamermesh, 1990). However, this is a hard parameter to empirically identify.

[3]Blanchard (1997) and Blanchard and Giavazzi (2003) stress labor market institutions. Azmat, Manning and Van Reenen (2012) put more weight on privatization, at least in the network industries. Krueger (2018) emphasizes changes in worker power, such as increased monopsony. We discuss this more below in Section 2.5.

In this paper, we propose and empirically explore an alternative hypothesis for the decline in the labor share that is based on the rise of "superstar firms". If a change in the economic environment advantages the most productive firms in an industry, product market concentration will rise and the labor share will fall as the economy becomes dominated by superstar firms with high markups and lower labor shares. This would occur if consumers have become more sensitive to quality-adjusted prices due to greater product market competition (e.g., through globalization) or improved search technologies (e.g., if consumers or corporate buyers become more sensitive to price due to greater availability of price comparisons on the Internet, as in Akerman, Leuven and Mogstad, 2017). Our "winner take most" mechanism could also arise due to the growth of platform competition in many industries or scale advantages related to the growth of intangible capital (e.g. Walmart's massive investment in proprietary software to manage their logistics and inventory control—see Bessen, 2017). Central to our empirical analysis, this superstar firm framework implies that the reallocation of economic activity among firms with differing heterogeneous productivity and labor shares is key to understanding the fall in the aggregate labor share.

This paper's contribution is threefold. First, we provide microeconomic evidence on the evolution of labor shares at the firm and establishment level using U.S. Census panel data covering six major sectors: manufacturing, retail trade, wholesale trade, services, utilities and transportation, and finance. Our micro-level analysis is distinct from most existing empirical evidence that is largely based on macroeconomic and industry-level variation. Those aggregate approaches, while valuable in many dimensions, obscure the distinctive implications of competing theories, particularly the contrast between models implying heterogeneous changes (such as our superstar firm perspective) compared to homogeneous changes in the labor share across firms in an industry.[4] Second, we formal-

[4]Exceptions are Bockerman and Maliranta (2012) who use longitudinal plant-level data to decompose changes in the labor share in Finnish manufacturing into between and within plant components. Focusing on U.S. manufacturing, Kehrig and Vincent (2017) also use U.S. Census of Manufactures micro data to decompose labor share changes. We discuss their contribution below.

ize a new "superstar firm" model of the labor share change. The model is based on the idea that industries are increasingly characterized by a "winner take most" feature where a small number of firms gain a very large share of the market. Third, we present a substantial body of evidence from the last 30 years using a variety of U.S. and international datasets that is broadly consistent with the superstar firm hypothesis.

Specifically, we establish the following seven facts that support our model's predictions for how the rise of superstar firms can lead to a fall of labor's share: (i) There has been a rise in sales concentration within four-digit industries across the vast bulk of the U.S. private sector. Part of this is due to increased specialization on core competencies and partly it is due to firms just getting bigger. For example, the share of U.S. employment in firms with over 5,000 employees rose from 28% in 1987 to 34% in 2016.[5]; (ii) industries with larger increases in product market concentration have experienced larger declines in the labor share; (iii) the fall in the labor share is largely due to the reallocation of sales between firms rather than a general fall in the labor share within incumbent firms; (iv) the reallocation-driven fall in the labor share is most pronounced in precisely the industries which exhibited the largest increase in sales concentration; (v) the industries that are becoming more concentrated are those with faster growth of productivity and innovation; (vi) larger firms have higher markups and the size-weighted aggregate markup has risen more than the unweighted average markup; and (vii) these patterns are not unique to the U.S. but are also present in other OECD countries. Although we do not provide precise causal identification of our superstar firm model, the fact pattern presented here supports a firm-level perspective on the changes in the labor share.[6]

Our formal model, detailed below, generates superstar effects from increases in the

[5]Census Bureau Business Dynamics Statistics (e.g. https://www.census.gov/ces/dataproducts/bds/data_firm2016.html). As we show below, employment shares underestimate the growth in superstar firms which often have high sales with relatively few workers. And, because firms are increasingly specialized in their main industries, as we document below using Compustat data, total sales underestimates the growth of concentration in specific industries.

[6]See Furman and Orszag (2015) for an early discussion. Berkowitz, Ma and Nishioka (2017) also stress how an increase in market power could generate a decline in the labor share and find some evidence in support of this in Chinese micro-data.

toughness of product market competition, which raise the market share of the most productive firms in each sector at the expense of less productive competitors. Though our model formalizes the market toughness channel, we underscore that a number of closely related mechanisms can deliver similar superstar effects. First, strong network effects are a related explanation for the dominance of companies such as Google, Facebook, Apple, Amazon, AirBNB and Uber in their respective industries. Second, rapid falls in the quality-adjusted prices of intangible capital such as software could give large firms an advantage if there is a large overhead/fixed cost element.[7] For example, Walmart has made substantial technology investments to enable it to monitor supply chain logistics and manage inventory to an extent that, arguably, would be infeasible for smaller competitors (Bessen, 2017). An alternative perspective on the rise of superstar firms is that they reflect a diminution of competition, due to a weakening of U.S. antitrust enforcement (Dottling, Gutierrez and Philippon, 2018). Our findings on the similarity of trends in the U.S. and Europe, where antitrust authorities have acted more aggressively on large firms (Gutierrez and Philippon, 2018), combined with the fact that the concentrating sectors appear to be growing more productive and innovative, suggests that this is unlikely to be the primary explanation, although it may important in some specific industries (see Cooper et al, 2019, on healthcare for example).

Our paper is also closely related to Barkai (2016), who independently documented a negative industry-level relationship between changes in labor share and changes in concentration for the United States. Barkai presents evidence at the aggregate level that profits appear to have risen as a share of GDP, and that the pure capital share (capital stock multiplied by the required rate of return) of GDP has fallen, a pattern consistent with our superstar firm model and the empirical analysis we will present on rising aggregate markups. Barkai's analysis uses exclusively industry-level and macro data. A major contribution of our micro-level approach is that we can explore the firm-level contributions

[7]See Bauer and Lashkari (2018); Crouzet and Eberly (2018), Karabarbounis and Neiman (2018) and Koh et al (2016) for variants of this argument.

to these patterns and link them to our model, particularly the implications and evidence on between-firm (output reallocation) versus within-firm contributors to falling industry- and aggregate-level labor share. We thus view our contribution and that of Barkai (2016) as complementary. Our work also corroborates (and helps to interpret) that of de Loecker, Eeckhout and Unger (2018) who argue that the (weighted average) markup of price over variable cost has been rising in the U.S. (where, ceteris paribus, a rise in the markup means a fall in the labor share). As with these papers, our model also implies rises in aggregate markups due to a reallocation of market share towards superstar firms, which have both lower labor shares and high markups. We confirm these patterns in our Census data.

Compared to our earlier work in the *American Economic Review Papers and Proceedings* (Autor et al, 2017b), this paper formalizes the superstar theory, presents firm-level decompositions of the labor share; explores the correlation of the labor share with concentration on the one hand and the factors influencing concentration on the other; analyzes markups directly; and provides a quantitative characterization of superstar firms using Compustat data.[8]

The structure of the paper proceeds as follows. Section 2.2 sketches our model. Section 2.3 presents the data and Section 2.4 the empirical support for the model's predictions. Section 2.5 presents additional descriptive facts of superstar firms, and Section 2.6 provides concluding remarks. Online Appendices detail the formal model (Appendix A), markup calculation (Appendix B), superstar firm characteristics (Appendix C) and data (Appendix D).

## 2.2 A Model of Superstar Firms

We provide a formal model in online Appendix A, but to provide intuition for why the fall in labor share may be linked to the rise of superstar firms, consider a production

[8]A point of overlap is that we again present concentration trends. Even here however, we have updated the earlier data in several ways, most importantly by incorporating the full 2012 Economic Census.

function $Y_i = z_i L_i^{\alpha^L} K_i^{1-\alpha^L}$ where $Y_i$ is value-added, $L_i$ is variable labor, $K_i$ is capital and $z_i$ is Hicks-neutral efficiency (TFPQ) in firm $i$.[9] Consistent with a wealth of evidence, we assume that $z_i$ is heterogeneous across firms (Melitz, 2003; Hopenhayn, 1992). More productive, higher $z_i$, firms will have higher levels of factor inputs and greater output.

Factor markets are assumed to be competitive (with wage $w$ and cost of capital $\rho$), but we allow for imperfect competition in the product market. From the static first order condition for labor we can write the share of labor costs ($wL_i$) in nominal value-added ($P_iY_i$) as:[10]

$$S_i \equiv \left(\frac{wL_i}{P_iY_i}\right) = \frac{\alpha^L}{m_i} \tag{2.1}$$

where $m_i = (P_i/c_i)$ is the mark-up, the ratio of product price $P_i$ to marginal cost $c_i$. The firm $i$ subscripts indicate that for given economy-wide values of $(\alpha^L, w, \rho)$, a firm will have a lower labor share if its mark-up is higher. Superstar firms (those with high $z_i$) will be larger as they produce more efficiently, charge lower prices and so capture a higher share of industry output. If they have have higher price-cost markups, they will also have lower labor shares. Indeed, a wide class of models of imperfect competition will generate larger price-cost mark-ups for firms with a higher market share, $\omega_i = P_iY_i/\sum_i (P_iY_i)$. This is because mark-ups ($m_i$) are generally falling in the absolute value of the elasticity of demand $\eta_i$, and according to Marshall's "Second Law of Demand", consumers will be more price-inelastic at higher levels of consumption and lower levels of price.[11] Most utility functions will have this property, such as the Quadratic Utility Function which generates a linear demand curve. In this case $m_i = \eta_i/(\eta_i - 1)$. Another example is the homogeneous product Cournot model, which generates $m_i = \frac{\eta_i}{\eta_i - \omega_i}$. The empirical literature also

[9]We treat output and value-added interchangeably here as we are abstracting away from intermediate inputs. We distinguish intermediate inputs in the empirical application.

[10]Employer product market power was emphasized by Kalecki (1938) as the reason for variations in labor shares over the business cycle.

[11]Mrazova and Neary (2017) discuss the implications of a wide class of utility functions (generating "demand manifolds") including those which are not consistent with Marshall's Second Law.

tends to find higher mark-ups for larger, more productive firms.[12] A leading exception to this is when preferences are CES (the Dixit-Stiglitz form with a constant elasticity of substitution between varieties), in which case mark-ups are the same across all firms of whatever size and productivity ($m = \eta / (\eta - 1)$). In Autor et al (2017), we show that even in such a CES model, labor shares could be lower for larger firms if there are fixed costs of overhead labor that do not rise proportionately with firm size.[13]

Because labor shares are lower for larger firms in standard models, an exogenous shock that reallocates market share towards these firms will tend to depress the labor share in aggregate. Intuitively, as the weight of the economy shifts toward larger firms, this will lower the average labor share even with no fall in the labor share at any given firm. In online Appendix A, we formalize these ideas in an explicit model of monopolistic competition, which we use to illustrate some key results. The model is a generalization of Melitz and Ottaviano (2008), augmented with a more general demand structure and (most importantly) a more general productivity distribution. In the model, entrepreneurs entering an industry are *ex ante* uncertain of their productivity $z_i$. They pay a sunk entry cost $\kappa$ and draw $z_i$ from a known productivity distribution with density function, $\lambda(z)$. Firms that draw a larger value of $z$ will employ more inputs and have a higher market share. Our demand functions obey Marshall's Second Law, so we obtain the first result that larger firms will have lower labor shares.

As is standard (e.g. Arakolis et al, 2018), we characterize the "toughness" of the market in terms of a marginal cost cut-off $c^*$. Firms with marginal costs exceeding this level

[12]See the discussion in Arkolakis et al (2018). In the time series, the empirical trade literature finds incomplete pass through of marginal cost shocks to price with elasticities of less than unity, which implies higher mark-ups for low cost firms. A smaller literature estimating cross sectional mark-ups finds larger mark-ups for bigger firms (e.g. de Loecker and Warzynski, 2012). Below, we empirically confirm this is true on our U.S. Census data.

[13]Denote fixed overhead costs of labor $F$ and variable labor costs so $V$, $L = V + F$. In this case $S_i = \frac{\alpha^L}{m} + \frac{wF}{P_i Y_i}$. Since high $z_i$ firms are larger, they will have a lower share of fixed costs in value-added ($wF/P_i Y_i$) and lower observed labor shares (see Bartelsman, Haltiwanger and Scarpetta, 2013). We emphasized this mechanism in the original working paper version of this paper (Autor et al, 2017a), but we regard the broader model of online Appendix A as more rigorous and more realistic.

will earn negative profits and exit. Globalization, which increases effective market size, or greater competition (meaning higher substitutability between varieties of goods) will tend to make markets tougher and reduce the cut-off, $c^*$, causing low productivity firms to shrink and exit. The reallocation of market share towards more productive firms will increase the degree of sales concentration and will be a force decreasing the labor share because a larger fraction of output is produced by more productive ("superstar") firms. This is our second result.

Since the change in market toughness will also tend to reduce the mark-up for any individual firm, labor shares at the firm level will *rise*. In order to obtain an *aggregate decline* in the labor shares when markets get tougher, the "between firm" reallocation effect must dominate this "within firm" effect. Our third result is to show that the aggregate labor share will indeed fall following this change in the economic environment if the underlying productivity density is log-convex $\lambda(z)$, meaning that the productivity distribution is more skewed than the Pareto distribution. Conversely, the aggregate labor share will rise if the density is log-concave and will remain unchanged if the density is log-linear. Interestingly, the standard assumption (e.g. Melitz and Ottaviano, 2008) is that the pdf of productivity is Pareto. Since this is an example of a log-linear density function, it delivers the specialized result that the within and between effects of a change in the economic environment perfectly offset each other, so the aggregate labor share is invariant to changes in market toughness. Since the underlying distribution of productivity draws $\lambda(z)$ is unobservable, the impact of a change in market toughness on the aggregate labor share is fundamentally an empirical issue. While the prediction that rising market toughness could generate an increase in concentration and the profit share is counter-intuitive, the ambiguous relationship between concentration, profit shares, and the stringency of competition is well known in Industrial Organization.[14]

[14]The interpretation of the relationship between profit margins and the concentration level is a classic issue in industrial organization. In the Bain (1951) "Structure-Conduct-Performance" tradition, higher concentration reflected greater entry barriers which led to an increased risk of explicit or implicit collusion. Demsetz (1973), by contrast, posited a "Differential Efficiency" model closer to the one in online Appendix A, where

The model in online Appendix A implies that after an increase in market toughness: (i) the market concentration of firm sales will rise, meaning that the market shares of the largest firms will rise; (ii) in those industries where concentration rises the most, labor shares will fall the most (assuming that the underlying distribution of productivity draws is log-convex); (iii) the fall in the labor share will have a substantial reallocation component between firms, rather than being a purely within-firm phenomenon; (iv) in those industries where concentration rises the most, the reallocation from firms with high to low labor shares will be the greatest; (v) the industries that are becoming more concentrated will be those with the largest productivity growth; (vi) due to high-markup firms expanding, the aggregate markup will rise ; and (vii) similar patterns of changes in concentration and labor's share will be found across countries (to the extent that the shock that benefits superstar firms is global). We take these predictions to a series of newly constructed micro-datasets for the U.S. and around the world.

Our stylized model is meant to illustrate our intuition for the connection between the rise of superstar firms and decline in labor's share. Similar results could occur from any force that makes the industry more concentrated—more "winner take most"—such as an increased importance of network effects, as long as high market share firms have lower labor shares.[15] A high level of concentration does not necessarily mean that there is persistent dominance: one dominant firm could swiftly replace another as in standard neo-Schumpeterian models of creative destruction (Aghion and Howitt, 1992). But dynamic models could create incumbent advantages for high market share firms. Such a phenomenon could occur through innovation incentives, as in the Gilbert and Newbery (1982) model, where incumbents are more likely to innovate than entrants. A more worrying explanation of growing concentration would be if incumbent advantage were

increases in competition allocated more output to more productive firms. In either case, however, concentration would be associated with higher profit shares of revenue and, in our context, a lower labor share. See Schmalensee (1987) for an effort to empirically distinguish these hypotheses.

[15] Another way of generating this would be if the underlying distribution of entrepreneurial ability become more skewed. As it is unclear why the primitives would change in this manner, we prefer to model a change in the economic environment as this offers more hope of empirically identifying the mechanisms at play.

enhanced through erecting barriers to entry (e.g., the growth of occupational licensing highlighted by Kleiner and Krueger, 2013, or a weakening of antitrust enforcement as argued by Gutierrez and Philippon, 2016 and 2018). Explanations for growing concentration from weakening antitrust enforcement have starkly different welfare implications than explanations based on innovation or toughening competition. We partially—though not definitively—assess these alternative explanations by examining whether changes in concentration are larger in dynamic industries (where innovation and productivity is increasing) or in declining sectors.

## 2.3 Data

We next describe the main features of our data. Further details on the datasets are contained in online appendix D.

### 2.3.1 Data Construction

The data for our main analysis come from the U.S. Economic Census, which is conducted every five years and surveys all establishments in selected sectors based on their current economic activity. We analyze the Economic Census for the three decade interval of 1982 - 2012 for six large sectors: manufacturing, retail trade, wholesale trade, services, utilities and transportation and finance.[16] The covered establishments in these six sectors comprise approximately 80 percent of total private sector employment. To implement our industry-level analysis, we assign each establishment in each year to a 1987 SIC-based time-consistent industry code. We are able to observe 676 industries, 388 of which are in

[16]Within these six sectors, several industries are excluded from the Economic Census: rail transportation is excluded from transportation; postal service is excluded from wholesale trade; funds, trusts and other financial vehicles are excluded from finance; and schools (elementary, secondary, and colleges), religious organizations, political organizations, labor unions and private households are excluded from services. The Census also does not cover government-owned establishments within the covered industries. We also drop some industries in Finance, Services, and Manufacturing that are not consistently covered across these six sectors. See online Appendix D for details.

manufacturing.

For each of the six sectors, the Census reports each establishment's total annual payroll, total output, total employment, and, importantly for our purposes, an identifier for the firm to which the establishment belongs. Annual payroll includes all forms of paid compensation, such as salaries, wages, commissions, sick leave, and also employer contributions to pension plans , all reported in pre-tax dollars. The Census of Manufactures also includes a wider definition of compensation that includes all fringe benefits, the most important of which is employer contributions to health insurance, and we also present results using this broader measure of labor costs.[17] The exact definition of output differs based on the nature of the industry, but the measure intends to capture total sales, shipments, receipts, revenue, or business done by the establishment. In most sectors, in constructing the NIPA, the BEA uses the Economic Censuses to construct gross output and then works through data sources on materials use to construct value added. The finance sector is the most problematic in this regard.[18] Accordingly, we place finance at the end of all tables and figures and advise caution in interpreting the results in this sector.

In addition to payroll and sales which are reported for all sectors, the Economic Census for the manufacturing sector further includes information on value-added at the establishment level. Value-added is calculated by subtracting the total cost of materials, supplies, fuel, purchased electricity, and contract work from the total value of shipments, and then adjusting for changes in inventories over that year. This enables us to present a more in-depth analysis of key variables in manufacturing.

Because industry definitions have changed over time, we construct a consistent set of industry definitions for the full 1982-2012 period (as is documented in Appendix D). We build all of our industry-level measures using these time-consistent industry definitions, and thus our measures of industry concentration differ slightly from published statistics.

[17]Additional compensation costs are only collected for the subset of Census establishments in the Annual Survey of Manufacturers (ASM) and are imputed by the Census Bureau for the remainder.

[18]For the banking sector, for example, BEA calculates value-added from interest rate spreads between lending and deposit rates.

The correlation between our calculated measures and those based on published data is almost perfect, however, when using the native but time-varying industry definitions.[19]

We supplement the U.S. Census-based measures with various international datasets. First, we draw on the 2012 release of the EU KLEMS database (see O'Mahony and Timmer, 2009, http://www.euklems.net/), an industry level panel dataset covering OECD countries since 1980. We use the KLEMS to measure international trends in the labor share and also to augment the measurement of the labor share in the Census by exploiting KLEMS data on intermediate service inputs.[20]

Second, we use data on industry imports from the UN Comtrade Database from 1992-2012 to construct adjusted measures of imports broken down by industry and country. To compare these data to the industry data in the Census, we convert six-digit HS product codes in Comtrade to 1987 SIC codes using a crosswalk from Autor, Dorn and Hanson (2013), and we slightly aggregate industries to obtain our time-consistent 1987 SIC-based codes. Our approach yields for each industry a time series of the dollar value of imports from six country groups.[21]

Third, to examine the relationship between sales concentration and the labor share internationally, we turn to a database of firm-level balance sheets from 14 European countries that covers the 2000-2012 period. This database, compiled by the European Central Bank's Competitiveness Research Network (CompNet), draws on various administrative and public sources across countries, and seeks to cover all non-financial corporations.[22] CompNet aggregates data from all firms to provide aggregate information on the labor

[19]One minor difference emerges because we drop a handful of establishments that do not have the LBDNUM identifier variable, which is needed to track establishments over time. In Appendix D, we also compare our results with the alternative set of consistent industry definitions developed by Fort and Klimek (2016) who used a NAICS-based measure, obtaining similar results to our own.

[20]We choose the 2012 KLEMS release because subsequent versions of EU KLEMS are not fully backward compatible and provide shorter time series for many countries.

[21]The six country groups are: Canada; eight other developed countries (Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain and Switzerland); Mexico and CAFTA; China; all low income countries other than China; and the rest of the world.

[22]See Lopez-Garcia, di Mauro and CompNet Task Force (2015) for details.

share and industry concentration for various two-digit industries. Although great effort was made to make these measures comparable across countries, there are some important differences that affect the reliability of cross-country comparisons.[23] Consequently, we estimate specifications separately for each country and focus on a within-country analysis.

Fourth, to implement firm-level decompositions internationally, we use the BVD Orbis database to obtain panel data on firm-level labor shares in the manufacturing sectors of six European countries for private and publicly-listed firms. BVD Orbis is the best publicly available database for comparing firm panels across countries (Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych and Yesiltas, 2015).[24]

Finally, in order to describe the characteristics of "superstar" firms and characterize their international scope, we supplement the analysis of Census data with the Standard & Poor's Compustat database. This database reports economic information for firms listed on a U.S. stock exchange. We focus on the largest 500 firms and explore the characteristics of the largest firms within that group. Further details on data construction are reported in Appendix D, and the Compustat analysis is found in Appendix C.

### 2.3.2 Initial Data Description

Figure 2.1 plots labor's share of value-added since the 1970s in 12 developed countries. A decline in the labor share is evident in almost all countries, especially in the later part of the sample period.[25] Focusing in on the United States, Figure 2.2 presents labor's share of value-added in U.S. manufacturing. The figure includes three measures of labor's share.

[23]Most importantly, for our purposes, countries use different reporting thresholds in the definition of their sampling frames. For example, the Belgian data cover all firms, while French data include only firms with high sales, and the Polish data cover only firms with more than five employees. Consequently, countries differ in the fraction of employment or value-added included in the sample.

[24]Unfortunately, due to partial reporting of revenues, BVD Orbis cannot be used to comprehensively construct sales concentration measures.

[25]Of the 12 countries, Sweden and the UK seem the exceptions with no clear trend. Bell (2015) suggests that the UK does have a downward trend in the labor share when the data are corrected for the accounting treatment of payments into (under-funded) private pension schemes for retirees. Payments into these schemes, which benefit only those workers who have already retired, are counted as current labor compensation in the national accounts data, therefore overstating the non-wage compensation of current employees.

We first construct the labor share using payroll, which is the standard labor cost measure that is available for all sectors, as the numerator and value-added as the denominator. We modify this baseline measure to include a broader measure of compensation that includes non-wage labor costs (such as employer health insurance contributions), which are only provided in the Census of Manufactures and not the other parts of the Economic Census. Lastly, we also plot payroll normalized by sales, rather than value-added, as this is the measure that can be constructed outside of Manufactures. Figure 2.2 shows that all three series show a clear downward trend, though of course their initial levels differ.

To what extent is manufacturing different from other sectors? Because robust firm-level measures of value-added are not available from the Economic Census outside of manufacturing, we use the cruder measure of the ratio of payroll to sales. This measure, which can be computed for all six broad sectors covered in the Census, is plotted by sector in the six panels of Figure 2.3. Finance stands out as the only sector where there is a clear upward trend in the labor share. As above, this is also the sector in which measures of inputs and outputs are most problematic. In all non-financial sectors, there has been a fall in the labor share since 2002—indeed the labor share is lower at the end of the sample than at the beginning in all sectors except services, where the labor share fell steeply between 2002 and 2008 then partly rebounded. The 1997-2002 period stands out as a notable deviation from the overall downward trend, as the labor share rose in all sectors except manufacturing in this period, and even here the secular downward trend only temporarily stabilized. One explanation for this temporary deviation is that the late 1990s was an unusually strong period for the labor market with high wage and employment growth. Appendix D compares Census data to NIPA. The fall in the labor share of value added is clearer in NIPA than Census payroll to sales ratios. The online appendix shows that all non-finance sectors saw a net fall in labor share over the full 1982 - 2012 time period, and even in finance, the labor share is stable from from the mid 1980s to the Great Recession (before then falling).

We next turn to concentration in the product market, which in the superstar firm model should be connected with the decline in the labor share. We measure industry concentration as (i) the fraction of total sales that is accrued by the four largest firms in an industry (denoted CR4), (ii) the fraction of sales accrued by the 20 largest firms (CR20), and (iii) the industry's Herfindahl-Hirschman Index (HHI).[26] For comparison, we also compute the CR4 and CR20 concentration measures based on employment rather than sales. Following Autor et al (2017b), Figure 2.4 plots the average sales- and employment-based CR4 and CR20 measures of concentration across four-digit industries for each of the six major sectors using updated data from the Census. The online appendix shows a corresponding plot for the Herfindahl-Hirschman Index (denoted HHI). Both figures show a remarkably consistent pattern. First, there is a clear upward trend over time: according to all measures, industries have become more concentrated on average. Second, the trend is stronger when measuring concentration in sales rather than employment. This suggests that firms may attain large market shares with relatively few workers— what Brynjolfsson, McAfee, Sorrell and Zhou (2008) term "scale without mass." Third, a comparison of Figure 2.4 and Figure **??** shows that the upward trend is slightly weaker for the HHI, presumably because this metric is giving some more weight to firms outside the top 20 where concentration has risen by less.

One interesting question is whether these increases in concentration are mainly due to superstar firms expanding their scope over multiple industries, as in the case of Amazon, or rather are due to a greater firm focus on core industries. We found that the largest firm (by sales) in the four digit industry in the Census operated on average in 13 other four digit industries in 1982, but this number fell to under 9 by 2012. Similarly, conditional on a firm being among the top four firms in a four-digit industry in 1982, it was on average among the top four in 0.37 other industries (i.e. statistically speaking, being the top firm in

[26]Since we calculate concentration at the industry level, we define a firm as the sum of all establishments that belong to the same parent company and industry. If a company has establishments in three industries, it will be counted as three different firms in this analysis. About 20% of manufacturing companies span multiple four-digit industries.

one industry gave a firm a 37% chance of being among the top four in another industry). By 2012, this fraction had fallen from 0.37 to 0.24. Thus, the data suggests that companies like Amazon, which are becoming increasingly dominant across multiple industries, are the exception. Overall, firms are becoming more concentrated in their leading line of business but less integrated across other activities. Table 2.1 provides further descriptive statistics for sample size, labor share, and sales concentration in each of the six sectors.

Before more formally exploring the implications of the model, we present evidence of the cross-sectional relationship between firm size and labor share. As discussed in Section 2.2, our conceptual framework is predicated on the idea that because "superstar" firms produce more efficiently, they are both both larger and have lower labor shares. To check this implication, Figure 2.5 reports the bivariate correlation between firms' labor shares, defined as the ratio of payroll to sales, and firms' share of their respective industry's annual sales. Consistent with our reasoning, there is a negative relationship between labor share and firm size across all six sectors, and this relationship is statistically significant in five of the six sectors.

## 2.4 Empirical Tests of the Predictions of Superstar Firm Model

### 2.4.1 Rising Concentration Correlates with Falling Labor Shares

#### Manufacturing

Table 2.2 presents the results of regressing the change in the labor share on the change in industrial concentration for our sample window of 1982 through 2012. We begin with the manufacturing sector as these data are richest, but then move on to results from the other sectors. In each of the six sectors, we separately estimate OLS regressions in long differences (indicated by $\Delta$) of the form

$$\Delta S_{jt} = \beta \Delta \text{CONC}_{jt} + \tau_t + u_{jt}, \tag{2.2}$$

where $S_{jt}$ is the labor share of four-digit SIC industry $j$ at time $t$ , $\text{CONC}_{jt}$ is a measure of concentration, $\tau_t$ is a full set of time dummies, and $u_{jt}$ is an error term. We allow for the standard errors to be correlated over time by clustering at the industry level. All cells in Table 2.2 report estimates of $\beta$ from equation (2.2). The first three columns present five-year long differences, and the last three columns present ten-year long-differences. Since the left- and right-hand side variables each cover the same time interval in each estimate, the coefficients have a comparable interpretation in the five-year and ten-year specifications.

Our baseline specification in row 1 detects a striking relationship between changes in concentration and changes in the share of payroll in value-added. Across all three measures of concentration (C4, C20, and HHI), industries where concentration rose the most were those where the labor share fell by the most. These correlations are statistically significant at the 5 percent level in all but the last column (where it is significant at the 10% level).[27] The subsequent rows of Table 2.2 present a variety of robustness tests of this basic association. In row 2, we use a broader measure of the labor share—using "compensation" instead of payroll—that includes employer contributions to fringe benefits such as private health insurance, which account for a growing fraction of labor costs (Pessoa and Van Reenen, 2013). Row 3 uses an adjusted value-added measure (for the denominator of labor share) which uses KLEMS data to attempt to account for intermediate service inputs that are not included in the Census data (see Appendix D for details). In row 4, we define concentration based on value-added rather than sales. Row 5 presents a stringent robustness test by including a full set of four-digit industry dummies, thus obtaining identification exclusively from acceleration or deceleration of concentration and labor shares relative to industry-specific trends. The strong association between rising concentration and falling labor share is robust to all of these permutations.

Our core measure of concentration captures exclusively domestic U.S. concentration

[27]The HHI estimates, which give weight to firms outside of the top 20, are least precise. Our superstar model focuses on the leading firms in each sector rather than the entire distribution of firms.

and hence may overstate effective concentration for traded-goods industries, particularly in manufacturing, where there is substantial international market penetration.[28] If firms operate in global markets and the trends in U.S. concentration do not follow the trends in global concentration, then our results may be misleading. We address this issue in several ways. Since import penetration data are not available on a consistent basis across our full time period, we focus on the 1992-2012 period where these data are available. For reference, row 6 of Table 2.2 re-estimates our baseline model for the shortened period and finds a slightly stronger relationship between labor share and concentration. Row 7 next adds in the growth in imports over value-added in each five year period on the right hand side, and finds that the coefficient on concentration falls only slightly. In Section (2.5), we further investigate the role of trade in explaining the fall in the labor share.

Karabarbounis and Neiman (2014) stress the role of the falling cost of the prices of investment goods in driving down the labor share. To examine this idea, row 8 includes the start-of-period level of the capital to value-added ratio on the right hand side of the regression. Under the Karabarbounis and Neiman (2014) hypothesis, we would expect capital-intensive industries to have the largest falls in the labor share. Consistent with this logic, the coefficient on capital intensity is negative and significant. The coefficient on concentration is little changed from row 1, however, suggesting that the superstar mechanism linking rising concentration to falling industry-level average labor shares is not a simple manifestation of capital intensity or capital deepening.

Finally, note that our measure of concentration is based on firm sales (or value added), but it is also possible to construct concentration indices based on employment. The relationship of the labor share with these alternative measures of concentration is presented in the final row of Table 2.2. Interestingly, the coefficients switch sign and are positive (although with one exception, insignificant). This is not a problematic result from the perspective of our conceptual framework; measures based on outputs, reflecting a firm's

[28]This is a minor concern in non-manufacturing sectors, where there are comparatively few imports.

position in the product market, is the appropriate measure of concentration, not employment. Indeed, many of the canonical superstar firms such as Google and Facebook employ relatively few workers relative to their market capitalization. Thus, their market value is based on intellectual property and a cadre of highly-skilled workers. Measuring concentration using employment rather than sales fails to capture this revenue-based concentration among IP and human capital-intensive firms.

### All Sectors

We now broaden our focus to include the full set of Census sectors (alongside manufacturing): retail, wholesale, services, utilities and transportation, and finance. We apply our baseline specification to these sectors, with two modifications: first, the sample window is shorter for finance and utilities and transportation (1992-2012) because of lack of consistent data prior to 1992 in these sectors; second, because we do not have value-added outside of manufacturing, we use payroll over sales as our dependent variable. To assess whether this change in definition affects our results, we repeat the manufacturing sector analysis from Table 2.2 in Table 2.3 using payroll normalized by sales rather than value-added, the results of which are reported in row 1. In the first three columns, for example, All the coefficients remain negative, statistically significant, and quantitatively similar.[29]

Figure 2.6 plots the coefficients that result from the estimation of equation (2.2) separately for each sector using the CR20 as the measure of concentration and looking at changes over five year periods (column 2 of Table 2.3). It is clear from both Figure 2.6 and Table 2.3 that rising concentration is uniformly associated with a fall in the labor share both outside of manufacturing as well as within it. The coefficient on the concentration measure is negative and significant at the 5 percent level or lower in each sector. When we

[29]Figure 2.3 shows that the mean fall in payroll as a share of sales in manufacturing is 7 percentage points, which is less than half of the 16.5 percentage point fall for payroll normalized on value-added (Figure 2.2). Similarly, the coefficient on concentration in the share of value-added equation is just over twice as large as the that in the share of sales equation (e.g. -0.148 for the CR4 in column (1) of Table 2 compared to -0.062 in Table 3).

pool all six sectors and estimate equation (2.2) with sector-specific fixed effects (final row of Table 2.3, labeled "combined"), we again find a strong negative association between rising concentration and falling labor share.

Table 2.3 also reports several variants of this regression using alternate measures of concentration as well as stacked ten-year changes rather than five-year changes. The negative relationship is robust across specifications: negative in all 36 specifications in rows 1 to 6 of Table 2.3 and significantly so at the 10 percent or greater level in 28 cases.[30] We also examined specifications using the change in the CR1 (that is, the market share of the single largest firm in the industry) as the concentration measure. As expected given the other results, we find that the change in the CR1 is negatively associated with changes in the labor share in all specifications in all six segments.[31] Since most employment and output is produced outside of manufacturing, these results underscore the pervasiveness and relevance of the concentration-labor-share relationship for almost the entire U.S. economy.

**Further robustness tests**

We have implemented a large number of robustness tests on these regressions and discuss several of them here. First, we repeated the robustness tests applied to manufacturing in Table 2.2 to the full set of six sectors to the extent that the data permit. For example, following the model of row 5 of Table 2.2, we added a full set of four-digit industry trends to the five-year first-difference by-sector estimates in Table 2.3. All coefficients were negative across the three measures of concentration and 14 of the 18 were significant at the 5

[30]To assess whether the results are driven by the number of firms in the industry rather than their concentration, we additionally included the count of firms as a separate control variable in changes and initial levels. Although the coefficient on concentration tends to fall slightly in such specifications, it remains generally significant, suggesting that it is the distribution of market shares that matters and not simply the number of firms (though obviously these are correlated).

[31]For the five year difference specifications the coefficient (standard error) on the CR1 in manufacturing was -0.124 (0.041) for payroll over value added, -0.146 (0.054) for compensation over value added, and -0.060 (0.014) for payroll over sales. For payroll over sales it was -0.018 (0.019), -0.035 (0.016), -0.114 (0.064), -0.097 (0.043) and -0.252 (0.091) for retail, wholesale, services, utilities and transportation and finance respectively (the combined value pooled across all six sectors was -0.074 (0.016)).

percent level.

Second, the superstar firm model is most immediately applicable to higher-tech industries, which may have developed a stronger "winner takes most" character, while it is less obviously applicable to declining sectors. To explore this heterogeneity, we divide our sample of industries into the high-tech versus other sectors. Consistent with expectations, we find that the coefficient on firm concentration predicts a larger fall in the labor-share in high-tech sectors (classified in a variety of ways) than in the complementary set of non high-tech sectors.[32]

Third, we note that our main estimating equation (2.2) imposes a common coefficient over time on the concentration measures and takes heterogeneity between years into account only through the inclusion of time dummies. Results in the online appendix show that using either definition of the labor share (i.e. value-added or sales) in manufacturing, the relationship between the change in the labor share and the change in concentration is significantly negative in all periods except for 1982-1987, and generally strengthens over the sample period. Although the numbers of individual industries within each of the the five non-manufacturing sectors are fewer than in the manufacturing sector and therefore provide noisier measurement, the same broad patterns emerge: a negative relationship is evident across most years and tends to become stronger over time.

[32]We followed Decker, Haltiwanger, Jarmin and Miranda (2018) by using the definition of high-tech in Hecker (2005). Here, an industry is deemed high-tech if the industry-level employment share in technology-oriented occupations is at least twice the average for all industries. This occupation classification is based on the 2002 BLS National Employment Matrix that gives the occupational distribution across four-digit NAICS codes. We use the NAICS-SIC crosswalk and identify the SIC codes that map entirely to the high tech four-digit NAICS codes, yielding 109 four-digit "high tech" SIC codes. Re-running our primary model with this classification, we found that the coefficient on concentration is negative and significant in both sub-samples, but is almost twice as large in absolute magnitude in the high-tech sub-sample. In a pooled specification, the interaction between the high tech dummy and the CR20 is negative and significant (-0.067 with a standard error of 0.031).

### 2.4.2 Between-Firm Reallocation Drives Fall in Labor Share

#### Methodology

The third implication of the superstar firm model is that the fall in the labor share should have an important between-firm (reallocation) component, as firms with a low labor share capture a rising fraction of industry sales or value-added. To explore this implication, we implement a variant of the Melitz and Polanec (2015) decomposition which was originally developed for productivity decompositions but it be applied readily to the labor share.[33] We write the level of the aggregate labor share as

$$S = \sum \omega_i S_i = \bar{S} + \sum \left(\omega_i - \bar{\omega}\right)\left(S_i - \bar{S}\right), \tag{2.3}$$

where the size-weight, $\omega_i$, is firm $i$'s share of value-added in an industry, $\omega_i = P_i Y_i / \sum_i P_i Y_i$, $\bar{S}$ is the unweighted mean labor share of the firms in the industry, and $\bar{\omega}$ is the unweighted mean value-added share.[34]

Consider the change in the aggregate labor share between two time periods, $t = 0$ and $t = 1$. Abstracting from entry and exit, we write the Olley-Pakes decomposition as: [35]

$$\Delta S = S_1 - S_0 = \Delta \bar{S} + \Delta \left[\sum (\omega_i - \bar{\omega})(S_i - \bar{S})\right]. \tag{2.4}$$

Following Melitz and Polanec (2015), we augment this decomposition with terms that account for exit and entry:

[33]The Melitz and Polanec (2015) generalizes the Olley and Pakes (1996) productivity decomposition to allow for firm entry and exit.

[34]The weight $\omega_i$ used in these calculations is the denominator of the relevant labor share measure. Thus, within manufacturing, when we consider decompositions of the payroll-to-value-added ratio, we use the value-added share as the firm's weight. In all other decompositions, we use the payroll-to-sales ratio, and use the firm's share of total sales as the firm's weight.

[35]Note that five year changes in the Census data form the bulk of our analysis.

$$\Delta S = \Delta \bar{S}_S + \Delta \left[ \sum (\omega_i - \bar{\omega}) \left( S_i - \bar{S} \right) \right]_S + \omega_{X,0} \left( S_{S,0} - S_{X,0} \right) + \omega_{E,1} \left( S_{E,1} - S_{S,1} \right). \quad (2.5)$$

Here, subscript $S$ denotes *survivors*, subscript $X$ denotes *exiters* and subscript $E$ denotes *entrants*. The variable $\omega_{X,0}$ is the value-added weighted mean labor share of exiters (by definition all measured in period $t_0$) and $\omega_{E,1}$ is the value-added weighted mean labor share of entrants (measured in period $t_1$). The term $S_{S,t}$ is the aggregate labor share of survivors in period $t$ (i.e. firms that survived between periods $t_0$ and $t_1$), $S_{E,1}$ is the aggregate value-added share of entrants in period $t_1$, and $S_{X,0}$ is the value-added share of exiters in period $t_0$. One can think of the first two terms as splitting the change in the labor share among survivors into a within-firm component, $\Delta \bar{S}_S$, and a reallocation component, $\Delta \left[ \sum (\omega_i - \bar{\omega}) \left( S_i - \bar{S} \right) \right]_S$, which reflects the change in the covariance between firm size and firm labor shares for surviving incumbents. Meanwhile, the last two terms account for contributions from exiting and entering firms.

**Main Decomposition Results**

In Figure 2.7, we show an illustrative plot for the Melitz-Polanec decomposition calculated for adjacent five-year periods for manufacturing payroll over value-added, cumulated over two 15-year periods: 1982-1997 and 1997-2012. The labor share declined substantially in both periods: -10.42 percentage points between 1982 and 1997 and -5.65 percentage points between 1997 and 2012. Consistent with the superstar firm framework, the reallocation among incumbents ("between") was the main component of the fall: -8.24 percentage points in the early period and -4.90 percentage points in the later period. While the within-firm component is negative over both periods, the reallocation component among incumbents is three (1982-1997) to ten (1997-2012) times as large as the within-firm component. Notably, the within-incumbent contribution to the falling labor

share is only 0.4 percentage points during 1997-2012, meaning that for the unweighted average incumbent firm, the labor share fell by under half a percentage point over the entire 15 year period.

The reallocation term captures changes in activity among incumbent firms, but there is an additional reallocation effect coming from entry and exit. Exiting firms contribute to the fall in the labor share over both periods, by -2.4 and -2.8 percentage points, respectively, in the early and later time interval. The fact that the high labor share firms within a sector are disproportionately likely to exit is logical since such firms are generally the less profitable. Conversely, the contribution from firm entry is positive in both periods: 2.7 and 2.4 percentage points in the early and later period respectively. New firms also tend to have elevated labor shares, presumably because they set relatively low output prices and endure low margins in a bid to build market share (see Foster, Haltiwanger and Syverson 2008, 2016 for supporting evidence from the Census of Manufacturers). Since the contribution of entry and exit is broadly similar, these two terms approximately cancel in our decomposition exercise.

Table 2.4 reports the decompositions of labor share change in manufacturing for each of the individual five-year periods covered by the data. In the first five columns we detail the payroll to value-added results. Reallocation among incumbent firms contributes negatively to the labor share in every five-year period whereas within-firm movements contribute *positively* in two of the six time periods (1987-1992 and 2007-2012). The right panel of Table 2.4 repeats these decompositions using the broader measure of compensation over value-added, and shows that the patterns are even stronger for this metric: almost all of the fall in the labor share can be explained by a between-incumbent reallocation of value-added. The last row shows, for example, that the compensation share fell by 18.5 percentage points between 1982 and 2012 and that essentially all of this change is accounted for by reallocation among incumbent firms. By contrast, the unweighted labor share for incumbents fell by only 0.24 percentage points.

The finding that the reallocation of market share among incumbent firms contributes negatively to the overall labor share generalizes to all of the six sectors that we consider.[36] Figure 2.8 plots the Melitz-Polanec decomposition for each sector cumulated now over the entire sample period for which data is available (e.g., 1982-2012 for manufacturing, but only 1992-2012 for finance and utilities/transportation). Table 2.5 reports the decompositions over five-year periods underlying the sample totals plotted in Figure 2.8. Recall that we do not have firm-level value-added data outside of manufacturing, so this analysis decomposes payroll over sales using a firm's sales share as its weight. As in Figure 2.7 for payroll over value added within manufacturing, the total contribution of market share reallocation among incumbent firms (4.54 percentage points) is almost three times as large as the within-firm component (1.71 percentage points) for payroll over sales. Also echoing the findings in manufacturing, we find that the between-incumbent reallocation effect contributes strongly to the decline in the payroll share in each of the other five sectors except services where the entry component dominates. By contrast, the within-incumbent contribution is *positive* in all sectors except for manufacturing. Indeed, this is exactly what is predicted by the model in Section 2.2, as in that model, the unweighted average labor share is the flip side of the unweighted average markup. Proposition 2 shows that for sufficiently skewed firm productivity distributions (specifically, a log-convex distribution), an increase in the toughness of competition reduces margins for individual firms, but reallocates so much market share to firms with high markups and low labor shares that the aggregate labor share falls and the aggregate markup rises.

### Robustness of the Decomposition Analysis

We have subjected the decomposition findings to a large number of robustness tests, some of which are reported below (and others considered in the online appendix). A key feature of the above decomposition analysis is that it is performed at the level of the entire firm

[36]The level of the payroll to sales ratio differs substantially across sectors due in part to differences in intermediate input costs (see Figure 2.3), and we thus implement decompositions separately by sector.

(within a sector). While this is appealing because it closely aligns with the model, there is a potential complication as entry and exit can occur through firm merger and acquisition activity rather than de novo start-ups or closing down of establishments.[37] Additionally, since firms may span multiple industries, some of the reallocation we measure in the baseline decomposition may reflect shifts of firm activity across four-digit industries.

In order to explore the importance of the specific firm definition in driving the decomposition results, we repeat the decomposition analysis at the both the establishment level and at the firm-by-four-digit SIC industry.[38] In both cases, we find qualitatively similar patterns to our main estimates, reflecting the fact that the overwhelming number of firms have only a single establishment. In both cases, exit makes a larger contribution, but the sum of entry and exit is still small compared to the reallocation term.[39]

We also perform the decomposition at 15-year intervals rather than five-year intervals. The pattern of findings persists, even though the definition of a "survivor" is now changed to comprise only firms that survive at least 15 years (rather than the baseline of five years).

In order to more concretely assess the magnitude of the between-industry reallocation in our baseline firm-level decomposition, we perform an extended decomposition that explicitly distinguishes between-industry versus within-industry but between-firm components. We first use a standard shift-share technique to decompose the overall change in the labor share into between-industry $\sum_j \left(\widetilde{S}_j \Delta\omega_j\right)$ and within-industry $\sum_j \left(\widetilde{\omega}_j \Delta S_j\right))$ components:

[37]For example, when a firm is taken over, its establishments are reallocated to those of the the acquiring firm, this leads to an "exit" of the acquired firm even though its establishments do not exit the economy. On the other hand, if an incumbent firm creates a new greenfield establishment, this will not be counted as firm entry.

[38]This is the same definition used in Tables 2.2 and 2.3 linking changes in labor shares to changes in industry-level concentration.

[39]Additionally, motivated by concerns over the accuracy of firm identifiers in the Census panel (see Haltiwanger, Jarmin and Miranda, 2013), we applied a looser definition of what constitutes an ongoing firm by using the identity of ongoing establishments. Specifically, if an ongoing establishment experiences a change in firm identifier, we reclassify the firm to be the same if the "new" firm contains all the establishments of a previously exiting firm. Our results are again almost identical to those in Tables 2.7 and 2.8.

$$\Delta S = \sum_j \left(\widetilde{S}_j \Delta\omega_j\right) + \sum_j \left(\widetilde{\omega}_j \Delta S_j\right). \tag{2.6}$$

Here, $\widetilde{S}_j$ is the time average of the (size-weighted mean) labor share in industry $j$ ($S_j$) over the two time periods, and $\widetilde{\omega}_j$ is the industry size share (e.g. value added share of industry $j$ in total manufacturing value added), $\omega_j$ , averaged across the two time periods. We then use the industry specific version Equation (2.5) to split up within-industry $\sum_j \left(\widetilde{\omega}_j \Delta S_j\right)$ contribution into its four parts (technical details are in Appendix D).

We show the components of this five way decomposition in the online appendix. Looking across all sectors, it is clear that the main qualitative finding that the fall in the labor share is dominated by a *within-industry* between-firm reallocation is robust to this alternative decomposition. In some segments, the *between-industry* contribution increases the labor share (e.g. services, utilities and transportation, and finance). In the others, it is relatively small compared to the reallocation term that operates between firms within an industry. For example, in the wholesale sector, the between-industry term is -0.3 as compared to -5.7 for reallocation between firms. In manufacturing, the between-industry term is -0.4 for payroll over sales; -2.2 for payroll over value-added and -2.9 for compensation over value-added, as compared to a total (between-firm reallocation contribution) change of -6.7 (-5.5); -16.1 (-7.9), and -18.5 (-10.3) respectively. These results are also in line with Kehrig and Vincent (2018), who extensively analyze changes in the labor share in manufacturing using full distributional accounting techniques. Like us, they find that the between-firm reallocation term dominates in accounting for the aggregate fall in the labor share.

### 2.4.3 Between-Firm Reallocation is Strongest in Concentrating Industries

We have established that across most of the U.S. private-sector economy, there has been a fall in the labor share and a rise in sales concentration; that the fall in the labor share is

greatest in the four digit industries where concentration rose the most; and that the fall in labor share is primarily accounted forby between-firm reallocation of value-added sales rather than within-firm declines in labor share. Figure 2.9 examines the fourth prediction of the superstar firm model: the reallocation component of falling labor share should be most pronounced in the industries where concentration is differentially rising. This occurs in our model because superstar firms capture market share through their high relatively high productivity, meaning that they are aggressive competitors. If, contrary to our superstar firm hypothesis, rising concentration reflects weakening competition, we would instead expect to see a general rise in mark-ups, a rise in profit shares, and a fall in labor shares that is common across firms within an industry.

We explore the model's prediction in Figure 2.9 by plotting the relationship within each sector between changes in industry concentration and each of the four components of the Melitz-Polanec decomposition. In the figure, the upper bars report the coefficient estimates and standard errors from regressions of the reallocation component of the fall in the labor share (based on Table 2.5) on the change in the CR20. The bars directly underneath report the estimates that result from regressing the within-incumbent component of the change in the labor share on the change in concentration. The remaining two bars show the corresponding estimates for the firm entry and exit components. The online appendix reports the corresponding regressions underlying Figure 2.9 alongside analogous estimates using our two alternative measures of concentration. The pattern of results in Figure 2.9 is consistent across all sectors: the tight correlations between rising concentration and falling labor share reported earlier in Figure 2.6 are driven by the reallocation component. Specifically, the between-incumbent reallocation component shows up as negative and significant in all sectors, indicating that rising concentration predicts a fall in labor-share through between-incumbent reallocation. Conversely, the coefficients on the within-firm component are small, generally insignificant, and occasionally positive. Firm entry and exit correlate with concentration differently across sectors, but these

components always play a small role compared to the between-incumbent reallocation component. The results provide further evidence, consistent with the superstar firm hypothesis, that concentrating industries experienced a differential reallocation of economic activity towards firms that had lower labor shares.

A further extension we considered was to implement our decompositions of changes in the labor share into between- and within-firm components using alternative techniques such as a traditional shift-share analysis, as in Bailey, Hulten and Campbell (1992), or a modified shift-share approach where the covariance term is allocated equally to the within- and between-components, as in Autor, Katz and Krueger (1998). We implemented a variety of such approaches and performed decompositions such as those underlying Figure 2.8. We continue to find a large role for the between-firm reallocation component of the fall in the labor share but the within-firm component becomes more important as well. In contrast to Figure 2.9, we also find for the shift-share decompositions that concentration loads significantly on the within-firm component. These shift-share decompositions give greater weight to the within-firm changes of *initially larger* firms than do the Olley-Pakes and Melitz-Polanec methodologies, where the within component is simply the unweighted mean of within-firm changes. The shift-share models therefore suggest that within-firm declines in labor share make some contribution to the aggregate decline in labor share, but this within-firm contribution primarily comes from larger firms. In short, increases in concentration are associated with decreases in labor share among the largest firms.[40]

### 2.4.4 Markup Analysis

Our imperfect competition approach emphasizes that at the firm level, the labor share depends on the ratio of the output elasticity of labor to the markup (equation 2.1), while

[40]The covariance term in the shift-share analysis $\left(\sum [\Delta\omega_i \Delta S]_S\right)$ is a non-trivial component although it does not seem related to increases in concentration. This appears to be related to outliers, to which the double difference in the covariance term is particularly sensitive.

the economy-wide labor share depends on how market shares are distributed across these heterogeneous firms. A corollary of this approach is that for stable elasticities, markups should move in the opposite direction of labor shares. The formal model in online Appendix A shows that the conditions under which the aggregate labor share falls are the same as those for obtaining a rise in the markup.

**Measuring Markups**

To empirically test this implication of the model, we must estimate markups, which is more challenging than measuring the labor share. Following the literature (e.g. de Loecker, Eeckhout and Unger, 2018) we can estimate markups by re-arranging and generalizing (equation 2.1):

$$m_{it} = \left(\frac{\alpha_{it}^{v}}{S_{it}^{v}}\right) \tag{2.7}$$

where $S_{it}^{v} = \left(\frac{W_{it}^{v} X_{it}^{v}}{P_{it} Y_{it}}\right)$ is the share of any variable factor of production $X_{it}^{v}$ (with factor price $W_{it}^{v}$) in total sales: $\alpha_{it}^{v}$ is the output elasticity with respect to factor $v$. This is a very general result and assumes only that firms cost minimize; it therefore allows for non-constant returns, general technologies, etc. (see Hall, 1988, 2018). Although factor shares ($S_{it}^{v}$) are in principle observable, elasticities ($\alpha_{it}^{v}$) are not. One simple way to recover the elasticity is to assume that the production function exhibits constant returns to scale, in which case we can measure $\alpha_{it}^{v}$ by the share of factor $v$ in total costs ($\sum_f W_{it}^{f} X_{it}^{f}$). In this case the markup formula becomes:

$$m_{it} = \left(\frac{P_{it} Y_{it}}{\sum_f W_{it}^{f} X_{it}^{f}}\right) \tag{2.8}$$

where $f$ indicates we are summing up over the costs of all factors $f$ whether quasi-fixed (like capital) or quasi-variable (like labor). Equation (2.8) is simply the ratio of sales to total costs, which is used for measuring the markup by Antras, Fort and Tintelnot (2018)

among others. We call this the "accounting approach" as it does not rely on econometric estimation. A second approach to recovering markups is to estimate $\alpha_{it}^{v}$ from a production function as recommended by de Loecker and Warzynski (2012). This relaxes the constant returns assumption implicit in the accounting approach but does require econometric estimation of a production function.

A practical data challenge for both the accounting or econometric approaches is that in the Economic Census, data on capital are unavailable outside of manufacturing, and data on intermediate input usage are sparse. Consequently, we focus on the Census of Manufactures where richer data are available. Appendix B details how we estimate plant-level production functions using methods due to *inter alia* Levinsohn and Petrin (2003) and Ackerberg, Caves and Frazer (2015). In all cases we allow all parameters to freely vary across the 18 two-digit SIC manufacturing industries and (in some specifications) we also allow the parameters to vary over time and across plants (e.g. using a translog production function). The plant markups are aggregated to the firm level using value added weights (for multi-plant firms).

**Results**

We summarize the results of this exercise here, with further details provided in online Appendix B. Before exploring trends, figures in the online Appendix confirms that larger firms have higher markups, no matter how they are estimated, a finding that is consistent with standard IO models. In Figure 2.10, we present the trends in aggregate markups (where firm markups are weighted by value added) in red triangles across four alternative ways of calculating markups. Alongside the weighted markup, we also presents the median markup (green diamonds) and unweighted average markup (blue circles). Panel A uses the accounting in equation 2.8 and Panel B calculates markups using the Levinsohn and Petrin (2003) method of estimating a Cobb-Douglas production function. Panel C does the same as Panel B, but uses the Ackerberg, Caves and Frazer (2015) method of

estimating a Cobb-Douglas. Panel D continues using the Ackerberg, Caves and Frazer (2015) method but generalizes Panel C to estimate a translog production function.

Although the exact level of the markup differs across the four panels of Figure 2.10, the broad patterns are quite similar. First, the weighted average mark-up always exceeds the unweighted markup (and the unweighted mean is above the median), reflecting the fact that larger firms have higher markups, as noted above. Second, aggregate markups have risen considerably over our sample period. For example, in Panel B the weighted markup has risen from about 1.2 in 1982 to 1.8 in 2012, similar to the finding in De Loecker, Eeckhout and Unger (2018) using publicly listed firms in Compustat across all sectors.[41] Third, across all methods, the aggregate markup has risen much more quickly than that of the typical firm. Indeed, median markups are flat or even falling in some specifications (e.g. Panel D). This implies that rising average markups are driven by the changing market shares and markups of the largest firms, a pattern consistent with the decomposition analysis of labor shares discussed above. This pattern again underscores the centrality of superstar firms for the evolution of the markup, which is consistent with the findings in de Loecker et al (2018) and Baqaee and Fahri (2018). We further explore the evolution of markups and subject our findings to many other robustness tests in online Appendix B.[42]

### 2.4.5 Concentrating Industries have Higher Innovation and Productivity Growth

The fifth prediction of the superstar model from Section 2.2 is that rising concentration is more prevalent in dynamic industries that exhibit faster technological progress, since

[41] Their Figure 11(a) suggests that the manufacturing markup rose from about 1.2 to 1.6. They also present production function based estimates of markups for Compustat, but they do not implement this method in the Census data, so our results are novel with respect to theirs. They do implement the accounting approach in the Census of Manufactures, although they use a slightly different method of calculating capital costs, employing estimates of cost shares from Foster, Haltiwanger and Syverson (2008). By contrast we use the approach of Antras et al (2017). Despite our methodological differences, it is reassuring that the markup estimates tell the same story.

[42] Edmond et al. (2015) argue that input-weighted markups are a better welfare related measure than the output-weighted markups shown here (as in de Loecker et al, 2018). Using input weights gives the same qualitative patterns as Figure 2.10, but the increase in the aggregate markup is smaller over time.

our superstar firm framework emphasizes technological and competitive forces as driving the trend towards greater concentration and a reallocation of output towards high-productivity and low labor share firms. Before examining the industry-level relationship between the change in concentration and productivity, we first present underlying firm-level evidence that larger firms are more productive. For all firms in manufacturing, we measure firm-level productivity using the estimates of TFP that result from the estimated production function described above in Section 2.4.4. The online appendix shows that large firms in manufacturing are more productive, regardless of how we measure TFP. The stylized fact that larger firms have higher TFP and lower labor shares is consistent with the model in online Appendix A and underpins the industry-level prediction relating concentration and dynamism.

Moving to the industry level, we explore the relationship between dynamism and industry concentration by employ two commonly used measures of technical change, patent-intensity and productivity growth, along with other relevant industry characteristics. Table 2.6 displays regressions where the dependent variable is the (five-year) growth in concentration and the explanatory variables are proxies for industry dynamism.[43] Panel A focuses on the manufacturing sector where the data are richer, while Panel B reports results for all six sectors.

The first row shows that there is a significant and positive relationship between the growth of concentration and the growth of patent intensity across all three measures of concentration. The second row of Table 2.6 shows that industries that had faster growth in labor productivity (as measured by value-added per worker) had larger increases in concentration. This regression is similar to the reciprocal of the labor share (payroll over value added) regressions that we presented in subsection 2.4.1. There are at least two differences, however. First, the denominator of labor productivity is the number of workers whereas the denominator of the labor share measure is total payroll. Second, and more

[43] All regressions are weighted by the initial size of the industry, include year dummies and cluster the the standard errors by industry as in Tables 2.2 and 2.3.

importantly, value-added is deflated by an industry-specific producer price index in the productivity measure in Table 2.6, but it is simply equal to the nominal labor share in Table 2.2. This is important as increased concentration may be associated with higher prices, meaning the correlation with the nominal, non-deflated labor productivity measures could be driven by higher markups rather than increased productivity. In fact, there seems to be little systematic correlation between increased concentration and higher prices (see Ganapati, 2018; Peltzman, 2018), but a rather strong relationship with real labor productivity. Of course, this relationship could still just be due to faster growth in input growth in these concentrating industries. Indeed, we do find the concentrating industries have faster growth in the capital-worker ratio, as is shown in the third row of Table 2.6. Nevertheless, even when we control for output increases arising from five possible factor inputs (labor, structures capital, equipment capital, energy inputs and non-energy material inputs) in our TFP measure in the fourth row, we find a significantly positive correlation between concentration growth and TFP growth.[44]

In Panel B of Table 2.6, we repeat these specifications for all six sectors. Due to the absence of value-added data outside of manufacturing, we measure productivity as output per worker. Despite this limitation, we find a positive relationship across all 18 regressions, with 12 coefficients significant at the five percent level, two at the ten percent level, and the remaining four insignificant. In net, we find that the industries exhibiting rising concentration are also those that are more dynamic as measured by innovative output and productivity growth.[45]

[44]This TFP measure is measured as a Solow-style residual based on deducting the cost-weighted inputs from deflated output. We also replicated these regressions using TFP measured from industry specific production functions identical to those we used when estimating price-cost markups as detailed in subsection (2.4.4) and Appendix B. The qualitative results were unsurprisingly similar, since all TFP measures are strongly and positively correlated with each other.

[45]This evidence is consistent with the cross OECD evidence in Autor and Salomons (2018), who find that the labor share fall was greater in those industries where TFP growth had been most rapid. In our data if we regress the change in the labor share on five-factor TFP growth we obtain a coefficient (standard error) of -0.078 (0.018) in a specification the same as row 1 of Table 2.2 without concentration and of -0.092 (0.021) if we add four digit industry trends (i.e. in a specification the same as row 5 of Table 2.2 without concentration).

The above correlation between concentration and productivity supporting the superstar mechanism implies that the reallocation of sales and value-added towards the most productive firms in each sector should contribute to overall productivity growth. Yet it is widely acknowledged that aggregate productivity growth in the U.S. and Europe slowed in the early 1970s, rebounded modestly in the mid-1990s, and then slowed again in the mid-2000s (Syverson, 2017). Thus, if the superstar mechanism is operative, this implies that there are countervailing forces that mute this effect. One possibility is that there has been a slowdown of productivity diffusion from industry leaders to laggards.[46] A second possibility is that underlying productivity differences between superstar firms and others are not economically large, but that changes in the economic environment have nevertheless yielded substantial reallocation of market shares towards competitors with modest productivity advantages. This would generate superstar effects without large gains in aggregate productivity. Reconciling the aggregate productivity puzzle remains an important topic for further study that we do not claim to resolve here.

### 2.4.6 Superstar Firm Patterns are International

The final empirical implication of the superstar framework that we test is that the patterns that we document in the U.S. should be observed internationally. Karabarbounis and Neiman (2013) and Piketty (2014) have documented that the fall in the labor share is an international phenomenon, although the speed and timing of the changes differ across

[46] Andrews, Criscuolo, and Gal (2015) examine firm-level data in 24 OECD countries between 2001 and 2013 and find that while productivity growth has been robust at the global productivity frontier (referring to the most productive firms in each two-digit industry), productivity differences have widened between these frontier firms and the remainder of the distribution. These authors attribute this widening to a slowdown in technological diffusion from frontier firms to laggards, and infer that leading firms have become better able to protect their competitive advantages, which in turn contributes to a slowdown in aggregate productivity growth. Andrews, Criscuolo, and Gal (2015) do not look directly at labor shares, but a slowdown in technological diffusion could be a reason for the growth of superstar firms. We investigated this possibility by examining a measure of technology diffusion based on the speed of patent citations. Consistent with the hypothesis of Andrews, Criscuolo, and Gal, (2015), we find that in industries where the speed of diffusion (as indicated by a drop in the speed of citations) has slowed, concentration has risen by more and labor shares has fallen by more. For example, in industries where the percent of total citations received in the first five years was 10 percentage points lower, concentration rose by an extra 3.3 percentage points.

countries. Using industry and firm-level data from various OECD countries, we document that the superstar firm patterns relating rising concentration to falling labor shares found in the U.S. are prevalent throughout the OECD. Our superstar firm framework emphasizes global technological forces for the trend towards greater concentration and a reallocation of output towards high-productivity and low labor share firms. As discussed in the Introduction, the precise mechanisms through which this occurs may include platform competition, adoption of more intangible capital by leading firms, or by toughening market competition, as formalized in the model in online Appendix A. An alternative interpretation of these patterns is offered by Dottling et al (2018), who argue that *weakening* U.S. antitrust enforcement has led to an erosion of product market competition. The broad similarity of the trends in concentration, markups and labor shares across many countries that we document below casts some doubt on the centrality of these institutional explanations. Indeed, as Dottling et al (2018), emphasize antitrust enforcement has, if anything, strengthened in the European Union—and yet the labor share in OECD countries has seemingly fallen *despite* this countervailing force.

**Concentration in the OECD**

Obtaining comprehensive data on changes in sales concentration over time across countries is challenging. The most comprehensive source for such an analysis is "Multiprod", a firm level database that the OECD produces in cooperation with the national statistical agencies in many countries. By design, these data are broadly similar to the U.S. Economic Census. Bajgar et al (2018a) find that between 2001 and 2012, industry-level concentration levels rose within the ten European countries where comprehensive data are available. They estimate that the share of the top decline of companies (measured by sales) increased on average by two percentage points in manufacturing and three percentage points in non-financial market services. Because some of these European economies are small and heavily integrated in the broader EU economy, the authors also look at

an alternative market definition based on considering Europe as a single market. Under this definition, they also find that concentration levels have risen, akin to findings for the United States. [47]

### Correlation of Industry Labor Shares

Figure 2.1 documented the pervasive decline in the labor share across several OECD countries. Looking below these time series relationships, we perform a cross-national industry-level and firm-level analysis by exploring the correlation of the labor share (measured in levels) across the 32 industries that comprise the market sector. The online appendix reports these correlations for each country over the 1997-2007 period (where the data are most abundant). The correlation is high in all cases, with average correlation coefficients between 0.7 and 0.9. Panel B correlates the *change* in labor shares by country pairs and reports the average correlation for each country as well as the fraction of the country's pairwise correlations that are negative. The correlations in changes are weaker than those in levels, as expected, but the bulk of the evidence still indicates that declines in the labor share tend to occur in the same industries across countries: the average correlation is positive for each country, and there is a positive correlation across industries between country pairs in over three-quarters of all cases (51 of 66). The correlation matrices underlying these summary tables are reported in the online appendix.

[47] Dottling, Gutierrez and Philippon (2017) have argued the opposite—that concentration has been falling in the EU. Bajgar et al (2018b) trace the discrepancy to Dottling et al's (2017) use of BVD Orbis data to calculate concentration rather than the near-population Multiprod data used by the OECD. While Orbis does a reasonable job of tracking sales in the largest firms, it has quite incomplete coverage of small and medium sized firms in many countries, especially in the late 1990s and early 2000s, which then improves thereafter. Consequently, Orbis overestimates overall industry sales growth after the early 2000s as it includes the increase in industry sales arising through expanding sample coverage. When using Orbis for both the numerator and denominator of concentration, Bajgar et al (2018b) reproduce Dottling et al's finding of falling EU concentration. But when using the more consistent industry size measure from population data as the denominator for industry sales, they reverse this result and report rising concentration.

### Industry Labor Shares and Concentration

We next examine the relationship between the change in industry-level labor shares and concentration across countries. Although we do not have access to the equivalent of the Census Bureau firm-level data for all countries outside of the United States, we can draw on cross-national, industry-level data for a shorter period from the COMPNET database. COMPNET, developed by the European Central Bank, is originally a firm-level data set constructed from a variety of country-specific sources through the Central Banks of the contributor nations. The public use version of these data that we analyze are collapsed to the industry-year level. COMPNET reports measures of both the labor share and of industry-level concentration, defined as the fraction of industry sales produced by the top ten firms in a country. We estimate equation (2.2) in five-year (2006-2011) and ten-year (2001-2011, for countries where data are available) long differences separately across all of the 14 countries in the database. These estimates, reported in Appendix Table **??**, finds that in 12 of 14 countries, there is a negative relationship over the five-year first-difference between rising concentration and falling labor share, as predicted by the superstar firm model. In the longer ten-year difference model in column 2 (for which fewer countries are available), all countries but Belgium also show a negative relationship. These coefficients are imprecise, however, and the majority are insignificant. In the 10-year difference specification, five of the 10 coefficients are negative and significant at the 10% level or greater, while four additional countries have negative but insignificant coefficients.

### Firm-Level Decompositions

To explore the role of between-firm reallocation in falling labor share in cross-national data, we turn to data from BVD Orbis, which is currently the best available source for comparable, cross-national *firm-level* data. Orbis is a compilation of firm accounts in electronic form from essentially all countries in the world. Accounting regulations and Orbis coverage differ across countries, however, so we confine the analysis to a set of six OECD

countries for which reasonable quality data are available for the 2000s. For these countries, we decompose changes in labor share into between- and within-firm components, using the earliest five-year periods available for which Orbis has comprehensive data. These are the years 2003-2008 for the UK, Sweden and France, and 2005-2010 for Germany, Italy and Portugal. In all six countries, we see a decline in the aggregate labor share of value-added over this period. The online appendix reports the Olley-Pakes decomposition for the manufacturing sector for all six countries.[48] As in the more comprehensive U.S. data, it is the between-firm reallocation component that is the main contributor to the decline in the labor share in all countries. This reallocation component is always negative and in all cases larger in absolute magnitude than the within-firm component. In three of six countries, this within-firm component is positive.

### Markups in Different Countries

There has also been considerable recent work on markups using firm-level data across countries (Calligaris et al, 2018; de Loecker and Eeckhout, 2018). These findings appear consistent with the patterns that we document for the U.S., with markups being the flip-side of the pattern of the labor share. On average across countries, the weighted average markup has risen. This pattern appears largely driven by a reallocation of sales and value-added towards firms with high markups (low labor shares).

### Summary on International Evidence

Although the international data are not as rich and comprehensive as those available for the United States, the pattern of cross-national findings mirrors the evidence from the more detailed U.S. data: (i) concentration has generally risen across the OECD, as with the US; (ii) the decline in the labor share has occurred in broadly similar industries across countries; (iii) the industries with the greatest increases in concentration exhibited the

[48] We focus on manufacturing as measurement of the labor share is more reliable for this sector.

sharpest falls in the labor share; and (iv) the fall in the labor share is primarily accounted for by the reallocation of value-added or sales between firms rather than within-firm labor share declines; and of course, (v) the rise in markups can be read as the flip-side of the fall in labor shares. We read the international evidence as broadly consistent with the hypothesis that a rise in superstar firms has contributed to the decline in labor's share throughout the OECD.

### 2.4.7 Magnitudes

The previous sections have shown evidence that is qualitatively in line with the seven empirical predictions of the superstar firm framework, most importantly by documenting the central role of between-firm reallocation in (proximately) driving the labor share decline. Ideally, we would like to answer the question of how *much* of the fall of the labor share is due to the underlying change in competitive conditions that gives rise to superstar firms. In the absence of an explicit quantitative macro model, it is difficult to precisely answer this question.[49]

To shed some light on the magnitudes, we performed two simple exercises. First, we take a model-based approach. We take logs of the size-aggregated version of equation (2.1) and write the aggregate labor share change as a function of the change in the weighted average markup and a residual term, $\varsigma$, $\Delta \ln S = -\Delta \ln m + \varsigma$. The Cobb-Douglas production function underlying equation (2.1) implies that $\varsigma = \Delta \ln \alpha^L$ , i.e. the part of the labor share unexplained by the markups is due to the changing output elasticity of labor.[50] We can implement this approach only for manufacturing, where we have the data necessary to properly measure mark-ups (see subsection 2.4.4 above). Using Table

[49]Karabarbounis and Neiman (2018) rigorously quantitatively evaluate alternative macro-models of the labor share decline.

[50]See Nekarda and Ramey (2013) for what determines the labor share under more general models. For example, if the production function is CES then $\varsigma = \Delta \ln \alpha^L + \Delta \left(\frac{1}{\sigma} - 1\right) \ln \left(PY/B^L L\right)$ where $\sigma$ is the elasticity of substitution between labor and capital and $B^L$ is a labor augmenting efficiency parameter. If there are overhead labor costs, the residual will also include the ratio between the marginal wage and the average wage.

2.4, the proportionate fall in the labor share of value added ($\Delta \ln S$) is 40 percent (a 16.1 percentage point change divided by a 41 percent initial level). The percentage change in the markup ($\Delta \ln m$) depends on which measure we use. Using the accounting method in Panel A of Figure 2.10, there is a 17 percent rise in the markup (0.22/1.31) implying that we account for about two-fifths (17/40) of the labor share change.[51] By contrast, using the production function based measures of the markup, we account for essentially *all* of the labor share change (e.g. in Panel B, the growth of the markup is 50 percent (0.6/1.2), greater than the change in the labor share).

A second approach follows directly from our regression models. We can use the estimates of equation (2.2) to assess what would have been the change in the labor share had concentration not risen. The predicted aggregate change in the labor share over the whole 2012-1982 period is $\Delta\widehat{S} = \sum_k \left(\omega_k \hat{\beta}_k \Delta \text{CONC}_k\right)$ where $k = 1, ..., 6$ indicates the broad sector, $\hat{\beta}_k$ is the estimated coefficient from equation (2.2), and $\omega_k$ is the relative size of the sector (value added weights from the NIPA). Excluding the financial sector, the predicted change in the labor share of sales (using the change in the CR20's from (2.4) is 0.97 percentage points, as compared to an overall fall in the labor share of 1.86 percentage points. By this measure, rising concentration can account for about half of the fall in the labor share ($52\% = 0.97/1.86$).[52] Looking at this calculation sector-by-sector, we predict that the labor share of sales should have fallen in all sectors, especially in the post 2000 period. For example, although we account for only a tenth of the fall in the labor share of sales in manufacturing over the whole period, we account for over a third of the 1997-2012 change.[53]

[51] Although part of the aggregate change in the markup may be due to markup growth at smaller firms, we showed in subsection 2.4.4 that the vast majority of the aggregate markup growth is due to the superstar mechanism—that is, changes at the upper tail.

[52] If we additionally include the financial sector in these aggregate calculations, we account for even more of the overall change. Here, we predict an even larger labor share fall (-1.6 percentage points) since there has been a large increase in concentration in finance. As noted above, we are cautious about using this sector given the data concerns over the Census sales measures, and hence we prefer the more conservative non-financial estimates.

[53] This is partly due to a faster rise in concentration after 1997 (see Figure 2.4) and partly due to the coeffi-

All these estimates are highly speculative. The first, markup-based approach, probably overestimates the superstar contribution because the labor share implicitly enters some of the calculations of the markup. The second, regression-based approach, may underestimate the superstar effect as concentration is a coarse proxy. Nevertheless, both methods suggest that the key empirical relationships that we highlight in the paper appear economically large as well as statistically significant.

## 2.5 Further Descriptive Evidence on Superstar Firms

The previous section documented empirical support for the main empirical predictions of the superstar firms framework derived in Section 2.2. This section further explores the relationship between the rise of superstar firms and other economic phenomena of the last several decades.

### 2.5.1 Import Exposure and Superstar Firms

Using data from both manufacturing and non-manufacturing industries, Elsby, Hobijn and Sahin (2013) find a negative industry-level association between the change in the labor share and growth of total import intensity.[54] They conclude that the offshoring of the labor-intensive components of U.S. manufacturing may have contributed to the falling domestic labor share during the 1990s and 2000s. Following their work, we explore the relationship between changes in labor's share and changes in Chinese import intensity. The online appendix reports regressions of changes in industry-level outcomes in U.S. manufacturing on changes in Chinese imports intensity using both OLS models and 2SLS mod-

cient on concentration rising over time. From 1997 to 2012, the CR20 in manufacturing went up by around 6 percentage points and the labor share fell by around 6 percentage points. The average coefficient relating the change in concentration to the change in labor share in manufacturing over this period was $-0.345$, implying that concentration explained $\frac{-0.34\times 6}{6} \times 100 = 34\%$ of the fall in the labor share in manufacturing over this period.

[54]They define total import intensity using the 1993-2010 input-output tables as the percentage increase in value-added needed to satisfy U.S. final demand were the U.S. to produce all goods domestically.

els that apply the Autor, Dorn and Hanson (2013) approach of instrumenting for import exposure using contemporaneous import growth in the same industries in eight other developed countries. We further report results both including and excluding the post-2007 Great Recession. The results corroborate the well-documented finding that industries that were more exposed to Chinese imports had greater falls in sales, payroll and value-added than other sectors (significantly so in almost all cases). The next three columns find a positive correlation between the growth of Chinese import penetration and the rise of industry concentration, although this relationship is imprecisely estimated. The last two columns find that an increase in Chinese imports predicts a *rise* in industry labor share (though this relationship is often insignificant). While this result is unexpected in light of Elsby, Hobijn and Sahin (2013), it is implied by the estimates in columns (1) through (3). Specifically, because the negative effect of rising Chinese import exposure on industry payroll is smaller in absolute magnitude than its negative effect on industry value-added and industry sales, the labor share of sales and value-added tends to rise with growth of industry import exposure.[55]

### 2.5.2 Compustat Analysis: Publicly Listed Superstar Firms

Although it has the advantage of being comprehensive, Census data have the disadvantage that we are not permitted to illustrate the key fact patterns with specific examples (since the identity of individual companies is confidential). In addition, our Census data do not report on the international activity of these superstar firms. To provide these examples and explore the international scope of these superstar firms, we turn to Compustat

[55] A key difference with Elsby, Hobijn and Sahin (2013) is that they pool data from both manufacturing and non-manufacturing industries whereas we analyze the impact of trade exposure on manufacturing only. Using their approach, we are able to replicate the finding of a negative association between rising imports and falling labor share. But this negative relationship is eliminated when we include a dummy variable for the manufacturing sector. This pattern likely reflects the facts that (1) the fall in the labor share has been greater in manufacturing than in other sectors; and (2) manufacturing is more subject to import exposure then non-manufacturing. Within manufacturing, cross-industry variation in import exposure appears to have little explanatory power for the fall in the labor share. Additionally, rising import exposure cannot readily explain why labor's share has fallen outside of manufacturing.

data, which contains company accounts of firms listed on stock markets. The details of these data and analysis are provided in online Appendix C. We summarize findings here. Focussing on the largest 500 U.S. based firms in Compustat, as defined (primarily) by their worldwide sales, we highlight four stylized facts.

First, the average size of such firms has increased substantially over time. For example, between 2015 and 1972 the average firm tripled in size as measured by real sales, and it rose by a factor of six in terms of market value.[56] The average employment in the top 500 also grew. But echoing the finding that large firms increasingly have "scale without mass", employment growth at the mean was only 50 percent, which is far smaller than the growth in sales or market value. Second, concentration has risen among these top 500 superstar firms, especially since 2000. For example, the share in total sales of the 50 largest firms among the top 500 rose from 39 percent in 1999 to 48 percent in 2015 (and was 43 percent in 1973). The gap between firms at the 95th percentile of the sales distribution and others further down the distribution has risen particularly strongly. Third, the increase in concentration has been accompanied by an increase in the persistent dominance of top firms, with churn rates falling (consistent with Decker et al, 2018, on the Census LBD). For example, the probability that a firm in the top 500 (by sales) was also in that category five years earlier rose from 66 percent to 80 percent between 2000 and 2015. Similarly, the ten-year survival rate of firms in the top 500 rose from 55 percent in 2005 to 68 percent in 2015.

A fourth finding relates to the growing global engagement of U.S. firms. We estimate that the share of sales outside of the U.S. for superstar manufacturing firms doubled between 1972 and 2012, from 30 percent to 60 percent, and tripled for superstar non-manufacturing firms, from 10 percent to 30 percent. This pattern raises the question of whether rising global engagement could itself be a driving force behind the fall of the labor share. This is particularly hard to explore in Compustat data because only a mi-

[56]We report real 2015 prices deflated from their nominal values using the Consumer Price Index.

nority of firms reports payroll data in Compustat (it is not a mandatory reporting item). Looking among the firms that do report payroll, we find that globally engaged firms have somewhat higher labor shares and that the average labor share of globally engaged firms has fallen since 1982 (see also Hartman-Glaser, Lustig and Zhang, 2017). However, as is shown in Appendix C, the labor share has fallen only slightly more among globally-engaged than non-globally engaged firms. This pattern echoes our broader finding that the fall in the labor share, and the rise in concentration, are prevalent across non-traded sectors in Census data rather than being limited to the heavily traded manufacturing sector. This suggests that globalization, construed narrowly, is unlikely to be the key driver of falling labor shares—though we recognize that a fuller analysis of this question awaits a conceptual and empirical frame that encompasses the full set of general equilibrium forces in play.[57]

### 2.5.3 Worker Power and the Rise in Concentration

There has been much recent discussion of whether the declining labor share reflects falling worker power (Krueger, 2018). Declining union power would be one potential mechanism contributing to the decline in the labor share, although the broad decline of labor shares in non-manufacturing (where unions have little presence), and in countries where union power has not fallen so steeply as in the US, would go against this story. Alternatively, some papers have suggested that the growth of superstar firms confers more monopsony power to employers, driving down both wages and employment. In row 5 of Panel A in Table 2.6 we find that the relationship between concentration and average wages (payroll per worker) in manufacturing is in fact positive, although insignificant. This suggests that concentrating sectors in manufacturing are those where the share of labor is falling, but

[57]Specifically, general equilibrium effects emanating from the overall expansion of global operations and offshoring may impact the financial structures of non-globally engaged firms, a force for which our descriptive analysis cannot account. At a practical level, although Compustat reflects the activities of foreign affiliates in its consolidated accounts, it does not include activities that are offshored and outsourced (e.g. Apple's manufacturing agreements with the independent Taiwanese company FoxConn).

the average wage is not.[58]

The sixth row of Panel A in Table 2.6 shows that concentrating industries in manufacturing have moved towards relying significantly more on materials inputs, which is consistent with greater intermediate goods outsourcing. We suspect that these concentrating industries are also relying more on intermediate service outsourcing, especially for low paid workers as seen for example in Germany (Goldschmidt and Schmieder, 2017). Unfortunately, the Census data do not report direct information on service inputs. We return to the issue of service outsourcing in our concluding remarks.

## 2.6 Conclusions

In this paper we have considered a new "superstar firm" explanation for the widely remarked fall in the labor share of value-added. We hypothesize that markets have changed such that firms with superior quality, lower costs, or greater innovation reap disproportionate rewards relative to prior eras. We shows that, consistent with a simple model, these superstar firms have higher markups and a lower share of labor in sales and value-added. As superstar firms gain market share across a wide range of sectors, the aggregate labor share falls. Our model, combined with technological or institutional changes advantaging the most productive firms in many industries, yields predictions that are supported by Census micro-data across the bulk of the U.S. private sector. First, sales concentration levels rise across large swathes of industries. Second, those industries where concentra-

[58]Payroll per worker is a crude measure of the price of labor as it does not account for composition effects (e.g. skills and demographics). Moreover, *local* labor market concentration is likely a better measure of monopsony power than national product market concentration. Several papers have found a negative link between local labor market concentration and local wages (e.g. Azar et al, 2018; Benmelech, Bergman and Kim, 2018; and Rinz, 2018). Although our conclusion that national sales concentration rates have risen is now widely reported (see Barkai, 2016; Gutierrez and Philippon, 2018), the trends in local concentration are less clear cut. For example, Benmelech et al (2018) find increases in local concentration whereas Rinz (2018) and Rossi-Hansberg, Sarte and Trachter (2018) find a decrease. A challenge for analyzing local measures of concentration is obtaining reliable data on local sales. The LBD used by Rinz (2018) and Benmelech et al (2018) contains employment but not sales data. The NETS database used by Rossi-Hansberg et al (2018) contains a large number of imputed establishment-level sales values.

tion rises the most have the sharpest falls in the labor share. Third, the fall in the labor share has an important reallocation component between firms—the unweighted mean of labor share has not fallen much in manufacturing and has actually risen in most of non-manufacturing. Fourth, this between-firm reallocation of the labor share is greatest in the sectors that are concentrating the most. Fifth, aggregate markups have been rising, but unweighted firm markups have not. Sixth, the industries that are becoming more concentrated are also growing relatively more productive and innovative. Seventh, these broad patterns are observed not only in U.S. data, but also internationally in other OECD countries. A final set of results shows that the growth of concentration is disproportionately apparent in industries experiencing faster technical change as measured by the growth of patent-intensity or total factor productivity, suggesting that technological dynamism, rather than simply anti-competitive forces, is an important driver—though likely not the sole driver—of this trend.

The work in this paper is documents a set of robust and cohesive firm-level, industry-level, and cross-national facts that we believe any explanation of falling labor shares must accommodate. We have presented a formal model where the market-share consequences of productivity differences between firms is magnified when the competitive environment becomes more strenuous, turning leading firms into dominating superstars. One source for the change in the environment could be technological: high tech sectors and parts of retail and transportation as well have an increasingly "winner takes most" aspect. Our evidence is consistent with this explanation but does not constitute a definitive causal test of it. An alternative story is that leading firms are now able to lobby better and create barriers to entry, making it more difficult for smaller firms to grow or for new firms to enter. In its pure form, this "rigged economy" view seems unlikely as a complete explanation since the industries where concentration has grown are those that have been increasing their innovation most rapidly. A more subtle story, however, is that firms initially gain high market shares by legitimately competing on the merits of their innovations or supe-

rior efficiency. Once they have gained a commanding position, however, they use their market power to erect various barriers to entry to protect their position. Nothing in our analysis rules out this mechanism, and we regard it as an important area for subsequent research and policy (see Tirole, 2017; Wu, 2018). Future work therefore needs to understand more precisely the economic and regulatory forces that lead to the emergence of superstar firms.

The rise of superstar firms and decline in the labor share also appears to be related to changes in the boundaries of large dominant employers, with such firms increasingly using domestic outsourcing to to contract a wider range of activities previously done in-house to third party firms and independent workers. These activities may include janitorial work, food services, logistics, and clerical work (Weil, 2014; Katz and Krueger 2016; Goldschmidt and Schmieder, 2017). This apparent 'fissuring' of the workplace (Weil, 2014) can directly reduce the labor share by excluding a large set of workers from the wage premia paid by high-wage employers to rank-and-file workers. This fissuring may also reduce the bargaining power of both in-house and outsourced workers in occupations subject to outsourcing threats and increased labor market competition (Dube and Kaplan, 2010; Goldschmidt and Schmieder, 2017). The fissuring of the workplace has been associated with a rising correlation of firm wage effects and person effects (skills) that accounts for a significant portion of the increase in U.S. wage inequality since 1980 (Song et al., 2019). Linking the rise of superstar firms and the fall of the labor share with the trends in inequality between employees should also be an important avenue of future research.

## References

Ackerberg, Daniel, Kevin Caves and Garth Frazer. 2015. "Identification Properties of Recent Production Function Estimators" *Econometrica* 83(6) 411-2451

Aghion, Philippe and Peter Howitt. 1992. "A Model of Growth. Through Creative Destruction." *Econometrica* 60(2) 323-351.

Akerman, Anders, Edwin Leuven, and Magne Mogstad. 2017. "Information Frictions, Broadband Internet and the Relationship between Distance and Trade." mimeo, University of Chicago.

Andrews, Dan, Chiara Criscuolo, and Peter Gal. 2015. "Frontier Firms, Technology Diffusion and Public Policy." OECD Future of Productivity Background Paper.

Antras, Pol. 2004. "Is the U.S. Aggregate Production Function Cobb-Douglas? New Estimates of the Elasticity of Substitution." *The B.E. Journal of Macroeconomics*, 4(1): 1-36.

Antras, Pol, Teresa Fort and Felix Tintelnot. 2017. "The Margins of Global Sourcing: Theory and Evidence from U.S. Firms." *American Economic Review* 107 (9): 2514-64.

Arkolakis, Costas, Costinot, Arnaud, David Donaldson and Andreas Rodriguez-Clare. 2018 "The Elusive Pro-competitive effects of trade" *Review of Economic Studies* 86(1) 46-80

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review*, 103(6): 2121–2168.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017a. "The Fall of the Labor Share and the Rise of Superstar Firms." NBER Working Paper 23396

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2017b. "Concentrating on the Fall of the Labor Share." *American Economic Review Papers and Proceedings*, 107(5): 180–185

Autor, David, Lawrence Katz, and Alan Krueger. 1998. "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, 113 (4):1169-1214.

Autor, David and Anna Salomons. 2018. "Is Automation Labor-Displacing? Productivity Growth, Employment, and the Labor Share" *Brookings Papers on Economic Activity*, 1-63.

Azar, Jose, Ioana Marinescu, Marshall Steinbaum and Bledi Taska. 2018. "Concentration in U.S. Labor Markets: Evidence from online vacancy data" NBER working Paper 24395

Azmat, Ghazala, Alan Manning, and John Van Reenen. 2012. "Privatization and the Decline of Labor Share: International Evidence from Network Industries." *Economica*, 79, 470-492.

Bailey, Martin, Charles Hulten and David Campbell. 1992. "Productivity Dynamics in Manufacturing Plants" *Brookings Papers on Economic Activity: Microeconomics*, 187-267.

Bain, Joseph. 1951. "Relation of Profit Rate to Industry Concentration: American Manufacturing 1936-1940" *Quarterly Journal of Economics* 65(3), 293-324.

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo and Jonathan Timmis. 2018a. "Industry Concentration in Europe and North America", Paris: OECD

Bajgar, Matej, Giuseppe Berlingieri, Sara Calligaris, Chiara Criscuolo and Jonathan Timmis. 2018b. "To Use or Not to Use (and How to Use): Coverage and Representativeness of ORBIS data", Paris: OECD

Baqaee, David and Emmanuel Fahri. 2017. "Productivity and Misallocation in General Equilibrium" NBER Working Paper 24007

Barkai, Simcha. 2016. "Declining Labor and Capital Shares." Working Paper, University of Chicago.

Bartelsman, Eric, John Haltiwanger and Stefano Scarpetta. 2013. "Cross-Country Differences in Productivity: The Role of Allocation and Selection." *American Economic Review*, 103(1): 305-34.

Bauer, Arthur and Lashkari, Daniel. 2018. "Information Technology and Returns to Scale" Yale mimeo

Bell, Brian. 2015. "Wage Stagnation and the Legacy Costs of Employment" *Centrepiece*, Fall, 20-23.

Berkowitz, Daniel, Hong Ma, and Shuichiro Nishioka. 2017. "Does Capital-Labor Substitution or Do Institutions Explain Declining Labor Shares?" University of Pittsburgh.

Benmelech, Efraim, Nittai Bergman and Hyunseob Kim. 2018. "Strong employers and Weak Employees: How does employer concentration affect wages?" NBER Working Paper 24307

Bernard, Andrew B., Stephen J. Redding, Peter K. Schott. 2010. "Multi-Product Firms and Product Switching" *American Economic Review*, 100, 70-97.

Bessen, James. 2017. "Information Technology and Industry Concentration" Boston University Law and Economics Paper No. 17-41

Blanchard, Olivier. 1997. "The Medium Run." *Brookings Papers on Economic Activity*, 2, 89-158.

Blanchard, Olivier and Francesco Giavazzi. 2003. "Macroeconomic Effects of Regulation And Deregulation In Goods and Labor Markets." *Quarterly Journal of Economics*, 118, 879-907.

Bockerman, Petri and Mika Maliranta. 2012. "Globalization, Creative Destruction, and Labor Share Change: Evidence on the Determinants and Mechanisms from Longitudinal Plant-Level Data," *Oxford Economic Papers*, 64, 259-80.

Boldrin, Michele and David K. Levine. 2008. *Against Intellectual Monopoly*. Cambridge University Press.

Bridgman, Benjamin. 2014. "Is labor's loss capital's gain? Gross versus net labor shares." *Bureau of Economic Analysis*, Washington DC.

Brynjolfsson, Erik, Andrew McAfee, Michael Sorell, and Feng Zhu. 2008. "Scale Without Mass: Business Process Replication and Industry Dynamics" Harvard Business School Technology and Operations Management Unit Research Paper No 07-016.

Calligaris, Sarah, Chiara Criscuolo and L. Marcolin. 2018. "Mark-ups in the digital era", OECD Science, Technology and Industry Working Papers, 2018/10

Cooper, Zack, Stuart Craig, Martin Gaynor and John Van Reenen. 2019. "The Price Ain't Right? Hospital prices and health spending on the privately insured" *Quarterly Journal of Economics* 134 (1), 51–107

Crouzet, Nicolas and Janice Eberly. 2018. "Understanding Weak Capital Investment: the Role of Market Concentration and Intangibles" , Northwestern mimeo

Dao, Mai, Mitali Das, Zsoka Koczan, and Weicheng Lian. 2017. "Why is Labor Receiving a Smaller Share of Global Income? Theory and Empirical Evidence," IMF Working Paper.

Decker, Ryan, John Haltiwanger, Ron Jarmin and Javier Miranda. 2018 "Changing Business Dynamism and Productivity: Shocks vs. Responsiveness" NBER Working Paper 24236

De Loecker, Jan and F. Warzynski. 2012. "Markups and Firm Export Status" *American Economic Review* 102(6) 2437-2471

De Loecker, Jan and Jan Eeckhout. 2018. "Global Market Power", CEPR Discussion Paper 13009

De Loecker, Jan, Jan Eeckhout and Gabriel Unger. 2018 "The Rise of Market Power", mimeo

De Loecker, Jan, Pinelopi Goldberg, Amit Khandelwal, Nina Pavcnik. 2016. "Prices, Markups, and Trade Reform" *Econometrica*, 84(2) 445-510

Demsetz, Harold. 1973. "Industry Structure, Market Rivalry and Public Policy." *Journal of Law and Economics*, 16: 1-10.

Dottling, Robin, German Gutierrez and Thomas Philippon. 2017. "Is There an Investment Gap in Advanced Economies? If So, Why?" NYU mimeo

Dube, Arindrajit and Ethan Kaplan. 2010. "Does Outsourcing Reduce Wages in the Low-Wage Service Occupations? Evidence from Janitors and Guards." *Industrial and Labor Relations Review*, 63(2), 287-306.

Edmond, Chris, Virgiliu Midrigan, and Daniel Yi Xu. 2015 "Competition, Markups, and the Gains from International Trade" *American Economic Review* 105(10): 3183–3221

Elsby, Mike, Bart Hobijn, and Aysegul Sahin. 2013. "The Decline of the U.S. Labor Share." *Brookings Papers on Economic Activity*, 1-42.

Fort, Theresa and Shawn Klimet. 2016. "The effects of Industry Classification changes on U.S. employment composition" Dartmouth mimeo

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, 98(1): 394-425.

Foster, Lucia, John Haltiwanger, and Chad Syverson. 2016. "The Slow Growth of New Plants: Learning about Demand?" *Economica*, 83(329): 91-129.

Furman, Jason and Peter Orszag. 2015. "A Firm-Level Perspective on the Role of Rents in the Rise in Inequality", Working Paper.

Ganapati, Sharat. 2018. "Oligopolies, Prices, Output, and Productivity", Georgetown mimeo

Gilbert, Richard and David Newbery. 1982. "Preemptive Patenting and the Persistence of Monopoly," *American Economic Review*, 72(3) 514-526.

Gollin , Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy*, 110(2), 458-74

Goldschmidt, Deborah and Johannes F. Schmieder. 2017. "The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure." *Quarterly Journal of Economics*, 132(3), 1165-1217.

Gordon, Robert. 2016. *The Rise and Fall of American Growth: The U.S. Standard of Living Since the Civil War*. Princeton: Princeton University Press.

Gutierrez, German and Thomas Philippon. 2016. "Investment-less growth: An empirical investigation" NBER Working Paper 22897

Gutierrez, German and Thomas Philippon. 2018. "How EU markets became more competitive than U.S. markets" NBER Working Paper 24700

Gopinath, Gita, Şebnem Kalemli-Ozcan, Loukas Karabarbounis and Carolina Villegas-Sanchez (2017) "Capital Allocation and Productivity in South Europe", *Quarterly Journal of Economics*, 132(4), 1915–1967

Hamermesh, Daniel S. 1993. *Labor Demand*. Princeton University Press.

Hall, Robert. 1988. "The Relationship between Price and Marginal Cost in U.S. Industry" *Journal of Political Economy*, 96(5), 921-947

Hall, Robert. 2018. "New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms", NBER Working Paper 24574

Haltiwanger, John, Ron Jarmin and Javier Miranda (2013) "Who Creates Jobs? Small vs. Large vs. Young" *Review of Economics and Statistics*, 347-361

Hartman-Glaser, Barney, Hanno Lustig, and Mindy Zhang. 2017. "Capital Share Dynamics when Firms Insure Workers." forthcoming, *Journal of Finance*.

Hopenhayn, Hugo. 1992. "Entry and Exit in Long Run Equilibria." *Econometrica*, 60(5) 1127-1150

Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India."*Quarterly Journal of Economics*, 124, 1403-1448.

Jones, Charles. 2005. "The Shape of Production Functions and the Direction of Technical Change" *Quarterly Journal of Economics*, 120 (2), 517-549

Kaldor, Nicholas. 1961. "Capital Accumulation and Economic Growth," in *The Theory of Capital*, ed. F. A. Lutz and D. C. Hague, New York: St. Martin's Press: 177-222.

Kalecki, Michal. 1938."The Determinants of the Distribution of National Income." *Econometrica*, 6(2), 97-112.

Kalemli-Ozcan, Sebnem, Bent Sorensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcan Yesiltas. 2015. "How to Construct Nationally Representative Firm Level data from the ORBIS Global Database" NBER Working Paper 21558.

Karabarbounis, Loukas and Brent Neiman. 2013. "The Global Decline of the Labor Share." *Quarterly Journal of Economics*, 129(1), 61-103.

Karabarbounis, Loukas and Brent Neiman. 2018. "Accounting for Factorless Income" NBER Working Paper 24404

Katz, Lawrence F. and Alan B. Krueger. 2016. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995-2015." NBER Working Paper No. 22667.

Kehrig, Matthias and Nicolas Vincent. 2018. "The Micro-Level Anatomy of the Labor Share Decline", NBER Working Paper 25275

Kleiner, Morris and Alan Krueger. 2013. "Analyzing the Extent and Influence of Occupational Licensing on the Labor Market." *Journal of Labor Economics*, 31, S173–S202.

Klette, Toe and Samuel Kortum. 2004. "Innovating Firms and Aggregate Innovation" *Journal of Political Economy*, 112(5) 986-1018.

Koh, Dongya, Raul Santaeulalia-Lopis, and Yu Zheng. 2016. "Labor Share Decline and Intellectual Property Products Capital." SSRN Working Paper.

Krueger, Alan B. 2018. "Reflections on Dwindling Worker Bargaining Power and Monetary Policy", Presentation at Jackson Hole Conference

Lawrence, Robert. 2015. "Recent Declines in Labor's Share in U.S. Income: A Preliminary Neoclassical Account." NBER Working Paper No. 21296.

Lopez-Garcia, P., F. di Mauro, and the CompNet Task Force. 2015. "Assessing European competitiveness: the new CompNet micro-based database." European Central Bank, Working Paper Series No 1764.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics*, 9, 2, 508-523.

Melitz, Marc. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica*, 71, 1695–1725.

Melitz, Marc and Saso Polanec. 2015. "Dynamic Olley-Pakes Productivity Decomposition with Entry and Exit." *RAND Journal of Economics*, 46 (2): 362-375.

Melitz, Marc and Gianmarco Ottaviano. 2008. "Market Size, Trade, and Productivity" *Review of Economic Studies* (2008) 75, 295–316

Mrazova and Peter Neary. 2017. "Not So Demanding: Demand Structure and Firm Behavior" *American Economic Review*, 107(12): 3835–3874

Nekarda, Christopher and Valerie Ramey. 2013. "The Cyclical Behavior of the Price-Cost Markup" NBER Working Paper 19099

Oberfield, Ezra and Devesh Raval. 2014. "Micro Data and Macro Technology." Working Paper, Princeton University.

Olley, Steven and Pakes, Ariel. 1996. "The Dynamics of Productivity in the Telecommunications Industry." *Econometrica*, 64(6): 1263-1298.

O'Mahony, Mary and Marcel Timmer. 2009. "Output, Input and Productivity Measures at the Industry Level: the EU KLEMS Database," *Economic Journal*, 119(538) F374-503.

Paunov, Caroline and Dominique Guellec. 2017. "Digital Innovation and the Distribution of Income", OECD mimeo

Peltzman, Samuel. 2018. "Productivity and Prices in Manufacturing during an era of rising Concentration", University of Chicago mimeo

Pessoa, Joao P. and John Van Reenen. 2013. " Decoupling of Wage Growth and Productivity Growth? Myth and Reality", CEP Discussion Paper 1246.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Harvard University Press.

Rinz, Kevin. 2018. "Labor Market concentration, Earnings Inequality and Earnings Mobility" CARRA Working Paper 2018-10

Rognlie, Matthew. 2015. "Deciphering the Fall and Rise in the Net Capital Share: Accumulation or Scarcity?" *Brookings Papers on Economic Activity*, 1-69.

Rossi-Hansberg, Estaban, Pierre-Daniel Sarte and Nicholas Trachter. 2018. "Diverging Trends in National and Local Concentration" NBER Working Paper 25066

Schmalensee, Richard. 1987. "Collusion Versus Differential Efficiency: Testing Alternative Hypotheses." *Journal of Industrial Economics*, 35(4), June, 399-425.

Smith, Mathew, Daniel Yagan, Owen Zidar and Eric Zwick. 2017. "Capitalists in the Twenty-First Century" Chicago mimeo

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2019. "Firming up Inequality." *Quarterly Journal of Economics* 134(1), 1–50.

Syverson, Chad. 2017. "Challenges to Mismeasurement Explanations for the U.S. Productivity Slowdown." *Journal of Economic Perspectives*, Spring, 165-186

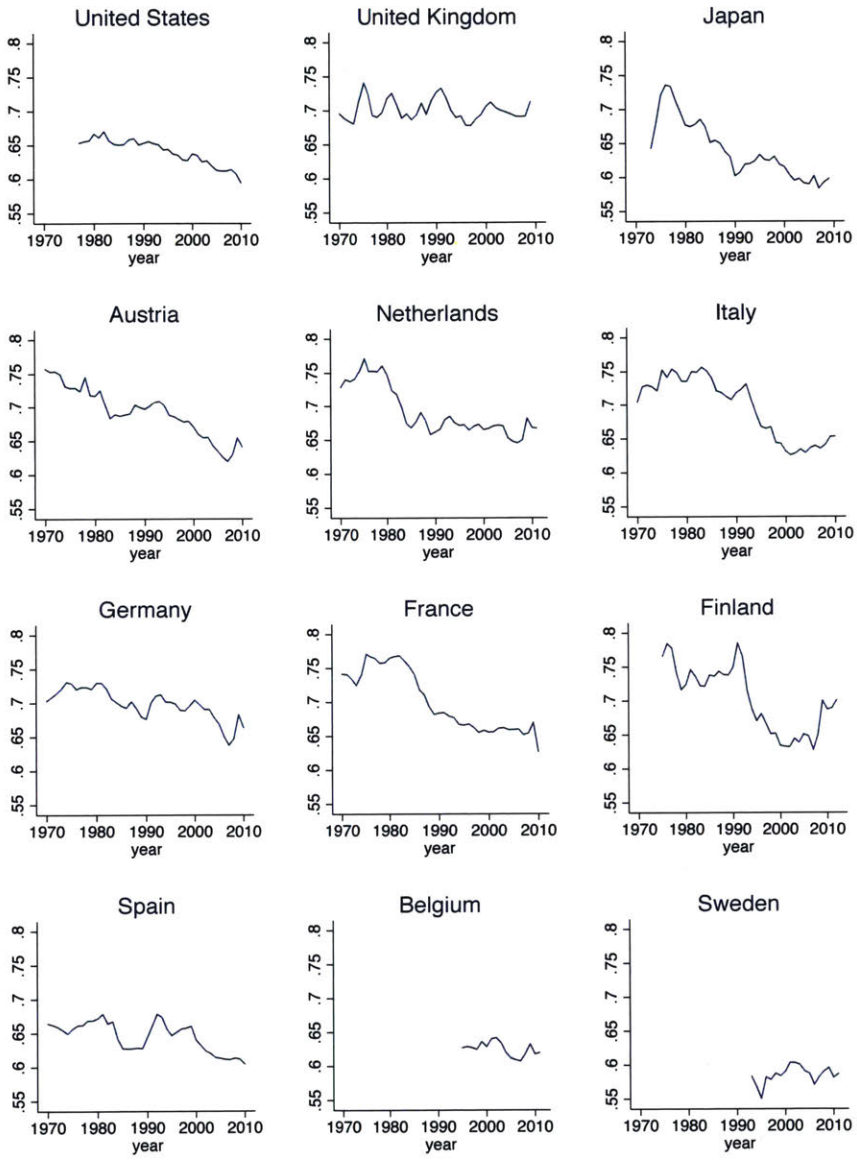
Tirole, Jean. 2017. *Economics for the Common Good,* Princeton: Princeton University Press

Wu, Timothy. 2018. *The Curse of Bigness: Antitrust in the New Gilded Age,* New York: Columbia Press

Weil, David. 2014. *The Fissured Workplace.* Cambridge, MA: Harvard University Press.

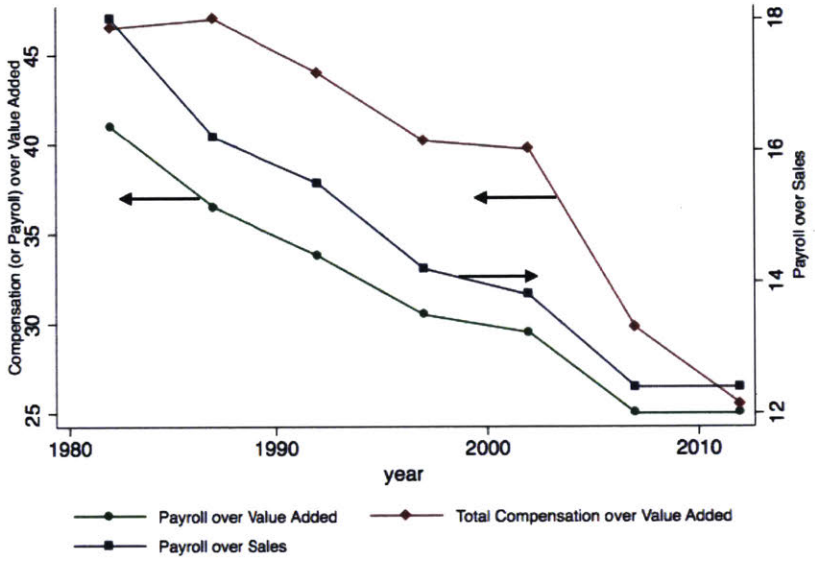
# Figures and Tables

Figure 2.1: International Comparison: Labor Share by Country



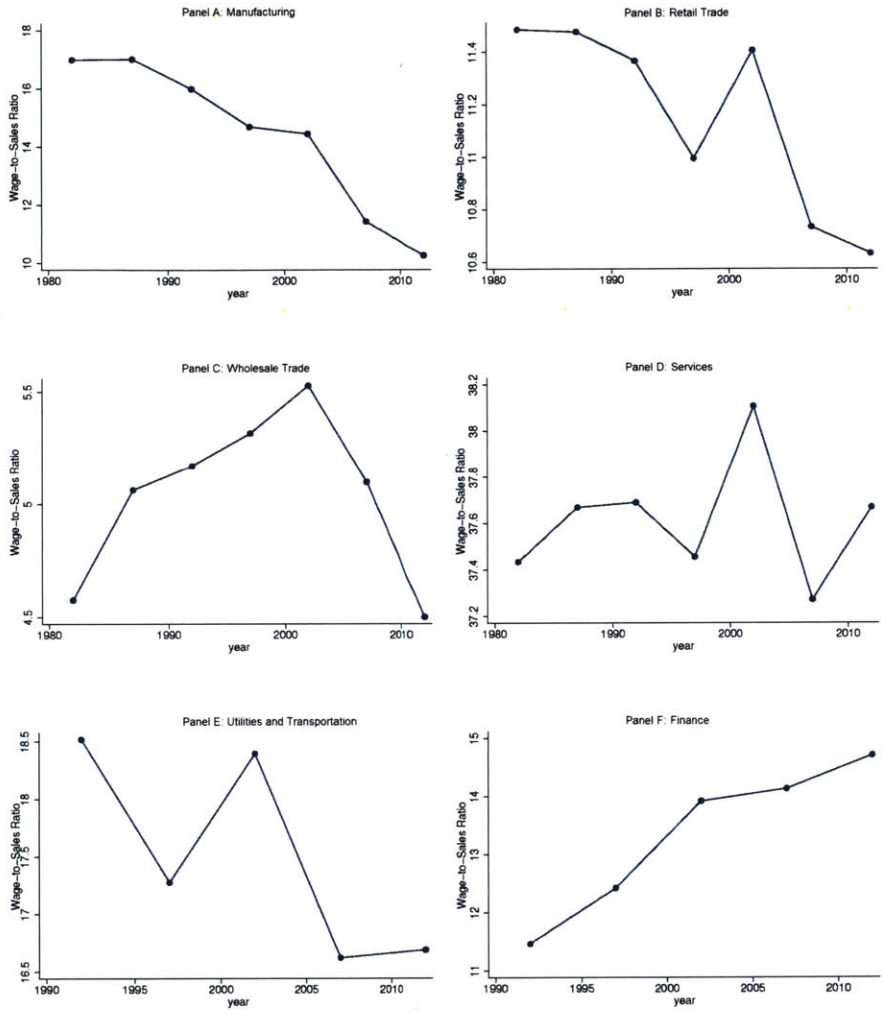

**Notes.** Each panel plots the ratio of labor compensation to gross value-added for all industries. Data is from EU KLEMS July 2012 release.

Figure 2.2: The Labor Share in Manufacturing



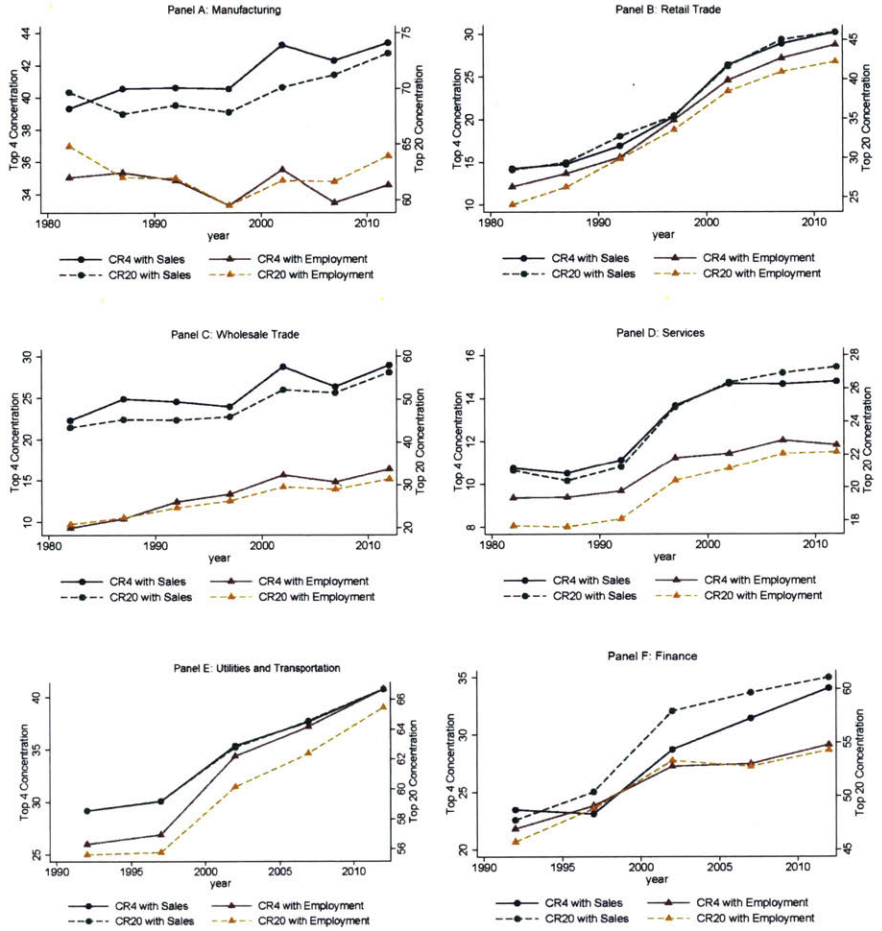

**Notes.** This figure plots the aggregate labor share in manufacturing from 1982-2012. The green circles (plotted on the left axis) represent the ratio of wages and salaries (payroll) to value-added. The red diamonds (also plotted on the left axis) include a broader definition of labor income and plots the ratio of wages, salaries and fringe benefits (compensation) to value-added. The blue squares (plotted on the right axis) show wages and salaries re-normalized by sales rather than value-added.

Figure 2.3: Average Payroll-to-Sales Ratio



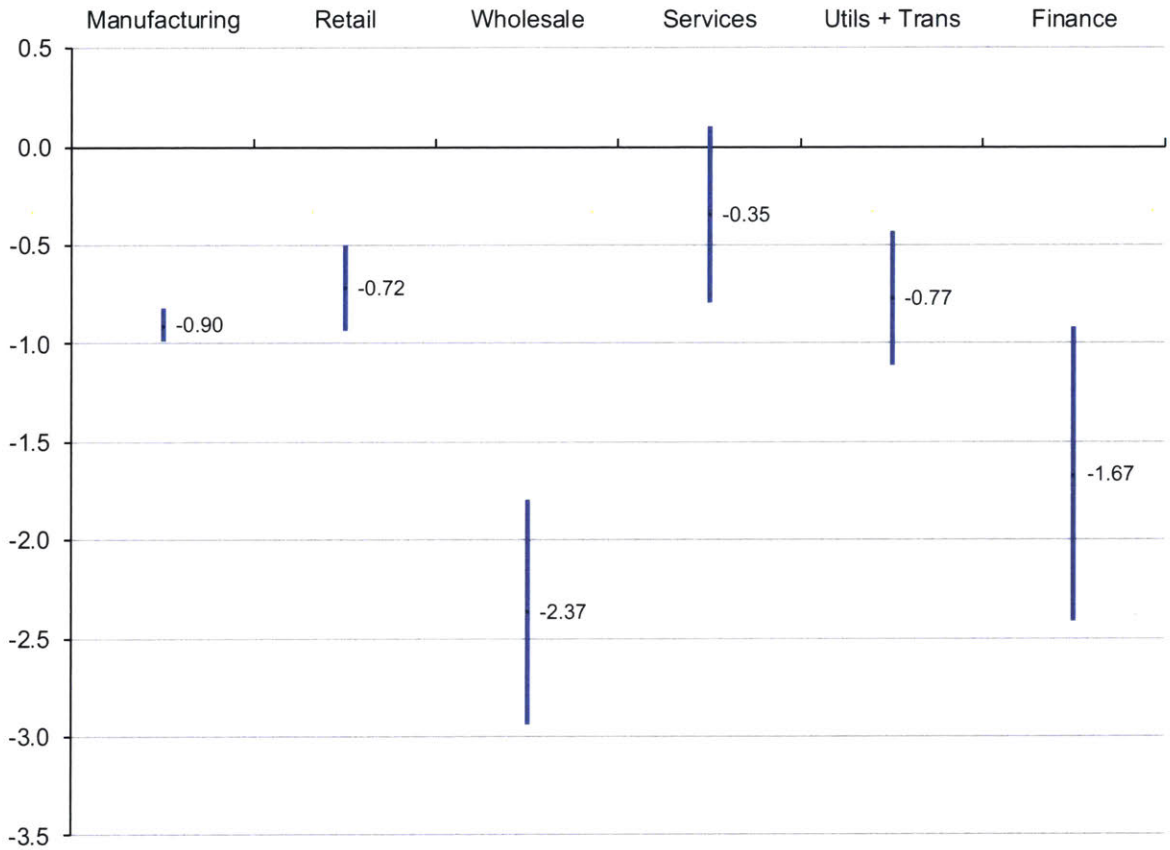

**Notes.** Each panel plots the overall payroll-to-sales ratio in one of the six major sectors covered by the U.S. Economic Census. Add to notes at the end "These figures update Autor et al (2017) to include more recently released Census data.

Figure 2.4: Average Concentration Across Four Digit Industries by Major Sector



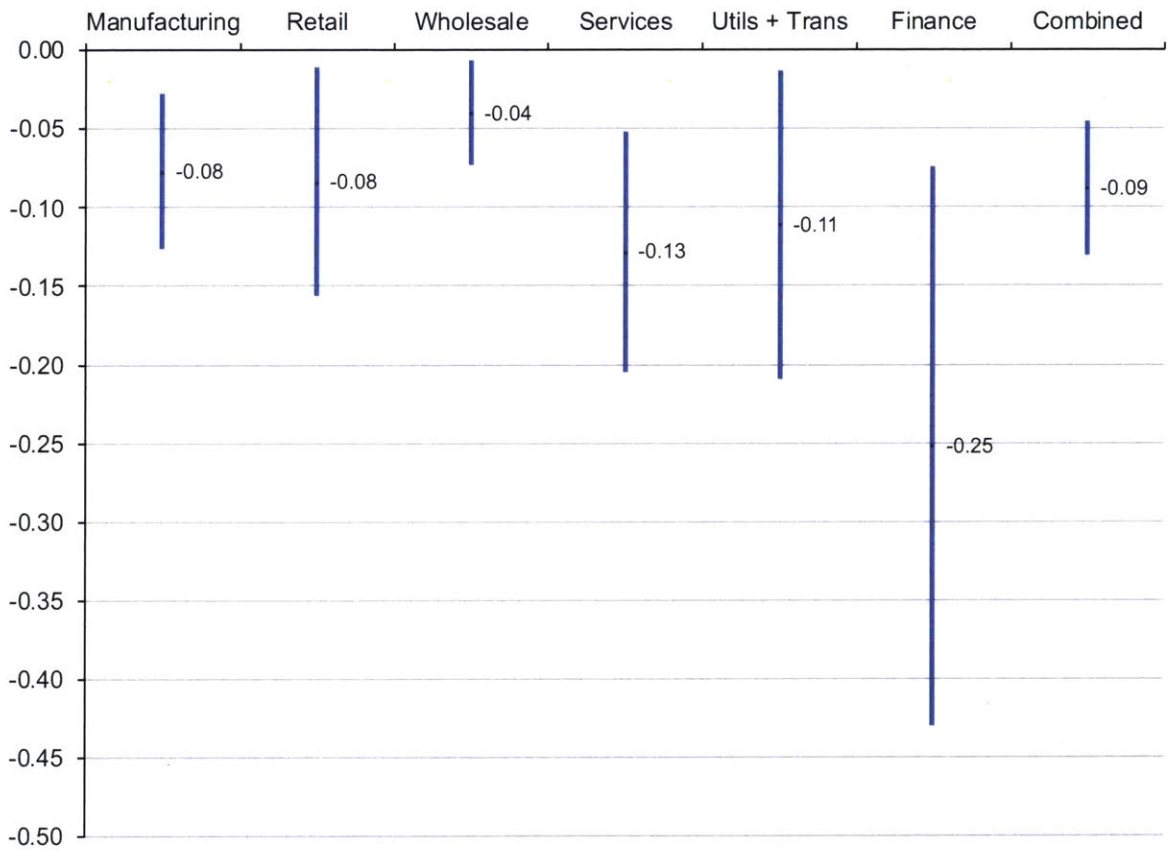

**Notes.** This figure plots the average concentration ratio in six major sectors of the U.S. economy. Industry concentration is calculated for each time-consistent four-digit industry code, and then averaged across all industries within each of the six sectors. The solid blue line (circles), plotted on the left axis, shows the average fraction of total industry sales that is accounted for by the largest four firms in that industry, and the solid red line (triangles), also plotted on the left axis, shows the average fraction of industry employment utilized in the four largest firms in the industry. Similarly, the dashed green line (circles), plotted on the right axis, shows the average fraction of total industry sales that is accounted for by the largest 20 firms in that industry, and the dashed orange line (triangles), also plotted on the right axis, shows the average fraction of industry employment utilized in the 20 largest firms in the industry.

Figure 2.5: The Relationship Between Firm Size and Labor Share



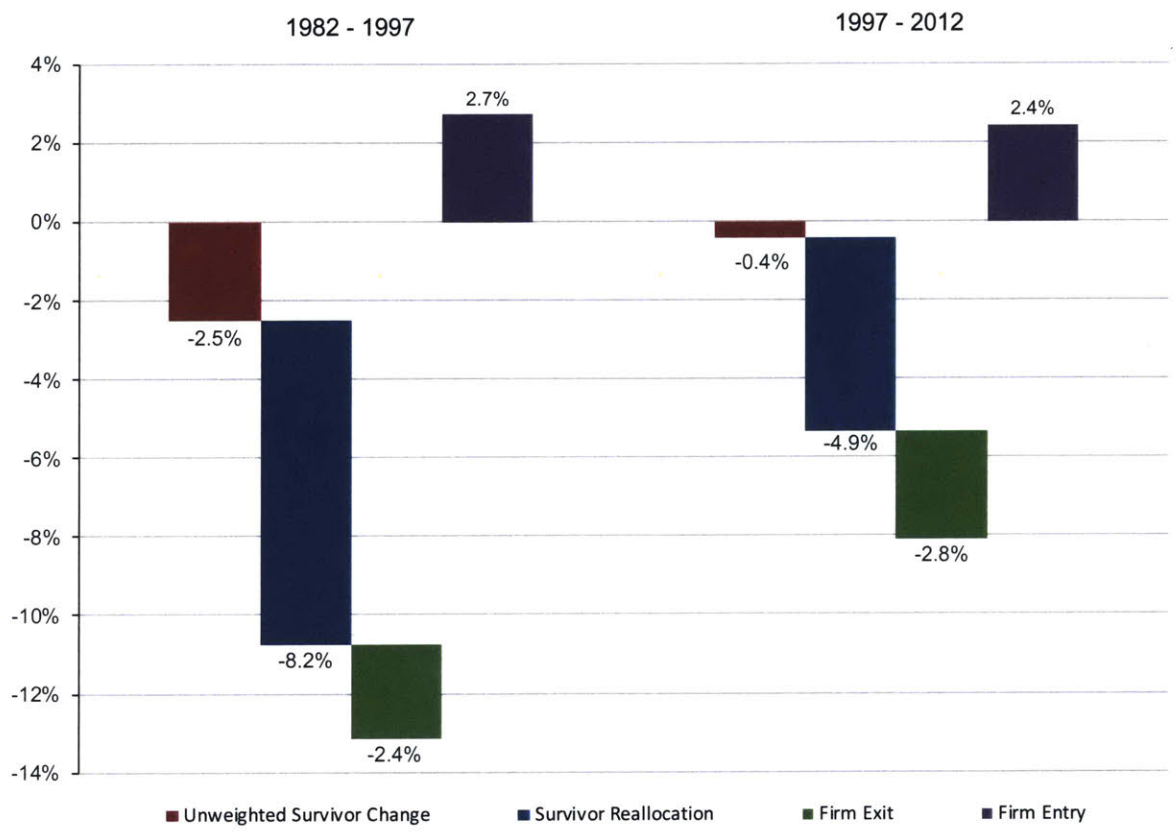

**Notes.** The dots indicate the coefficient estimates of a regression of a firm's labor share on its share of overall sales in its four-digit industry. The regressions include all years available for that sector, and year fixed effects. The labor share is defined as the payroll-to-sales ratio in each sector. The blue lines represent the 95% confidence intervals.

Figure 2.6: The Relationship Between the Change in Labor Share and the Change in Concentration Across Six Sectors



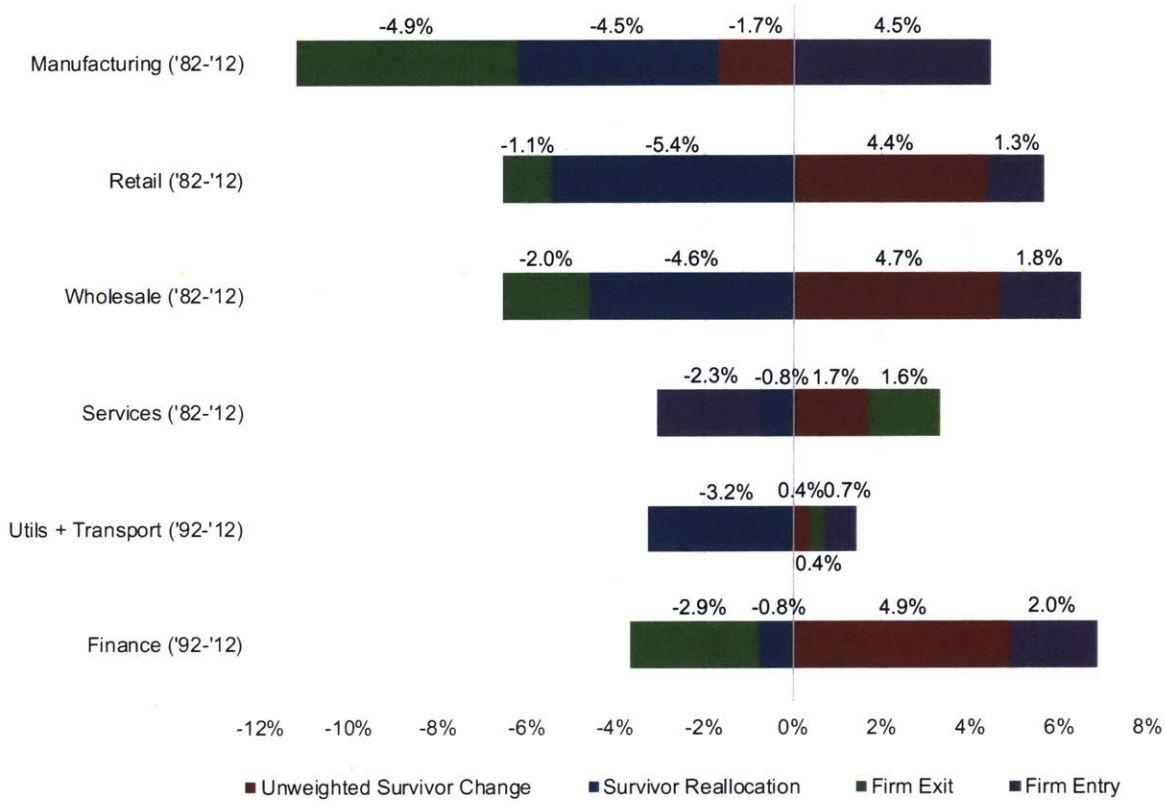

**Notes.** The figure indicates OLS regression estimates from ΔLabor Share (payroll over sales) on ΔCR20 (stacked five-year changes from 1982-2012 with dummies for each time period). Dots indicate coefficient estimates and lines indicate 95% confidence intervals. This is taken from panel A column (2) of Table 2.3 which also tabulates the full regression results using alternative measures of concentration and specifications.

Figure 2.7: Melitz-Polanec Decomposition of the Change in Labor Share in Manufacturing



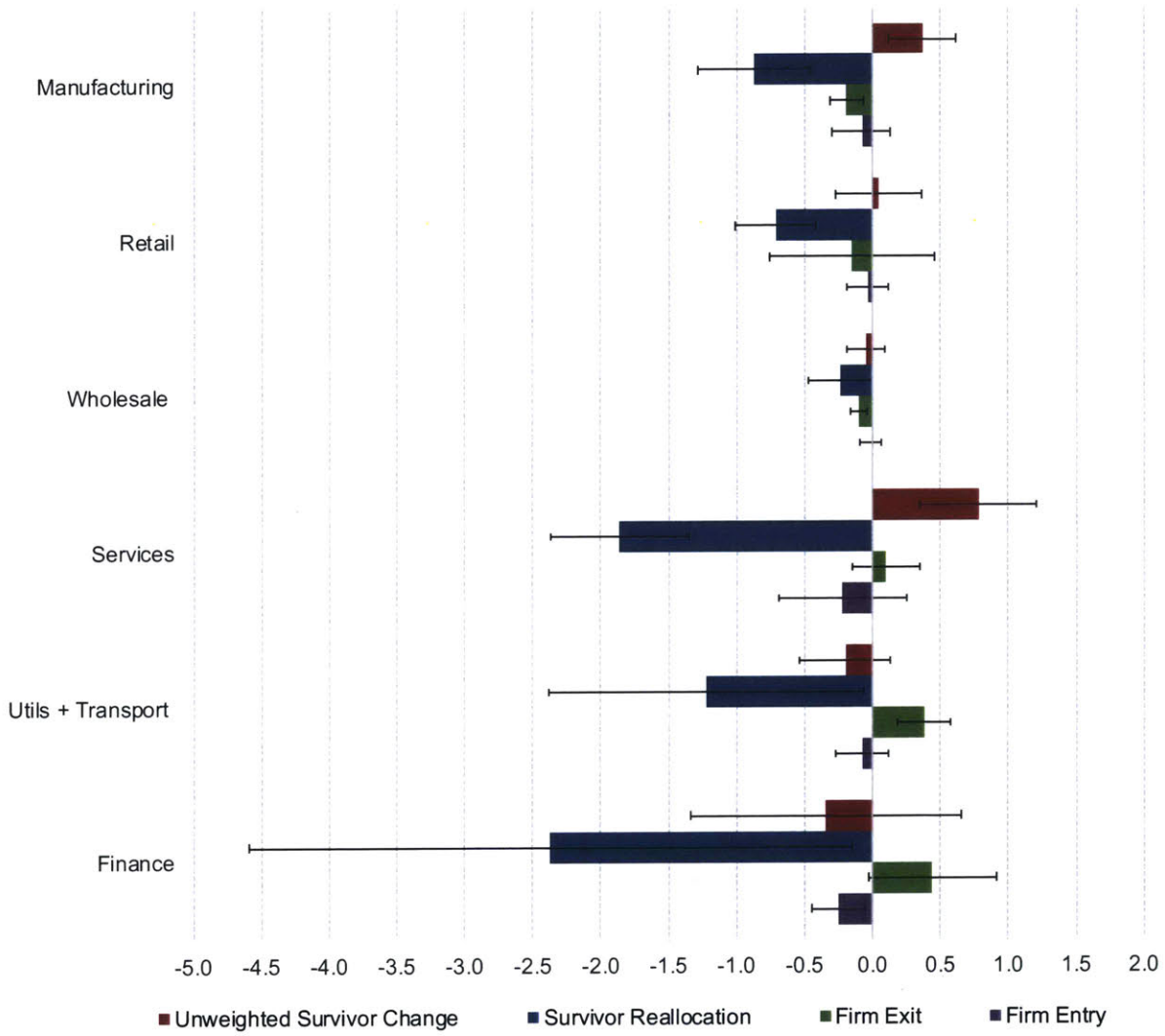

**Notes.** Each bar represents the cumulated sum of the Melitz-Polanec decomposition components calculated over adjacent five-year intervals. The left panel shows the sum of the decompositions from 1982-1987, 1987-1992 and 1992-1997 and the right panel shows the sum of the decompositions from 1997-2002, 2002-2007, and 2007-2012. Table 2.4 reports the underlying year-by-year estimates.

Figure 2.8: Melitz-Polanec Decomposition of the Change in Labor Share in all Six Sectors



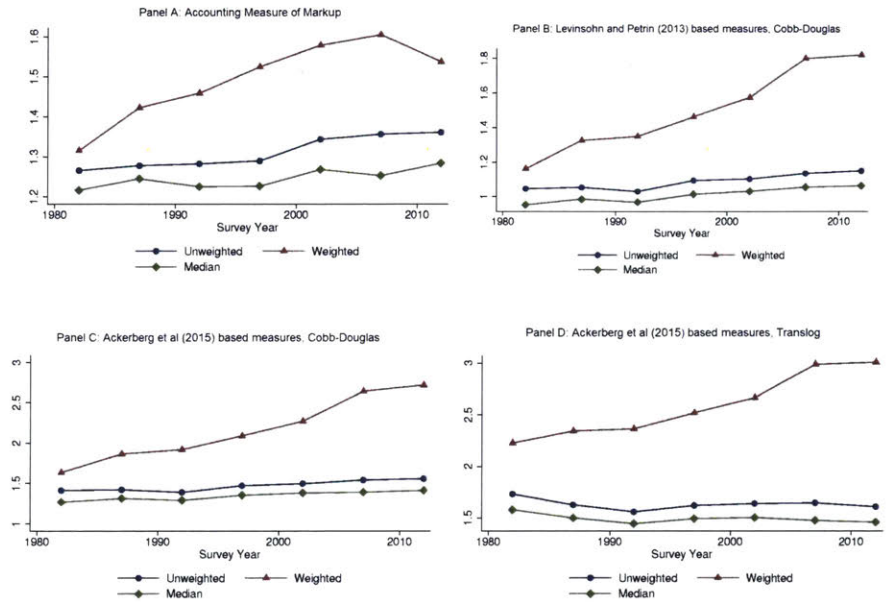

**Notes.** Each bar represents the cumulated sum of the Melitz-Polanec decomposition components calculated over adjacent five-year intervals. Table 2.5 reports the underlying year-by-year estimates.

Figure 2.9: Regressions of the Components of the Change in Labor Share on the Change in Concentration




**Notes.** Each bar plots ten times the regression coefficient resulting from regressions of the Melitz-Polanec decomposition components on the change in CR20 concentration. Regressions include year dummies and standard errors are clustered at the four-digit industry level. Each industry is weighted by its initial share of total sales. Whisker lines represent 95% confidence interval.

Figure 2.10: Markup changes




**Notes.** These are all estimates of the markup of price over marginal cost in manufacturing using the first order condition described in the text (equation 2.7). Panel A uses Antras et al (2017) "accounting" method; Panels B-D use production function methods following de Loecker and Warzynski (2012). In these panels we estimate industry specific production functions (two-digit SIC). In Panels B and C the production function is assumed to be Cobb-Douglas and in Panel D it is assumed to be translog. Panels B uses the Levinsohn-Petrin (2003) approach and Panels C and D use the Ackerberg et al (2015) approach. Each panel presents three period specific estimates of the markup. The lower lines present the unweighted mean (blue circles) and median (green diamond) firm level markups. The upper line (red triangles) present the mean markups weighted by a firm's value added.

# Tables

Table 2.1: Summary Statistics

| | Mean (1) | SD (2) | Minimum (3) | Maximum (4) |
|---|---|---|---|---|
| | | *A. Manufacturing (388 industries, 2,328 obs)* | | |
| Number of establishments | 197,530 | 10,635 | 169,107 | 216,730 |
| Number of Firms | 151,936 | 10,386 | 129,080 | 171,233 |
| Payroll to Sales Ratio | 15.2386 | 8.3752 | 0.872 | 48.582 |
| Change in Payroll to Sales Ratio | -0.9611 | 1.9821 | -17.616 | 14.614 |
| CR4 | 40.6642 | 22.5451 | 3.344041 | 100 |
| Change in CR4 | 0.7476 | 6.4473 | -39.725 | 39.505 |
| CR20 | 68.7607 | 23.2561 | 8.376 | 100 |
| Change in CR20 | 0.7566 | 4.3078 | -32.526 | 24.002 |
| | | *B. Retail Trade (58 industries, 348 obs)* | | |
| Number of establishments | 1,598,458 | 74,292 | 1,562,915 | 1,722,947 |
| Number of Firms | 1,115,863 | 17,814 | 1,104,697 | 1,152,079 |
| Payroll to Sales Ratio | 11.258 | 5.7401 | 2.748 | 29.112 |
| Change in Payroll to Sales Ratio | -0.0588 | 0.9862 | -11.703 | 10.259 |
| CR4 | 19.9905 | 18.9734 | 0.635 | 79.133 |
| Change in CR4 | 2.5071 | 4.8131 | -23.844 | 32.407 |
| CR20 | 35.0778 | 26.4192 | 1.824 | 99.983 |
| Change in CR20 | 2.6928 | 4.2785 | -35.006 | 49.889 |
| | | *C. Wholesale Trade (56 industries, 336 obs)* | | |
| Number of establishments | 411,651 | 22,275 | 400,878 | 442,693 |
| Number of Firms | 324,899 | 20,452 | 306,174 | 355,052 |
| Payroll to Sales Ratio | 5.0694 | 3.1859 | 0.45 | 14.093 |
| Change in Payroll to Sales Ratio | -0.1811 | 0.8854 | -3.742 | 4.372 |
| CR4 | 24.6336 | 13.4093 | 4.32 | 65.046 |
| Change in CR4 | 0.3548 | 6.8544 | -30.894 | 35.26 |
| CR20 | 46.4094 | 17.3136 | 11.326 | 83.67 |
| Change in CR20 | 1.0315 | 7.0595 | -26.108 | 33.956 |

| | Mean (1) | SD (2) | Minimum (3) | Maximum (4) |
|---|---|---|---|---|
| | | *D. Services (95 industries, 570 obs)* | | |
| Number of establishments | 2,039,671 | 412,831 | 1,769,458 | 2,698,102 |
| Number of Firms | 1,725,578 | 287,188 | 1,586,300 | 2,256,011 |
| Payroll to Sales Ratio | 37.4223 | 10.9437 | 5.489 | 74.268 |
| Change in Payroll to Sales Ratio | -0.352 | 2.4102 | -14.288 | 19.654 |
| CR4 | 12.1406 | 11.4397 | 0.316 | 77.131 |
| Change in CR4 | 0.7283 | 4.409 | -32.727 | 35.399 |
| CR20 | 22.7854 | 17.1222 | 0.848 | 100 |
| Change in CR20 | 0.9533 | 4.7568 | -27.768 | 31.461 |
| | | *E. Finance (31 industries, 124 obs)* | | |
| Number of establishments | 676,357 | 101,246 | 637,839 | 842,694 |
| Number of Firms | 456,175 | 65,420 | 432,753 | 561,940 |
| Payroll to Sales Ratio | 12.8464 | 9.1203 | 1.152 | 39.701 |
| Change in Payroll to Sales Ratio | -0.7437 | 3.5948 | -20.704 | 17.068 |
| CR4 | 26.0744 | 15.1231 | 2.634 | 97.387 |
| Change in CR4 | 2.0704 | 6.2006 | -21.075 | 34.552 |
| CR20 | 53.0273 | 19.7478 | 6.102 | 100 |
| Change in CR20 | 3.6006 | 5.8551 | -25.22 | 31.261 |
| | | *F. Utilities and Transportation (48 industries, 144 obs)* | | |
| Number of establishments | 286,939 | 30,476 | 292,474 | 345,951 |
| Number of Firms | 203,626 | 17,563 | 213,349 | 228,854 |
| Payroll to Sales Ratio | 18.0455 | 8.4094 | 4.484 | 53.536 |
| Change in Payroll to Sales Ratio | -0.658 | 2.3697 | -11.528 | 10.021 |
| CR4 | 31.0864 | 19.7924 | 3.042 | 91.645 |
| Change in CR4 | 1.9307 | 8.5871 | -27.318 | 27.699 |
| CR20 | 59.6948 | 24.2405 | 9.221 | 100 |
| Change in CR20 | 1.203 | 6.4252 | -25.247 | 25.538 |

**Notes.** The number of establishments and number of firms reflect totals for the entire sector. All other variables are the weighted averages of the underlying four-digit industries, where the weight is the industry's share of sales in the initial year. Changes refer to five year averages. Data period is 1982-2012 for manufacturing, services, wholesale trade and retail trade, 1992-2012 for finance and 1992-2007 for utilities and transportation. CR4 and CR20 are defined in terms of sales. In future drafts, this table will include summary statistics on the payroll to value-added share in manufacturing. Those summary statistics have not yet been disclosed by the census.

Table 2.2: Industry Regressions of Change in Share of Labor on Change in Concentration, Manufacturing

| | 5-year Changes | | | | | | 10-year Changes | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CR4 | | CR20 | | HHI | | CR4 | | CR20 | | HHI | |
| | (1) | | (2) | | (3) | | (4) | | (5) | | (6) | |
| 1 Baseline | -0.148 | *** | -0.228 | *** | -0.213 | ** | -0.132 | *** | -0.153 | *** | -0.165 | * |
| | (0.036) | | (0.043) | | (0.085) | | (0.040) | | (0.055) | | (0.093) | |
| 2 Compensation Share | -0.177 | *** | -0.266 | *** | -0.256 | ** | -0.139 | *** | -0.151 | ** | -0.183 | |
| of Value Added | (0.045) | | (0.056) | | (0.110) | | (0.053) | | (0.071) | | (0.125) | |
| 3 Deduct Service Intermediates | -0.339 | *** | -0.514 | *** | -0.502 | *** | -0.261 | *** | -0.353 | *** | -0.303 | |
| from Value Added | (0.064) | | (0.074) | | (0.175) | | (0.056) | | (0.065) | | (0.275) | |
| 4 Value Added-based | -0.219 | *** | -0.337 | *** | -0.320 | *** | -0.210 | *** | -0.251 | *** | -0.289 | *** |
| Concentration | (0.028) | | (0.045) | | (0.060) | | (0.037) | | (0.054) | | (0.075) | |
| 5 Industry Trends | -0.172 | *** | -0.290 | *** | -0.243 | ** | -0.196 | *** | -0.240 | *** | -0.220 | * |
| (Four-Digit Dummies) | (0.043) | | (0.047) | | (0.100) | | (0.059) | | (0.088) | | (0.128) | |
| 6 1992-2012 Sub-Period | -0.187 | *** | -0.309 | *** | -0.261 | ** | | | | | | |
| | (0.043) | | (0.061) | | (0.102) | | | | | | | |
| 7 Including Imports | -0.163 | *** | -0.285 | *** | -0.233 | *** | | | | | | |
| (1992-2012) | (0.036) | | (0.052) | | (0.089) | | | | | | | |
| Coefficient on Fraction | 18.809 | *** | 20.467 | *** | 20.957 | *** | | | | | | |
| of Imports | (3.027) | | (3.213) | | (3.187) | | | | | | | |
| 8 Control for initial capital | -0.146 | *** | -0.231 | *** | -0.214 | *** | -0.122 | *** | -0.148 | *** | -0.161 | * |
| /Value Added | (0.035) | | (0.042) | | (0.084) | | (0.040) | | (0.053) | | (0.092) | |
| Capital/Value Added | -1.242 | *** | -1.295 | *** | -1.278 | *** | -2.535 | *** | -2.648 | *** | -2.669 | *** |
| coefficient | (0.308) | | (0.324) | | (0.292) | | (0.595) | | (0.598) | | (0.563) | |
| 9 Employment-Based | 0.036 | | 0.024 | | 0.160 | ** | 0.018 | | 0.029 | | 0.082 | |
| Concentration Measure | (0.036) | | (0.033) | | (0.075) | | (0.035) | | (0.040) | | (0.083) | |

**Notes.** The number of establishments and number of firms reflect totals for the entire sector. All other variables are the weighted averages of the underlying four-digit industries, where the weight is the industry's share of sales in the initial year. Changes refer to five year averages. Data period is 1982-2012 for manufacturing, services, wholesale trade and retail trade, 1992-2012 for finance and 1992-2007 for utilities and transportation. CR4 and CR20 are defined in terms of sales. In future drafts, this table will include summary statistics on the payroll to value-added share in manufacturing. Those summary statistics have not yet been disclosed by the Census Bureau.

Table 2.3: Industry Regressions of the Change in the Payroll-to-Sales Ratio on the Change in Concentration, Different Sectors

| | Stacked 5-year Changes | | | Stacked 10-year Changes | | |
|---|---|---|---|---|---|---|
| | CR4 | CR20 | HHI | CR4 | CR20 | HHI |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| 1 Manufacturing | -0.062 *** | -0.077 *** | -0.112 *** | -0.035 | -0.034 | -0.088 ** |
| n=2,328; 1,164 | (0.013) | (0.025) | (0.026) | (0.021) | (0.033) | (0.037) |
| 2 Retail | -0.034 * | -0.084 ** | -0.041 | -0.043 ** | -0.067 ** | -0.068 *** |
| n=348; 174 | (0.020) | (0.037) | (0.025) | (0.018) | (0.029) | (0.023) |
| 3 Wholesale | -0.038 *** | -0.040 ** | -0.084 ** | -0.037 ** | -0.036 * | -0.064 |
| n=336; 168 | (0.014) | (0.017) | (0.041) | (0.018) | (0.019) | (0.048) |
| 4 Services | -0.091 | -0.128 *** | -0.350 *** | -0.093 | -0.137 *** | -0.377 ** |
| n=570; 258 | (0.057) | (0.039) | (0.084) | (0.070) | (0.042) | (0.156) |
| 5 Utilities/Transport | -0.110 *** | -0.111 ** | -0.320 *** | -0.064 | -0.096 ** | -0.226 ** |
| n=144; 48 | (0.031) | (0.050) | (0.082) | (0.044) | (0.038) | (0.098) |
| 6 Finance | -0.221 ** | -0.252 *** | -0.567 ** | -0.236 ** | -0.274 *** | -0.723 ** |
| n=124; 62 | (0.084) | (0.091) | (0.208) | (0.095) | (0.084) | (0.295) |
| 7 Combined | -0.077 *** | -0.088 *** | -0.150 *** | -0.060 *** | -0.076 *** | -0.118 *** |
| n=3,850; 1,901 | (0.017) | (0.022) | (0.028) | (0.018) | (0.023) | (0.032) |

**Notes.** * $p \leq 0.10$, ** $p \leq 0.05$,*** $p \leq 0.01$. Number of observations (n = x;y) are indicated below each sector for the first 3 columns (x) and the last 3 columns (y). Each cell displays the coefficient on a concentration measure from a separate OLS regression (standard errors in parentheses clustered by industry). Data are aggregated up to time consistent four-digit industries. In manufacturing, retail, services and wholesale, we pool data from 1982-2012; in finance, we pool data from 1992-2012; and in utilities + transport, we pool data from 1992-2007. The combined regression in row 7 includes six sector fixed effects. Regressions are weighted by the share of sales of the four-digit industry in total sector sales in the initial year and each regression includes fixed effects for each five-year period.

Table 2.4: Decompositions of the Change in the Payroll-to-Value-Added Ratio, Manufacturing

| | Δ Unweighted Mean (1) | Incumbent Reallocation (2) | Exit (3) | Entry (4) | Total (5) |
|---|---|---|---|---|---|
| | | *A. Payroll Share of Value Added* | | | |
| 1982-1987 | -3.03 | -1.75 | -0.59 | 0.86 | -4.52 |
| 1987-1992 | 2.60 | -5.26 | -0.90 | 0.98 | -2.58 |
| 1992-1997 | -2.08 | -1.24 | -0.89 | 0.89 | -3.32 |
| 1997-2002 | 0.00 | -0.76 | -1.00 | 0.69 | -1.08 |
| 2002-2007 | -3.06 | -1.53 | -1.12 | 1.23 | -4.48 |
| 2007-2012 | 2.64 | -2.61 | -0.63 | 0.51 | -0.09 |
| 1982-1997 | -2.52 | -8.25 | -2.38 | 2.73 | -10.42 |
| 1997-2012 | -0.42 | -4.90 | -2.76 | 2.43 | -5.65 |
| 1982-2012 | -2.93 | -13.15 | -5.14 | 5.15 | -16.07 |
| | | *B. Compensation Share of Value Added* | | | |
| 1982-1987 | -0.78 | -5.66 | -0.47 | 0.98 | -5.93 |
| 1987-1992 | 3.73 | -5.69 | -1.00 | 1.05 | -1.91 |
| 1992-1997 | -2.78 | -1.90 | -0.93 | 0.97 | -4.64 |
| 1997-2002 | -2.07 | 1.11 | -1.09 | 0.79 | -1.25 |
| 2002-2007 | 1.26 | -6.21 | -1.20 | 1.55 | -4.60 |
| 2007-2012 | 0.40 | -0.32 | -0.77 | 0.53 | -0.15 |
| 1982-1997 | 0.17 | -13.25 | -2.40 | 3.01 | -12.48 |
| 1997-2012 | -0.41 | -5.42 | -3.06 | 2.88 | -6.00 |
| 1982-2012 | -0.24 | -18.67 | -5.46 | 5.89 | -18.48 |

**Notes.** This table shows the results of a decomposition of the change in the labor share using the dynamic Melitz-Polanec (2015) methodology as described in the text. We divide the change in the overall labor share (columns 1 and 6) into four components: "Change in Unweighted Mean" is the change in the labor share due to a general fall in the share across all incumbent firms; "Incumbent Reallocation" is incumbent reallocation from the change due to the growing relative size of 1978 labor share incumbent firms (and the interaction of the growth in their size and the growth in their labor share); "Exit" is the contribution to the change from the exit of high labor share firms; and "Entry" is the contribution from the entry of low labor share firms. All calculations use micro-data from the quinquennial Census of Manufacturing. "15 year period" is the cumulated sum of each five year change over three five-year periods: e.g. -10.42% in column (1) for 1982-1997 is comprised of the sum of each 5 year period (-4.52%, -2.58%, -3.31%). "Overall" is the cumulated sum over the entire 1982-2012 period.

Table 2.5: Decompositions of the Change in the Payroll to Sales Ratio, All Sectors

| | Δ Unweighted Mean (1) | Incumbent Reallocation (2) | Exit (3) | Entry (4) | Total (5) | Δ Unweighted Mean (1) | Incumbent Reallocation (2) | Exit (3) | Entry (4) | Total (5) |
|---|---|---|---|---|---|---|---|---|---|---|
| | *A. Manufacturing* | | | | | *B. Retail* | | | | |
| 1982-1997 | -0.44 | -1.75 | -2.58 | 2.46 | -2.30 | 2.25 | -2.72 | -0.63 | 0.62 | -0.49 |
| 1997-2012 | -1.27 | -2.80 | -2.37 | 2.00 | -4.43 | 2.14 | -2.72 | -0.47 | 0.67 | -0.36 |
| 1982-2012 | -1.71 | -4.54 | -4.94 | 4.46 | -6.73 | 4.39 | -5.44 | -1.10 | 1.29 | -0.85 |
| | *C. Wholesale* | | | | | *D. Services* | | | | |
| 1982-1997 | 2.59 | -1.96 | -0.90 | 1.00 | 0.74 | 1.01 | -1.31 | 2.04 | -1.71 | 0.02 |
| 1997-2012 | 2.06 | -2.64 | -1.07 | 0.82 | -0.82 | 0.72 | 0.55 | -0.46 | -0.59 | 0.21 |
| 1982-2012 | 4.66 | -4.59 | -1.97 | 1.82 | -0.08 | 1.73 | -0.76 | 1.57 | -2.30 | 0.23 |
| | *E. Utilities and Transportation* | | | | | *F. Finance* | | | | |
| 1992-2002 | 1.48 | -2.18 | 0.14 | 0.43 | -0.12 | 2.74 | 0.20 | -1.36 | 0.87 | 2.46 |
| 2002-2012 | -1.11 | -1.07 | 0.22 | 0.26 | -1.71 | 2.17 | -0.95 | -1.54 | 1.11 | 0.79 |
| 1992-2012 | 0.37 | -3.25 | 0.36 | 0.69 | -1.83 | 4.92 | -0.75 | -2.89 | 1.98 | 3.25 |

**Notes.** This table shows the results of a decomposition of the change in the labor share using the dynamic Melitz and Polanec (2015) methodology as described in the text and notes to the previous Table. All analyses use micro-data from the quinquennial Censuses in the relevant industry.

Table 2.6: Characteristics of Concentrating Industries

| | CR4 (1) | | CR20 (2) | | HHI (3) | |
|---|---|---|---|---|---|---|
| | | | *A. Manufacturing Only* | | | |
| Patents Per Worker | 0.09 | ** | 0.057 | *** | 0.056 | ** |
| | (0.006) | | (0.022) | | (0.022) | |
| Value-Added Per Worker | 0.126 | *** | 0.074 | *** | 0.067 | *** |
| | (0.028) | | (0.020) | | (0.025) | |
| Capital per Worker | 0.067 | ** | 0.057 | *** | 0.024 | |
| | (0.029) | | (0.014) | | (0.026) | |
| 5-Factor TFP | 0.055 | *** | 0.024 | * | 0.028 | * |
| | (0.019) | | (0.013) | | (0.017) | |
| Payroll Per Worker | 0.013 | | 0.005 | | 0.016 | |
| | (0.018) | | (0.011) | | (0.010) | |
| Material Costs Per Worker | 0.120 | *** | 0.074 | *** | 0.068 | *** |
| | (0.028) | | (0.018) | | (0.023) | |
| | | | *B. All Sectors* | | | |
| Manufacturing Sales Per Worker | 0.125 | *** | 0.067 | *** | 0.069 | *** |
| | (0.027) | | (0.018) | | (0.016) | |
| Retail Sales Per Worker | 0.049 | | 0.098 | | 0.027 | |
| | (0.048) | | (0.067) | | (0.023) | |
| Wholesale Sales Per Worker | 0.16 | *** | 0.207 | *** | 0.031 | ** |
| | (0.058) | | (0.042) | | (0.013) | |
| Services Sales Per Worker | 0.082 | | 0.125 | *** | 0.041 | ** |
| | (0.055) | | (0.036) | | (0.019) | |
| Utilities/Transportation Sales Per Worker | 0.415 | *** | 0.304 | *** | 0.117 | *** |
| | (0.096) | | (0.092) | | (0.023) | |
| Finance Sales Per Worker | 0.27 | * | 0.216 | * | 0.144 | *** |
| | (0.143) | | (0.111) | | (0.052) | |
| Combined | 0.155 | *** | 0.147 | *** | 0.053 | *** |
| | (0.031) | | (0.026) | | (0.011) | |



**Notes.** This table displays regressions where the dependent variable is the change concentration. Each cell represents a separate regression. Panel A regresses concentration in the manufacturing sector on six explanatory variables. Panel B regresses concentration on sales per work in all sectors. All regressions in Panel A are weighted by value-added, and all regressions in Panel B are weighted by sales. Independent and dependent variables are standardized so coefficients reflect correlations. Regressions in are estimated as five-year differences.

THIS PAGE INTENTIONALLY LEFT BLANK

# Chapter 3

# It's What You Know That Matters: Task-Based Networks and Industry-Specific Shocks

## 3.1 Introduction

The Great Recession brought huge declines in employment and earnings within the U.S. However, not all workers were affected similarly. Previous research has shown that men, the young, and the less educated bore more of the costs of the downturn, and these differences in exposure across demographic groups are largely explained by their distributions across industries and occupations (Hoynes, Miller and Schaller (2013)). Indeed, while the Great Recession was a broad shock affecting many segments of the county, industries varied significantly in their employment experience over this period. Some industries, such as utilities or professional services, saw employment increases from 2007-2010, while other industries, such as construction or leather manufacturing, saw declines of over 10 percent.

In this project, I propose a framework that speaks to these differences in industry

labor market volatility and differences in workers' experiences in response to shocks. While economic shocks generally hit industries and the product markets they serve, workers' skills are more closely tied to the occupations in which they work (Kambourov and Manovskii 2009). Importantly, occupations are distributed unevenly across industries – some occupations, such as HR managers, are employed in many industries while other occupations, such as pharmacists, are highly concentrated in particular industries. This uneven distribution of occupations across industries means that some workers are potentially employed in only a few industries while other workers are potentially employed across several industries. An oil price shock may mean something very different to a petroleum engineer (a highly concentrated occupation) than to her secretary (a highly dispersed occupation). In this paper, I explore how these differences in the distributions of tasks and occupations affect the labor market response of industries to idiosyncratic shocks.

More formally, in the model, workers apply their skills to perform the bundle of tasks (an occupation), and industries utilize these bundles of tasks to produce output. When a given industry $j$ is hit by a negative shock, that industry decreases their demand for labor, leading to a drop in the price of the types of labor that they employ. Other industries that also employ that type of worker will benefit from the drop in their wage, which encourages them to absorb the displaced workers from industry $j$, who are now less expensive. If shocked industry $j$ is a small employer of those tasks and there are many other industries willing to hire those workers, the movement in the wage required to clear the labor market is small. Conversely, if industry $j$ is a large employer of those tasks, then the wage must move enough that industry $j$ is willing to still employ most of those workers. I assume in the model that wages are fully flexible, and derive the prediction that in industries that are the primary employer of their tasks, wages are more sensitive to shocks and employment is less sensitive. Since the market clears on the wage, more of the original workers are retained but at a lower wage.

To clarify the intuition of both the model and the empirical strategy, consider the example of two industries – Residential Care Facilities (RCFs) and Dental Offices (DOs). The main occupations employed in RCFs are social workers and medical assistants. Social workers make up 18% of the Residential Care Facilities industry, but only 8% of all social workers work in that industry. Similarly, medical assistants make up another 10 percent of the industry, but only 2% of medical assistants work in that industry. Averaging across all of the occupations utilized in the residential care facilities industry, this industry employs an average of about 3% of all people working in occupations used by RCFs. Contrast that with Dental Offices, whose main occupations are dental assistants and dental hygienists. Dental assistants make up 28% of the industry, but 91% of all dental assistants work in Dental Offices. Across all occupations used by DOs, an average of 62% of those people are employed in the Dental Offices industry. The model captures the intuition that, since RCFs are a small employer of their occupations, when the industry is hit by a negative shock and reduces its demand for labor, other non-shocked industries can easily absorb those workers. However, when Dental Offices are hit, there are fewer other industries to absorb the dental assistants, meaning that either the wage must fall more to induce DOs to retain those workers, or there is more unemployment.

In addition to exploring the occupation distribution, I also decompose these occupations into the tasks that they perform. I define tasks at several levels of disaggregation, but tasks can be thought of as core work activities that, in general, are shared by multiple occupations. Take again the previous example. Medical assistants, a main occupation of residential care facilities, and dental assistants, a main occupation of dental offices, actually perform very similar tasks, and therefore individuals could switch between these roles. Indeed, the main task of each of these occupations is to maintain health records. Therefore, in my empirics, I move from the occupation level to the task level, where instead of calculating the fraction of total employment of each occupation that is in an industry, I calculate the fraction of total task hours that are in that industry. In this case, in

the example above, the difference in the concentration of the two industries shrinks but is still apparent – residential care facilities employs on average 1% of the task-hours for its tasks while dental offices employ an average of 7%.

Using local labor market variation and government spending shocks, I find evidence that an industry's employment is more responsive to demand shocks when the industry accounts for a larger share of that task's local employment. Additionally, I find little evidence for the predicted patterns on industry wages, though nominal rigidities may be muting the effect somewhat.

This work is related to several literatures. First, several recent studies have explored the micro foundations of aggregate productivity shocks and explored the role of various inter-industry and inter-firm linkages in transmitting shocks through the system (Gabaix 2011, Acemoglu et. al. 2015, Acemoglu et. al. 2014). Gabaix (2011) showed that firm level shocks can have macroeconomic consequences only when the firm size distribution has a fat right tail. Similarly, Acemoglu, Carvalho, Ozdaglar and Tahbaz-Salehi (2012) demonstrated that industry level shocks can propagate through the economy and translate into aggregate fluctuations when there exists sizable asymmetry in the roles that sectors play as suppliers to others. Most relatedly, Acemoglu et al (2014) show that, in addition to the direct effect on employment of a trade shock on a given industry, other downstream and upstream industries are also affected through the input-output linkages. They find that incorporating the inter-industry linkages doubles the employment effect in manufacturing, where the trade shocks hit directly, and produces sizable effects outside manufacturing as well. In this project, I explore how overlaps in a different input, namely in the types of tasks they employ, affect the ways in which industries respond to shocks and what this means for the incidence of these shocks.

This project is also related to a growing literature on sectoral reallocation. Empirical studies show that there are large and persistent effects of sectoral shocks on worker outcomes (Jacobson et. al. 1993). For example, Davis and Van Wachter (2011) show that men

lose an average of 1.4 years of pre-displacement earnings upon job loss, and these losses are larger for older workers. Additionally, Dix-Caniero (2014) shows that sector mobility costs are heterogenous across workers and sectors.

Lastly, this project is also related to the literature on insurance in labor markets. For example, Yagan (2014) showed that the ability to migrate provides little insurance against labor market shocks. Two closely related recent papers show, however, that the local industry composition does provide insurance. Specifically, Muller, Stegmaier and Yi (2015) show that in areas where the labor market is more flexible and skills are more easily transferred across jobs or firms, workers earnings losses in response to trade shocks are smaller. Their paper uses matched employer-employee data in Germany and constructs distances to other sectors based on the observed frequency of wage declines between sectors. Additionally, using data from the Current Population Survey, Macaluso (2016) shows that unemployed workers from occupations that are locally less common are less likely to be employed 1 year later.

The rest of this paper is organized as follows. Section 3.2 formalizes the intuition that the dispersion of tasks across industries affects employment volatility using a simple model, Section 3.3 describes the data that I use to construct the task distribution, industry shocks, and local labor market outcomes, Section 3.4 derives and explains the empirical specifications, Section 3.5 shows the the empirical tests of the framework, and Section 3.6 concludes.

## 3.2 Theory

### 3.2.1 Model Setup

Assume that firms produce goods using inputs from other industries ($x_{ij}$) and a combination of tasks ($l_{tj}$). Acemoglu and Autor (2011) argue that the task framework is a natural

framework for interpreting patterns related to the labor market.[1] Assume that there is a constant-returns to scale Cobb-Douglas production function given by:

$$y_j = \prod_{i=1}^{N} x_{ij}^{\alpha_{ij}} \prod_{t=1}^{T} l_{tj}^{\beta_{tj}}$$

where $\sum_{i=1}^{N} \alpha_{ij} + \sum_{t=1}^{T} \beta_{tj} = 1$. Note that embedded in this model is the assumption that tasks can be completely unbundled, as in Acemoglu and Zilibotti (2001) and Acemoglu and Autor (2011).[2] I also assume here for simplicity that people cannot change their type, so a worker cannot switch and begin performing another task. Lastly, I assume that these tasks are general in that they are used in a variety of different industries but that industries differ in the importance of each task in production.

I assume that labor is supplied inelastically and the labor market clearing condition for each task market is given by:

$$\sum_j l_{jt}(w_t) = L_t = \sum_j \frac{\beta_{tj}\tilde{y}_j}{w_t} \tag{3.1}$$

where $L_t$ is a fixed supply of labor in task $t$ and $w_t$ is the wage for task $t$. Lastly, I assume that wages are fully flexible and, therefore, that task-wages are also equalized across industries.

### 3.2.2 Simple Case: Shutting Down the Input-Output Network

To begin, consider a modified production function without intermediate goods, given by

$$y_j = \prod_{t=1}^{T} l_{tj}^{\beta_{jt}} \tag{3.2}$$

[1]This setting highlights that while workers are endowed with skills which they apply to tasks to product output. While the tasks that individuals perform may be determined by their endowment of skill, the firm would care ultimately about the tasks that they are able to perform.

[2]Empirically, this can be handled in part by using occupation as a coarse definition of task.

where $\sum_{t=1}^{T} \beta_{tj} = 1$. Denote nominal output as $\tilde{y}_j = p_j y_j$ and consider a shock to demand in industry $j$ that leads to a drop in the output of industry $j : d\tilde{y}_j = \epsilon$. Due to the Cobb-Douglas structure, the effect on labor demand for task $t$ in the shocked industry $j$ is given by:

$$\frac{\partial l_{jt}}{\partial \tilde{y}_j} = \frac{\beta_{jt}}{w_t} \tag{3.3}$$

meaning that in response to a positive demand shock in the industry, the industry will increase their demand for labor in the tasks that they utilize. The effect that this increase in demand for task $t$ in industry $j$ has on overall employment of task $t$ depends on the structure of the labor market along two dimensions – the flexibility of wages and the dispersion of demands for tasks across industries. Since I am assuming that the wage is fully flexible and labor is supplied inelastically, all adjustments will occur along the wage margin and the increase in labor demand in industry $j$ will increase the wage in task $t$. However, the size of the wage increase will depend on the distribution of that task across industries. Since $l_{tk} = f(\tilde{y}_k, w_t)$ for each industry $k$, the effect on labor demand for task $t$ in each industry $k$ of a shock to industry $j$ is given by the total derivative:

$$\frac{dl_{kt}}{d\tilde{y}_j} = \frac{\partial l_{kt}}{\partial \tilde{y}_k}\frac{d\tilde{y}_k}{d\tilde{y}_j} + \frac{\partial l_{kt}}{\partial w_t}\frac{dw_t}{d\tilde{y}_j} \tag{3.4}$$

Note that this equation captures both the direct effect of the shock on the shocked industry and the general equilibrium effect of a change in wages on other industries. The total effect on each labor market $t$ from a shock to industry $j$ is calculated by summing across industries:

$$\Delta l_t = \sum_{k=1}^{n} \frac{dl_{kt}}{d\tilde{y}_j}\Delta\tilde{y}_j = \Delta\tilde{y}_j(\sum_{k=1}^{N} \frac{\beta_{kt}}{w_t}\frac{d\tilde{y}_k}{d\tilde{y}_j} - \sum_{k=1}^{N} \frac{\beta_{kt}\tilde{y}_k}{(w_t)^2}\frac{dw_t}{d\tilde{y}_j})$$

Since I assume that the labor market clears on the wage, there is no change in employment whenever the increase in labor demand coming from the shocked industry (the LHS of

Equation 3.5) is exactly offset by drops in labor demand from other industries coming from the increase in the wage (the RHS of Equation 3.5):

$$\sum_{k=1}^{N} \frac{\beta_{kt}}{w_t} \frac{d\tilde{y}_k}{d\tilde{y}_j} = \sum_{k=1}^{N} \frac{\beta_{kt}\tilde{y}_k}{(w_t)^2} \frac{dw_t}{d\tilde{y}_j} \tag{3.5}$$

Substituting in the industry's first order conditions, and imposing the labor market clearing condition (3.1), Equation 3.5 can be simplified to an expression for the elasticity of the task-level wage with respect to a demand shock for industry $j$:

$$\epsilon_{tj}^{w} = \frac{l_{jt}}{L_t} \tag{3.6}$$

Equation 3.6 shows that the elasticity of task-level wages with respect to a shock to industry $j$ is given by the fraction of the task that is in industry $j$. To solidify intuition for this result, consider the boundary case in which the market is completely segmented, each industry utilizes only 1 task, and industry $j$ faces a negative shock. In this case, each industry employs all of its only task, and $\epsilon_{tj}^{w} = 1$, meaning that the wage moves completely to clear the market. As the level of task-sharing across industries increases, the elasticity of task level wages falls, as other industries that are not directly shocked are willing to absorb the labor in task $t$ from industry $j$, and thus the wage does not need to fall by as much to maintain full employment. Note that the elasticity of task-level wages depends only on the share of the task that is employed in the shocked industry, not the way in which the task is distributed across all other industries. This is because wages for a given task are equalized across industries, and therefore, it does not matter which of the non-shocked industries re-absorbs the labor, only that there is sufficient demand outside the shocked industry to employ the workers at a given wage movement.

While overall *task* level employment is assumed to be constant, the task sharing across industries implies that in response to industry shocks, *industry* employment will change as workers switch industries in response to the shocks. The change in employment in

shocked industry $j$ is the sum of the change in task-level employment:

$$\frac{dl_j}{d\tilde{y}_j} = \sum_{t=1}^{T} \frac{dl_{tj}}{d\tilde{y}_j} = \sum_{t=1}^{T} \left( \frac{\partial l_{jt}}{\partial \tilde{y}_j} + \frac{\partial l_{tj}}{\partial w_t} \frac{dw_t}{d\tilde{y}_j} \right)$$

Again, plugging in the firm's first order conditions and the elasticity of task-level wages in Equation 3.6, we get and expression for the elasticity of industry $j$ employment with respect to a shock to industry $j$ as a function of the task distribution:

$$\epsilon^l_{jj} = 1 - \sum_{t=1}^{T} \frac{l_{tj}}{l_j} \frac{l_{tj}}{L_t} \tag{3.7}$$

Equation 3.7 shows that the response of industry employment to a shock to its own industry is decreasing in the average share of the tasks that they employ. Again, to solidify intuition, think of the case where tasks are completely segmented across industries and there is a negative shock to industry $j$. In this case, the average task share of the shocked industry is 1, meaning that the change in employment is 0. This result follows directly from the assumption of inelastic labor supply and no task switching, so task level employment is constant and, therefore, industries that are made up of a single task have constant employment. As industries increase the amount of task sharing, the responsiveness of industry employment to shocks increases. This is because the movement in the wage that is needed to clear the market is smaller, and therefore, the wages of industry $j$'s workers do not fall by as much, meaning that the industry must fire more of them in response to the shock. In the case of a positive shock to industry $j$, the industry increases their labor demand, leading to an increase in the wage of their tasks.

Lastly, the relationship between the cross-industry task distribution and the elasticity of task level wages implies that industry wages are also going to respond as a function of the task distribution. Average wages in industry $j$ are given by:

$$\bar{w}_j = \sum_{t=1}^{T} \frac{l_{tj}}{l_j} w_t$$

and the elasticity of the average industry wages is given by

$$\epsilon^w_{jj} = \sum_{t=1}^{T} \left( \frac{l_{jt}}{l_j} \frac{l_{tj}}{L_t} \frac{w_t}{\bar{w}_j} \right) = \sum_{t=1}^{T} \frac{w_t}{\bar{w}_j} \epsilon^w_{tj} \tag{3.8}$$

Equation 3.8 shows, perhaps unsurprisingly, that the elasticity of industry wages is simply a weighted average of the elasticity of task-level wages, where the weights reflect the relative importance (wage) of the workers in that task. Average wages in an industry fall more in response to a negative shock when the wages of workers in the tasks that industry employs are more elastic ($\epsilon^w_{tj}$ is higher) and when the workers in highly elastic tasks are relatively better paid ($\frac{w_t}{\bar{w}_j}$ is higher for high $\epsilon^w_{tj}$ tasks), since the same percent change to their wage has a bigger impact on the overall average.

### 3.2.3 Full Model: Adding Back the Input-Output Matrix

In this section, I consider again the full model that incorporates intermediate goods in the production function, and in doing so, I incorporate the input-output network. This additional network structure will be important if there is an overlap between the industries that share tasks with industry $j$ and industries that either supply to or buy from industry $j$. The intuition is that in response to a positive demand shock to industry $j$, industry $j$ increases its production, creating additional demand for its inputs (See Acemoglu, Akcigit, and Kerr (2015) and Acemoglu et al (2014) for details).[3] Recall from Equation 3.6 that the elasticity of task level wages in response to a demand increase to industry $j$ is decreasing in the fraction of the task that is employed in industry $j$ because there are many unshocked industries from which $j$ can hire workers at only a slight increase over their original wage. However, if the other industries employing those tasks are also industry $j$'s suppliers, they will also increase their demand for task $t$ in response to industry

[3]Since this paper empirically examines demand shocks, I focus on demand shocks througout this section. However, as in Acemoglu, Akcigit, and Kerr (2015), there would be symmetric predictions for a supply shock on downstream industries.

$j$'s increasing demand for their product. This overlap between the industries that share tasks and that supply to each other will mitigate the dampening effect of task dispersion on wage movements. Because of demand shocks that are correlated through the input-output network, industries that would have had workers leave to join $j$ instead retain most of those workers, meaning that wages must rise by more. In the case of a negative demand shock, this force operates in the other direction, meaning that wages must fall by more. In this way, the input-output network can undo the "insurance" provided by the task network against shocks, if the the task and input-output networks are correlated.[4]

Formally, again, consider a demand shock to industry $j$ s.t. $d\tilde{y}_j = \epsilon$. An increase in production in industry $j$ will translate into a proportional increase in the use of all inputs, so not only will labor demand in task $t$ increase as in Equation 3.3, but demand for goods from industry $k$ will also increase, and this increase will be given by:

$$\frac{d\tilde{x_{kj}}}{d\tilde{y}_j} = \alpha_{kj} \tag{3.9}$$

where $\tilde{x_{kj}}$ is the nominal value of industry $k$'s output used in industry $j$. As I discussed in Section 3.2.2, the crucial empirical object is the elasticity of task level wages with respect to a shock to industry $j$, $\epsilon^w_{tj}$. This is always given by:

$$\epsilon^w_{tj} = \sum_{k=1}^{N} \frac{l_{kt}}{L_t} \epsilon^{\tilde{y}}_{kj}$$

In the simple case discussed above, $\frac{d\tilde{y_k}}{d\tilde{y_j}} = 0$ for all $k \neq j$, resulting in Equation 3.6. However, this is no longer the case, and $\epsilon^y_{jk}$ contains the effect that comes from the input-output

[4]Note that this idea that correlated shocks undo the strength of the task network is perhaps broader than the input-output network. Any two industries that have correlated demand shocks (i.e. products that are affected by the same shift in consumer preferences, products that are affects by the same tax changes, etc.) would have similar predictions. In the extreme, if two industries have perfectly correlated demand shocks and share exactly the same shocks, there is no effect of task sharing between the two industries. In the model, it is as though the two industries were in fact the same sector and we are in the extreme benchmark case in which the industry employs all of the task and there would be no employment change and all adjustment would occur through the wage.

matrices as in Equation 3.9, and therefore, will be non-zero for industry other than $j$.

As in Acemoglu et al (2014), the first order effect of shock to industry $j$ on output in industry $k$ is:

$$\epsilon^{y}_{jk} = \frac{\tilde{x}_{kj}}{\tilde{y}_k}$$

where $\frac{x_{kj}}{y_k}$ is the fraction of industry $k$'s output that is purchased by industry $j$. Intuitively, this comes from the fact the industry $j$ increases their demand for the products of industry $k$ in proportion to the amount that they use, which creates a positive demand shock for industry $k$. Plugging this into the equation for $\epsilon^{w}_{tj}$, I get that, with the first order effect of the input-output table,

$$\epsilon^{w}_{tj} = \frac{l_{jt}}{L_t} + \sum_{k=1}^{N} \frac{l_{kt}}{L_t} \frac{\tilde{x_{kj}}}{\tilde{y}_k} = \mathbf{L}'_j \left(\mathbf{e}_j + \mathbf{A}_j\right)$$

where $\frac{\tilde{x}_{kj}}{\tilde{y}_k}$ is the fraction of industry $k$'s output the used in industry $j$, $\mathbf{L}$ is a column vector with fraction of task $t$ that is in industry $j$, $\mathbf{e}_j$ is a column vector with an indicator for industry $j$, and $\mathbf{A}_j$ is the $j$th column of an $\mathbf{A}$ matrix whose $kj$ entry is $\frac{\tilde{x}_{kj}}{\tilde{y}_k}$. Extending this to consider the higher order effects (i.e. the effect of the shock on the suppliers suppliers., etc), this becomes:

$$\epsilon^{w}_{tj} = \mathbf{L}'_j \left(\mathbf{e}_j + \mathbf{A}_j + \mathbf{A}_j^2 + ...\right) = \mathbf{L}'_j \left(\mathbf{I} - \mathbf{A}\right)^{-1}_j = \sum_{k=1}^{N} \frac{l_{kt}}{L_t} m_{kj}$$

where $(\mathbf{I} - \mathbf{A})^{-1}_j$ is the $j$th column of the Leontief inverse of the matrix $A$ and $m_{kj}$ is the $kj$ entry of the leontief inverse. Note that with the input-output network included, the distribution of tasks across other industries matters (i.e. $\frac{l_{kt}}{L_t}$ enters for $k \neq j$). If the tasks are concentrated in industries that produce a large fraction of industry $j$'s output (or the industries that supply those suppliers, etc.), the movement in the wage that is needed to clear the market is larger, undoing the strength of the task-sharing patterns described

above.

In the empirical section that follows, I will focus on the empirical tests in the simple case, and I leave the inclusion of the input-output network to future drafts.

## 3.3 Data Description

In order to explore the patterns derived in Section 3.2, I utilize a variety of data sources: (1) Census data and the O*NET database to construct task-based distances, (2) Quarterly Workforce Indicators (QWI) for detailed data on county-level employment and wages by industry, (3) the BEA Benchmark Input-Output tables to construct the industry-level government spending shocks shocks and the industry product market linkages, and (4) the UN Comtrade database to construct import shocks at the industry level.

### 3.3.1 Task Composition of Industries

In order to construct the task composition of each industry, I use a combination of the O*NET occupation data and the census data. First, using the five percent sample of the 2000 census downloaded from IPUMS, I calculate the fraction of employment in each industry that is in each occupation. Because I use the census, the finest level of industry disaggregation that is available is the 1990 Census Industry codes, which has 213 industries based loosely on the 1987 SIC codes.

To map the occupations and ultimately industries into tasks, I utilize the O*NET information on task categories. O*NET data provides information about required skills, education or experience that workers need for detailed occupations, and importantly for my purposes, the types of work activities that workers in a given occupation perform. Throughout my analysis, I use primarily 332 "Intermediate Work Activities" (IWAs), although I check the robustness of the results using only 36 "General Work Areas" (GWAs), examples of which can be seen in Table 3.1.

These activities span occupations but are used in different intensities across occupations. For each 6-digit standard occupation code (SOC), industry experts report the fraction of establishments in an industry that use a given work activity at a given frequency. Using this information, I construct a weighted measure of the frequency with which each 6-digit occupation uses each task.[5] I then aggregate these usage measures from 6-digit SOC codes to the Census occupation codes and calculate the percent of each occupation that is in each task.[6] Combining the information on the industry-occupation distribution and the occupation-task distribution, I calculate the fraction of hours in each industry that are devoted to each task, given by:

$$L_{it} = \sum_{o=1}^{M} E_{io} T_{ot}$$

where $L_{it}$ is the employment in industry i in task t, $E_{io}$ is the number people employed in industry i and occupation $o$, and $T_{ot}$ is the fraction of employed hours in occupation $o$ that is utilized in task $t$.

Table 3.1 and Figure 3.1 describe the resulting task distribution. Table 3.1 shows the variation at the task level, listing examples of Intermediate Work Areas, General Work Areas and occupations that are either particularly concentrated or are very dispersed across industries. These skills and occupations make intuitive sense, with hyper-specific skills like corpse embalming being concentrated, while maintaining tools and equipment is common to many industries. Figure 3.1 displays the industry level variation. The vertical axis plots the average fraction of tasks that are in each industry, using the intermediate

[5]Specifically, the 7 possible frequency of use categories are hourly or more, several times daily, daily, more than a week, more than monthly, more than yearly and yearly or less. Consider $x_{if}$, the fraction of establishments in occupation i that use task $t$ at frequency $f$. I construct the occupation's use of task t as $\sum_f x_{if} w_{if}$ where $w_{if}$ is 0 for less than yearly, $\frac{2}{365}$ for more than yearly, $\frac{1}{12}$ for monthly, $\frac{1}{7}$ for weekly, 1 for daily, 4 is for several times daily, and 8 is for hourly or more.

[6]In mapping the occupations from the 2010 SOC codes to the Census codes, I had to drop some occupations that are not covered in O*NET. These occupations are the "other" categories, for which O*NET does not provide specific information. I also had to combine Census occupation codes that did not map uniquely to a O*NET occupation. This resulted in 350 final occupation codes which I use throughout the analysis.

work areas as the task definition. For example, if the industry employed only 1 task and that industry employed 20% of task's total hours, this metric would be 0.2. This is plotted against the same metric defined using occupations instead of tasks. There are two important takeaways from this figure. First, the correlation between these two series is high but the average occupation fraction is generally higher than the average task fraction. Indeed, there are some industries such as beauty shops that employ a significantly higher fraction of their occupations than tasks. For example, in the case of beauty shops, this is because the main occupations are hairdressers and cosmetologists, which share tasks such as cleaning tools or scheduling appointments with many other occupations that are more dispersed across industries.[7] Second, this figure also shows that there is significant cross-industry variation in the task concentration, even within broad industry classes.

It is important to note that while this O*NET data is very rich, it has several limitations. First of all, the task categories are detailed, and therefore there is a high possibility for measurement error. Since the task level is very fine, the task categories may be ambiguous and repetitive in a way that confuses respondents (Autor (2013)). This concern is likely less important when aggregating to IWAs or GWAs. Additionally, the O*NET data is so rich that I had to make selection choices in aggregating the data. While constructing time-use measures seems reasonable, I have also checked the robustness of my analysis using the reported relevance or importance of the task to a given occupation as a measure industry task-fraction. I find that the task distribution is very similar across these different metrics.

### 3.3.2 Local Labor Market Data

Data on employment and wages at the county level come from the Quarterly Workforce Indicators (QWI), which is a large dataset produced by the Census Bureau and derived

[7]This also suggests that using tasks as the basis for the analysis will handle some of the concerns about individuals switching occupations, for example, since if the two occupations share many tasks, they would already be counted as being similar.

from state administrative records.[8] The dataset includes only information on UI covered employment, which represents 92% of all civilian wage and salary jobs in the US (Abowd and Vilhuber (2011)). While the majority of employment within a state is covered in the data, the sample of states available in the data varies over time as the Census Bureau built relationships with state agencies sequentially over time. In order to have a balanced panel, I restrict my analysis to the set of states with data available in 2000, and restrict my empirical analysis to 2001-2015. This data selection excludes 15 states, or around 25% percent of 2015 employment.[9] I aggregate the county-level data to commuting zones, again restricting to the set of commuting zones that are entirely contained in the states in the restricted sample, leaving me with 605 commuting zones. The QWI is released in 4-digit 2007 NAICS detail, which I convert to 201 1990 Census industry codes.[10] Lastly, while the QWI is available quarterly, my analysis is done annually so I take the annual average of the quarterly data.[11] Appendix Table B1 shows simple summary statistics for the QWI data.

I also explored the relationship between the state-level time series resulting from the QWI with the state-level time series coming from the County Business Patterns (CBP), another commonly used data source for local labor market analysis. Overall, at the state-by-industry level, the correlation between the level of the two series is 0.99 and the correlation of the annual changes is 0.68. I use the QWI instead of the CBP for my baseline

[8] A variety of administrative and survey data sources contribute to the construction of the QWI. These include Unemployment Insurance earnings data, the quarterly census of firms (QCEW), the 2000 Census, Social Security Administrative records, and Business Dynamic statistics (BDS).

[9] I have also done a version where I instead restrict to the set of states available beginning in 2001. This includes 6 additional states, but results are very similar.

[10] Some 4-digit industry-by-county cells are suppressed for privacy concerns. I impute this data using the following algorithm, which relies on the fact that total quarterly payroll is almost never suppressed. First, for each county-by-year cell, I use the 3-digit industry data to calculate the amount employment missing in the sub-industries and then I attribute that employment to the sub-industry based on that industry's share of quarterly payroll. If the 3-digit industry data is also suppressed, I do the same procedure using the data at the 2-digit industry level. This procedure results in some very noisy time series at the industry-by-CZ level. Therefore, in my baseline specifications, I hp-filter the data with a filter parameter of 0.5, which smooths out some of the high frequency noise in the time series. However, in the appendix, I report the estimates on the raw data as well and find qualitatively similar patterns.

[11] Results are also robust to using the quarter-to-quarter changes.

analysis for two reasons. First, it has fewer missing or suppressed county by industry cells, which allows me to get more accurate wage and employment time series at the commuting zone level, and second, the QWI contains additional information on labor market dynamics not contained in the CBP.

### 3.3.3 Industry Shocks

To construct industry level government spending shocks, I follow Acemoglu, Akcigit and Kerr (2015) in using the 1992 BEA Input-Output matrix to construct the share of the sales for each industry that go to the federal government. I then interact this with the lagged total change in federal spending, and construct industry level government spending shock $FED_{it}$ which is given by:

$$FED_{it} = \frac{Sales^{1992}_{i,FED}}{Sales^{1992}_{i}} \Delta(FED_{t-1})$$

for each industry $i$ and time period $t$. The share of the industry consumed by the government in 1992 varies from 0 (10% of industries) to 60%. Examples of industries with high government consumption shares are guided missiles, space vehicles, and parts (60%), Ordnance (58%), Research, Development and Testing Services (20%), and Computer and Data Processing Services (15%).

In an alternative specification, I replace total federal spending with state-level federal spending on defense. This data was constructed by Nakamura and Steinsson (2013) and originally came from an electronic database of DD-350 military procurement forms from the US Department of Defense. The forms list the prime contractor and the location that the majority of the work is performed. Because I use their pre-constructed data, this time series ends in 2006. I also replace the share of each industry that goes to the federal government with the share that goes to national defense only.

## 3.4 Empirical Application

In order to empirically test for the patterns derived in Section 3.2, I extend the framework to simultaneously consider idiosyncratic shocks to all industries and derive expressions for the relationship between industry employment and wages and the underlying task distribution. Considering idiosyncratic shocks to all industries at once, and abstracting from the input-output linkages, Equation 3.7 becomes

$$\frac{\Delta l_j}{l_j} = \frac{\Delta \tilde{y}_j}{\tilde{y}_j} - \frac{\Delta \tilde{y}_j}{\tilde{y}_j} \sum_{t=1}^{T} \frac{l_{tj}}{l_j} \frac{l_{jt}}{L_t} - \sum_{k \neq j}^{N} \left( \sum_{t=1}^{T} \frac{l_{tk}}{L_t} \frac{l_{jt}}{L_t} \right) \frac{\Delta \tilde{y}_k}{\tilde{y}_k} \tag{3.10}$$

where $t$ is the task and $j$ is the industry. The first term is simply the direct effect of industry $j$'s own demand shock on employment. The second term captures the relationship between the task structure and industry $j$'s response to its own shock – industry $j$'s employment change in response to its own shock is decreasing in the share of the total task that industry $j$ employs, averaging this fraction across all tasks that industry $j$ employs and weighting each task by its share of that industry's employment. The third term captures the indirect effect of shocks to other sectors with whom industry $j$ shares tasks. When other industries receive positive demand shocks, they increase their demand for their tasks, leading to a increase in the wage for those tasks in all industries. That increase in wages reduces industry $j$'s demand for those tasks.

Similarly, when extended to consider shocks to all industries, Equation 3.8 becomes:

$$\frac{\Delta \bar{w}_j}{w_j} = \frac{\Delta \tilde{y}_j}{\tilde{y}_j} - \frac{\Delta \tilde{y}_j}{\tilde{y}_j} \sum_{t=1}^{T} \frac{l_{jt}}{l_j} \frac{w_t}{\bar{w}_j} \frac{(L_t - l_{jt})}{L_t} + \sum_{k \neq j}^{N} \left( \sum_{t=1}^{T} \frac{l_{jt}}{l_j} \frac{w_t}{\bar{w}_j} \frac{l_{kt}}{L_t} \right) \frac{\Delta \tilde{y}_k}{\tilde{y}_k} \tag{3.11}$$

Again, the first term captures the direct positive effect of the demand shock on the average wage in industry $j$. The second term captures the way in which the task network mitigates the initial shock – when other industries employ a large share of the tasks (and consequently industry $j$ employs a small share), the movement in the average wage should

be smaller. Lastly, the third term captures the change in the average wage in industry $j$ from shocks to other industries with whom industry $j$ shares tasks.

I explore the empirical predictions of this framework using local labor market variation, and throughout the empirical analysis, my main estimating equations are the direct empirical analogs of Equations 3.10 and 3.11 and take the following forms:

$$\Delta E_{ics} = \beta_0 + \beta_1 Shock_{is} + \beta_2 B(Shock_{is}) + \beta_3 C(Shock_{is}) + \gamma_{cs} + \gamma_{is} + \gamma_{ic} + \epsilon_{ics} \quad (3.12)$$

$$\Delta \bar{W}_{is} = \alpha_0 + \alpha_1 Shock_{is} + \alpha_2 Y(Shock_{is}) + \alpha_3 Z(Shock_{is}) + \gamma_{cs} + \gamma_{is} + \gamma_{ic} + \epsilon_{ics} \quad (3.13)$$

where

$$B(Shock_{is}) = Shock_{is} \sum_{t=1}^{T} \frac{l_{ti,2000}}{l_{i,2000}} \frac{l_{it,2000}}{L_{t,2000}}$$

$$C(Shock_{is}) = \sum_{k \neq j}^{N} \left( \sum_{t=1}^{T} \frac{l_{tk,2000}}{L_{t,2000}} \frac{l_{it,2000}}{L_{t,2000}} \right) Shock_{ks}$$

$$Y(Shock_{is}) = Shock_{is} \sum_{t=1}^{T} \frac{l_{jt,2000}}{l_{j,2000}} \frac{w_{t,2000}}{\bar{w}_{j,2000}} \frac{(L_{t,2000} - l_{jt,2000})}{L_{t,2000}}$$

$$Z(Shock_{is}) = \sum_{k \neq j}^{N} \left( \sum_{t=1}^{T} \frac{l_{jt,2000}}{l_{j,2000}} \frac{w_{t,2000}}{\bar{w}_{j,2000}} \frac{l_{kt,2000}}{L_{t,2000}} \right) Shock_{ks}$$

The first subscript $i$ denotes the industry, the second subscript $c$ denotes the commuting-zone, the third subscript $s$ denotes the time period, and $t$ denotes the task. The lettered terms are almost identical to the terms in Equations 3.10 and 3.11 but fix the task-distribution in 2000 rather than letting it vary over time. It is important to fix the task-

distribution as this may change endogenously over time as industries are hit by shocks. The parameters of interest in the employment equation are $\beta_2$ and $\beta_3$, which are the coefficients on the task-network terms derived above. As Equation 3.10 shows, the hypothesis is that $\beta_2 < 0$ and $\beta_3 < 0$. Similarly, the parameters of interest in the wage equation are $\alpha_2$, which is predicted by Equation 3.11 to be negative, and $\alpha_3$, which is predicted to be positive.

The fixed effects $\gamma$ are essential for identification. First, $\gamma_{cs}$ are year-by-local labor market fixed effects, allowing for employment growth to vary flexibly across commuting zones. I also include industry-by-local labor market fixed effects ($\gamma_{ic}$), which, since the equation is already differenced, allows for industries to be on different linear trends in different commuting zones. Lastly, I also include industry-by-year fixed effects ($\gamma_{is}$). In some specifications, I included the full set of detailed industry-by-year fixed effects, which completely absorb the main effect of the shock on local employment (i.e. $Shock_{is}$), but allow the other terms to remain. In other specifications, I include only parent industry (2 digit industry)-by-year fixed effects and thus get an estimate for the main effect. With the inclusion of all of these two-way fixed effects, identification comes from within-industry across-labor market variation. While I assume that industries have the same task structure across labor markets, commuting-zone level variation in the task share of an industry comes from the local industry mix of an area. Differences in the local industry mix create variation in the task share in two ways. First, across labor markets, industries make up varying shares of total local employment. When an industry is a larger share of the local labor market, it is likely to employ a larger fraction of its tasks. Second, even for a given industry size, the task composition of the other industries in the area differ in the degree to which they share tasks. If the other industries in the area utilize similar tasks, the industry is likely to be less concentrated.[12]

[12]Note that it is possible that industries that are a larger portion of the local labor market will respond differently to shocks because of their relative size, not because of the level of task dispersion. For example, the general equilibrium effects of higher industry wages may be stronger when the industry is a larger share of local employment, creating a positive relationship between the local task share and employment response

Figure 3.2 shows two examples of this local variation. The top map shows the average fraction of commuting zone employment that is in the motor vehicles manufacturing industry. Unsurprisingly, motor vehicle manufacturing in commuting zones around Michigan employ a large share of its tasks, while car manufacturers on the coasts employ a smaller fraction of those tasks. The lower panel plots the same metric but for the advertising industry. This industry is on average has lower task concentration than car manufacturing, and relatively more concentrated around New York City than the Midwest.[13]

Utilizing this rich local labor market variation in the dispersion of tasks across industries to identify the relationship between the task structure and industry outcomes is appealing for both theoretical and statistical reasons. The mechanism driving the predictions of the model is the change in the task level wage, which is the result of workers within a task moving across across industries. Since within local labor market transitions are move likely than cross local labor market moves, the local labor market level is the most natural level of geographic variation to utilize. Additionally, the rich variation at the local labor market level yields additional power, which is particularly important because I am limited to the Census industry disaggregation, which limits the amount of industry variation. This identification strategy, however, relies on two important assumptions. First, I must assume that there is not anything else at the local labor market level that affects industry sensitivity to shocks and that is correlated with the task concentration of that industry. For example, if local governments fear that task-concentrated industries are particularly

that is unrelated to the task-share across industries. However, since the local industry size and the average task-share are positively correlated, this would bias against finding the patterns predicted by the model. To briefly explore this distinction between the size and the task distribution, in Appendix Table B4, I show a version where I separately include the fraction of the local labor market employed in the industry interacted with the national change in output of that industry (an extension of Table 3.2). I find evidence that, as expected, there is a positive relationship between the industry's size in the local market and its co-movement with the national trends. However, the estimates relationship between the task share and the national trend is still statistically significant and negative.

[13]Note that the CZs that have no data are CZs that are partially contained in states that are not in the QWI sample in 2000 and therefore are not included in the analysis.

hard-hit by shocks and therefore target local stabilization policies to that industry, this assumption would be violated and my results would be biased. It would appear as though task-concentrated industries have less volatile employment as the framework predicts, but it would be occurring not through the greater wage movement but rather through the application of policy. Second, in exploring how different local labor markets respond to a national shock, I must assume that demand shocks at the local level are not differentially correlated with the national demand shocks in a way that covaries with local task structure.

Throughout the next section, I utilize several different shocks in estimating Equation 3.12 and 3.13. I begin by simply using national changes in industry output as a measure of changing national demand. While this is certainly not a shock to the industry, this specification answers the question "Does a national trend in an industry translate differently to a local labor market depending on the local task composition of the area?" Since the regression includes local labor market-by-year fixed effects, I already account for anything in that year that affects overall growth in a commuting zone, such as population growth in the commuting zone or local-level policy changes. Additionally, the regression includes industry-by-commuting zone fixed effects, which allows employment in industries to be on different linear trends in different commuting zones. Therefore, this specification is identified under the assumption that the national change in output reflects demand changes, not only supply changes, and that the local task composition does not directly affect the way that the local industry absorbs productivity changes.

In addition to this more descriptive analysis, I consider government spending shocks, which capture the idea that industries that have a high share of their output purchased by the government will be more "shocked" by changes in government spending. There are two important identifying assumptions in this context. First, it must be the case that changes in total government spending are uncorrelated with productivity shocks hitting the sub-industry, and that, for example, the government does not spend more overall

because of the condition of the ship building industry. This assumption is discussed more in length in Acemoglu, Akcigit and Kerr (2015) and is likely reasonable. Since I am also using local variation, I need an additional assumption in my setting that was not relevant in their analysis. I must also assume that government spending at the industry level is not geographically concentrated in a way that is correlated with the task network. For example, if government spending is concentrated in large labor markets and large labor markets also tend to have lower average task shares, this would bias my results.

A complementary method for constructing government spending shocks that is less subject to this concern is to use state-level defense spending rather than national total government spending, as in Nakamura and Steinsson (2014). This state level spending better captures the targeting of government spending across regions. I convert this state-level spending to commuting-zone spending by assuming that each county received a share of the defense spending that is proportional to its population share and then summing this across counties within the commuting zone. Of course, the assumption that government defense spending is equally distributed by population within a state across counties according to population is not correct, but may be a sufficient approximation.

## 3.5 Empirical Results

In the subsections below, I show the estimated relationships between industry shocks, industry outcomes, and the role that the task-distribution plays in explaining that.

### 3.5.1 National Changes

As a first test of the model's predictions, Table 3.2 shows the estimates of Equation 3.12 and 3.13 using national changes in industry output as the "shock" variable. This specification includes rolling 5-year changes and focuses on tasks defined at the intermediate level. Because output is only readily available at this level of industry disaggregation

within manufacturing, this analysis restricted to manufacturing industries. The left panel of the table shows the results for employment and the right table shows the results for wages. The first two columns show the specification without the full set of industry-by-year fixed effects. Unsurprisingly, there is a strong positive relationship between national industry output and local employment. More interesting, though, the coefficient on the own task share is also negative and significant – when industries employ a larger fraction of their tasks, the co-movement of output and employment is lower. This is exactly what the framework predicted – when industries employ a larger fraction of their tasks, the labor market clears more on the wage than employment, meaning that, in the case of a negative shock, the industry retains more of their workers but at a lower wage, and in the case of a positive shock, the larger increase in the wage offsets the desire to hire. These same pattens can be seen, and even more strongly, in column 3, in which I include the full set of industry and year fixed effects. In column 2 and 4, I include the full specification and include the third term that comes from output increases in other industries with whom one shares tasks. Again, as the framework predicted, this term enters negatively – when other industries with whom an industry share tasks are growing, the industry's employment grows less. The magnitudes of these relationships are are meaningful – using my preferred specification in column 4, I get that a one standard deviation increase in the average fraction of tasks decreases the elasticity of employment with respect to national output changes by 0.05. This is a non-trivial amount given that the estimated average elasticity is estimated to be around 0.13.

In the appendix, Table 3.2 shows that these patterns are largely robust to several modifications. Specifically, the same negative relationship between the task share and employment fluctuations is evident when considering a smaller set of more general tasks as well as more aggregate occupations. Interestingly, however, it is not robust to using national employment, rather than national output, as the measure of aggregate demand increases. In fact, in this case, the signs flip and having a higher share of tasks is associated with

local employment changes that are more correlated with national changes. This may be in part because by using national employment changes, I am already conditioning on the change in demand translating into employment, when the model predicts that the trade-off between translating into employment and into wage changes is precisely where the task network matters.

The right panel of Table 3.2 shows the analogous results for average industry wages, and here the results are more mixed. The first two columns show, unsurprisingly, that average wages increase when national output in the industry increases. However, the terms interacting this with the task distribution do not support the predictions of the model. When an industry has a small share of the tasks, the movement in the wage should be smaller. However, the estimates on the $Y$ term in Table 3.2 are slightly positive, although very tiny. The $Y$ term is large when the industry employs a small share of the tasks, therefore a positive coefficient implies that the wage moves more when the industry has a small share of the tasks. The estimates on the third term, which captures between shocks to other industries and wages in the primary industry, are positive, which is the anticipated sign, but are insignificant.

How does one square the significant results on employment with the insignificant results on the wage when the mechanism in the model driving the results on employment is the wage movement? One possibility is that the movement in the average monthly wages paid are too small to detect, as a change in the wages of a few workers within the industry may not move the average very much. Another possibility is that the average wage masks changes in wages that are offset by changes in selection. If in response to positive shocks, industries hire more workers but those workers are less good, their wages will be lower, bringing down the average.[14]

[14]Indeed, average wages for new hires are significantly lower than the average wages for incumbent workers. See the summary statastics table in the appendix.

### 3.5.2 Government Spending Shocks

Table 3.3 shows the relationship between local employment and demand shocks using the government spending shock described in Section 3.4. This analysis differs from the analysis in the previous section in two ways. First, the regressions include sectors outside of manufacturing. Second, the shocks are solely demand shocks, rather than a combination of both supply and demand movements at the national level. The left panel of Table 3.3 shows the results using total national government spending while the right panel shows the results using state-level defense spending.

The first column of each panel shows that, as suspected, there is positive relationship between the government spending shock and industry employment – when governments increase their demand for the industry, the industry employs more workers. This effect is positive and marginally significant with national spending and imprecisely estimated with state defense spending.[15] More interesting, however, is the negative estimate on the second term, which shows a negative relationship between the average task share and the movement in employment. This is as the model predicted – industries that have a higher average task share have employment that is less responsive to government demand shocks. The third column in each panel adds the effect coming from the task share of other industries, and, at least in the specification with national government spending movements, there is the predicted negative relationship here, meaning that when other industries with whom the industry shares tasks are also shocked, the employment growth in the industry attenuated. The intuition for this in the model is that the other industries put upward pressure on the wages of those tasks. Column 4 is an alternate specification that includes the full set of industry-by-year fixed effects. This soaks up all of the variation in the government spending shock, leaving only the interaction term, and yet, it yields similar patterns. Appendix Table B5 shows that these patterns are robust to the task

[15]Note that the state defense spending estimates are estimates on a smaller sample period, and consequently, on a subset of the commuting zones who were present early enough to have 5-year changes in 2006.

definition that I use, as well as looking at 1-year changes or weighting by the industry's share of employment. Again, consider the magnitudes of these estimates. A one standard deviation increase in the average task share of the industry decreases the elasticity of employment with respect to a government spending shock by 0.015. The average estimates elasticity is 0.019, making this an 80 percent decrease from the average effect.

While the results on employment broadly support the predictions of the model, the results on wages shown in Table 3.4 are again noisy and inconclusive. There is no effect of government spending on the average industry wages, and this is true irrespective of the task structure. As I discussed above, this is likely the result of a low power to detect movements in average wages.

## 3.6 Conclusion

In this paper, I examined how the dispersion of tasks across industries affects the volatility of industry wages. I proposed a model in which industries share tasks and in which wages at the task level are equalized across industries. The movement in the task-level wage that is needed to clear the market after an idiosyncratic shock to a given industry is increasing in the fraction of that task that the shocked industry employs. When wages are fully flexible and the market clears on the wage, this results in the prediction that industry employment is less responsive to shocks when the fraction of the tasks employed by that industry is high. After deriving the predictions, I proceeded to test the specific empirical predictions of the model using local labor market variation. Using government spending shocks, I found that employment in industries that on average employ a higher fraction of their industries is less responsive to demand shocks. However, I did not find any significant movements in average industry wages, which I suspect is at least in part because average industry wages are too crude of a metric to capture small changes.

## References


Abowd, John M. and Lars Vilhuber "National Estimates of Gross Employment and Job Flows from the Quarterly Workforce Indicators with Demographic and Industry Detail" *Journal of Econometrics*, Vol. 161, 2011. 82-99

Acemoglu, Daron and David H. Autor, "Skills, Tasks and Technologies: Implications for Employment and Earnings," *Handbook of Labor Economics*, 2011, 4B, 1043–1171.

Acemoglu, Daron, David Autor, David Dorn, Gordon H. Hanson and Brendan Price. "Import Competition and the Great U.S. Employment Sag of the 2000s" *Journal of Labor Economics*, forthcoming

Acemoglu, Daron, Ufuk Akcigit, William Kerr. "Networks and the Macroeconomy: An Empirical Exploration" *NBER Macroeconomics Annual 2015*, Volume 30, Eichenbaum and Parker. 2016

Acemoglu, Daron, Vasco Carvalho, Asuman Ozdaglar and Alireza Tahbaz-Salehi "The Network Origins of Aggregate Fluctuations" *Econometrica*, Vol. 80, No. 5. September 2012. 1977-2016.

Acemoglu, Daron and Fabrixio Zilibotti. "Productivity Differences," *Quarterly Journal of Economics*, 2001, 116 (2), 563–606.

Autor, David H., "The "task approach" to labor markets : an overview," *Journal for Labour Market Research*, vol. 46(3), pages 185-199

Autor, David, David Dorn and Gordon Hanson. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States" *American Economic Review* Vol. 103, No 6, 2013. 2121-2168.

Autor, David H. and Michael J. Handel "Putting Tasks to the Test: Human Capital, Job Tasks and Wages" *Journal of Labor Economics*, Vol. 31, No. 2, April 2013, S59-S96.

Autor, David .H., Frank Levy and .Richard J. Murnane. "The Skill Content of Recent Technological Change: An Empirical Exploration". *Quarterly Journal of Economics* Vo.118, No. 4, November 2003, 1279–1333

Autor, David and Brendan Price "The Changing Task Composition of the US Labor Market: An Update of Autor, Levy, and Murnane (2003)", 2013

Ambrus A, Mobius M, Szeidl A. "Consumption Risk-sharing in Social Networks." *American Economic Review.* 2014. Vol. 104, No. 1, 149-182

Blinder, Alan .S. "How many US jobs might be offshorable?" *World Economics*. Vol. 10, No.2, 2009. 41–78

Bloom, Nicholas, Mark Schankerman and John Van Reenen. "Identifying Technology Spillovers and Product Market Rivalry" *Econometrica* Vol. 81, No. 4. July 2013, 1347-1393

Chang, Yongsung and Jay H. Hong "Do Technological Improvements in the Manufacturing Sector Raise or Lower Employment." *American Economic Review*. Vol. 96, No. 1. March 2006, 352-368.

Cortes, Guido Matias and Giovanni Gallipoli "The Costs of Occupational Mobility:

An Aggregate Analysis" mimeo

Fafchamps, Marcel and Flore Gubert. "Risk Sharing and Network Formation." *American Economic Review*. Vol. 97, No. 2. May 2007, 75-79.

Fisher, Jonas D.M and Ryan Peters. "Using Stock Returns to Identify Government Spending Shocks." *The Economic Journal*, 120, May 2009. 414–436

Gabaix, Xavier "The Granular Origins of Aggregate Fluctuations" *Econometrica*, Vol. 79, No. 3. May 2011, 733-772

Gathmann, Christina and Uta Schonberg. "How General Is Human Capital? A Task-Based Approach" Journal of Labor Economics. Vol. 28, No. 1 2010.

Hoynes, Hilary W., Douglas L. Miller, and Jessamyn Schaller. 2012. "Who Suffers during Recessions?" *Journal of Economic Perspectives* Vol. 26, No. 3. 27-48.

Jaffe A. "Technological Opportunity and Spillovers of R&D: Evidence From Firms' Patents, Profits and Market Value,"*American Economic Review*, Vol. 76, No. 5, December 1986, 984–1001.

Macaluso, Claudia. "Skill-remoteness, Mismatch and Employment Fluctuations in Heterogenous Labor Markets." mimeo.

Muller, Steffen, Jens Stegmaier and Moises Yi. "Local Labor Markets and the Incidence of Trade Shocks." mimeo.

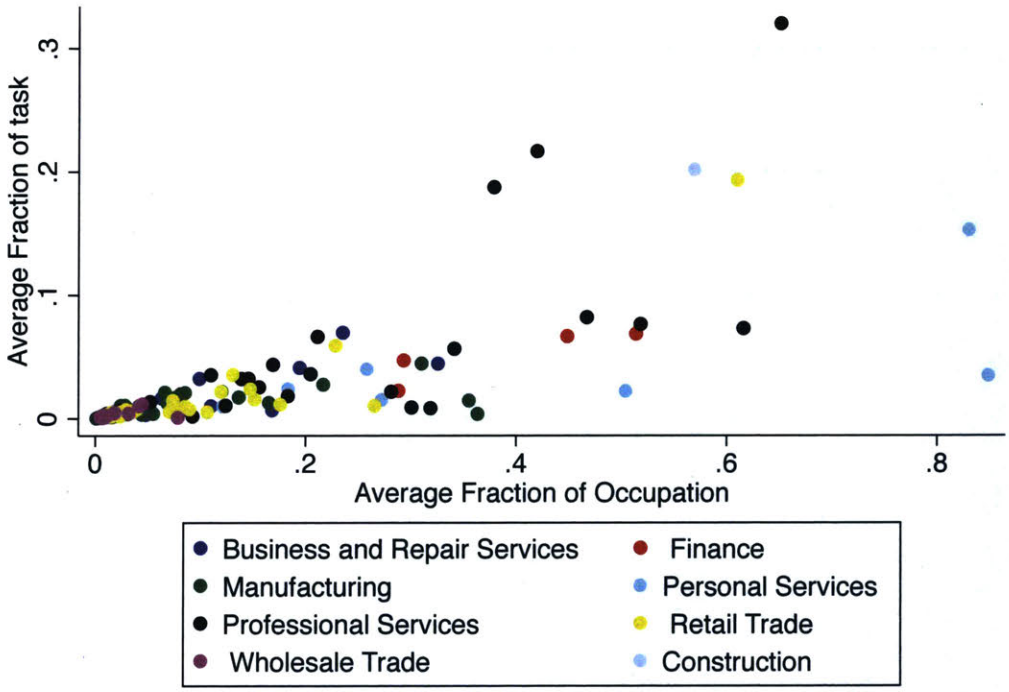
Nakamura, Emi and Jon Steinsson. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions" *American Economic Review*, Vol. 104, No 3. 2014, 753-792.

Shaw, Kathryn. "A Formulation of the Earnings Function Using the Concept of Occupational Investment" The Journal of Human Resources, Vol. 19, No 3. Summer1984, 319-340.

Kambourov, Gueorgui and Iouri Manovskii. "Occupational Specificity of Human Capital" *International Economic Review* Vol. 50, No. 1, 2009. 63-115.

## Figures and Tables

Figure 3.1: Description of Task Dispersion Across Industries




Source: Author's Calculations using O*NET and Census data

Figure 3.2: Cross Commuting Zone Variation in Industry Taks Concentration

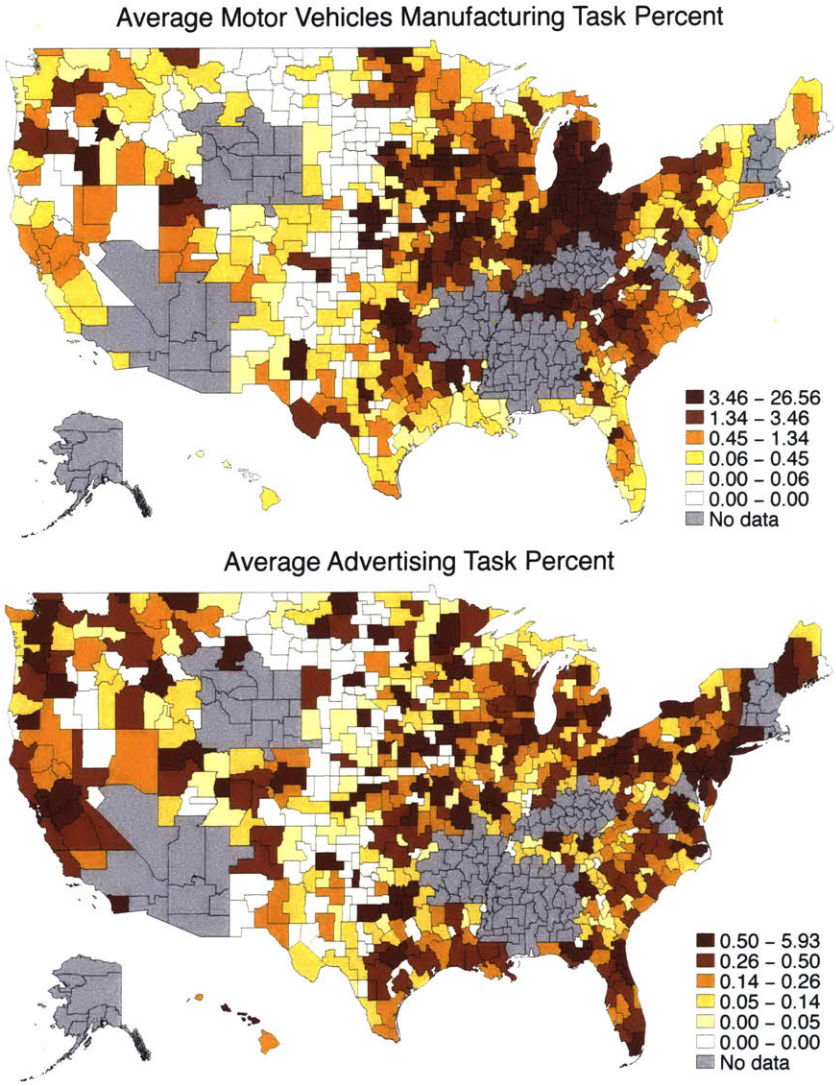


Table 3.1: Examples of Concentrated Tasks and Occupations

| Intermediate Work Areas (IWA) | |
|---|---|
| Concentrated | Un-Concentrated |
| Embalm corpses | Maintain tools/equipment |
| Test for environmental hazards | Examine materials for accuracy |
| Train animals | Prepare financial documents |
| Make legal decisions | Develop plans to preserve natural resources |
| Groom and style hair | Engrave objects |
| **General Work Areas (GWA)** | |
| Concentrated | Un-Concentrated |
| Getting Information | Establishing Interpersonal Relationships |
| Monitor Processes, Materials, or Surroundings | Coaching and Developing Others |
| Documenting/Recording Information | Staffing Organizational Units |
| Performing General Physical Activities | Organizing, Planning, and Prioritizing Work |
| Communicating with Supervisors/Peers/Subordinates | Developing Objectives and Strategies |
| **Occupations** | |
| Concentrated | Un-Concentrated |
| Desk Clerks | Record Keepers |
| Insurance Sales Agents | HR Managers |
| Police Officers/Detectives | Managers in Marketing/PR |
| Cement Masons | Maintenance and Repair Workers |
| Clergy | Shipping Clerks |

Table 3.2: Relationship Between Local Industry Employment and Wages and National Output

| | Dependent Variable: Δ ln(Employment) | | | | Dependent Variable: Δ ln(Wages) | | | |
|---|---|---|---|---|---|---|---|---|
| $\Delta logY_{it}$ | 0.126*** | 0.132*** | | | 0.014** | 0.014** | | |
| | (0.023) | (0.024) | | | (0.006) | (0.006) | | |
| Own Task Share (B) | -0.916*** | -2.474*** | -1.046*** | -2.694*** | | | | |
| | (0.195) | (0.372) | (0.205) | (0.336) | | | | |
| 1-Own Task Share (Y) | | | | | 0.001* | 0.001* | 0.001 | 0.001* |
| | | | | | (0.000) | (0.001) | (0.001) | (0.001) |
| Other Task Share (C)/(Z) | | -1.857*** | | -1.973*** | | 0.022 | | 0.106 |
| | | (0.330) | | (0.283) | | (0.054) | | (0.073) |
| State x Year Dummies | parent | parent | yes | yes | parent | parent | yes | yes |
| Industry x State Dummies | yes | yes | yes | yes | yes | yes | yes | yes |
| Industry x Year Dummies | yes | yes | yes | yes | yes | yes | yes | yes |
| No. Observations | 234720 | 234720 | 234720 | 234720 | 234669 | 234669 | 234669 | 234669 |
| $R^2$ | 0.480 | 0.480 | 0.483 | 0.484 | 0.396 | 0.396 | 0.399 | 0.399 |
| Dep. var. mean | -0.087 | -0.087 | -0.087 | -0.087 | 0.141 | 0.141 | 0.141 | 0.141 |
| Mean $\Delta logY_{it}$ | 0.102 | 0.102 | | | 0.102 | 0.102 | | |
| Mean Own Task Share (B)/(Y) | 0.002 | 0.002 | 0.002 | 0.002 | 0.586 | 0.586 | 0.586 | 0.586 |
| Mean Other Task Share (C)/(Z) | 0.087 | 0.087 | 0.087 | 0.087 | 0.081 | 0.081 | 0.081 | 0.081 |

Notes: The sample includes annual data for 71 manufacturing industries and 605 commuting zones from 2001-2011. Regression includes rolling 5 year changes and tasks are defined using Intermediate Work Activities (IWA). Standard errors are twoway clustered at the state and industry level.

Table 3.3: Relationship Between Local Industry Employment and Government Spending Shocks

| | Dependent Variable: Δ ln(Employment) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | National Gov. Spending | | | | State Defense Spending | | | |
| Gov. Spending | 0.011* | 0.016** | 0.019** | | 0.000 | 0.001 | 0.001 | |
| | (0.006) | (0.007) | (0.008) | | (0.001) | (0.001) | (0.001) | |
| Own Task Share (B) | | -0.638*** | -0.755*** | -1.042*** | | -0.223** | -0.221** | -0.185** |
| | | (0.177) | (0.153) | (0.213) | | (0.097) | (0.098) | (0.089) |
| Other Task share (C) | | | -0.416*** | -0.562*** | | | 0.009 | 0.033** |
| | | | (0.110) | (0.129) | | | (0.015) | (0.014) |
| State x Year Dummies | parent | parent | parent | yes | parent | parent | parent | yes |
| Industry x State Dummies | yes | yes | yes | yes | yes | yes | yes | yes |
| Industry x Year Dummies | yes | yes | yes | yes | yes | yes | yes | yes |
| No. Observations | 749544 | 749544 | 749544 | 749544 | 274238 | 274238 | 274238 | 274238 |
| $R^2$ | 0.373 | 0.373 | 0.373 | 0.381 | 0.701 | 0.701 | 0.701 | 0.705 |
| Mean Dep. Var | -0.011 | -0.011 | -0.011 | -0.011 | -0.011 | -0.011 | -0.011 | -0.011 |
| Mean Gov. Shock (A) | 0.391 | 0.391 | 0.391 | 0.391 | 0.139 | 0.139 | 0.139 | 0.139 |
| Mean Own Task Share (B) | | 0.007 | 0.007 | 0.007 | | 0.002 | 0.002 | 0.002 |
| Mean Other Task Share (C) | | | 0.284 | 0.284 | | | 0.070 | 0.070 |

Notes: The sample with total government spending includes annual data for 141 industries and 605 commuting zones from 2001-2014. The sample with state defense spending goes from 2001-2006. Standard errors are twoway clustered at the state and industry level. Regressions includes state-by-year fixed effects, industry-by-state fixed effect, and parent-industry-by-time fixed effects.

Table 3.4: Relationship Between Local Industry Wages and Government Spending Shocks

| Dependent Variable: $\Delta$ ln(Wages) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | National Gov. Spending | | | | State Defense Spending | | | |
| Gov. Spending | 0.000 | 0.000 | 0.000 | | -0.000 | -0.000 | -0.000 | |
| | (0.001) | (0.001) | (0.001) | | (0.000) | (0.000) | (0.000) | |
| 1-Own Task Share (Y) | | -0.000 | -0.000 | -0.000 | | -0.000 | -0.000 | 0.000 |
| | | (0.000) | (0.000) | (0.000) | | (0.000) | (0.000) | (0.000) |
| Other Task share (Z) | | | -0.006 | 0.017 | | | -0.005 | -0.003 |
| | | | (0.011) | (0.014) | | | (0.004) | (0.003) |
| State x Year Dummies | parent | parent | parent | yes | parent | parent | parent | yes |
| Industry x State Dummies | yes | yes | yes | yes | yes | yes | yes | yes |
| Industry x Year Dummies | yes | yes | yes | yes | yes | yes | yes | yes |
| No. Observations | 749386 | 749386 | 749386 | 749386 | 274182 | 274182 | 274182 | 274182 |
| $R^2$ | 0.285 | 0.285 | 0.285 | 0.291 | 0.623 | 0.623 | 0.623 | 0.626 |
| Mean Dep. Var | 0.146 | 0.146 | 0.146 | 0.146 | 0.146 | 0.146 | 0.146 | 0.146 |
| Mean Gov. Shock (A) | 0.391 | 0.391 | 0.391 | 0.391 | 0.139 | 0.139 | 0.139 | 0.139 |
| Mean Own Task Share (B) | | 1.099 | 1.099 | 1.099 | | 0.260 | 0.260 | 0.260 |
| Mean Other Task Share (C) | | | 0.302 | 0.302 | | | 0.074 | 0.074 |

Notes: The sample with total government spending includes annual data for 141 industries and 605 commuting zones from 2001-2014. The sample with state defense spending goes from 2001-2006. Standard errors are twoway clustered at the state and industry level. Regressions includes state-by-year fixed effects, industry-by-state fixed effect, and parent-industry-by-time fixed effects.

# A Appendix for Chapter 1

## LEHD data

This project uses a 23-state subset of the Longitudinal Employer-Household data set, or LEHD. The LEHD is an unbalanced panel of states, and Table A1 reports the years for which each state is included in this analysis. The blue line in Figure A1 shows the number of individuals in those states in each year. I exclude individuals for whom I am missing industry and firm information. For each state, I drop the first two years of available individual-level data, as lagged incomes are not well-defined for workers in those years. In addition, I drop the first two years in which an individual appears in the entire sample. For the majority of individuals, this is the same restriction as dropping the first two years of the state. However, for workers who appear in the sample in the later years (e.g., workers moving into the sample, young workers, etc.), this is an additional restriction that ensures that the lagged earnings, and thus marginal propensities to consume, or MPCs, are well-defined for all workers. The green line shows that this two-year lag restriction excludes about 20 percent of the sample in each year. Additionally, in order to abstract from education and retirement decisions, I also restrict my attention to workers between 25 and 62 years old. This excludes around another 12 percent of the original sample.

One potential concern with the rolling-panel structure of the LEHD is that the magnitude of the measurement error in constructing the total earnings series is changing over time. Both in constructing lagged incomes for the MPC estimation and in defining income processes over the business cycle, I rely on the full set of earnings across all states for the individual. As states enter the sample over time, total earnings for individuals employed across multiple states may jump artificially. In order to address these issues, I supplement the baseline analysis with analysis run on a balanced panel of states using the subset of states with data available by 1993. These states, listed in bold in Table A1, represent 70 percent of the workers in the full sample. While I only report the robustness for the key findings, this subsample produces very similar patterns throughout the entire analysis.

The LEHD provides a comprehensive snapshot of employment in each quarter, but it does not provide information on labor market activity for workers in periods when they are not employed within this sample. Therefore, I must take a stance on the labor market activity of workers who leave employment during the sample period. This assumption enters in both my measurement of labor market outcomes and my calculation of an individual's MPC, in so far as it affects the level of an individual's average earnings in the two previous years.[16]

Throughout the main analysis, I assume that prime-age workers who leave employment in my sample transition to unemployment and make no labor market earnings in

[16] Recall that I exclude the first two years that an individual arrives in the state sample, but I do not exclude observations after subsequent nonemployment spells. Rather, I assume that these are true periods of zero labor market earnings.

those quarters. This assumption would be violated if individuals moved to a job either outside the LEHD coverage (i.e., to the military, federal employment, or self-employment) or to a state that is not in my sample. While this is certainly true for some fraction of workers, this would be a problem for my analysis if workers of different characteristics were differentially likely to move outside the LEHD sample over the business cycle. For example, if younger workers were disproportionately likely to move to states outside of my sample during recessions, then I would overstate the unemployment of young workers, in recessions, and thus, I would erroneously conclude that the earnings of young workers were more sensitive to recessions, when in fact they are not. In addition to overstating the sensitivity of these workers' earnings to GDP, I may exaggerate the difference in MPCs between workers of different ages, as the lagged earnings of younger workers would be biased downward (because they had positive earnings in other states, rather than the zero earnings that I assume). Note that since I mostly focus on the cross-section of employment, and since I remove the initial employment periods, this bias in the measurement of lagged earnings would only appear for workers who moved out of my sample to noncovered employment and then returned to my sample in a future period. However, both of these patterns together would lead me to potentially overstate the differential sensitivity of workers of different MPCs to business cycles.

While I cannot fully address these concerns, the following two pieces of evidence suggest that the differential mobility of workers outside the LEHD sample over the business cycle will not substantially affect my results. First, Table A2 explores the cross-state mobility patterns across the 23 states in my sample. These states are geographically and economically diverse, and therefore, they should reflect the mobility patterns in the United States as a whole. The results in Columns 4 and 5 show that workers with higher MPCs are more likely to move on average, but their mobility is *less* sensitive to GDP, meaning that in recessions, the change in the probability of moving is smaller than for those with higher MPCs. This pattern suggests that differential geographic mobility over the business cycle, if anything, biases the estimate of the heterogeneity in exposure downward.

Second, using the matched monthly basic Current Population Survey, or CPS, I explore the probabilities that workers of different MPCs transition from private sector employment into the military, self-employment, or the federal government, all of which are sectors beyond the scope of the LEHD. The CPS features a rolling panel structure wherein individuals are interviewed for four months, have eight months off, and then are interviewed again for four months. I flag an individual as making a transition to noncovered employment at time $t$ if they are newly self-employed, in the military, or in federal government at time $t$ and were in the private sector in any previous survey. Table A3 shows how the probability of moving to noncovered employment varies over the business cycle. Column 1 shows that on average, high-MPC workers are less likely to move to employment outside the LEHD sample; Column 2 shows that on average, transitions from LEHD employment to noncovered employment are less likely during recessions; while Column

3 shows that there is no differential sensitivity in mobility by a worker's MPC.[17]

Together, Tables A2 and A3 suggest that the assumption that workers who leave the sample make no labor market earnings is reasonable and if anything, the tables imply that my estimates of the heterogeneity in exposure to business cycles by a worker's MPC are underestimates.

## PSID data

The Panel Study of Income Dynamics, or PSID, surveys households annually from 1968 to 1997 and every other year from 1997 to 2015. Each household in the PSID is interviewed once a year, primarily between March and June. In each year of the survey, households are asked about the demographics and labor market status of the household head and spouse. For the length of the survey from 1968 to 2015, households are also asked about how much their household spent on food in the average week. While this is not specific to the week of the interview, it likely refers to the recent period. The income measures, however, refer to the annual income for the previous tax year (i.e., 2007 income is collected in the 2008 survey), and I use the panel structure of the data to get a measure of annual labor market earnings in the year of the interview.[18] This means that the income measure will capture all income in that calendar year, while the consumption and labor market variables will refer to the survey month. If unemployed workers find a high-paying job after the survey month, their annual income will not reflect the income that they were facing (or potentially expecting) at the time in which their consumption was measured. This timing inconsistency likely biases down the estimate of the marginal propensity to consume, as it underestimates the true income loss faced by the household at the time of the survey, when consumption is measured.

Table A4 shows the summary statistics for the PSID sample. Columns 1 and 2 show the snapshot of employed workers in the PSID in the LEHD sample states and nationally, respectively. These two samples are very similar to each other, showing that the LEHD sample is largely nationally representative. A comparison of Column 1 in Table A4 and Column 3 of Table 1.1 in the main text also shows that the PSID sample is similar to the LEHD sample on all demographics. The only exception is that the average income in the PSID is higher than the average income in the LEHD. This is likely because the PSID is restricted to household heads and spouses who are employed full time, while the LEHD includes all workers who had any covered earnings in that quarter. Lastly, Column 3 shows summary statistics for the sample used in the baseline estimation of MPCs. This

[17] Since the CPS does not record the lagged income over the previous two years, I impute the marginal propensity to consume of the individual using only demographic characteristics (race, age, and gender). Within the PSID, the marginal propensities to consume that result from this modified imputation are similar to those that result from the baseline estimation used in the LEHD.

[18] From 1999 to 2015, when the survey becomes every other year, the data include variables on twice-lagged incomes.

sample differs from that in Column 2 along several dimensions. Most significantly, it includes workers who were employed in $t-2$ but makes no restrictions on employment status in time $t$. It also adds the SEO sample, which oversamples low-income households.

A key limitation of the PSID data is that the main measure of expenditure is food. Figure A2 shows that the fraction of total expenditures that is spent on food is changing across the income distribution. The upward slope at the low end of the income distribution reflects the phase out of food stamps, which subsidize the consumption of food for lower-income households. The downward slope across the rest of the income distribution suggests that estimates of the heterogeneity in marginal propensities to consume food by the individual's lagged earnings may not generalize well to estimates for total consumption. Therefore, while I show the robustness of my main findings to the use of food consumption only, I instead use the information on food consumption, as well as the richness of the PSID, to impute more comprehensive measures of consumption.

The main method that I use to impute total consumption in the PSID closely follows the methodology laid out in Blundell, Pistaferri, and Preston (2008). They propose a method to impute expenditures in the PSID using information in the Consumer Expenditure Survey, or CEX.[19] The approach involves estimating a demand system for food as a function of nondurable consumption, demographic variables, and relative prices. Under the assumption that food demands are monotonic, this demand function can be inverted to get an estimate of total consumption in the PSID. In order to deal with measurement error in expenditures, Blundell, Pistaferri, and Preston (2008) instrument the nondurable consumption with the average (by cohort, year, and education) hourly earnings of the husband and wife.[20]

I modify and simplify the Blundell, Pistaferri, and Preston (2008) analysis along several dimensions. First, I estimate the reverse of the equation (i.e., estimate total consumption as a function of food consumption).[21] Second, I estimate this relationship for both durable and nondurable consumption combined.[22] Third, I estimate this by OLS, rather than using the instruments as the original paper does. Fourth, I use an updated sample, extending their sample to 2013. Table A5 shows the summary statistics for the CEX sample used in the imputation. The sample looks very similar to the PSID sample on de-

[19] The CEX includes much more comprehensive measures of consumption. Indeed, it covers around 95 percent of all expenditures, excluding housekeeping, personal care products, and nonprescription drugs. In the interview survey data, consumption is recorded for each month in the three months preceding the interview. This is then aggregated to create measures of total quarterly consumption for each household in an array of spending categories.

[20] Guvenen and Smith (2014) extend this procedure to include the earlier years of the sample. To do so, they modify the specification slightly, replacing time dummies with the food and fuel inflation rate. They also relax the quadratic in age that is used in Blundell, Pistaferri, and Preston (2008).

[21] Results are similar when I estimate the reverse of this equation.

[22] Specifically, I estimate this equation using total consumption, which is the sum of nondurable and durable consumption. However, results are very similar when I estimate it separately for nondurable and durable consumption and then aggregate.

mographics, except that on average, the workers have slightly lower incomes. Using this sample in the CEX, I estimate the following equation

$$\ln C_{it} = Z_{it}\beta + p_t\gamma + g(f_{it}, X_{it}; \theta) + u_{it} \tag{A1}$$

where $C_{it}$ is total household consumption, $Z_{it}$ are household demographics, $p_t$ are relative prices (i.e., Consumer Price Index, or CPI, series), and $X_{i,t}$ are demographic characteristics and time dummies that shift the relationship between food consumption and overall consumption. These time dummies allow the food share to shift over time and can be used because the PSID and CEX have overlapping time frames.[23]

The top panel of Figure A3 shows the time series of actual total consumption in the CEX, which is shown with the red diamonds, and imputed total consumption within the PSID. The blue circles show average total consumption in the PSID (which is the baseline used throughout the analysis), and the green squares show total consumption resulting from separately imputing durable and nondurable consumption and then summing them together. The level of total imputed consumption in the PSID is slightly higher, but all three lines show similar levels and time series patterns. In addition to approximating the time series patterns, the imputed consumption in the PSID shows patterns in the cross-section that are similar to those patterns in the CEX. The middle panel of Figure A3 shows that imputed total consumption in the PSID closely captures the relationship between food consumption and total consumption across the income distribution in the CEX, and the bottom shows that the PSID imputation closely captures the age profile of consumption in the CEX.

An alternative methodology for imputing total consumption in the PSID is to follow Attanasio and Pistaferri (2014) and use the relationship between food consumption and overall consumption in the later years of the sample to impute the total consumption in the previous years of the sample.[24] This imputation approximates total consumption less food (i.e., net consumption) as a function of demographics and food consumption

$$\ln n_{it} = Z_{it}\beta + p_t\gamma + g(f_{it}; \theta) + u_{it} \tag{A2}$$

where $n_{it}$ is the net consumption of the household, $Z$ are various socioeconomic variables, $p$ are prices, and $f$ is food consumption. I estimate this equation restricting the sample to one observation per household and include controls individually for the demographics of the head and spouse. The implicit assumption in this imputation is that the preferences of

[23] I follow Blundell, Pistaferri, and Preston (2008) in my choice of controls. These include dummies for the number of children in the household; three education bins; a quadratic in age; region of residence dummies; an indicator for being white; and education, year, and children dummies. All of these are interacted with food consumption.

[24] The expanded consumption measure within the PSID includes home insurance, rent, electricity, heating, water and miscellaneous utilities, car insurance, car repairs, gasoline, parking, bus fares, taxi fares and other transportation, school tuition and other school expenses, child care, health insurance and out-of-pocket health costs, and food.

individuals are stable over time, and thus, the relationship between overall consumption and food consumption remains stable. This contrasts with the assumption in the CEX-based imputation that uses the same time period from two different samples.[25] Using the $\beta$, $\gamma$, and $\theta$ that result from estimating Equation A2 on the 1999 to 2013 subsample, I recover an estimate of total household consumption in each year using

$$\widehat{c_{it}} = f_{it} + e^{Z_{it}\widehat{\beta}+p_t\widehat{\gamma}+g(f_{it},\widehat{\theta})}$$

Figure A4 shows the performance of this imputation measure. The top figure shows the time series of average expanded consumption. The blue circles show actual expanded consumption in the PSID from 1997 to 2013 while the red squares show the imputed values over the entire sample period. Imputed consumption is slightly below actual consumption, on average, but the time series in the overlapping period is similar. The middle panel of Figure A4 shows that like in the CEX, the implied share of food in expanded consumption is falling with income, and this is true for both actual and imputed consumption. Lastly, the bottom panel shows that the age profile of the PSID-based imputation lines up closely with the age profile in the expanded reported consumption in the PSID in the later years of the survey. While I use the CEX-based imputation method, which covers a broader swath of consumption, as the baseline measure of consumption throughout the analysis, I show robustness of my main relationship between worker MPCs and earnings sensitivity to the aggregate using this alternate consumption measure in Appendix Table A11.

## Additional estimates of marginal propensities to consume

### Robustness of baseline MPC estimates

In the baseline estimation of marginal propensities to consume, I restrict the sample to those individuals who are employed in year $t-2$. I restrict to $t-2$ rather than $t-1$ so that I can include the later years of the PSID sample when the survey is collected every two years.[26] The changes in both income and consumption are also defined over a two-year period. From the entire PSID sample, I exclude observations that do not meet the panel structure necessary to define two-year changes in income and consumption, restrict attention to those between ages 25 and 62 in year $t$, and drop observations with missing race

[25] I closely follow Attanasio and Pistaferri (2014) in parametrizing controls in $Z$. Like them, I include a third-degree polynomial in total food consumption; dummies for age, education, marital status, race, state, and employment status; the hours worked by the household head; homeownership status; family size and the number of children in the household, and consumer price indices to capture relative prices (overall CPI, CPI for food at home, CPI for food away from home, and CPI for rent). I also include household income as a consumption shifter and the spouses' labor market variables as controls in $Z$.

[26] These later years are particularly important both because they overlap with the time period of the LEHD and because they represent a significant fraction of the years for which the CEX and the PSID overlap – and thus the years for which I have the expanded consumption measure.

or education. In addition, in each regression, I exclude observations where the two-year change in log consumption or log income is more than 4.[27] I define an individual's lagged income as the labor market earnings for the individual in years $t-1$ and $t-2$. I group this average into five approximately equally sized bins: $< \$22,000$, $\$22,000 - \$35,000$, $\$35,000 - \$48,000$, $\$48,000 - \$65,000$, and $> \$65,000$. The measure of lagged income is intended to capture differences in permanent earnings capacity across groups. The right panel of Figure A6 shows that the patterns across lagged incomes are not sensitive to the particulars of how lagged earnings are defined; the same patterns for estimated MPCs appear when using additional income lags or fixing earnings at a given age, which may capture a more permanent measure of income.

Table A6 and Figure A5 display supporting statistics for the baseline estimates discussed in detail in the main text. Specifically, Figure A5 show the first-stage and reduced-form estimates associated with Figure 1.3 in the main text. The left panel shows substantial variation in the effect of unemployment on the level of labor income, with the largest falls, unsurprisingly, being among the highest earners. The right panel shows that there is less, although still substantial, variation in the level of the consumption drop across households. Table A6 shows the regression estimates for Equation 1.5 that produce the distribution of MPCs shown in Figure 1.4 in the main text. The left column reports regression coefficients using food consumption only; the middle panel shows estimates using the PSID-based imputation measure, which is described in Appendix A; and the third panel shows the baseline estimates using the CEX-based imputation of total consumption, which is used as the baseline consumption measure throughout this analysis. Unsurprisingly, these multivariate estimates echo the patterns displayed in Figure 1.3, in which black, lower-income, and young workers have higher MPCs. Lastly, Figure A8 shows that these baseline estimates are similar, on average, to other similar estimates of average MPCs in the literature.

The following set of tables and figures explores additional patterns for these baseline MPC estimates. The left panel of Figure A6 shows the estimates of MPCs for alternate individual characteristics that are not reported in the LEHD and thus not included in the baseline set of $x$ variables. Less wealthy households and those who do not own homes have higher MPCs. This finding is in line with the extensive theoretical and empirical literature demonstrating that MPCs vary with household wealth, and it bolsters the supposition that patterns across demographic groups likely capture, in part, differences in wealth holdings across these groups.

One key feature of the baseline MPC estimates are that they capture the MPC of household consumption out of *individual* income, rather than household income. Since individual income is what I observe in the LEHD, this is the correct object for exploring the relationship between individual income movements and consumption patterns. However, it

[27] This restriction on outliers is similar to that in Hendren (2017), who excludes individuals with more than a threefold change in food consumption, and Gruber (1997), who excludes observations with a greater than 1.1 log change in food consumption.

is true in the data that household income is less volatile that individual income, suggesting income smoothing through household formation. Indeed, Figure A6 shows that married workers have slightly lower MPCs. To the extent that this degree of household consumption and income smoothing is similar within demographic groups, the baseline estimates of MPCs will account for this. For example, it may be that one reason that black women have higher MPCs is that they are more likely to be married to spouses with volatile incomes. However, if within demographic groups workers with partners with less volatile incomes sort into riskier jobs, I will be overstating the relationship between demographic group MPCs and income sensitivities, since it will be precisely the unobservable high-MPC workers who are in the less risky jobs. Appendix Figure A7 explores the degree to which the MPC out of individual income differs from the MPC out of household income. Since individual unemployment shocks have smaller effects on total household income than individual income, MPCs out of household income are uniformly larger than MPCs out of individual income. However, while the levels are different, the tight linear relationship between the two estimates reveals that they are highly correlated at the individual level. Indeed, the correlation between these two estimates is 0.89. This strongly suggests that assortative matching in the marriage market is not important in driving the MPC patterns across demographic groups. However, I explore the role of household formation further in Appendix A.

Another reason that MPCs may vary across demographic groups are that the persistence of the income shock differs across groups, meaning that forward-looking agents respond today to expected changes in future income. Figure A10 explores the persistence of the unemployment shock on incomes by demographic group.[28] The left panel of Figure A10 shows, however, that the expected duration of the shock does not differ much by race, and the right panel shows that it does not differ much by age. Each point in the plot shows the effect of unemployment on the change in income in each period relative to the unemployment event, normalized by the baseline effect. By one year after the unemployment shock, worker income growth is essentially back to its preunemployment level for all demographic groups.

Lastly, Figure A11 explores the sensitivity of the MPC estimates to using alternate consumption measures. The top panel shows that while the level differs substantially, the individual-level MPC estimates are highly correlated across consumption measures. The blue squares show scatterplots relating the MPCs estimated using food consumption directly reported in the PSID with the baseline MPCs estimated using CEX-based total consumption. The red circles similarly show the relationship of the baseline MPCs with MPCs estimated using an alternate PSID-based imputation of expended consumption. (see Section A for an explanation of this alternate consumption measure) The tight fit of the lines shows that all measures are highly correlated at the individual level. The bottom

[28] Note that the duration of the shock will not matter for workers who are fully borrowing-constrained. These workers will have an MPC of 1 out of lost income, and this will be true for both permanent and transitory shocks. See Appendix A for an extended discussion.

two panels show the age and income profiles of the MPCs using the alternate consumption measures, respectively. The food-based MPCs demonstrate a slightly different age pattern, with the MPC dropping steadily over the life cycle, while the broader consumption measures feature a U-share in age. Otherwise, the three outcome measures show similar patterns across demographics.

**Stability of MPC estimates**

While the above method for estimating MPCs closely follows existing methods in the literature, my subsequent imputation of these MPCs by demographic in the LEHD necessitates several important additional assumptions that warrant further discussion. First, in imposing that MPCs only vary by worker demographic, I assume that individual MPCs are invariant to the sign and magnitude of the income shock. As I discussed above, in a standard model in which agents can borrow and save, an individual's MPC depends on the persistence of the shock, not on the magnitude or sign. However, with liquidity constraints, the marginal propensity to consume may depend on the size of the shock as well as the sign (Kaplan and Violante (2014)). To explore the importance of this assumption, the left panel of Figure A9 shows the overall estimates of the marginal propensity to consume that result from re-estimating Equation 1.5 using different identifying income shocks. First, the left-most point shows the OLS version of Equation 1.5. The coefficient is close to 0, suggesting a substantial downward bias and the need for an instrument to identify the causal relationship between consumption and income movements. However, while the use of an instrument matters critically, across the x-axis, estimates of the MPC are relatively stable to the type of income shock used as the instrument. For comparison, the second point shows the baseline MPC estimated using the unemployment shock. The next four estimates show the MPC estimated using either the change in state GDP or the national unemployment rate of the worker's industry. For an individual worker, these aggregate changes are plausibly exogenous to their own earnings and affect their earnings both positively and negatively and on both the intensive and extensive margins. While noisier, the average MPC estimates are similar. This is true whether I include all workers (as in points 3 and 4) or restrict to those workers who remain employed (as in points 5 and 6). Those who remain employed across years experience a smaller income change yet a similar MPC.[29] Lastly, the farthest-right point shows the average MPC estimated using an indicator for whether the worker becomes employed between $t-2$ and $t$.[30] The average MPC is slightly higher with the positive income "shock," but this is an artifact of the different estimation samples – the hires estimation restricts to the nonemployed, who, on

[29] I find that those who remain employed across surveys experience a drop in total hours worked when the unemployment rate is high, suggesting moves to part-time employment.

[30] This specification includes only those who were not employed in $t-2$; thus the control group is the set of individuals who remain nonemployed between $t-2$ and $t$. Patterns are robust to including only those who are unemployed, rather than nonemployed, in $t-2$.

average, have higher MPCs than the employed. When averaged on the same sample, the estimates are similar. Lastly, Appendix Table A7 shows that not only are the averages similar for these different shocks, but the alternate MPC estimates are also highly correlated at the individual level. Together, these estimates suggest that, empirically, the MPCs to different business-cycle labor income shocks do not differ substantially depending on the sign or magnitude of the particular shock. The stability of worker MPCs is also something I address within the quantitative exercise in Section 1.7 in the main text.

A second key stability assumption embedded in the imputation of MPCs in the LEHD is that for an individual income shock of a given magnitude, the consumption response is constant over the business cycle. I explore this in my setting by adding an interaction of changes in income with the state unemployment rate, thereby allowing the MPC to vary over the business cycle. The right panel of Figure A9 plots the bivariate versions of Equation 1.5, modified to allow the MPC to vary with the state unemployment rate. The blue circles plot the MPC at the average unemployment rate in the state, and the red squares show the implied MPC at 4 percentage points above the average unemployment rate.[31] Generally, the MPC is somewhat lower in recessions, but the differences are economically and statistically insignificant.

Third, I impose that at the individual level, the marginal propensity to consume is a function only of the characteristics that I include in X (i.e., age, earnings history, gender, and race). While this is obviously an approximation, this assumption would be a problem for my analysis if within each demographic bin there is sorting across jobs such that it was precisely the higher-MPC workers within the group who were at cyclically insensitive jobs. If this were the case, I would be inaccurately capturing the heterogeneity in exposure of workers to business cycles by their MPC. While my data do not allow me to fully address this, I explore the sensitivity of my MPC estimates to including job-level characteristics. If sorting across jobs of different characteristics were important in explaining MPC heterogeneity within demographic group, then these terms should have additional explanatory power. Table A9 shows the correlation between the baseline MPC estimates and those estimates including various job-level characteristics, and Table A8 shows the underlying coefficients. In the first specification, I include the individual's tenure with their current employer. This variable is intended to capture some amount of private information on the riskiness of the individual's job, as workers with longer job tenures are less likely to lose their jobs (Farber (1999)). The coefficient in Column 2 of Table A8 is negative but small and statistically insignificant. Column 3 adds the lagged variance of an individual's earnings. The coefficient on this is also negative, suggesting that conditional on the average past earnings, a higher variance of earnings is associated with a lower MPC. Columns 4 and 5 include the variance of an individual's lagged earnings, capturing the fact that individuals with a higher earnings variance may differ in their MPCs. This

[31] Results are similar when interacting the change in earnings with an indicator for national recession years.

variable is calculated using the matched monthly basic CPS from 1976 to 2013 and is the sample average of the change in log earnings between interview 4 and interview 8, which are a year apart. This variable is intended to capture the expected variability of earnings of the job. For both industry and occupation, these variances enter positively, suggesting that workers with higher MPCs are in riskier industries. Lastly, I include dummies for the census region of residence, allowing MPCs to vary geographically. Across columns, none of the coefficients on the added variables are statistically significant, nor, as shown in Table A9, does the resulting MPC heterogeneity change much when these variables are added.

### Alternate MPC estimation methods

In this section, I present alternate methods for estimating MPCs by individual demographics. I show that while these estimates are noisier, they are qualitatively consistent with the PSID-based MPC estimates presented above and used throughout the analysis.

The first alternative method for estimating MPCs is to follow a similar strategy to that outlined above but to estimate it entirely within the Consumer Expenditure Survey. While the CEX is not a full longitudinal panel like the PSID, it does feature a rolling panel structure in which individuals are interviewed five times, three months apart.[32] A clear advantage of the CEX over the PSID is that it includes more comprehensive measures of consumption that do not rely on imputation. Indeed, it covers around 95 percent of all expenditures, excluding housekeeping, personal care products, and nonprescription drugs. In the interview survey data, consumption is recorded for each month in the three months preceding the interview. This is then aggregated to create measures of total quarterly consumption for each household in an array of spending categories.

While the CEX has extensive consumption data, the disadvantage of the CEX is that it includes sparse information on individuals' labor market experiences, making it difficult to identify income shocks. However, while incomplete, the CEX does include enough labor market information to roughly identify periods of nonemployment (Gruber (1998)). In particular, in survey 2 and survey 5 (taken nine months apart), the CEX asks how many weeks in the last year the individual has worked for pay, including vacations and sick leave, as well as how much income the individual and household earned in the preceding 12 months.[33] I define an individual as employed in survey 2 (what I will refer to as the base period) if they report having worked 52 weeks in the past 12 months. This

[32] The CEX also has a second component in which households are interviewed in two consecutive weeks. This higher-frequency data collection is intended to capture detailed information on the consumption of smaller high-frequency items such as food or household supplies. As the labor market information is not available at such high frequencies, this exercise will focus on the summary measures of these items included in the interview survey.

[33] Note that because the labor market and earnings information is collected nine months apart but refers to a 12-month lookback period, there is a mechanical overlap in these two variables.

ensures that the individual is employed both at the time of interview and in the previous three months, when baseline consumption is measured. I define an individual as unemployed in interview 5 if they report having worked fewer than 38 weeks in the previous 12 months.[34] Importantly, while this definition does capture some labor market disruption, it does not distinguish between voluntary nonemployment, retirement, and unemployment. I restrict the sample to those individuals between ages 25 and 55 to limit the frequency of schooling and retirement decisions, but the concern still remains in this sample.[35] Additionally, this definition does not guarantee that the individual is unemployed either at the the time of the interview or in the three previous months over which consumption is measured.

Restricting to the individuals employed in survey 2, I run the following regression, which is similar to Equation 1.5 in the main text:

$$\Delta C_{c,i,t} = \beta_0 + \beta_1 \Delta Y_{i,t} + \beta_2 X_{it} + \gamma_{st} + \epsilon_{i,t} \tag{A3}$$

where $C_{c,i,t}$ is consumption in category $c$ of individual $i$ in year $t$, $Y_{i,t}$ is annual labor income, and $X_{i,t}$ are individual-level controls (the change in the number of individuals in the household). I instrument $\Delta Y_{i,t}$ with an indicator for whether or not the individual is nonemployed in the survey 5.[36] This should isolate unexpected changes in income. As with the PSID estimates, the coefficient of interest is $\beta_1$, which measures the average MPC for consumption category $c$. I then explore the heterogeneity by interacting $\Delta Y_{i,t}$ with individual characteristics.

Figure A12 shows the estimates of MPC heterogeneity using Equation A3 for various measures of consumption. First, the left figure shows the estimated MPCs using food consumption, and the right figure shows the MPC estimates using total consumption.[37] A

[34] I explore the robustness of the results to this cutoff. Another obvious possibility would be to use the cutoff of 12 weeks, which essentially ensures that the individual is unemployed at the time of the fifth interview. However, this reduces the sample size substantially, and results are similar to those achieved using less restrictive cutoffs; therefore, I keep 39 weeks as the main specification. An additional possibility is to use an indicator for whether the household received any unemployment insurance compensation in the previous year. This variable is available in the CEX from 1980 to 2013. This does not isolate individuals who are unemployed at the time of the interview, but it does aid in isolating individuals who are unemployed rather than voluntarily nonemployed.

[35] I can compare the frequency of unemployment in the CEX using this definition and the PSID. The CEX is consistently higher, but the patterns are broadly similar over time.

[36] I take several additional steps in cleaning the CEX data. First, I drop observations with food or incomes that change by more than 200 percent in the nine months of the sample. I exclude from the analysis any households that have a change in the number of children between survey 2 and survey 5 (approximately 10 percent of the sample). I do this to abstract from nonemployment due to childbirth. Lastly, I further restrict the sample to include workers from age 25 to 55. This is because this is the age range over which the incidence of nonemployment is relatively stable. Above age 55, the probability of nonemployment rises, suggesting that there is the possibility of early retirement in this age range.

[37] Total consumption is defined to include nondurable consumption, education, health care, insurance, and other housing expenditures.

couple of patterns emerge. First, the overall magnitude of the MPCs for food consumption are very similar in the PSID and the CEX. Some of the descriptive patterns are also similar – for example, black men have the highest MPCs. These CEX-based estimates are noisier than those in the PSID, and the patterns from the PSID in age and income are less apparent here. The estimate MPC for total consumption, however, is lower in the CEX than in the PSID. While the estimates are noisier, the qualitative patterns across demographic groups are similar – young, less educated, and poorer households have higher MPCs.

Both of the previous methods for computing an individual's MPC, utilizing either the PSID or the CEX, used shocks to labor market earnings. This is relevant for this paper's analysis because it identifies the business cycle variation in income that I will ultimately explore. However, an alternative method is to use an unexpected positive income shock in the form of a tax rebate. To explore the relationship between these estimates, I replicate and extend Parker et al. (2013), who use the CEX to explore the change in household spending caused by the receipt of economic stimulus payments in 2008. In 2008 and 2009, the CEX included questions about the receipt of stimulus payments. The estimating equation in Parker et al. (2013) is

$$\Delta C_{i,t,c} = \beta_1 X_{i,t} + \beta_2 ESP_{i,t+1} + \beta_m \gamma_m + u_{i,t+1} \tag{A4}$$

where $C_{i,t,c}$ is expenditure in a category $c$ (i.e., food, nondurables, or total); $\gamma_m$ are month dummies; $X_{i,t}$ are control variables; and $ESP_{i,t+1}$ is the tax rebate payment. The timing of the ESP payment was determined by the last two digits of the recipients' Social Security numbers, meaning that it was effectively randomized across households. In order to explore heterogeneity in the resulting MPCs, I interact ESP with various household characteristics. In the baseline analysis, consumption is measured in the three months following the receipt of the payment. Estimates are also produced using 2SLS, where the ESP is instrumented with an indicator for the receipt of an ESP. This is because while the timing of the receipt of an ESP is exogenous, the amount of the ESP, conditional on receipt, is not, and it varies with income and family size.

Figure A13 shows the resulting estimates of $\beta_2$, the marginal propensities to spend out of tax rebates, for various individual characteristics and various measures of spending. I specified these characteristics to match the PSID analysis rather than Parker et al. (2013), though the patterns are qualitatively similar to those reported in Table 6 their paper.[38] These estimates are very noisy, and thus, it's hard to make any concrete conclusions. However, the patterns in the point estimates go in similar directions to those in the unemployment-based analysis; younger, less educated, and poorer households have higher marginal propensities to spend their tax rebates. The patterns for gender

[38] One important difference, though, between this analysis and the previous unemployment analysis is that in this case, the income shock is at the household level, rather than the individual level. Therefore, I can only explore heterogeneity by household characteristics, not individual characteristics; thus, I define education, race, and age by the reference person (who is the property owner).

are flipped, though this is possibly a data artifact stemming from my use of the gender of the reference person in the household rather than the individual.

Together, the analysis throughout Section A supports the robustness of the general patterns captured by the baseline estimates of marginal propensities to consume used throughout the text, as similar patterns emerge when using an array of different estimation procedures.

## Additional discussion of heterogeneity in worker exposure to business cycles

### Robustness of relationship between worker MPCs and earnings sensitivity to GDP

Figure 1.1 clearly demonstrates that there is a positive relationship between the average earnings cyclicality of a demographic group and the average marginal propensity to consume of that group. The additional results presented in this section support the robustness of this pattern. First, Figure A14 shows this pattern separately for the intensive margin of earnings (i.e., earnings conditional on remaining employed between $t-1$ and $t$) and the extensive margin of employment. The figure clearly shows that higher-MPC workers are more likely to become unemployed during recessions and earn less conditional on remaining employed. Indeed, the same demographic groups that are exposed on the intensive margin are also exposed on the extensive margin of earnings.

In order to explore the individual patterns underlying this positive relationship, Figure A15 shows Figure 1.1 broken down further by demographic groups. The subplots of Figure A15 highlight all the bubbles that have a certain characteristic to reveal how workers of this specific characteristic experience recessions. A key pattern emerges showing that while some of the variation is across groups, the positive relationship between MPCs and earnings elasticities is also apparent within each group. Indeed, for *any* subgroup explored in Figure A15, the positive relationship holds both within (i.e., within women and men) and across the groups (i.e., across women and men).

Tables A11 and A10 probe the robustness of the positive relationship between a worker's MPC and the exposure of their earnings to recessions to several data decisions. Table A11 shows that the estimated relationship is robust to various methods of imputing MPCs. Since the magnitude of the MPCs changes with the imputation method, so does the magnitude of the coefficient, but the proportional relationship is fairly stable. Indeed, across all specifications, a 1 standard deviation increase in the MPC of the individual is associated with an increase in the elasticity of earnings with respect to GDP of between 0.33 and 0.39, or between 50 percent and 80 percent.

Columns 2 and 3 of Table A10 show that the overall sign of the relationship is insensitive to the functional form imposed on the dependent variable. When estimating the relationship between earnings sensitivities and MPCs at the individual level, taking the log of earnings will restrict to those who remain employed. Therefore, in the baseline analysis, I estimate an overall earnings elasticity at the individual level by replacing the

change in log earnings with $\Delta E_{i,t} = \frac{E_{i,t}-E_{i,t-1}}{.5*E_{i,t}+.5*E_{i,t-1}}$, which bounds the earnings loss of the nonemployed at negative 2. Column 2 shows that the patterns are similar when using a log transformation (i.e. $\Delta E_{i,t} = \log(E_{i,t}+100) - \log(E_{i,t-1}+100)$), and Column 3 shows similar patterns when using a level transformation, in which the level change in earnings is normalized by the average level of earnings in that worker's MPC bin. Both of these are alternate functional forms that combine the intensive and extensive margins of earnings, with the baseline transformation producing the most conservative estimate.

The following columns of Table A10 show that the estimated relationship is robust to various other modifications. Baseline estimates consider movements in fourth-quarter earnings, but Column 5 shows that patterns are similar when considering annual incomes. Column 6 restricts to a balanced panel from 1995 to 2011 and finds that in this subset, the estimate is similar but slightly larger. Column 7 replaces aggregate GDP with state-level GDP and shows that heterogeneity patterns are similar, suggesting that these patterns hold not only across states but also within states.

Lastly, Column 8 of Table A10 shows adjusted standard errors for the baseline estimate. Individually clustered standard errors do not take into account the additional noise imposed by the imputation of a worker's MPC. Indeed, the worker MPC estimates rely on two imputations. I first impute total consumption in the PSID using the CEX, and then I impute the MPC in the LEHD using the MPC estimates from the PSID. In order to account for the additional noise injected at each of these steps, I implement multiple imputation techniques, as in Rubin (1987). Specifically, I take 500 draws in which I randomly sample with replacement both the CEX and the PSID. This produces 500 estimates of main-text Equation 1.5, which estimates the MPC for each demographic group $x$. In the LEHD, I then estimate main-text Equation 1.6 for each imputation, which results in 500 estimates of the degree to which workers of different MPCs are exposed to recessions. I combine these various estimates following the formulas derived in Rubin (1987):

$$\bar{\alpha}_2 = \sum_{i=1}^{500} \frac{\widehat{\alpha}_{2,i}}{500}$$

$$var(\alpha) = \sum_{i=1}^{500} \frac{var(\widehat{\alpha_{2,i}})}{500} + \sum_{i=1}^{500} \frac{(\widehat{\alpha}_{2,i} - \bar{\alpha}_2)^2}{499} + \sum_{i=1}^{500} \frac{(\widehat{\alpha}_{2,i} - \bar{\alpha}_2)^2}{499 * 500}$$

The point estimate is the average across imputation draws, while the variance of the estimate is the combination of the average within-draw variance and the between-draw variance. Column 8 of Table A10 shows that while the standard error is larger, as should be expected, the coefficient is still tightly estimated. This methodology is very computationally intensive due to the large sample size, and therefore, given the small qualitative difference, I proceed with the simple clustered standard errors and leave the estimation of adjusted standard errors to future drafts.

Table A12 explores the robustness of the results using the expanded information avail-

able in the American Community Survey, or ACS, subsample. This expanded information helps me test for the possibility of additional sorting patterns within demographic group that could potentially bias the estimated relationship between worker MPCs and earnings elasticities. For a direct comparison, Column 1 of Table A12 first shows the baseline estimate in the LEHD for only 2001 to 2011, the years of the ACS subsample. The overall relationship is smaller in the later part of the sample, and Column 2 shows that the estimate is nearly identical in the much smaller ACS subsample in those same years. More interestingly, Column 3 of Table A12 shows the robustness to an estimate using an individual-level MPC that varies along additional demographic characteristics – namely, the number of children in the household and an individual's marital status. One possibility is that workers of the same demographic group but with different household structures sort into jobs with different sensitivity to aggregate movements. The results in Column 3 show that this is not the case. While the magnitude of the coefficient changes slightly, there is still a strong positive relationship between an individual's MPC and the sensitivity of their earnings to GDP.

The MPC used above is the household consumption response to changes in individual-level earnings, but it's possible that high-MPC individuals are actually in low-MPC and low-sensitivity households if their partners have lower MPCs and earnings that are less sensitive to the aggregate shock. The individual MPCs that I estimate and use as the baseline already include the average effect of this by demographic group, but the concern is that workers within these demographic groups sort across jobs of different riskiness based on household MPC, which is an omitted variable in the baseline analysis. Moving explicitly to the household level is the cleanest way to address these concerns. In Columns 4 and 5 of Table A12, the unit of observation is the household and the dependent variable is the change in income of the household, rather than of the individual. In Column 4, the household MPC is calculated as the earnings-weighted average individual MPCs of the members of the household, while in Column 5, I specifically construct the MPC at the household level.[39] In both columns, I find that the earnings of higher-MPC households are more exposed to movements in aggregate GDP. Column 4 provides the most direct comparison to the baseline results and shows that the coefficient is invariant to aggregating to the household level.

### Relationship to Guvenen et al. (2017)

Using individual earnings data from the U.Ss Social Security Administration, Guvenen et al. (2017) document that the earnings of both the lowest and the very highest earners

[39] I construct a household MPC using household-level data in the PSID and using only household, rather than individual, characteristics. Specifically, I allow the household MPC to vary by the lagged earnings of the household, grouped into five earnings bins, a quadratic in the age of the oldest member of the household, an indicator for whether one or both members of the household are black, and the number of people in the household.

are more sensitive to aggregate income fluctuations. In Figure A16, I largely replicate this finding using my sample within the LEHD. The left panel shows the elasticity of worker earnings to aggregate GDP by the income decile of the worker. As in Figure 1 of Guvenen et al. (2017), there is a U-shaped relationship; the sensitivity of worker earnings is decreasing through much of the earnings distribution, but it spikes again at the very top of the income distribution.[40]

How does this relate to the relationship between worker MPCs and earnings sensitivity documented in Figure 1.1? This U-shaped relationship in earnings does not directly imply any relationship between worker MPCs and earnings sensitivity – the bottom of the income distribution has high estimated MPCs but a small share of overall earnings; the top of the income distribution has both low MPCs and the majority of the earnings in the economy; and in my estimation, lagged income only explains 40 percent of the overall variation in MPCs. For a direct comparison, the right panel of Figure A16 shows earnings elasticities by the MPC decile of the worker. There are two important observations. First, the overall pattern is upward-sloping, meaning that workers in the top decile have the highest income sensitivity. Second, there is a nonlinear pattern in the bottom deciles, likely reflecting the increased earnings sensitivity of the very higher earners.

**Additional heterogeneity patterns**

While the overall estimates of the differential earnings elasticities by MPC are the key inputs to the calculation of the matching multiplier, or MM, in this section, I decompose this overall relationship to explore possible reasons that high-MPC workers are more exposed to aggregate fluctuations. One key factor that could underlie the positive relationship between worker MPCs and the sensitivity of earnings to the aggregate could be the sorting of workers across industries, firms, and jobs according to their characteristics and skills. Firms differ substantially in their exposure to recessions; industries such as manufacturing or construction are more cyclically sensitive than other industries, such as health care or education, and even within an industry, firms differ in their exposure to business cycles, with young and small firms being particularly cyclically sensitive (Fort et al. (2013), Crouzet and Mehrotra (2017).) In addition, a large and growing literature has documented substantial and increasing sorting of workers across firms along several dimensions. For example, young firms are more likely to hire young workers (Ouimet and Zarutskie (2014)), women have a higher willingness to trade higher pay for job security (Wiswall and Zafar (2018)), and high-wage workers congregate together in high-paying

[40] The magnitude of the relationship documented in the left panel of Figure A16 is somewhat smaller than that documented in Guvenen et al. (2017) , who show that workers in the 80th percentile have an earnings elasticity that is a full 2 points lower than the bottom decile. Since the LEHD does not include the uncensored earnings of the very high earners, the movement among the highest earners in Figure A16 is also smaller.

firms (Barth et al. (2016), Song et al. (2015)).[41] Together, the sorting of workers across firms could mean that high-MPC workers are more exposed to business cycles if the skills of high-MPC workers are utilized by more cyclically sensitive firms.

I explore the role that this sorting of workers plays in explaining the aggregate relationship displayed in Figure 1.1 by adding various fixed effects to Equation 1.6. Adding a full set of four-digit industry-by-year fixed effects will sweep out any earnings sensitivities that are common within an industry; thus, the remaining relationship is identified using purely within-industry differences in the exposure of workers to shocks. The top row of Column 1 in Table A13 reproduces the overall heterogeneity estimate from main-text Table 1.2. The top row of the second column in Table A13 shows that while the estimated slope drops when industry fixed effects are added, the sorting of workers of different MPCs across industries explains only about 12 percent of the differential exposure of high-MPC workers to recessions. Column 3 in Table A13 shows the estimated slope after adding a full set of firm-by-year fixed effects, which isolates the within-firm heterogeneity in worker exposure to recessions.[42] These fixed effects have very little further effect on the estimates, suggesting that the sorting across firms within industries explains very little of the total relationship.

Together, the previous two results suggest that the large majority of heterogeneity in business cycle exposure by MPC occurs within the firm rather than between firms. Within the firm, occupation is a key dimension along which workers of different MPCs may sort and along which earnings sensitivities may differ. While the LEHD does not include worker occupation, the ACS does. Therefore, for the subset of individuals whom I observe in both the ACS and the LEHD, I estimate a specification where I include a full set of industry-by-year-by-occupation fixed effects.[43] Column 4 of Table A13 reproduces the results using only year fixed effects on this modified subsample of people matched to the ACS and shows that the estimates of the overall heterogeneity are somewhat smaller in this subsample.[44] More importantly, Column 5 shows that this coefficient drops by about 40 percent when restricting the variation to within- industry occupation pairs.

If the sorting of workers across occupations, industries, or firms only explains up to 40 percent of the heterogeneity in exposure, what economic forces are driving this systematic

[41] Several of these sorting patterns are present in this sample as well. Specifically, I find that workers at firms fewer than five years old are younger, have lower earnings histories, and have higher average MPCs. Similarly, I find that small firms employ, on average, a workforce that has higher MPCs and is more male, white, and low-income.

[42] I define the firm as the combination of the firm and state, which is defined by the State Employer Identification Number. I define both the firm and the industry for the job held in $t-1$.

[43] Because of the diminished sample size, I do not fully restrict to within the firm and include firm-by-year-by-occupation fixed effects. However, since cross-industry sorting is almost as important as cross-firm sorting, this specification should mostly capture the within-firm heterogeneity.

[44] Estimates on the full LEHD sample restricting to 2001 to 2011, which is the time period covered by the ACS subsample, are very similar to those in Column 4 of Table A13, confirming that the difference in the overall estimate is driven not by the different subsampling but by the different estimation period.

positive relationship between a worker's MPC and their earnings cyclicality? The lower rows of Table A13 explore this by decomposing the MPC into two components – the component that comes from differences in the worker demographics (age, race, and gender) and the component that comes from differences in earnings histories. The final four rows of Table A13 go even further and explicitly break down the patterns by the four worker characteristics that enter the MPC calculation – age, gender, race, and earnings histories. Several important patterns appear. First, comparing across rows, all components of the worker MPC contribute to the overall positive relationship between MPCs and earnings sensitivities to GDP. The earnings of women, black workers, lower-income workers, and younger workers, all of whom have higher MPCs on average, are all more exposed to GDP. However, comparing the importance of the various fixed effects across columns reveals that while demographic differences become less important within industries and occupations, differences in lagged incomes actually become *more* important within the firm and further within occupations.

Indeed, the importance of lagged income in explaining within-firm patterns suggests that workers could actually be high MPC *because* they are exposed to aggregate shocks. Workers who experience unemployment during recessions are likely to have lower lifetime earnings and fewer assets, thereby raising their MPC. This pattern, wherein heterogeneity in earnings volatility across workers within the firm is largely explained by variation in their earnings history, is consistent empirically with a large literature showing that an unemployment spell increases the risk of future unemployment spells (Stevens (1997)), as well as consistent theoretically with a job ladder model as in Jarosch (2014), wherein workers search for both more productive and more secure jobs, and thus, as they climb the job ladder they sort into higher-paying and more secure jobs.

To summarize, while what will matter for the overall consumption response will be the total relationship between earnings sensitivities and GDP and MPCs, the patterns in Table A13 highlight that the substantial positive relationship is largely a within-firm phenomena that is only partially explained by the sorting of workers across occupations within the firm. Instead, the within-firm patterns are explained largely by a worker's past earnings.

### MPCs and hiring over the business cycle

Since the baseline analysis is restricted to the set of incumbent workers, the differential exposure of workers to recessions is the combination of differential separation probabilities or wage growth across worker MPCs. However, it does not include differences in exposure to recessions through the hiring margin, which is an equally if not more important margin of adjustment for firms during recessions. For example, while the construction and manufacturing sectors experienced spikes in their separation rates during the Great Recession, most other industries adjusted their employment levels with large swings in their hiring rates. If high-MPC workers are less likely to be hired during reces-

sions, there will be an additional contribution to the matching multiplier coming from the unemployed.

In order to explore the heterogeneous exposure to recessions among the unemployed, I combine information in the LEHD and the ACS. The LEHD is a data set of employment, not of the labor force as a whole. The American Community Survey is a sample of the entire population, but it does not include the within-person time series variation that is necessary to identify hiring patterns over the business cycles. Thus, I explore the differential sensitivity of hiring by worker MPC using the overlap between the ACS and the LEHD. Specifically, I identify an individual as unemployed in year $t$ in sample state $s$ using their reported unemployment status and residence in the ACS. I then match that set of unemployed workers to the LEHD in the following year $t+1$ to determine whether they are employed a year later and if so, what their earnings are.[45] I then explore how this re-employment probability varies over the business cycle by estimating Equation 1.6 on this set of unemployed workers. I explore the extensive margin of hires using an indicator for whether the individual transitions to employment between $t$ and $t+1$ as the outcome variable, and I explore the intensive margin of earnings using the earnings conditional on re-employment. For worker MPCs, I deviate from the baseline analysis and use the MPC estimated using an instrument for being hired, rather than an instrument for becoming unemployed. This estimate may better capture the variation that is most relevant for the unemployed. However, since this MPC estimate and the baseline MPC is very close, the results are similar when using the baseline MPC instead.

Column 1 of Table A14 shows that re-employment probabilities over the business cycle are similar for high- and low-MPC workers. Similarly, Column 3 shows that the earnings of the re-employed are also not more sensitive to GDP. This finding is possibly driven by changes in the composition of the unemployed over the business cycle. In recessions, the unemployed are higher quality workers on average, and thus, the re-employment wages conditional on worker quality could fall in recessions while the overall re-employment wage goes up. If the magnitude of the selection on worker quality over the business cycle is smaller for higher-MPC workers, you would observe this pattern wherein the re-employment wages of high-MPC workers are, on average, less sensitive to the business cycle.

An alternate method for defining the unemployed is to roughly define the set of unemployed within the LEHD. While this has the disadvantage of relying on a cruder definition of the unemployed, it has the advantage of including a much larger sample. Within the LEHD, I define the unemployed as the set of individuals who were employed in year $t$ in a sample state in some previous year $t-k$ but who do not have any earnings in the LEHD at time $t$. This definition of the unemployed makes three potential errors. First,

[45] Worker MPCs are a function of an individual's earning history, but the ACS does not include the earnings history for an individual. Therefore, I calculate a worker's lagged earnings using the LEHD. For each unemployed individual, I match them to the LEHD in years $t-1$ and $t-2$ and assume that if they do not appear anywhere in the sample in those years, they did not earn any labor market earnings over that period.

it excludes the unemployed who are actively searching for employment in those states but who have never been employed in those states. Aside from new market entrants, this group is likely unimportant for aggregate earnings dynamics, as it is very unattached from the labor market. Second, it includes as unemployed those who actually found jobs in states or sectors outside of the LEHD sample (e.g., transitioned to federal employment or moved to a state outside the 23-state sample). Third, it includes as unemployed those who in fact left the labor market and are not actively searching. While I can further restrict the sample to prime-age workers between ages 25 and 50 to mitigate this concern, this is still a possible source of misallocation.

Despite the crudeness of the definition, Columns 2 and 4 of Table A14 show that the estimated heterogeneity in earnings elasticities among the unemployed are similar to those resulting from the ACS-based identification. These two methods are likely similar because, as was discussed in Section A, individuals of different MPCs do not show different propensities to move either across sectors or across states over the business cycle. Thus, while the pool of unemployed differs across the two pools, the earnings variation from which the estimates in the LEHD are identified are similar.

Lastly, Table A15 shows that these unemployment estimates have a very minor effect on the overall estimates of the matching multiplier. This is a direct consequence of the small amounts of heterogeneity in Table A14 and the very small share of the overall dollars in the economy earned by the unemployed.

## Details of commuting zone analysis

### Additional data definitions

**Local control variables:** I closely follow Kaplan, Mitman, and Violante (2016) in defining household wealth measures in each local labor market. Specifically, I define housing wealth as the total number of housing units in a county, which are published by the U.S. Census Bureau, multiplied by the Zillow Home Value Index for All Homes.[46] Data on household debt come from the Federal Reserve Bank of New York Consumer Credit Panel (CCP), which provides the dollar values of mortgage, auto, and revolving credit debt annually in each county from 1999 to 2011. I define household debt as the total value of both mortgage and nonmortgage debt. I construct data on financial assets by allocating total financial assets in a quarter from the Flow of Funds Balance Sheet of Households and Nonprofit Organizations to counties using the fraction of total financial assets in that county from the quarterly IRS Statistics of Income. Lastly, I aggregate these county-level

[46] This housing index is publicly available monthly for each U.S. county beginning in 1996. Housing units are available annually at the county level back until 2000. Prior to 2000, these counts are only released at the state-by-year level. I interpolate county housing units prior to 2000 by assuming that the fraction of houses in each county in the state is constant and by assigning total state housing units in each year to counties based on the 2000 distribution.

measures to the commuting zone level, restrict attention to fourth-quarter estimates, and divide by population estimates to obtain per-capita values.[47]

The fraction of the commuting zone that is employed comes from the ACS. All other control variables – namely, demographic controls for the area and the average size and age of establishments by commuting zone – are calculated within the LEHD in each year.

**Bartik shock:** I construct a Bartik-style shock at the commuting zone level using

$$Shock_{c,t} = \sum_i \frac{L_{i,c,t_0}}{L_{c,t_0}} \Delta \log E_{i,t,-c}$$

where $t_0 = 1999$ and $\Delta \log E_{i,t,-c}$ is the change log of total earnings in industry $i$ within the states that are in the LEHD subsample in year $t-1$ but excluding earnings in commuting zone $c$. I exclude the own commuting zone, since my LEHD sample is not national and thus any given commuting zone may represent a non-negligible fraction of the industry's total earnings. Additionally, because the LEHD is not balanced across states, the aggregated time series for any given industry are inconsistent due to the entry of states. Therefore, to be consistent over time, I define the change in earnings in an industry using only the incumbent states in each year. Unsurprisingly, this shock is very highly predictive of both changes in overall earnings in a commuting zone and movements in the GDP of the commuting zone's state.

In exploring the role of the local labor market multiplier in affecting local cyclicality using this Bartik shock, I *both* re-estimate $MM_c$ using this shock and then re-estimate Equation 1.8 replacing aggregate GDP with this shock.

### Robustness of results at the commuting zone level

The following section includes several results supporting the robustness of the findings presented in Table 1.4 and Table 1.5 in the main text. Before revisiting the full estimates of Equation 1.8, Figure A18 first shows year-by-year estimates of the relationship between $\widehat{MM}_c$ and the change in local employment. Specifically, the coefficients plotted in Figure A18 are the coefficients on $\widehat{MM}_c$ from the following regression, estimated separately in each year $t$

$$\Delta \log L_{c,t} = \kappa_1 \widehat{MM}_c + \kappa_2 \widehat{B}_c + \sum_x \kappa_x x_{c,t-1} + \epsilon_{c,t}$$

[47] CCP data are only released for counties with an estimated population of at least 10,000 consumers with credit reports in the fourth quarter of 2010. This restriction excludes 20 percent of counties. Since these are predominately small counties, I ignore these missing values in the aggregation from counties to commuting zones.

where $x_{c,t-1}$ includes controls for fraction of the commuting zone in each two-digit industry; the average age, race, gender, and lagged earnings of the area; and the fraction of the population that is employed. If a higher matching multiplier makes areas more sensitive to shocks, then we should see that during the years of the great recession, areas with a higher matching multiplier had a bigger fall in their employment (i.e., $\kappa_1$ is negative). The left panel of Figure A18 shows this is indeed the case. The value of $\kappa_1$ in 2008 is negative, meaning that between the fourth quarter of 2007 and the fourth quarter of 2008, areas that had a larger matching multiplier saw a greater fall in their employment than areas with a smaller matching multiplier. Indeed, in that time period, employment in areas with a matching multiplier 1 standard deviation above the mean fell by about 1 percentage point more than areas with a matching multiplier at the mean. However, the other points in the figure show that the opposite relationship holds in boom years. Indeed, in the boom years from 2002 to 2005, employed areas with a matching multiplier 1 standard deviation above the mean grew by about 0.7 percentage points more per year. In other words, exactly as I found in Table 1.4, the relationship between local earnings growth and the matching multiplier is cyclical – during the years of the Great Recession, areas with a higher matching multiplier experienced larger employment losses, but during the boom of the early 2000s, those areas were also experiencing faster employment growth. Furthermore, the differential sensitivities to GDP reported in Table 1.4 are not simply driven by different experiences in only recessions or only in booms; rather, they experience the entire business cycle differently.

To show this more clearly, the right panel of Figure A18 plots the annual coefficients not against years as in the left panel but against the change in GDP in those years. The point in the bottom left is 2008, when there was both a negative movement in aggregate GDP and a negative relationship between the matching multiplier and employment. The boom years of the mid 2000s are in the top right, where GDP was booming and so were areas with a higher matching multiplier. As was expected given the time series pattern, there is a clear positive relationship between the effect of the matching multiplier on employment and the movement of aggregate GDP. The slope of the relationship in the right panel is in essence a more nonparametric estimate of $\phi_1$, the coefficient on the interaction between $\widehat{MM}_c$ and GDP in Equation 1.8; when there is a positive aggregate shock, a higher matching multiplier is associated with more employment growth, but when the shock is negative, a higher matching multiplier is associated with lower employment growth.

Table A16 explores the robustness of the overall cross-commuting zone patterns presented in main-text Table 1.4. Column 1 simply replicates the baseline estimates for comparison. Column 2 shows the estimates of the relationship replacing aggregate GDP with a Bartik shock. The coefficient on $\widehat{MM}_c$ interacted with GDP is still positive and statistically and economically significant – a commuting zone with the average matching multiplier has an employment response to the shock that is 0.2 percentage points higher

than an area with a matching multiplier of 0, representing a 42 percent increase above the average response. Column 3 presents the estimates using an alternate imputation of MPC, which is based on the positive income shock from hires rather than the negative income shock of unemployment. While the magnitudes adjust, the sign of the coefficients remains the same. Column 4 shows the estimates replacing $\widehat{MM}_c$ with simply the difference between the actual and benchmark MPC, which is the numerator of $\widehat{MM}_c$. This alternate functional form reveals similar patterns and magnitudes. Lastly, Column 5 includes the baseline commuting zone demographic controls interacted with a full set of time dummies, controlling nonparametrically for differences across commuting zones along these dimensions. The results are very similar, suggesting that allowing for a different trend and cycle closely captures the heterogeneity along these dimensions.

Figure A17 zooms in on a case study of the Great Recession. The left panel shows the relationship between the employment change in the Great Recession and the matching multiplier, estimated over the full sample. The right panel plots the same y-axis but instead shows the relationship with the matching multiplier estimated on the full sample instrumented with the matching multiplier estimated pre-2007. Both figures demonstrate a strong negative relationship, with the slope on the right being slightly steeper than the slope on the left.

Lastly, Table A17 explores the sensitivity of the relative patterns for employment in tradable and nontradable industries presented in main-text Table 1.5. Column 1 reproduces the baseline estimates from Column 3 of Table 1.5. Column 2 shows that the estimates are similar when moving to more disaggregated four-digit NAICS codes. Column 3 shows that the coefficient on the triple interaction between the matching multiplier and tradables is robust to including financial controls. Finally, Column 4 shows estimates using only the numerator of $\widehat{MM}_c$, which measures the difference between the actual and benchmark MPCs. While the magnitudes of the coefficients naturally rescale with the change in the function, the relative patterns across tradable and nontradable industries remain similar.

## The Matching Multiplier and the labor share

While the matching multiplier is derived in a setting in which all output is earned by labor, in order to provide *empirical* estimates of the matching multiplier, I need to take into account the fact that in reality, not all output goes to worker wages. Since the focus of this paper is on quantifying a particular mechanism within the labor market, I do not explore potentially important heterogeneity in MPCs out of nonlabor income. Rather, I make modest adjustments to the simple framework to rescale the contribution of this particular mechanism. To see how I arrive at the expression in Equation 1.7, consider the case where output is given by $Y = E + K$, where $E$ are labor market earnings and $K$ are earnings from nonlabor income (e.g., profits, return on capital, etc). In this case, the aggregate MPC in the economy is given by:

$$MPC = \frac{dC}{dY} = \frac{dC}{dE}\frac{dE}{dY} + \frac{dC}{dK}\frac{dK}{dY}$$

or, defining the average earnings elasticity as $\gamma = \frac{dE}{dY}\frac{Y}{E}$, this becomes

$$MPC = \frac{dC}{dY} = \frac{E}{Y}\frac{dC}{dE}\gamma + \frac{dC}{dK}\frac{dK}{dY}$$

This simple total derivative highlights the importance of two terms that were not in the simple framework – the consumption response from changes in nonlabor income $\left(\frac{dC}{dK}\frac{dK}{dY} = MPC_{Y-E}\frac{Y-E\gamma}{Y}\right)$ and the labor share $\left(\frac{E}{Y}\right)$. The importance of the labor share is intuitive – a mechanism affecting labor market income matters more for the total economy when labor earns a higher share of total income.

To see how this yields the expression in Equation 1.7, consider the benchmark and actual MPCs in this extended scenario. The benchmark MPC, where all workers have the same earnings elasticity, is given by

$$MPC^b = \frac{E}{Y}\overline{MPC}\gamma + MPC_{Y-E}\frac{Y - E\gamma}{Y}$$

The actual MPC is instead given by

$$MPC^a = \frac{E}{Y}\sum_i MPC_i\gamma_i\frac{E_i}{E} + MPC_{Y-E}\frac{Y - E\gamma}{Y}$$

Using the definition of $MM$ in Equation 1.4 and these expanded expressions for $MPC^a$ and $MPC^b$, I arrive at Equation 1.7.

## Details for model in Section 1.7

### Proof of Proposition 1

**Proposition 1:** *Under the assumption that wages are sticky for $k+1$ periods, for any shock to parameters $(\theta, \tau, \theta_G)$, the total change in output from an initial flexible price allocation is given to first order by*

$$d\mathbf{Y} = \left(I - \mathbf{C_Y} J_k - (\mathbf{C_r} + \mathbf{G_r}) J_{T-k} (\mathbf{L_r})^{-1}\right)^{-1} \partial \mathbf{Y} \tag{A5}$$

*where subscripts denote partial derivatives (i.e., $\mathbf{C}_r$ is the partial derivative of consumption with respect to $r$) and $J_k$ and $J_{T-k}$ are diagonal matrices with 1s in the first $k$ or the last $T-k$ entries, respectively.*

*Proof.* Begin by totally differentiating the good market clearing condition (Equation 1.18) in each period $t$:

$$dY^t = \sum_{j=1}^{k} C_{t,y_j} dy_j + \sum_{j=1}^{T} (C_{r_j} + G_{r_j}) dr_j + \sum_{j=1}^{T} C_{t,\tau} d\tau_t + C_{t,\theta} d\theta + \sum_{j=1}^{T} G_{t,\tau_t} d\tau_t + G_{t,\theta_G} d\theta_G \tag{A6}$$

where $C_{t,x}$ is an $I$-length vector across individuals where each entry is the partial derivative of the individual consumption function $c_i(\{y_{i,t}\}_{t\leq k}, \{\tau_{i,t}\}_{t\in T}, \{r_t\}_{t\in T}, \beta_i, b)$ with respect to the variable $x$. Recall that in the rationed equilibrium, income is exogenous in all rationed periods, and thus enters the consumption function. Note that the first sum is only across periods 1 through $k$, the periods in which there is labor market rationing. Beyond that, the workers are back on their labor supply curves and their income is endogenously given by their decisions and prices. By the definition of the income process imposed by the rationing function in Equation 1.21,

$$dy_{i,t} = n_{i,t} - l_{i,t-1} = \gamma_i \frac{l_{i,t-1}}{L_t} dY_t$$

Denote $N_y$ as the $I$ vector where the $i$ entry is $dy_{i,t} = \gamma_i \frac{l_{i,t-1}}{Y_t}$. Plugging this in, we get:

$$dY^t = \sum_{j=1}^{k} \left(C'_{t,y_j} N_y dY_j\right) + \sum_{j=1}^{T} (C_{r_j} + G_{r_j}) dr_j + \sum_{j=1}^{T} C_{t,\tau} d\tau_t + C_{t,\theta} d\theta + \sum_{j=1}^{T} G_{t,\tau_t} d\tau_t + G_{t,\theta_G} d\theta_G \tag{A7}$$

Note that $C'_{t,y_j} N_y = \frac{dc_t}{dy_j} = C_{t,j}$, which is the aggregate response in time $t$ to a change in income at time $j$. Equation A13 holds for all periods $t$, and stacking equations, this becomes

$$dY = C_Y J_k dY + (C_r + G_r)dr + C_\tau d\tau + C_\theta d\theta + G_\tau d\tau + G_{\theta_G} d\theta_G \tag{A8}$$

where $C_Y$ is a matrix where the $i, j$ entry is the aggregate consumption response at time $i$ to an income shock in time $j$ and $J_k$ is a diagonal matrix with 1s only for the first $k$ periods. For the first $k$ periods, the interest rate is pinned down by the monetary policy rule (rather than by labor market clearing). Given the rule specified in Equation 1.22,

$$dr_t = 0 \quad \forall t \leq k$$

After $k$ periods, the interest rate goes back to the flexible price scenario and the change in the interest rate is pinned down by the labor market clearing condition. The total derivative of the labor market clearing condition is given by:

$$dY_t = \sum_{j=0}^{T} L_{t,r_j} dr_j + \sum_{j=1}^{T} L_{t,\tau} d\tau_t + L_{t,\theta} d\theta \quad \forall t > k$$

Stacking across periods and solving for $dr$, we get the expression for the change in the interest rate at time $t > k$:

$$dr = (L_r)^{-1} (dY - L_\tau d\tau - L_\theta d\theta)$$

Stacking these equations over time and defining

$$\partial Y = C_\tau d\tau + C_\theta d\theta + G_\tau d\tau + G_{\theta_G} d\theta_G + (C_r + G_r) J_{T-k} (L_r)^{-1} (L_\tau d\tau - L_\theta d\theta)$$

we can rewrite Equation A8 as

$$dY = C_Y J_k dY + (C_r + G_r) J_{T-k} (L_r)^{-1} dY + \partial Y \tag{A9}$$

where $dY$ and $\partial Y$ are $T \times 1$ matrices and $C_Y$ is a $T \times T$. $J_k$ is just a diagonal matrix with ones along the diagonal for the first $k$ entries, and $J_{T-k}$ is a diagonal matrix with 1s for the last $T - k$ entries. $C_r$, $G_r$ are $r_y$ are $T \times T$ matrices.

□

**More generalized monetary policy rule**

In the model in the main text, I assumed that in rationed periods, the central bank targeted a fixed interest rate. This simplifying assumption is important for providing a simplified mapping to the data, and it may provide a good approximation for understanding the multiplier in periods when the economy is at the zero lower bound. In this section, I derive the multiplier in the more general case where the interest rate moves as a function

of aggregate output:

$$r_t = r(Y_t) \tag{A10}$$

In this case, I get the following slightly modified expression for the multiplier. The additional term in Proposition 2 includes the change in the interest rate during the first $k$ periods that is induced by the central bank's rule. If the central bank lowers the interest rate when output is low, this term will generally dampen the multiplier, as consumption today falls in response to the interest rate.

**Proposition 2.** *Under the assumption that wages are sticky for $k+1$ periods, for any shock to parameters $(\theta, \tau, \theta_G)$, the total change in output from an initial flexible price allocation is given to first order by*

$$d\mathbf{Y} = \left(I - \mathbf{C_Y} J_k - (\mathbf{C_r} + \mathbf{G_r})\mathbf{r_y} - (\mathbf{C_r} + \mathbf{G_r}) J_{T-k} (\mathbf{L_r})^{-1}\right)^{-1} \partial \mathbf{Y} \tag{A11}$$

*where subscripts denote partial derivatives (i.e., $\mathbf{C}_r$ is the partial derivative of consumption with respect to r) and $J_k$ and $J_{T-k}$ are diagonal matrices with 1s in the first k or the last $T-k$ entries, respectively.*

*Proof.* This proof very closely follows the proof for Proposition 1. I begin by totally differentiating the good market clearing condition (Equation 1.18) in each period $t$:

$$dY^t = \sum_{j=1}^{k} C_{t,y_j} dy_j + \sum_{j=1}^{T} (C_{r_j} + G_{r_j}) dr_j + \sum_{j=1}^{T} C_{t,\tau} d\tau_t + C_{t,\theta} d\theta + \sum_{j=1}^{T} G_{t,\tau_t} d\tau_t + G_{t,\theta_G} d\theta_G \tag{A12}$$

where $C_{t,x}$ is an $I$-length vector across individuals where each entry is the partial derivative of the individual consumption function $c_i(\{y_{i,t}\}_{t \le k}, \{\tau_{i,t}\}_{t \in T}, \{r_t\}_{t \in T}, \beta_i, b)$ with respect to the variable $x$. Recall that in the rationed equilibrium, income is exogenous in all rationed periods, and thus enters the consumption function. Note that the first sum is only across periods 1 through $k$, the periods in which there is labor market rationing. Beyond that, the workers are back on their labor supply curves and their income is endogenously given by prices. By the definition of the income process imposed by the rationing function in Equation 1.21,

$$dy_{i,t} = n_{i,t} - l_{i,t-1} = \gamma_i \frac{l_{i,t-1}}{L_t} dY_t$$

Denote $N_y$ as the $I$ vector where the $i$ entry is $dy_{i,t} = \gamma_i \frac{l_{i,t-1}}{L_t} dY_t$. Plugging this in, we get:

$$dY^t = \sum_{j=1}^{k} \left( C'_{t,y_j} N_y dY_j \right) + \sum_{j=1}^{T} (C_{r_j} + G_{r_j}) dr_j + \sum_{j=1}^{T} C_{t,\tau} d\tau_t + C_{t,\theta} d\theta + \sum_{j=1}^{T} G_{t,\tau_t} d\tau_t + G_{t,\theta_G} d\theta_G \tag{A13}$$

Equation A13 holds for all periods $t$, and stacking equations, this becomes

$$dY = C_Y J_k dY + (C_r + G_r) dr + C_\tau d\tau + C_\theta d\theta + G_\tau d\tau + G_{\theta_G} d\theta_G \tag{A14}$$

where $J_k$ is a diagonal matrix with 1s only for the first $k$ periods. For the first $k$ periods, the interest rate is pinned down by the monetary policy rule (rather than by labor market clearing). The key difference from Proposition 1 is that the monetary policy rule now implies:

$$dr_t = r_{y_t} dY_t \quad \forall t \leq k$$

After $k$ periods, the interest rate goes back to the flexible price scenario and the change in the interest rate is pinned down by the labor market clearing condition. The total derivative of the labor market clearing condition is given by:

$$dY_t = \sum_{j=0}^{T} L_{t,r_j} dr_j + \sum_{j=1}^{T} L_{t,\tau} d\tau_t + L_{t,\theta} d\theta \quad \forall t > k$$

Stacking across periods and solving for $dr$, we get:

$$dr = (L_r)^{-1} (dY - L_\tau d\tau - L_\theta d\theta)$$

Stacking these equations over time and defining

$$\partial Y = C_\tau d\tau + C_\theta d\theta + G_\tau d\tau + G_{\theta_G} d\theta_G + (C_r + G_r) J_{T-k} (L_r)^{-1} (L_\tau d\tau - L_\theta d\theta)$$

we can rewrite Equation A14 as

$$dY = C_Y J_k dY + (C_r + G_r) r_y J_k dY + (C_r + G_r) J_{T-k} (L_r)^{-1} dY + \partial Y \tag{A15}$$

where $dY$ and $\partial Y$ are $T \times 1$ matrices, and $C_Y$ is a $T \times T$. $J_k$ is just a diagonal matrix with 1s along the diagonal for the first $k$ entries and $J_{T-k}$ is a diagonal matrix with ones for the last $T - k$ entries.

□

**Comparing flexible and sticky price multipliers**

Since this paper empirically explores the role that the distribution of income shocks across worker MPCs plays in affecting the economy's response to shocks, I focus on the case where labor is rationed and thus workers are off their labor supply curve. However, it is illustrative to consider the multiplier when prices are flexible and the interest rate adjusts such that workers remain on their labor supply curves.

**Proposition 3.** *Under the assumption that all prices are flexible, for any shock to parameters $(\theta, \tau, \theta_G)$, given initial conditions, the total change in output is given to first order by*

$$dY = \left(I - (C_r + G_r)(L_r)^{-1}\right)^{-1} \partial Y \tag{A16}$$

*where subscripts denote partial derivatives (i.e., $\mathbf{C}_y$ is the partial derivative of consumption with respect to $y$).*

*Proof.* As in the proof for Proposition 1, begin by taking the total derivative of the goods market clearing condition, noting that the consumption function for the individual is only a function of exogenous variables and prices as given by Equation 1.16:

$$dY_t = \sum_{j=0}^{t}(C_{r_j} + G_{r_j})dr_j + \sum_{j=1}^{T}(C_{r_j} + G_{r_j})dr_j + \sum_{j=1}^{T} C_{t,\tau}d\tau_t + C_{t,\theta}d\theta + \sum_{j=1}^{T} G_{t,\tau_t}d\tau_t + G_{t,\theta_G}d\theta_G \tag{A17}$$

The labor market clearing condition implies $dY_t = \sum_{j=0}^{T} L_{t,r_j}dr_j + \sum_{j=1}^{T} L_{t,\tau}d\tau_t + L_{t,\theta}d\theta$. Stacking these across years, you get:

$$dY = L_r dr + L_\tau d\tau - L_\theta d\theta \tag{A18}$$

where $L_r$ is also a $T \times T$ matrix. Assuming it is invertible, we can write $dr = (L_r)^{-1}(dY - L_\tau d\tau - L_\theta d\theta)$.

Define the partial equilibrium response as the movement in aggregate output *before* there are any movements of endogenous variables (e.g., $r$ and $w$),

$$\partial Y = C_\tau d\tau + C_\theta d\theta + G_\tau d\tau + G_{\theta_G}d\theta_G + (C_r + G_r)(L_r)^{-1}(L_\tau d\tau - L_\theta d\theta) \tag{A19}$$

Plugging Equation A19 and Equation A18 into Equation A17, we get:

$$dY = (C_r + G_r)(L_r)^{-1}dY + \partial Y$$

□

What is this multiplier in the flexible price case? It only captures the response of consumption and labor to changes in the interest rate, which adjusts endogenously to clear

the labor and output markets. When prices are flexible, workers are always on their labor supply curves, meaning that when their taxes go up, they will want to work more and consume less. However, since consumption demand fell, firms will want to hire fewer workers, and the wage will fall (or since the wage is normalized, the interest rate will fall to encourage people to work less today (and consume more today)). Changes in income are determined by changes in labor supply; thus, the change in the interest rate (and the change in the labor supply that it induces) are sufficient for understanding the response to shocks.

Indeed, in this standard RBC model with flexible prices, this multiplier is always weakly less than 1 (i.e., there is no amplification of the initial shock). Consider the case with two time periods. In this case, Equation A16 becomes

$$\begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} = \begin{bmatrix} C_{1,r} + G_{1,r} & 0 \\ 0 & C_{2,r} + G_{2,r} \end{bmatrix} \begin{bmatrix} L_{1,r}^{-1} & 0 \\ 0 & L_{2,r}^{-1} \end{bmatrix} \begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} + \begin{bmatrix} \partial Y_1 \\ \partial Y_2 \end{bmatrix}$$

or

$$\begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} = \begin{bmatrix} (C_{1,r} + G_{1,r})L_{1,r}^{-1} & 0 \\ 0 & (C_{2,r} + G_{2,r})L_{2,r}^{-1} \end{bmatrix} \begin{bmatrix} dY_1 \\ dY_2 \end{bmatrix} + \begin{bmatrix} \partial Y_1 \\ \partial Y_2 \end{bmatrix}$$

This matrix is block-diagonal , so we can write the first period responses independently of the second period as

$$dY_1 = (1 - (C_{1,r} + G_{1,r})L_{1,r}^{-1})^{-1}\partial Y_1$$

For each entry $i$, $C_r$ is weakly negative. The Euler equation implies that a higher interest rate will cause consumption today to fall, unless the individual is constrained, in which case they will not behave according to the Euler equation and will not respond to changes in $r$. $L_{1,r}$ is weakly positive – when the interest rate goes up, today is a better time to work (since the future payoff from working today is higher). Thus, $(C_{1,r} + G_{1,r})(L_{1,r})^{-1}$ is weakly negative, meaning that the multiplier is weakly less than 1.

## Quantitative details

### Estimation of the income process

I loosely follow Heathcote, Perri, and Violante (2010) in estimating the income process. I begin with the assumption that earnings are generated by the following model:

$$\log(y_{ia}) = \mu + z_{ia} + \epsilon_{ia}$$

$$z_{ia} = \rho z_{i,a-1} + u_{ia}$$

$$u_{ia} \sim (0, \sigma_u)$$

$$z_{i,0} \sim (0, \sigma_{z_0})$$

$$\epsilon_{ia} \sim (0, \sigma_\epsilon)$$

where $i$ is the individual, $a$ is the age, and $\mu$ is the average level shift common to all individuals and ages. Assume that the shocks are all i.i.d and thus uncorrelated with each other. The set of parameters that characterizes this income process is: $\theta = \rho, \sigma_u, \sigma_{z_0}, \sigma_\epsilon$. I identify these parameters using within-person variances and covariances in income over time. Specifically,

$$var(y_{i0}) = \sigma_{z0} + \sigma_\epsilon \tag{A20}$$

$$var(y_{ia}) = var(z_{ia}) + \sigma_\epsilon \tag{A21}$$

$$var(z_{ia}) = \rho^2 var(z_{i,a-1}) + \sigma_u \tag{A22}$$

$$Cov(y_{ia}, y_{i,a-j}) = cov(z_{ia}, z_{i,a-j}) \quad \forall j > 0 \tag{A23}$$

$$Cov(z_{ia}, z_{i,a-j}) = \rho^j var(z_{i,a-j}) \quad \forall j > 0 \tag{A24}$$

$$\frac{cov(y_{ia}, y_{i,a-2})}{cov(y_{i,a-1}, y_{i,a-2})} = \frac{\rho^2 var(z_{i,a-2})}{\rho var(z_{i,a-2})} = \rho \tag{A25}$$

The identification of $\sigma_\epsilon$ (the variance of the transitory component) comes from the difference between the variance and the covariance:

$$var(y_{i,a-1}) - \frac{1}{\rho} cov(y_{ia}, y_{i,a-1}) = var(z_{i,a-1}) + \sigma_\epsilon - var(z_{i,a-1}) = \sigma_\epsilon \tag{A26}$$

Once you know $\rho$ and $\sigma_\epsilon$, the identification of $var(z_{i0})$ comes immediately from Equation A20. Lastly, the identification of $\sigma_u$ (the variance of the persistent component) comes from

$$var(y_{i,a-1}) - cov(y_{ia}, y_{i,a-2}) - \sigma_\epsilon = \sigma_u \tag{A27}$$

I implement this estimation using a minimum distance estimator constructed using the following steps:

1. Construct a matrix of the covariances of earning across ages in the data.
2. Vectorize the matrix to include the set of unique covariances (i.e., take the upper triangular portion of C and turn it into a vector).
3. Define $f(\theta)$ as the corresponding vector based on the model covariance matrix.

4. The estimator solves the minimization problem:

$$\min_{\theta}[m - f(\theta)]'\Omega[m - f(\theta)]$$

where $\Omega$ is the identity matrix (Altonji and Segal (1996)).

In selecting the sample on which to estimate the income process, I make several important data decisions. In the baseline analysis, I restrict attention to those who report being unemployed or employed at the time of the PSID survey, are between the ages of 25 and 62, and for whom I can impute an MPC (i.e., observations with at least two lags of earnings). I include both the nationally representative sample and SEO subsample of the PSID. I match the analysis in the LEHD and use individual labor rather than household earnings or total post-tax-and-transfer earnings. In general, the earnings process it is less volatile when you include transfer income, but is just about as volatile when you add both transfer and capital income. In order to include periods of 0 earnings, I use the following transformation:

$$\log y_{it} = \log(y_{it} + \overline{UI})$$

where $\overline{UI}$ is the average annual unemployment insurance payment reported in the PSID. In the PSID, $\overline{UI} \approx 1758$. While adding the UI payments seems to be a reasonable benchmark, in Appendix Table A18, I explore the sensitivity of the discount rate calibration to using other transformations (i.e., adding 1,000 or 100). In general, the lower the additional factor, the more volatile is the income process, and thus the lower the discount rate that is needed to match the group average MPC.[48]

### Additional quantitative results

Table A19 summarizes the strength of the matching multiplier in the model. All estimates summarize the mechanism in period 1. The first two rows show the model estimates of the empirical statistic. In both cases, I consider a negative 1 percent shock, and therefore, upon impact, consumption drops in both cases. However, consumption drops by more in the case where the shock has the empirical distribution than in the benchmark scenario when the shock is distributed evenly.

While main-text Figure 1.5 and Table A19 focus on the matching multiplier in period 1, Figure A19 shows the dynamics of the consumption response. The left panel shows

[48] The literature generally takes the log of earnings and thus restricts to household heads with non-0 annual labor market earnings (Heathcote, Perri, and Violante (2010)). I do not follow this because I think it's important to include extended periods of unemployment as well as secondary earnings. This also more closely matches the empirical specification I use in the LEHD, where I include both secondary earners and volatility in the extensive margin of earnings. However, in Table A18, I show that when I use an income process estimated excluding observations with 0 earnings and restricting to household heads, I find similar conclusions about the importance of endogenous MPCs.

the level of the consumption drop in both the actual and the benchmark scenarios, while the right panel shows just the difference between these consumption responses. As in Table A19, in the period of the shock, consumption falls by more in the actual scenario than in the benchmark scenario. However, in the following year, this pattern reverses, and consumption in the benchmark scenario is actually higher than in the empirical case. This overshooting pattern is intuitive; since, on aggregate, in the benchmark scenario, less was consumed and more was saved in the first period, there is more to spend in future periods. This pattern persists until both scenarios eventually converge back to the steady-state level of consumption. Figure A20 shows that the dynamic patterns are similar even when the shock is more persistent – most of the consumption boost coming from the differential incidence of shocks comes in the first period.

Figure A21 explores the importance of the rationing period not using the length of the shock but rather using the sufficient statistics approach. When labor is rationed for only one period, agents only respond directly to changes in their income in the first period, and thus, the only non-0 column of the intertemporal MPC matrix ($C_Y$ in Proposition 1) is the first column. However, if wages are sticky for three periods, labor will be rationed for two periods, and the intertemporal MPC matrix will have non-0 entries in the first two columns. More generally, as the length of the rationing period grows, the columns of the intertemporal MPC matrix fill in. In Figure A21, I plot the matching multiplier as a percentage of the overall multiplier for different lengths of the rationing period. I assume that the interest rate is fixed and use Proposition 1 to calculate $MM$ as I increase the length of the rationing period. As shown in Auclert, Rognlie, and Straub (2018), I find that the level of the multiplier increases as intertemporal MPCs are taken into account, and Figure A21 shows that the relative importance of the matching multiplier in period 1 also increases. This is consistent with the finding in Table A19, which shows that as the length of the shock grows, so too does the importance of heterogeneity.

### MPCs and income processes

The quantitative analysis above demonstrates that in a heterogeneous agent model, the MPC depends on the income process. In this section, I build on the discussion in Jappelli and Pistaferri (2014) to demonstrate analytically how the MPC depends on both the income process facing the worker and the type of income shock experienced by the worker. In order to clarify the intuition with analytical expressions, I consider a more restrictive setting than that discussed in Section 1.7.

#### MPCs and income processes with quadratic utility

Consider an agent who maximizes the expected utility of consumption over an infinite time period subject to their intertemporal budget constraint:

$$\max_{c_t} \sum_{t=0}^{\infty} \beta_t E_0 \left[\log(c_t)\right]$$

s.t.

$$\sum_{t=0}^{\infty} E_0 \left[\frac{c_t}{(1+r)^t}\right] = \sum_{t=0}^{\infty} E_0 \left[\frac{y_t}{(1+r)^t}\right]$$

This setup embeds several important assumptions. First, it assumes that preferences are additively separable over time and across states (VNM expected utility). Second, it assumes log utility ($u(c_t) = log(c_t)$), which shuts down any precautionary savings motives. Third, it assumes that interest rates are constant ($r_t = r$). Finally, it assumes that the agent can borrow and save freely and is not subject to any constraints. If you further assume that $\beta = \frac{1}{1+r}$, you arrive at the stochastic consumption Euler equation:

$$c_1 = c_0 + \epsilon_{i1}$$

where $\epsilon_{i1}$ includes the effect on consumption of all new information about the uncertainty faced by the consumer. Here, the only uncertainty for future consumption is income. Plugging the Euler equation back into the intertemporal budget constraint, we get:

$$c_0 = \frac{r}{1+r} \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} E_0 y_t$$

$$\Delta c_1 = \frac{r}{1+r} \sum_{t=0}^{\infty} \frac{E_1 y_t - E_0 y_t}{(1+r)^t} \tag{A28}$$

Several things become apparent from Equation A28. First, there is a key distinction between anticipated and unanticipated income shocks. If income shocks are anticipated, then there is no revision of expectations and therefore no consumption response. However, if income shocks are unanticipated, consumption today will respond and the magnitude of the consumption response depends critically on the income process (i.e. how big the income revision is). For example, consider the following income processes:

1. **ARMA(1,1):** $y_{it} = \rho y_{it-1} + \nu_{it} + \theta \nu_{it-1}$

   In this case, Equation A28 becomes:

   $$\Delta c_{it} = \frac{r}{1+r} \frac{1+r+\theta}{1+r-p} v_t$$

   The more persistent the income process (i.e., the higher $\rho$ is), the larger the MPC to a transitory shock $v_t$. Similarly, the more persistent the transitory shock (i.e., the

higher $\theta$ is), the higher the MPC to the transitory shock.

2. **Permanent and transitory shocks:** $y_{it} = P_t + \nu_{it} = P_{it-1} + u_{it} + \nu_{it}$

$$\Delta c_{it} = \frac{r}{1+r} v_{it} + \sum_{t=0}^{\infty} \frac{1}{(1+r)^t} u_{it} = \frac{r}{1+r} v_{it} + u_{it}$$

The MPC to transitory shocks is $\frac{r}{1+r}$ and the MPC to permanent shocks is 1.

3. **Income process with unevenly distributed aggregate shocks:** $y_{it} = P_{ti} + \nu_{it} + \phi_i Y_t$ and $P_{it} = P_{it-1} + \nu_{it} + \rho_i Y_t$. (i.e., both the permanent and transitory shocks have an i.i.d and an aggregate component).

$$\Delta c_{it} = \frac{r}{1+r}(v_{it} + \phi_i Y_t) + u_{it} + \rho_i Y_t \tag{A29}$$

4. **General income process:** K components to the income process ($\pi_{it}$), all with different persistence:

$$\Delta c_{it} = \sum_{k=1}^{K} \phi^k \pi_{it}^k$$

In this simple model, the $\phi^k$ are all functions of the difference persistence of the income shock $\pi_{it}$. In a richer model, these reduced-form $\pi_{it}$ are a function of the underlying economic structure (credit constraints, risk preferences, access to insurance, etc.).

While the above framework certainly features too little heterogeneity and too restrictive preferences to be an accurate depiction of consumption decisions in the data, the discussion highlights that the estimate of the MPCs in theory is potentially sensitive to the type income shock. In the baseline analysis, I estimate the MPC to the unemployment shock, which is some unknown combination of the many possible income shocks $\pi_{it}^k$ in the generic income process above. The MPCs that I estimate are therefore some linear combination of the associated sensitivities ($\phi_k$). It's possible that a different income shock would contain different degrees of anticipation or persistence, and therefore, the estimated MPCs may differ. However, Figure A9 explores the sensitivity of MPCs to the type of income shock and finds that empirically, the estimates are very similar.

## Appendix Figures and Tables

Figure A1: LEHD Sample Selection

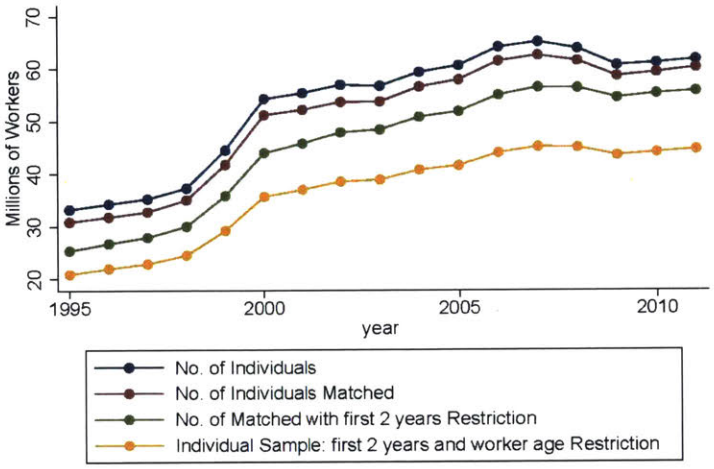
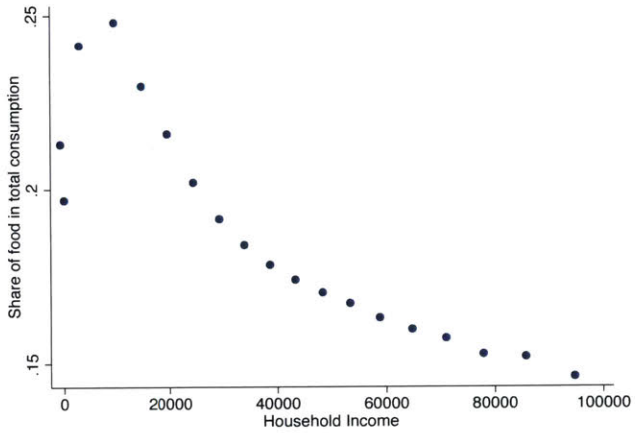
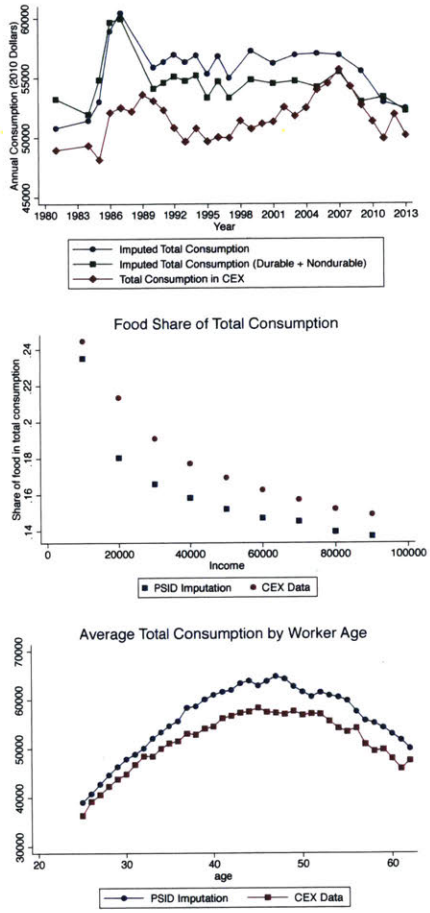


Figure A2: The Fraction of Food in Total Spending



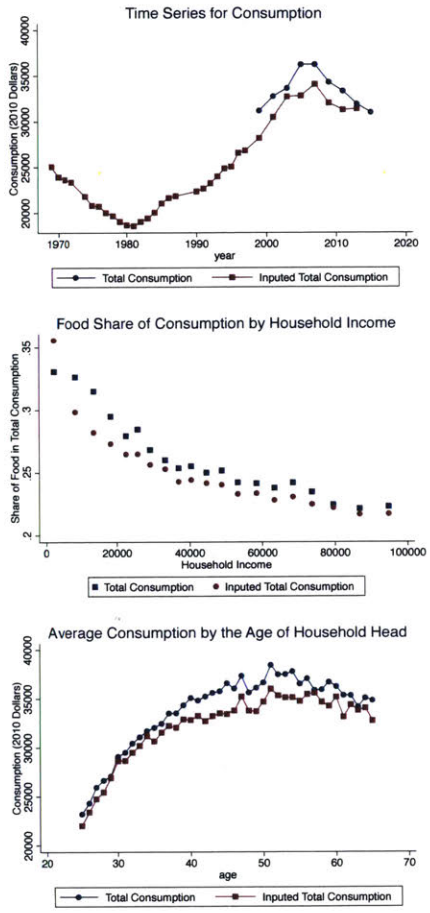

Notes: Data are from the Consumer Expenditure Survey and are pooled across 1984 to 2014 for all households with a head between the ages of 25 and 62. Household income is adjusted to 2010 dollars.

Figure A3: Imputed Total Consumption Using CEX-Based Imputation



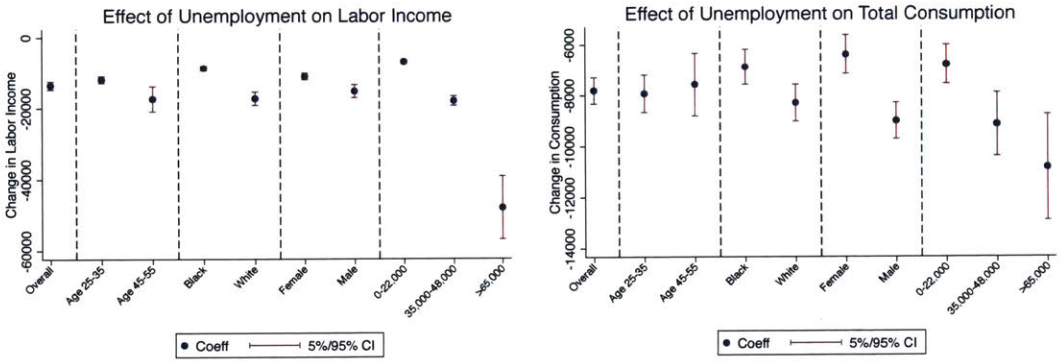

Notes: Average total consumption in the PSID is imputed using Equation A1. Averages are calculated using sample weights in the CEX and based on the nationally representative subsample in the PSID. Income refers to household income, adjusted to 2010 dollars.

Figure A4: Imputed and Actual Consumption Using Attanasio and Pistaferri (2014) Imputation



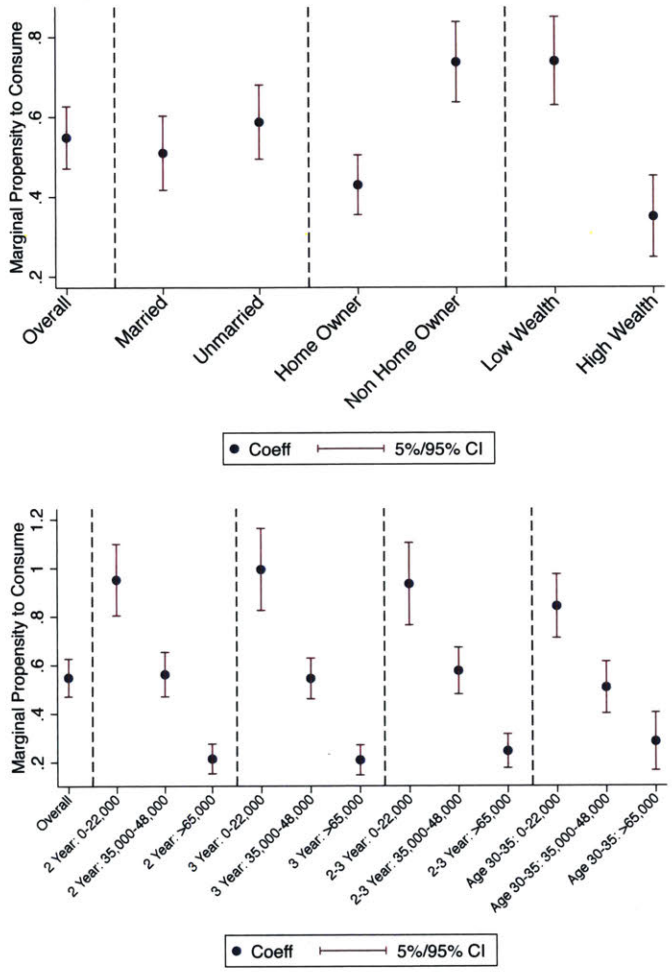

Notes: Red squares show consumption imputed using Equation A2 following the method outlined in Attanasio and Pistaferri (2014). Food shares and age profiles are calculated using data from 1998 to 2013, the period in which expanded food categories are available. Total consumption includes home insurance, rent, electricity, heating, water and miscellaneous utilities, car insurance, car repairs, gasoline, parking, bus fares, taxi fares and other transportation, school tuition and other school expenses, child care, health insurance and out-of-pocket health costs, and food. All averages are calculated based on the nationally representative subsample.

Figure A5: Heterogeneity in the Marginal Propensity to Consume: First Stage and Reduced Form



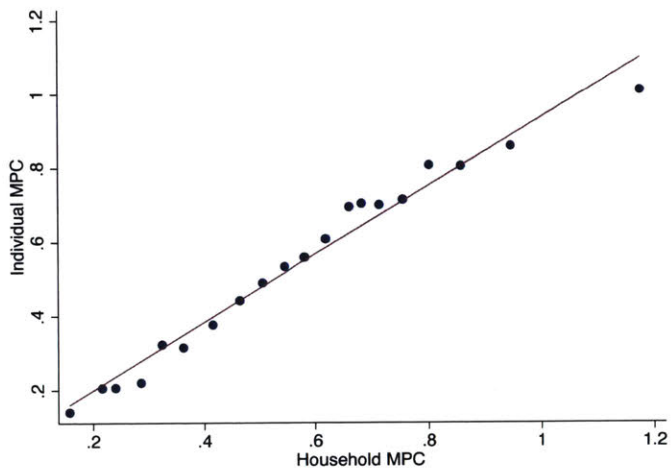

Notes: The left panel shows the first stage of unemployment on the level of labor earnings and the right panel shows the reduced form of unemployment on the level of consumption. These correspond to the instrumented regressions in Figure 1.3. Consumption is measured using total consumption, imputed using the method in Blundell, Pistaferri, and Preston (2008). Income is measured using individual labor income. The sample includes the set of workers who were employed two years before the current month. The sample in the PSID excludes observations with more than a 400 percent change in food consumption or income over a given two-year period. All regressions include year-by-state fixed effects and observations from 1981 to 2013.

Figure A6: Sensitivity of Baseline MPC Estimates to Specification of Demographic Variables



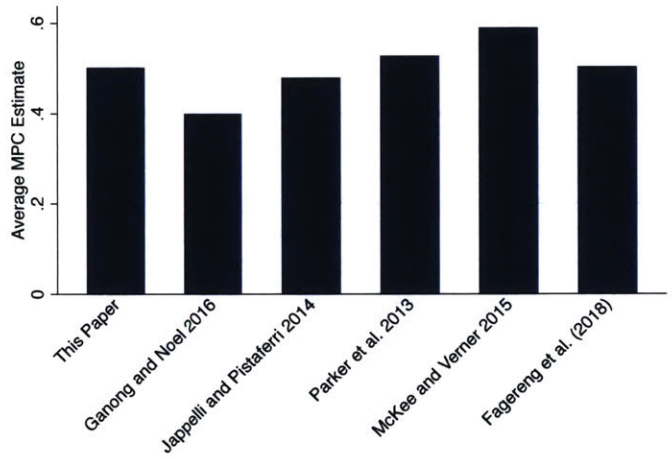

Notes: In the left panel, wealth is defined as the sum of assets less the value of debt plus the value of home equity. Assets include farm or business worth, checking or savings accounts, real estate assets, stocks, vehicles, and individual retirement accounts. Debts include business debt, real estate debt, credit cards, student loans, medical debt, legal debt, and family loans. Wealth is defined in 1984, 1989, and every year from 1999 to 2015. The right panel shows MPC estimates for different definitions of lagged earnings. "2 Year" defines income bins using the average earned in $t-1$ and $t-2$ (this is the baseline). "3 Year" defines income bins using the average earned in $t-1$, $t-2$, and $t-3$. "2-3 year" defines income bins using the average earned in $t-2$ and $t-3$. "Age 30-35" defines income bins using the income earned from ages 30 to 35, no matter what the current age of the worker. All regressions are otherwise as in Figure 1.3.

Figure A7: MPCs out of Individual and Household Incomes



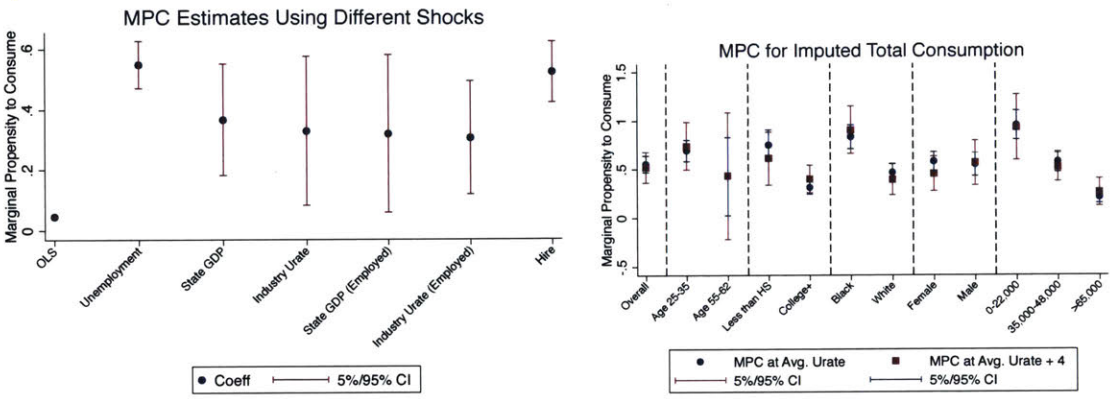

Notes: Individual MPCs on the y-axis are the baseline estimates plotted in Figure 1.4. Household marginal propensities to consume on the x-axis replace individual labor income with total household income, but otherwise, the specification is exactly as for the baseline MPC estimates. Each point represents an equal number of individuals. The sample includes all workers in the PSID estimation sample.

Figure A8: Estimates of the Average Marginal Propensity to Consume in the Literature



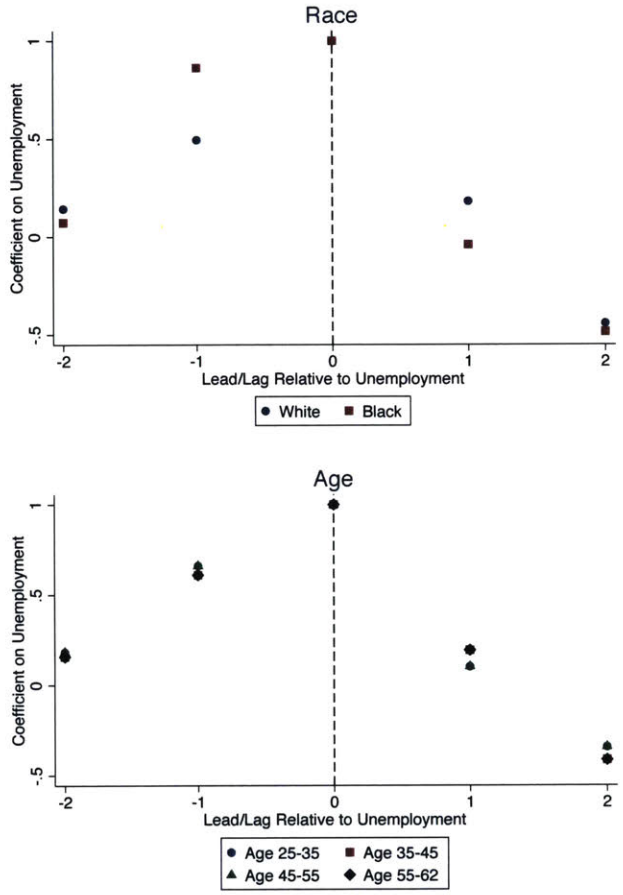

Notes: The estimates for this paper, labeled "baseline," are those plotted Figure 1.3. Ganong and Noel (2017) estimate the MPC at the onset of unemployment using balance sheet data from JPMorgan Chase & Co.. See Appendix Table 5 in their paper. Jappelli and Pistaferri (2014) use survey data in Italy to illicit MPCs out of transitory income shocks. Parker et al. (2013) identify the consumption response to the 2008 tax rebates. See Table 2 in their paper. McKee and Verner (2015) use Nielsen panel data to estimate the MPC out of unemployment insurance benefits.

Figure A9: The Stability of Marginal Propensity to Consume Estimates



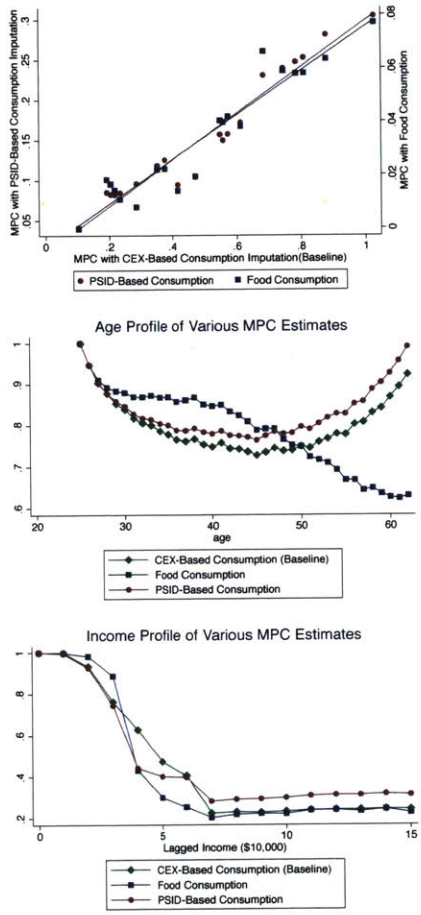

Notes: The instrument labeled "State GDP" in the left panel is defined as the percentage change in state GDP, defined by the Bureau of Economics Analysis. The industry unemployment rate is calculated from the Basic Monthly Current Population Survey, pooled over months within the year, and defined using time-consistent 1990 census industry codes. In Columns 1, 3, and 4, the sample includes the entire sample (no work restriction); in Column 2, the sample includes those employed in $t-2$; in Columns 5 and 6, the sample is restricted to those who are employed in $t-2$ and $t$; and in Column 7, the sample is restricted to those who are not employed in $t-2$. The unemployment rate in the right panel is defined as the unemployment rate in the state in which the individual was employed in $t-2$. Blue dots show the average MPC for the specified bin at the average unemployment rate in the sample. The red squares show the average MPC calculated at the average unemployment rate for each subsample at the average unemployment rate plus 4 percentage points. In both panels, regressions are based on two-year periods.

Figure A10: The Persistence of Unemployment Shocks on Income, by Demographics



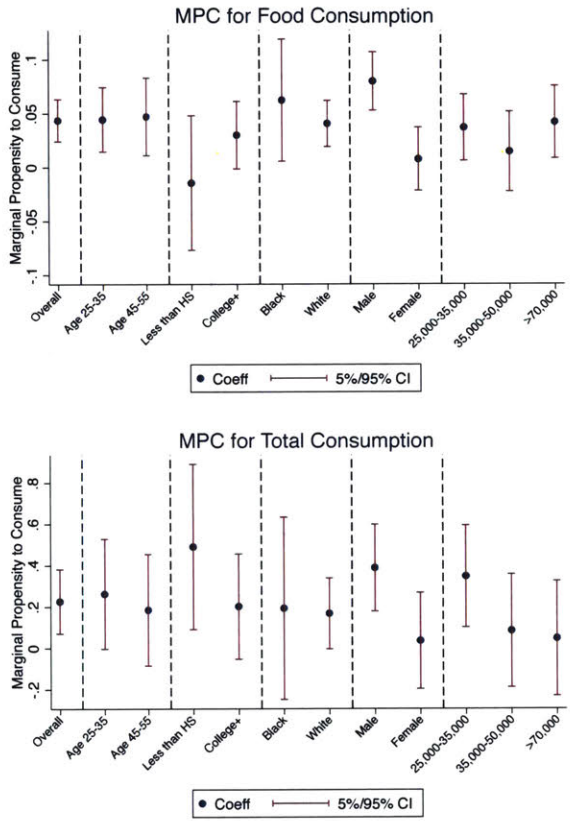

Notes: Each point shows the coefficient on unemployment at time $t$ from a regression on the change in income in time $t-k$ on unemployment in time $t$. Each coefficient is then rescaled by the coefficient on unemployment in time $t$ from a regression of the change in income in time $t$ on unemployment in time $t$. Therefore, each point presented the fraction of the $t = 0$ income loss in each group in each time period (i.e., if workers lose an average of \$20,000 when they become unemployed, 0.5 in time $t + 1$ means that workers' earnings are down by \$10,000 in $t+1$). Income is measured using individual labor income. In all estimates, the dependent variable is the change in income, the coefficient instrument for income changes is unemployment, and the sample includes the set of workers who were employed in $t - 2$ and excludes observations with more than a 400 percent change in food consumption or income over a given two-year period. The sample is restricted to years before 1997, when the PSID was conducted every year.

Figure A11: Marginal Propensities to Consume: Alternate Consumption Variables



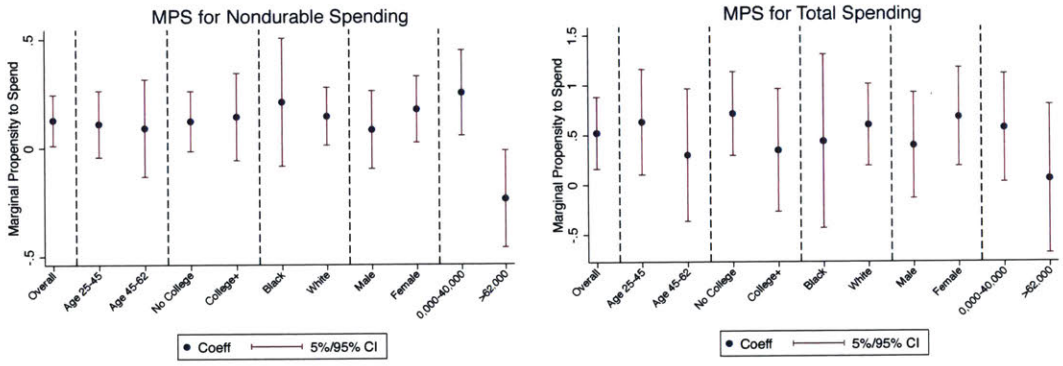

Notes: Each point in the top panel represents an equally sized data bin. Across all specifications, MPCs are estimated as in main-text Equation 1.5 with the baseline demographics. Income is measured using individual labor income. Food expenditure is measured using total food expenditure inclusive of food stamps. In all estimates, the instrument for income changes is unemployment and the sample includes the set of workers who were employed in $t-2$ and excludes observations with more than a 400 percent change in food consumption or income over a given two-year period. All regressions include year-by-state fixed effects. Regressions are estimated on the full sample, but the above figures are calculated on the nationally representative subsample of the PSID.

Figure A12: Marginal Propensities to Consume Estimated in the CEX



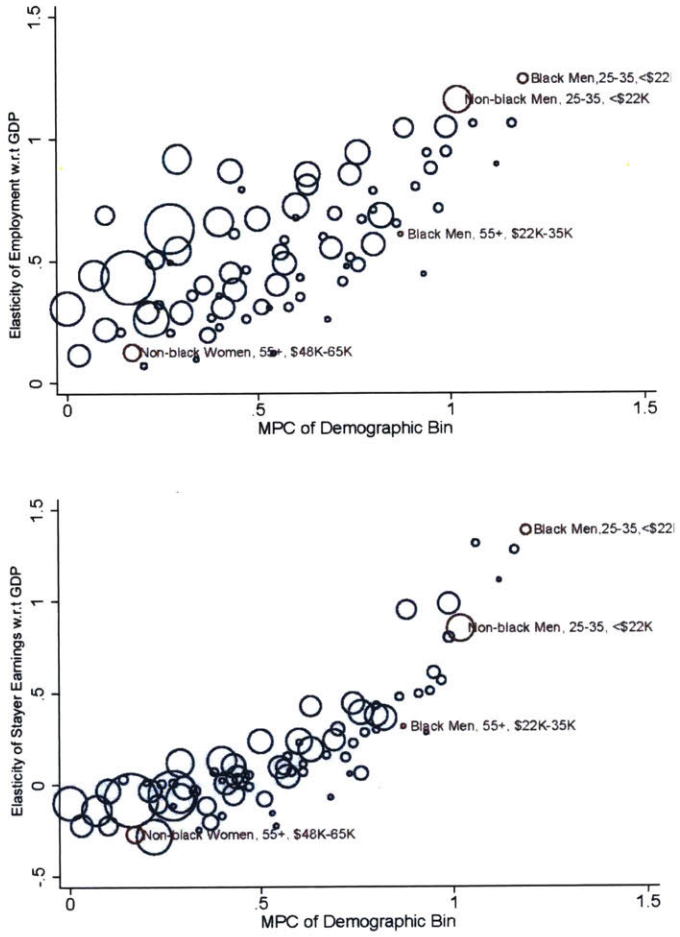

Notes: Income is measured using individual wage and salary earnings. Food expenditure is measured as the sum of food consumed at home and food eaten out. The instrument for income changes is reporting fewer than 38 weeks of employment in the previous year. The sample includes the set of workers who worked the prior 52 weeks in the second interview. Each regression excludes outlier observations, which are defined as observations with a change in income or relevant consumption of more than 200 percent. All regressions include year-by-quarter fixed effects, as well as month dummies.

Figure A13: Heterogeneity in Tax-Rebate-Based MPCs



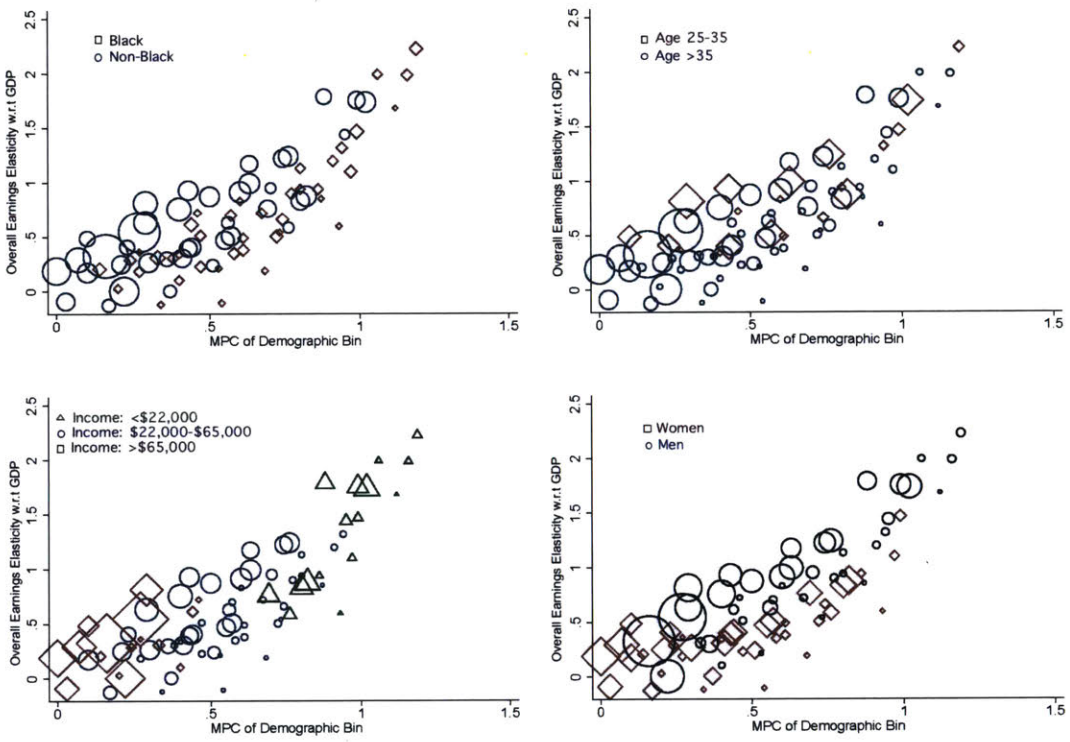

Notes: See Parker et al. (2013) for details. The overall estimate matches the two-stage least squares estimate in Table 2 of Parker et al. (2013). Standard errors are clustered at the household level.

Figure A14: Earnings Sensitivity to GDP and MPCs: Intensive and Extensive Earnings Margins



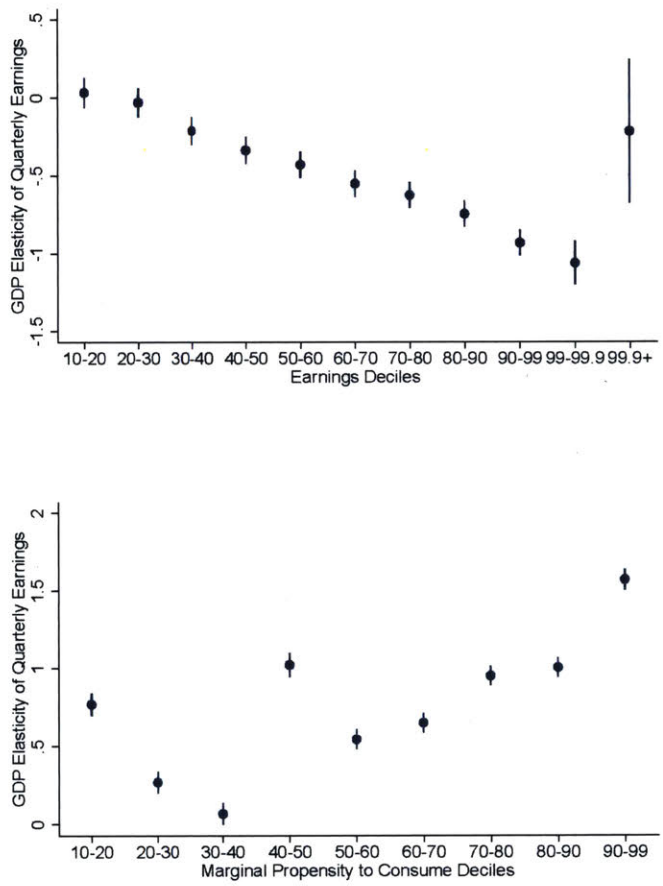

Notes: Sample includes the set of all workers employed in a sample state in year $t-1$ between 1995 and 2011. The dependent variable in the regression producing the y-axis estimates on the left graph is $\log(E_{i,t}) - \log(E_{i,t-1})$. The dependent variable in the regression producing the y-axis estimates in the right subplot is $L_t$, where $L_t$ is an indicator for being employed in time $t$. The size of each bubble represents the earnings share of that demographic group.

Figure A15: Earnings Sensitivity to GDP and MPCs: Demographic Group Decomposition



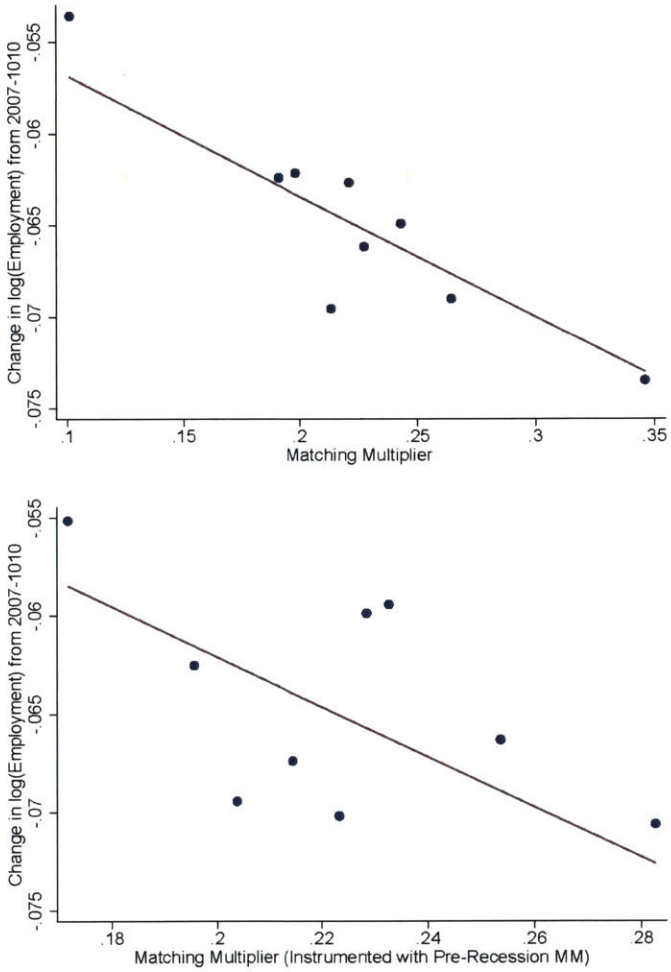

Notes: Sample includes the set of all workers employed in a sample state in year $t-1$ between 1995 and 2011. The dependent variable in the regression producing the y-axis estimates is the total change in log earnings for the demographic group. The size of each bubble represents the earnings share of that demographic group.

Figure A16: Earnings Sensitivity to GDP, by Decile of the MPC and Income Distribution



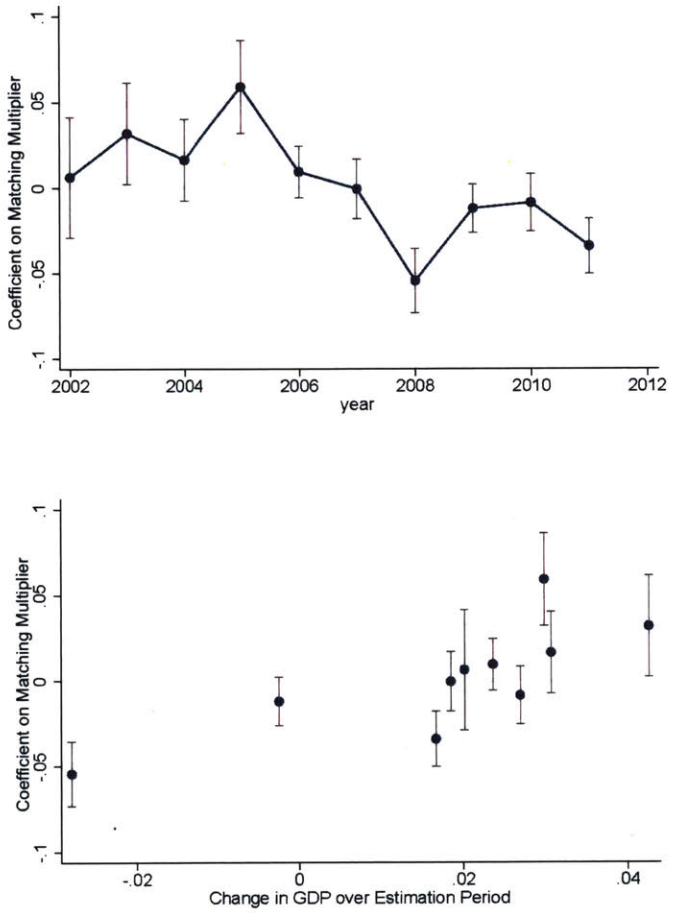

Notes: Sample includes a 5 percent random subset of all workers employed in a sample state in year $t-1$ between 1995 and 2011. The dependent variable in the regression is $\frac{E_{i,t}-E_{i,t-1}}{.5*E_{i,t}+.5*E_{i,t-1}}$. MPC decile bin cutoffs are defined on a sample pooled across all years and each bin represent an equal number of dollars of earnings, rather than individuals. Income deciles are also defined on a sample pooled across all years. Regressions include year fixed effects. Standard errors are clustered at the individual level. Blue bars reflect 95 percent confidence intervals.

Figure A17: Employment in the Great Recession and the Local Matching Multiplier



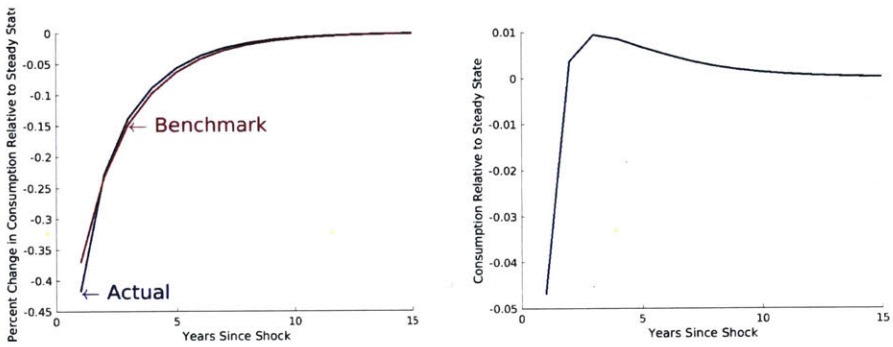

Notes: Each point represents a decile of the distribution of commuting-zones, which includes 270 commuting zones. Each commuting zone is weighted by its share of labor market earnings in 2007. Each plot includes controls for the share of employment in the 2-digit industry; the average age and lagged earnings of the area; as well as the fraction of the commuting-zone that is female, black, and in the labor force.

Figure A18: Employment Changes and the Local Matching Multiplier Over Time



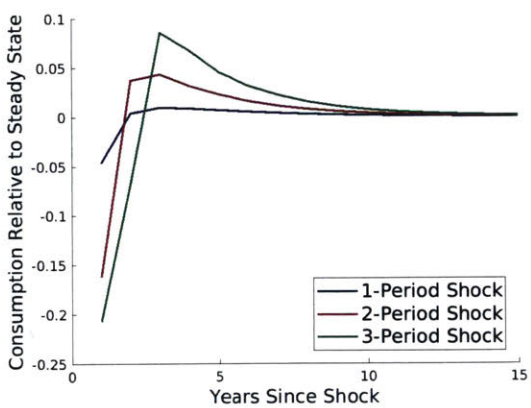

Notes: The left panel shows coefficients on $\widehat{MM}_c$ from estimates of $\Delta \log E_{c,t}$ on $\widehat{MM}_c$ in each year. The right panel shows the same annual estimates plotted against the change in GDP over that period. All regressions include controls for the share of employment in two-digit industries; the average lagged income and age; as well as the average fraction of the area that is black, female, and in the labor force. Red bars show the 90/10 confidence intervals.

Figure A19: Consumption Response to Aggregate Shock in Benchmark and Actual Scenarios



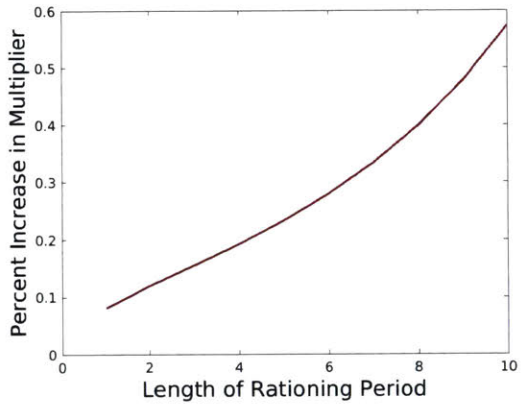

Notes: The left panel shows the consumption change relative to steady state in the actual and benchmark scenarios, expressed as a percentage relative to steady state. The right panel shows the difference between the consumption responses across these scenarios. In both figures, the interest rate is fixed at its steady-state value and the income shock is unanticipated and lasts for only one period.

Figure A20: Consumption Response by the Length of the Shock




Notes: The y-axis plots the difference in the consumption drop between the actual and the benchmark scenarios. The blue line shows the difference for a one-period shock, the red line shows the difference for a two-period shock, and the green line shows the difference for a three-period shock. In all cases, the interest rate is fixed at the steady-state value.

Figure A21: The Matching Multiplier and the Length of the Rationing Period




Notes: The y-axis plots the matching multiplier divided by the overall multiplier, showing the percentage increase in the overall multiplier coming from heterogeneity in the shock incidence. The x-axis shows the length of the rationing period. The matching multiplier and the overall multiplier are calculated using the sufficient statistic in Proposition 1, with the assumption that the interest rate is fixed in the future.

Table A1: Years in the Estimation Sample by State

| State | Sample Years |
|---|---|
| Arkansas | 2004-2011 |
| Arizona | 2006-2011 |
| **California** | **1993-2011** |
| **Colorado** | **1995-2011** |
| Washington DC | 2007-2011 |
| Delaware | 2000-2011 |
| **Florida** | **1994-2011** |
| Iowa | 2000-2011 |
| **Illinois** | **1992-2011** |
| Indiana | 2000-2011 |
| **Kansas** | **1995-2011** |
| **Maryland** | **1992-2011** |
| Maine | 1998-2011 |
| **Montana** | **1995-2011** |
| New Mexico | 1997-2011 |
| Nevada | 2000-2011 |
| Oklahoma | 2002-2011 |
| **Oregon** | **1993-2011** |
| Pennsylvania | 1999-2011 |
| South Carolina | 2000-2011 |
| Tennessee | 2000-2011 |
| **Washington** | **1992-2011** |
| West Virginia | 1999-2011 |

Notes: Sample years exclude the first two years for which there are individual-level data available. Bold states are those included in the balanced-panel subset of the data.

Table A2: Mobility Patterns: Cross-State

| Characteristic (X) | Worker Age | | Past Earnings | | Marginal Propensity to Consume | |
|---|---|---|---|---|---|---|
| $X_{t-1}$ | -0.002** | 0.004 | -0.045 | -0.095 | 0.177 | 0.284 |
| | (0.001) | (0.001) | (0.044) | (0.054) | (0.064) | (0.084) |
| $X_{t-1} * \Delta \log GDP_t$ | | -0.376 | | 0.000 | | -6.098 |
| | | (0.055) | | (0.000) | | (3.179) |
| No. Observations | 333 | 333 | 446 | 446 | 144 | 144 |
| R-Squared | 0.064 | 0.179 | 0.014 | 0.019 | 0.052 | 0.077 |

Notes: The dependent variable in each regression is the fraction of the workers in demographic bin $x$ who are employed at time $t$ and $t-1$ who change state of employment between $t-1$ and $t$. Worker age is binned into years, past earnings are binned into $1,000 bins, and MPCs are rounded to the nearest 0.01. The number of observations records the number of demographic bin years. All regressions include year fixed effects. Standard errors are clustered at the individual level.

Table A3: Mobility Patterns: Cross-Sector

| | | | |
|---|---|---|---|
| Marginal Propensity to Consume (MPC) | -0.0036*** | | -0.0032*** |
| | (0.0002) | | (0.0004) |
| $\Delta \log GDP$ | | 0.0087** | |
| | | (0.0043) | |
| MPC * $\Delta \log GDP$ | | | -0.0412 |
| | | | (0.0301) |
| No. Observations | 11104884 | 11104884 | 11104884 |

Notes: Data are from the Basic Monthly Current Population Survey from 1990 to 2011. The dependent variable is an indicator for moving from employment in the private sector to employment in self-employment, the military, or federal employment. MPC is imputed using PSID estimates based on age, gender, and race. The sample includes all adjacent periods in which an individual is employed. Standard errors are clustered at the individual level. Column 1 includes quarter fixed effects, and all other columns include year-by-month fixed effects.

Table A4: Summary Statistics for the Panel Study of Income Dynamics

| | LEHD Comparison Sample | National Sample of Employed | MPC Estimation Sample |
|---|---|---|---|
| ***Worker Characteristics*** | | | |
| Fraction Male | 0.548 | 0.54 | 0.514 |
| Average Worker Age | 42.7 | 42.3 | 41.3 |
| Average 2-year Lagged Income | 57,295 | 57,458 | 51,153 |
| Fraction College Educated | 0.365 | 0.373 | 0.308 |
| Fraction Black | 0.0469 | 0.053 | 0.265 |
| Average Income | 60,692 | 60,312 | 52,683 |
| Change in Income | 3,013 | 2,734 | 1,552 |
| ***Household Characteristics*** | | | |
| Household Size | 2.995 | 3.06 | 3.134 |
| Food Consumption | 9,033 | 9,163 | 8,795 |
| Total Consumption | 63,052 | 63,266 | 58,431 |
| Change in Food Consumption | 36.13 | 59.42 | 84.12 |
| Change in Consumption | 1,020 | 1,097 | 1,600 |
| **Number of Individuals** | 12,189 | 32,832 | 77,876 |

Notes: This table shows summary statistics for the PSID sample used in the analysis. The sample in all columns restricts to individuals ages 25 to 62 who are also observed in $t-2$ and $t+1$, who have nonmissing changes in food or income over two years, and whose consumption and income change over two years is less than 400 percent. The third column restricts to the set of individuals used in the estimation of MPCs and thus restricts to the set of workers employed in $t-2$. Column 2 instead restricts to those currently employed who are in the nationally representative subsample of the PSID. Column 1 further restricts to those living in the set of state-years available in the LEHD sample.

Table A5: Summary Statistics for the Consumer Expenditure Survey

| | CEX Sample | PSID Sample |
|---|---|---|
| ***Worker Characteristics*** | | |
| Fraction Male | 0.47 | 0.48 |
| Average Worker Age | 43.31 | 42.30 |
| Fraction College Educated | 0.34 | 0.37 |
| Fraction Black | 0.10 | 0.06 |
| Average Income | 45,620 | 50,775 |
| ***Household Characteristics*** | | |
| Household Size | 3.16 | 3.03 |
| Food Consumption | 8,366 | 8,884 |
| Total Consumption | 41,725 | 43,181 |
| **Number of Individuals** | 127,165 | 97,204 |

Notes: The first column shows summary statistics for the CEX sample used to impute total consumption. The second column shows the same statistics for the similarly constructed PSID sample. All nominal variables are adjusted to 2010 dollars.

Table A6: Coefficient Estimates for Individual Marginal Propensities to Consume

| Dep. Var. | Food Consumption | PSID Imputation | CEX Imputation |
|---|---|---|---|
| $< 22000$*Labor Income | 0.0273 | 0.2223 | 0.9171 |
| | (0.1173) | (0.2125) | (0.7844) |
| (22,000-35,000)*Labor Income | 0.0148 | 0.1235 | 0.6824 |
| | (0.1166) | (0.2117) | (0.7858) |
| (35,000-48,000) *Labor Income | -0.0157 | 0.0635 | 0.5572 |
| | (0.1173) | (0.2127) | (0.7901) |
| (48,000-65,000) *Labor Income | -0.0309 | 0.0047 | 0.3471 |
| | (0.1191) | (0.2155) | (0.7878) |
| $> 65,000$ *Labor Income | -0.0348 | -0.0305 | 0.1887 |
| | (0.1187) | (0.2143) | (0.7936) |
| Age*Labor Income | 0.0027 | 0.0045 | -0.0013 |
| | (0.0059) | (0.0106) | (0.0381) |
| $Age^2$ *Labor Income | -0.0334 | -0.0223 | 0.0415 |
| | (0.0701) | (0.1264) | (0.4464) |
| Female*Labor Income | -0.0224 | -0.0873 | -0.1347 |
| | (0.0139) | (0.0234) | (0.0706) |
| Black*Labor Income | 0.0417 | 0.2073 | 1.2314 |
| | (0.0589) | (0.1036) | (0.4141) |
| Black*Age*Labor Income | -0.0008 | -0.0042 | -0.0237 |
| | (0.0014) | (0.0025) | (0.0095) |
| Black*Female*Labor Income | -0.0052 | 0.0128 | -0.2532 |
| | (0.0296) | (0.0515) | (0.1717) |
| No. Observations | 123439 | 110403 | 69788 |
| Year FEs | X | X | X |

Notes: The table shows the regression estimates from PSID imputations. The dependent variable in the first column is total food consumption. The dependent variable in the second column is extended consumption imputed from the later years of the PSID. The dependent variable in the last column is total consumption imputed using the CEX data. The fitted values of the regression in Column 3 are plotted in Figure 1.4. Income is measured using individual labor income. The instrument for income changes is unemployment. The sample includes the set of workers who were employed two years before the current year. The sample in the PSID excludes observations with more than a 400 percent change in food consumption or income over a given two-year period. Lagged income is measured as the average labor marker earnings of the individual in t-2 and t-3. All regressions include state-by-year fixed effects. Columns 1 and 2 include years from 1971 to 2013 while Column 3 includes data from 1992 to 2013.

Table A7: The Correlation of MPCs Using Different Identifying Income Shocks

| Variables | Unemployment | Industry Urate | Hires | Jobloss |
|---|---|---|---|---|
| Unemployment Shock | 1.000 | | | |
| Industry Unemployment Rate | 0.111 | 1.000 | | |
| Hires Shock | 0.819 | -0.269 | 1.000 | |
| Unexpected Jobloss Shock | 0.869 | -0.097 | 0.681 | 1.000 |

Notes: Each row presents an MPC estimate using a different shock as an instrument for the change in earnings. The instruments across Rows 1 through 4 are unemployment, the unemployment rate in the industry in which the individual worked in $t-2$, being hired between $t-2$ and $t$, and unexpectedly losing your job. MPCs estimated using hires restrict the estimation sample to those not employed in $t-2$. MPCs estimated using unemployment or unexpected job loss restrict the sample to those who employed in $t-2$. MPCs estimated using the industry unemployment rate are estimated without a restriction on employment status in $t-2$. While the estimation sample differs across estimation methods, the MPCs are then imputed for the entire sample and the correlations are reported across the entire nationally representative subsample of the PSID.

Table A8: Estimates of the Marginal Propensity to Consume: Adding Job-Level Characteristics

| | | | | | | |
|---|---|---|---|---|---|---|
| Job Tenure | | -0.0008 | | | | |
| | | (0.0005) | | | | |
| Lagged Income Variance | | | -0.0676 | | | |
| | | | (0.1142) | | | |
| Industry Income Variance | | | | 1.5491 | | |
| | | | | (2.1224) | | |
| Occupation Income Variance | | | | | 0.8852 | |
| | | | | | (1.4540) | |
| North East Census Region | | | | | | 0.0276 |
| | | | | | | (0.0876) |
| North Central Census Region | | | | | | -0.0062 |
| | | | | | | (0.0908) |
| Southern Census Region | | | | | | 0.0025 |
| | | | | | | (0.0721) |
| $<$ 22000*Labor Income | 0.9171 | 0.7651 | 1.1373 | 0.9492 | 1.1150 | 0.9058 |
| | (0.7844) | (0.7979) | (0.8870) | (0.8201) | (0.9601) | (0.7910) |
| (22,000-35,000)*Labor Income | 0.6824 | 0.5433 | 0.6945 | 0.7181 | 0.9689 | 0.6707 |
| | (0.7858) | (0.7997) | (0.8817) | (0.8225) | (0.9574) | (0.7918) |
| (35,000-48,000) *Labor Income | 0.5572 | 0.4203 | 0.5426 | 0.6155 | 0.7544 | 0.5449 |
| | (0.7901) | (0.7987) | (0.8856) | (0.8266) | (0.9596) | (0.7961) |
| (48,000-65,000) *Labor Income | 0.3471 | 0.2073 | 0.3295 | 0.3981 | 0.5912 | 0.3388 |
| | (0.7878) | (0.8013) | (0.8875) | (0.8246) | (0.9651) | (0.7926) |
| $> 65,000$ *Labor Income | 0.1887 | 0.0527 | 0.1612 | 0.2389 | 0.4429 | 0.1760 |
| | (0.7936) | (0.8077) | (0.8940) | (0.8308) | (0.9650) | (0.7979) |
| Age*Labor Income | -0.0013 | 0.0071 | -0.0003 | -0.0070 | -0.0150 | -0.0007 |
| | (0.0381) | (0.0385) | (0.0433) | (0.0404) | (0.0460) | (0.0384) |
| $Age^2$ *Labor Income | 0.0415 | -0.0504 | 0.0321 | 0.1120 | 0.2060 | 0.0325 |
| | (0.4464) | (0.4474) | (0.5082) | (0.4725) | (0.5380) | (0.4495) |
| Female*Labor Income | -0.1347* | -0.1212 | -0.1456* | -0.1485** | -0.1188 | -0.1356* |
| | (0.0706) | (0.0747) | (0.0757) | (0.0755) | (0.0823) | (0.0720) |
| Black*Labor Income | 1.2314*** | 1.1508*** | 1.2590*** | 1.2504*** | 0.7959 | 1.2164*** |
| | (0.4141) | (0.4396) | (0.4543) | (0.4330) | (0.5080) | (0.4155) |
| No. Observations | 69788 | 69617 | 69788 | 69216 | 60516 | 69634 |
| Year FEs | X | X | X | X | X | X |

Notes: The dependent variable in all regressions is total consumption imputed using the CEX data. All regressions include state-by-year fixed effects. The expected variance of occupation- or industry-level earnings is calculated using the matched CPS monthly data and averaged over the sample period. The variance of lagged individual earnings is calculated as the log change of lagged earnings plus 1. This variance is winsorized at the top and bottom 5 percent to abstract from outliers. See text for more details.

Table A9: MPCs Estimated Using Job-Level and Geographic Variables

| Variables | Baseline | w/ Tenure | w/ Lagged Variance | w/ Ind. Variance | w/ Occ. Variance | w/ Region |
|---|---|---|---|---|---|---|
| Baseline | 1.000 | | | | | |
| w/ Tenure | 0.978 | 1.000 | | | | |
| w/ Lagged Variance | 0.974 | 0.956 | 1.000 | | | |
| w/ Ind. Variance | 0.995 | 0.971 | 0.968 | 1.000 | | |
| w/ Occ. Variance | 0.972 | 0.953 | 0.955 | 0.968 | 1.000 | |
| w/ Region | 0.999 | 0.977 | 0.974 | 0.994 | 0.972 | 1.000 |

Notes: The table shows the pairwise correlation between the baseline MPC estimate and estimates including the stated additional characteristic. The dependent variable is total consumption imputed using the CEX data. All regressions include state-by-year fixed effects and include individuals employed in time $t-2$. The expected variance of earnings at the industry or occupation level are calculated using the matched CPS monthly data and averaged over the sample period. These capture the average within-individual variance in earnings over a one-year period. Tenure is defined as the number of months the worker has been with the firm in which they were employed in time $t-2$. The lagged variance is the variance of an individual's earnings between $t-3$ and $t-2$. See Appendix text for more details on data construction, and see Appendix Table A8 for the regression estimates underlying these modified MPCs.

# Table A10: Robustness of Relationship Between MPCs and Earnings Elasticities: Alternate Specifications

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | | *Alt. Outcome* | | | | | | |
| | Baseline Estimate | log(Eit +100) | Levels | Aggregated Sample | Annual Income | 1993 State Subsample | State GDP | Multiple Imputation |
| $MPC_{i,t-1}$ | -0.245*** | -0.378*** | -0.258*** | -0.058*** | -0.077*** | -0.262*** | -0.237*** | -0.2261*** |
| | (0.001) | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) | (0.054) |
| $MPC_{i,t-1} * \Delta GDP_t$ | 1.300*** | 2.681*** | 2.073*** | 1.144*** | 0.732*** | 1.619*** | 0.888*** | 1.204*** |
| | (0.028) | (0.068) | (0.080) | (0.083) | (0.018) | (0.033) | (0.017) | (0.214) |
| No. Observations | 29,204,700 | 29,204,700 | 29,204,700 | 2,496 | 29,204,700 | 20,859,200 | 29,204,700 | 29,204,700 |
| R-Squared | 0.009 | 0.004 | 0.007 | 0.532 | 0.003 | 0.010 | 0.012 | 0.009 |
| Avg. MPC | .431 | .431 | .431 | .429 | .42 | .423 | .431 | .431 |

Notes: Column 1 shows regression estimates from Table 1.2, estimated on a 5 percent subsample of the data set. Columns 2 and 3 show different specifications of the earnings variable, both of which combine the intensive margin of earnings and the extensive margin of employment. Column 4 aggregates from the individual level to MPC bins (MPC rounded to the nearest 0.01) and runs a regression on the full aggregated sample. Column 5 defines income as the annual income in the calendar year, rather than quarterly income in the fourth-quarter. This analysis, however, still restricts the sample to the set of individuals employed in the fourth quarter of the previous year. Column 6 restricts to the subsample of states that are present in 1993, meaning that there is a balanced panel of states over time. Column 7 replaces aggregate GDP with state-level GDP. This specification includes state-by-year fixed effects, rather than simply year fixed effects. Lastly, Column 8 shows the estimates using multiple imputation, as described in the text. Across all columns, the number of observations is rounded to the nearest 100 to comply with U.S. Census Bureau disclosure requirements. Standard errors are clustered at the individual level in all columns except for Column 8.

Table A11: Robustness of Relationship Between MPCs and Earnings Elasticities: Alternate MPC Imputations

| *MPC Definition:* | Baseline Estimate | Food | Age Bins | PSID-Based Consumption | Hires Instrument |
|---|---|---|---|---|---|
| $MPC_{i,t-1}$ | -0.245 | -2.151 | -0.229 | -0.863 | -0.400 |
| | (0.001) | (0.009) | (0.001) | (0.002) | (0.004) |
| $MPC_{i,t-1} * \Delta \log GDP_t$ | 1.300 | 14.328 | 1.312 | 4.183 | 2.087 |
| | (0.028) | (0.310) | (0.027) | (0.082) | (0.113) |
| No. Observations | 29,204,700 | 29,204,700 | 29,204,700 | 29,204,700 | 29,204,700 |
| R-Squared | 0.009 | 0.006 | 0.008 | 0.013 | 0.002 |
| Avg. MPC | 0.431 | .029 | .421 | .169 | .246 |

Notes: Each column defines the marginal propensity to consume in a different way. Column 1 is the baseline MPC estimate; Column 2 is the same as in Column 1 but includes age bins rather than a quadratic in age. Column 3 uses the MPC defined only using food consumption. Column 4 uses the MPC calculated using the PSID-based imputation of expanded consumption. Column 5 uses the MPC estimated using hires, rather than unemployment, as the instrument for changes in income. The outcome variable in all regressions is the annual change in quarterly earnings across all jobs. All regressions include year fixed effects, and standard errors are clustered at the individual level. Across all columns, the number of observations is rounded to the nearest 100 to comply with U.S. Census Bureau disclosure requirements.

Table A12: Heterogeneity in Worker Exposure to Recessions: ACS Subsample With Household Characteristics

| | (1) LEHD 2001-2001 | (2) ACS Sample | (3) Extended MPC | (4) Household Avg. MPC | (5) Household MPC |
|---|---|---|---|---|---|
| $MPC_{i,t-1}$ | -0.25 | -0.22 | -0.177 | -0.187 | -0.077 |
| | (0.001) | (0.002) | (0.001) | (0.002) | (0.001) |
| $MPC_{i,t-1} * \Delta \log GDP_t$ | 0.952 | 0.926 | 0.802 | 0.896 | 0.447 |
| | (0.033) | (0.059) | (0.057) | (0.066) | (0.049) |
| Year FE | X | X | X | X | X |
| ACS Sample? | | X | X | X | X |
| No. Observations | 20,729,800 | 4,011,100 | 3,997,800 | 2,922,700 | 2,922,700 |
| R-Squared | 0.010 | 0.009 | 0.007 | 0.007 | 0.003 |
| Avg. MPC | .421 | .408 | .422 | .392 | 0.530 |

Notes: Column 1 restricts the baseline 5 percent LEHD subsample to 2001 to 2011, the years of the ACS sample. The dependent variable in Columns 1 through 3 is $\Delta E_{i,t} = \frac{E_{i,t}-E_{i,t-1}}{.5*E_{i,t}+.5E_{i,t-1}}$ at the individual $i$ level, and the dependent variable in Columns 4 and 5 is $\Delta E_{h,t} = \frac{E_{h,t}-E_{h,t-1}}{.5*E_{h,t}+.5E_{h,t-1}}$ at the household $h$ level. Observations using the ACS subsample are weighted by the multiple of the ACS person survey weight, and their earnings share in $t-1$ and observations in the LEHD are weighted by their earnings share. The MPC in Columns 1 through 3 is the baseline individual-level estimate used throughout the analysis. The household MPC in Column 4 is the earnings-weighted average individual MPC within the household. The MPC in Column 5 is constructed at the household level, allowing for heterogeneity by the following variables: the lagged earnings of the household, grouped into five earnings bins; a quadratic in the age of the oldest member of the household; an indicator for whether one or both members of the household are black; and the number of people in the household. Standard errors are clustered at the individual level in Columns 1 through 3 and at the household level in Columns 4 and 5.

Table A13: Decomposing the Relationship Between Worker MPCs and GDP Cyclicality

| | *Baseline LEHD Sample* | | | *ACS Sample* | |
|---|---|---|---|---|---|
| | Overall | Industry | Firm | Overall | Industry *Occupation |
| *Marginal Propensity to Consume* | 1.300 | 1.151 | 1.142 | 0.926 | 0.548 |
| | (0.028) | (0.030) | (0.035) | (0.059) | (0.073) |
| *Age* | -0.001 | 0.003 | 0.007 | -0.011 | -0.003 |
| | (0.001) | (0.001) | (0.001) | (0.002) | (0.002) |
| *Female* | -0.491 | -0.120 | -0.081 | -0.505 | -0.112 |
| | (0.015) | (0.016) | (0.018) | (0.032) | (0.041) |
| *Black* | 0.042 | 0.177 | 0.314 | -0.093 | 0.028 |
| | (0.026) | (0.026) | (0.029) | (0.060) | (0.064) |
| *Lagged Earnings* | -0.238 | -0.249 | -0.275 | -0.176 | -0.121 |
| | (0.010) | (0.010) | (0.012) | (0.018) | (0.024) |
| *MPC with Demographics Only* | 2.087 | 1.435 | 1.721 | 1.767 | 0.270 |
| | (0.113) | (0.112) | (0.117) | (0.153) | (0.176) |
| *MPC with Lagged Earnings Only* | 1.093 | 1.125 | 1.090 | 0.721 | 0.565 |
| | (0.029) | (0.031) | (0.037) | (0.060) | (0.077) |

Notes: Each entry reports an estimate of $\widehat{\alpha}_2$ from a separate estimation of Equation 1.6, where $MPC_{i,t-1}$ is replaced by the specified variable in the row (i.e., MPC, worker age, etc.). Each column refers to the set of fixed effects included in the regression. Columns 1 and 4 include year fixed effects. Column 2 includes industry-year fixed effects, Column 3 includes firm-year fixed effects, and Column 5 includes industry-year-occupation fixed effects. In Columns 1 through 3, the regression sample is a 5 percent subsample of the LEHD. In Columns 4 and 5, the regression sample includes the set of individuals in both the LEHD sample and ACS. The dependent variable in all regressions is $\Delta E_{i,t} = \frac{E_{i,t}-E_{i,t-1}}{.5*E_{i,t}+.5E_{i,t-1}}$. All standard errors are clustered at the individual level. Observations in Columns 1 through 3 are weighted by the earnings share in $t-1$, and observations in Columns 4 and 5 are weighted by the product of the ACS individual weight and the earnings share in $t-1$. The standard deviation of the baseline MPC is 0.257, and the standard deviation of the demographics and lagged-earnings-only MPCs are 0.06 and 0.249, respectively.

Table A14: Heterogeneity in Earnings Elasticity Among the Unemployed

| | Hires | | Conditional Earnings | |
|---|---|---|---|---|
| | ACS | LEHD | ACS | LEHD |
| $MPC_{i,t-1}$ | -0.430 | -0.266 | -1.824 | -2.065 |
| | (0.006) | (0.001) | (0.022) | (0.008) |
| $MPC_{i,t-1} * \Delta \log GDP_t$ | 0.309 | 0.527 | -0.284 | -1.958 |
| | (0.244) | (0.048) | (0.826) | (0.275) |
| No. Observations | 322,200 | 7,341,400 | 139,100 | 1,465,500 |
| R-Squared | 0.057 | 0.055 | 0.113 | 0.117 |

Notes: In Columns 1 and 3, the unemployed are identified using the ACS. In Columns 2 and 4, the unemployed are defined within the LEHD as the set of individuals who were employed in a sample state but are no longer employed. In Columns 1 and 2, the dependent variable is an indicator for becoming employed in $t+1$. In Columns 3 and 4, the dependent variable is log earnings conditional on hire. In all columns, MPC is defined using hiring as the instrument for income changes. All regressions include year fixed effects.

Table A15: National Estimates of the Matching Multiplier: Including the Unemployed

| | *Employed* | | *Unemployed* | | Unemp. | Total | % Increase |
|---|---|---|---|---|---|---|---|
| | $MPC^b$ | $MPC^a$ | $MPC^b$ | $MPC^a$ | share | MM | |
| ACS Sample | 0.23 | 0.29 | 0.79 | 0.79 | 0.04 | 0.19 | 56.93% |
| LEHD Sample | 0.23 | 0.29 | 0.83 | 0.81 | 0.04 | 0.19 | 55.80% |

Notes: Columns 1 and 2 are taken from Row 1 of main-text Table 1.3. Columns 3 and 4 show the actual and the benchmark MPCs only among the unemployed, estimated using a linear specification, as in Table A14. Column 5 gives the share of the total wage bill going to new hires from unemployment. Columns 6 and 7 give the overall estimates of the matching multiplier incorporating both the employed and the unemployed. The assumed labor share is two-thirds, and I assume that the overall MPC out of non-labor income is $\overline{MPC\beta}$.

Table A16: Employment Cyclicality and the Local Matching Multiplier: Robustness

| | Baseline | Bartik Shock | Hires-Based MPC | MPC Level | Saturated Demographics |
|---|---|---|---|---|---|
| $MM_c * Shock_t$ | 0.853 | 0.572 | 1.657 | 1.767 | 0.905 |
| | (0.303) | (0.203) | (0.857) | (0.776) | (0.299) |
| $B_c * Shock_t$ | -1.748 | -0.103 | 1.522 | 1.765 | -0.057 |
| | (1.223) | (0.075) | (1.227) | (1.092) | (0.407) ) |
| Year FE | X | X | X | X | X |
| Demographic Controls | X | X | X | X | X |
| No. Observations | 2245 | 1545 | 2028 | 2245 | 2245 |
| R-Squared | 0.751 | 0.776 | 0.812 | 0.749 | 0.799 |
| Avg. $MM_c$ | .199 | 0.123 | 0.076 | .089 | .0.199 |

Notes: Regression includes an unbalanced panel of 270 commuting zones from 2001 to 2011. All regressions include controls for the share of employment in the two-digit industry; the average age and lagged earnings of the area; and the fraction of the commuting zone that is female, black, and in the labor force. Each control is included independently and interacted with GDP. Column 2 shows results using the Bartik shock rather than GDP. See Appendix A for details on the construction of the shock. Column 3 shows estimates using hires as the identifying instrument for MPCs. Column 4 replaces $\widehat{MM}$ and $\widehat{B}$ with the level differences between $MPC_c^a$ and $MPC_c^b$ and the level of $MPC_c^b$, respectively. Column 5 includes demographic controls each interacted with a full set of year dummies. All regressions include year and commuting zone fixed effects. Observations are weighted by the share of employment in $t-1$, and standard errors are clustered at the commuting zone level.

Table A17: Tradable and Nontradable Employment and the Local Matching Multiplier: Robustness

| | Baseline | 4-digit NAICS | Additional Controls | MPC Level |
|---|---|---|---|---|
| $MM_c * \Delta \log GDP_t$ | 1.113 | 1.116 | 0.878 | 2.949 |
| | (0.294) | (0.260) | (0.324) | (0.693) |
| $MM_c * \Delta \log GDP_t * Tradable_i$ | -1.475 | -1.034 | -1.354 | -5.319 |
| | (0.658) | (0.567) | (0.714) | (1.953) |
| $B_c * \Delta \log GDP_t$ | -0.824 | -0.817 | -0.795 | 2.103 |
| | (0.479) | (0.361) | (0.507) | (1.095) |
| $B_c * \Delta \log GDP_t * Tradable_i$ | 1.327 | 0.777 | 1.536 | 1.129 |
| | (0.452) | (0.372) | (0.491) | (0.446) |
| Industry*Year FE | X | X | X | X |
| CZ*Industry FE | X | X | X | X |
| Demographic Controls | X | X | X | X |
| Financial Controls | | | X | |
| No. Observations | 50269 | 125308 | 41173 | 50269 |
| R-Squared | 0.394 | 0.329 | 0.401 | 0.394 |

Notes: Column 1 restates the baseline estimates from Column 3 in main-text Table 1.5. Column 2 shows results estimated on data disaggregated to four-digit NAICS codes. Column 3 shows estimates including additional commuting zone level controls, both interacted with GDP and with a tradable indicator. Column 4 shows results using the level differences in the $MPC^a$ and $MPC^b$, thus eliminating the rescaling in the calculation of $\widehat{MM}_c$. The dependent variable in each regression is winsorized at the 5th and 95th percentiles, as are the estimates of $\widehat{MM}_c$. Observations are weighted by the share of employment in $t-1$, and standard errors are clustered at the commuting zone level.

Table A18: Estimated Discount Rates ($\beta$) for Alternate Process Calibrations

| | $\log(y+UI)$ | $\log(y)$ | $\log(y+1000)$ | $\log(y+100)$ |
|---|---|---|---|---|
| *High School or Less* | | | | |
| Non-Black Men | 0.89 | 0.89 | 0.88 | 0.85 |
| Black Men | 0.65 | 0.65 | 0.63 | 0.55 |
| Non-Black Women | 0.86 | 0.83 | 0.83 | 0.78 |
| Black Women | 0.83 | 0.80 | 0.80 | 0.73 |
| *Some College or More* | | | | |
| Non-Black Men | 0.93 | 0.93 | 0.92 | 0.90 |
| Black Men | 0.74 | 0.75 | 0.74 | 0.65 |
| Non-Black Women | 0.89 | 0.88 | 0.89 | 0.86 |
| Black Women | 0.85 | 0.85 | 0.85 | 0.81 |

Notes: Each column reports the estimates of the discount factor $\beta$ calibrated to match the group average MPC with various estimates of the income process. Column 1 reports the baseline estimates where annual income in the PSID is measured as the log of income plus the average unemployment insurance payments. Column 2 estimates the income process restricting to only periods with positive earnings. Columns 3 and 4 show estimates using different transformations of log earnings.

Table A19: Model-Based Estimates of the Matching Multiplier

| | $\Delta C^b$ | $\Delta C^a$ | $MPC^b$ | $MPC^a$ | $Cov(\gamma_i, MPC_i)$ | Pct. Increase in $\Delta C$ |
|---|---|---|---|---|---|---|
| **Panel A: Benchmark Estimates** | | | | | | |
| Model Sufficient Statistic | -0.004 | -0.004 | 0.372 | 0.421 | 0.049 | 13.15% |
| Transitory 1 % Shock | -0.004 | -0.004 | 0.377 | 0.425 | 0.047 | 12.64% |
| **Panel B: Alternate Assumptions** | | | | | | |
| 2-period 1% Shock | -0.009 | -0.011 | – | – | – | 16.47 % |
| 3-period 1% Shock | -0.010 | -0.012 | – | – | – | 21.06 % |
| Anticipated Benchmark | -0.004 | -0.004 | 0.381 | 0.425 | 0.044 | 11.63% |

Notes: Column 1 shows the level decline in consumption relative to steady state in the first period in the benchmark case in which everyone gets a 1 percent income shock. Column 2 shows the level decline in consumption relative to steady state in the first period in the actual case in which everyone gets a $\gamma_i$ percent income shock. Column 3 shows the aggregate MPC in the benchmark case, and Column 4 shows the aggregate MPC in the actual case, both in the first period. Column 5 shows the differences between those two MPC (i.e. Column 5 less Column 4). Column 6 shows the percentage increase in the consumption response in the actual vs. benchmark case (i.e., $\frac{\Delta C^a - \Delta C^b}{\Delta C^b}$). All rows except for the one labeled "Endogenous Interest Rate" assume that the interest rate is fixed at the steady-state level.

# B Appendix for Chapter 3

Figure B1: Cross Commuting Zone Variation in Average Taks Concentration

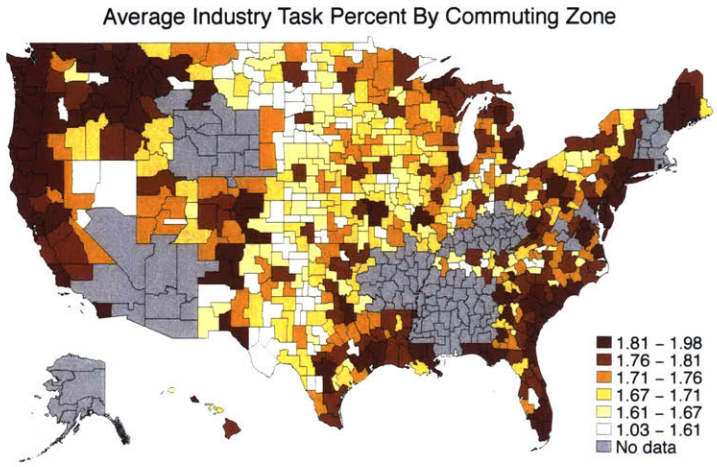



Table B1: Summary Statistics for Industry-by-Commuting Zone data in the QWI

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Employment | 955.54 | 4885.99 | 1 | 341020 | 1410159 |
| Avg. Monthly Wages | 3024.34 | 3114.26 | -72740.10 | 1018985.62 | 1410159 |
| Quarterly Separations | 656.85 | 4249.86 | 0 | 415424.76 | 1410159 |
| Quarterly Hires | 692.05 | 4379.67 | -665.12 | 411740 | 1410159 |
| Avg. Monthly Wages of New Hires | 1257.59 | 1490.94 | -64544 | 762150 | 1410159 |
| Avg. Monthly Wages of Incumbent Workers | 2371.17 | 1523.07 | -16510 | 114290 | 1410159 |
| 5-year % Change in Employment | -1.78 | 48.81 | -604.30 | 568.73 | 906536 |
| 5-year % Change in Avg. Monthly Wages | 12.98 | 21.27 | -464.03 | 677.77 | 906358 |
| 1-year % Change in Employment | -0.34 | 18.05 | -413.58 | 377.4 | 1305351 |
| 1-year % Change in Avg. Monthly Wages | 2.48 | 8.44 | -463.99 | 505.65 | 1305109 |
| 5-year Hires Rate | 3.95 | 8.4 | -72.01 | 3699.27 | 906536 |
| 5-year Separation Rate | 3.8 | 17.99 | 0 | 10454.97 | 906536 |

Table B2: Robustness: Relationship Between Local Industry Employment and National Industry Variation

| Dependent Variable: $\Delta$ ln(Employment) | | | | | | | |
|---|---|---|---|---|---|---|---|
| | Task Definition | | | | | | |
| | Int. | Gen. | Occ. | 1-year Change | Emp. Share Weighted | Raw Data | Nat. Emp. Change |
| Own Task Share (B) | -3.119*** | -2.694*** | -1.265*** | -0.241* | 2.074** | -6.920*** | 7.142*** |
| | (0.354) | (0.336) | (0.347) | (0.129) | (0.920) | (1.332) | (1.351) |
| Other Task Share (C) | -2.234*** | -1.973*** | -0.603** | -0.020 | 2.488*** | -3.626*** | 3.782*** |
| | (0.284) | (0.283) | (0.301) | (0.088) | (0.883) | (1.081) | (0.840) |
| State x Year Dummies | yes | yes | yes | yes | yes | yes | yes |
| Industry x State Dummies | yes | yes | yes | yes | yes | yes | yes |
| Industry x Year Dummies | yes | yes | yes | yes | yes | yes | yes |
| No. Observations | 234720 | 234720 | 234720 | 284488 | 231935 | 234720 | 268129 |
| $R^2$ | 0.484 | 0.484 | 0.483 | 0.282 | 0.599 | 0.305 | 0.441 |
| Dep. var. mean | -0.087 | -0.087 | -0.087 | -0.015 | -0.141 | 5.193 | -0.087 |
| Mean $\Delta log Y_{it}$ | | | | | | | |
| Mean Own Task Share (B)/(Y) | 0.002 | 0.002 | 0.004 | 0.000 | 0.005 | 0.002 | -0.001 |
| Mean Other Task Share (C)/(Z) | 0.086 | 0.087 | 0.084 | 0.014 | 0.077 | 0.087 | -0.043 |

Notes: The sample includes annual data for 71 manufacturing industries and 605 commuting zones from 2001-2011. Standard errors are twoway clustered at the state and industry level. Regressions includes state-by-year fixed effects, industry-by-state fixed effect, and parent-industry-by-time fixed effects.

Table B3: Robustness: Relationship Between Local Industry Wages and National Industry Variation

| Dependent Variable: $\Delta$ ln(Wages) | | | |
|---|---|---|---|
| | 1-year Change | Employment Weights | Employemnt Change |
| 1-Own Task Share (Y) | 0.001*** | -0.001 | 0.001 |
| | (0.000) | (0.001) | (0.001) |
| Other Task Share (Z) | 0.008 | -0.163 | 0.017 |
| | (0.028) | (0.166) | (0.033) |
| State x Year Dummies | yes | yes | yes |
| Industry x State Dummies | yes | yes | yes |
| Industry x Year Dummies | yes | yes | yes |
| No. Observations | 284432 | 231884 | 331837 |
| $R^2$ | 0.213 | 0.493 | 0.310 |
| Dep. var. mean | 0.026 | 0.143 | 0.141 |
| Mean $\Delta log Y_{it}$ | | | |
| Mean Own Task Share (B)/(Y) | 0.104 | 1.079 | -0.143 |
| Mean Other Task Share (C)/(Z) | 0.014 | 0.068 | -0.036 |

Notes: The sample includes annual data for 71 manufacturing industries and 605 commuting zones from 2001-2011. Standard errors are twoway clustered at the state and industry level. Regressions includes state-by-year fixed effects, industry-by-state fixed effect, and parent-industry-by-time fixed effects.

Table B4: Alternative Specifications: Relationship Between Local Industry Employment and National Output

| Dependent Variable: $\Delta$ ln() | | | | | | |
|---|---|---|---|---|---|---|
| | Size Interactions | | | Positive and Negative Movements | | |
| Own Task Share (B) | -1.467*** | -1.305*** | -0.888** | | | |
| | (0.312) | (0.322) | (0.408) | | | |
| $\Delta logY_{it}$*Own Emp. Share (B) | 3.719** | 3.472** | 2.729 | | | |
| | (1.610) | (1.626) | (2.216) | | | |
| Positive Own Task Share (B) | | | | -3.402** | -3.402** | -3.402** |
| | | | | (1.703) | (1.703) | (1.703) |
| Negative Own Task Share (B) | | | | -1.166 | -1.166 | -1.166 |
| | | | | (1.535) | (1.535) | (1.535) |
| State x Year Dummies | parent | parent | parent | parent | parent | parent |
| Industry x State Dummies | yes | yes | yes | yes | yes | yes |
| Industry x Year Dummies | yes | yes | yes | yes | yes | yes |
| No. Observations | 230931 | 230931 | 230931 | 230931 | 230931 | 230931 |
| $R^2$ | 0.478 | 0.478 | 0.478 | 0.478 | 0.478 | 0.478 |
| Mean $\Delta(log)$ | -0.087 | -0.087 | -0.087 | -0.087 | -0.087 | -0.087 |
| Mean Gov. Shock (A) | 0.102 | 0.102 | 0.102 | 0.102 | 0.102 | 0.102 |
| Mean Own Task Share (B) | 0.002 | 0.002 | 0.004 | 0.002 | 0.002 | 0.004 |
| Mean Other Task Share (C) | 0.086 | 0.087 | 0.084 | 0.086 | 0.087 | 0.084 |

Notes: The sample includes annual data for 71 manufacturing industries and 605 commuting zones from 2001-2011. Regression includes rolling 5 year changes. Standard errors are twoway clustered at the state and industry level.

Table B5: Robustness: Relationship Between Local Industry Employment and Government Spending Shocks

| Dependent Variable: $\Delta$ ln(Employment) | | | | | | |
|---|---|---|---|---|---|---|
| | Task Definition | | | | | |
| | Int. | Gen. | Occ. | 1-year Change | Emp. Share Weighted | Total Spending |
| Gov. Spending | 0.013** | 0.010** | 0.012* | 0.019** | 0.016* | 0.002 |
| | (0.006) | (0.005) | (0.007) | (0.008) | (0.009) | (0.002) |
| Own Task Share (B) | -0.535*** | -1.687*** | -0.140* | -0.755*** | -0.156 | -0.370** |
| | (0.115) | (0.261) | (0.071) | (0.153) | (0.113) | (0.161) |
| Other Task share (C) | -0.599*** | -1.444*** | -0.130*** | -0.416*** | -0.027 | -0.042 |
| | (0.089) | (0.105) | (0.033) | (0.110) | (0.068) | (0.036) |
| State x Year Dummies | parent | parent | parent | parent | parent | parent |
| Industry x State Dummies | yes | yes | yes | yes | yes | yes |
| Industry x Year Dummies | yes | yes | yes | yes | yes | yes |
| No. Observations | 858758 | 858758 | 858758 | 749544 | 740199 | 274238 |
| $R^2$ | 0.201 | 0.202 | 0.200 | 0.373 | 0.546 | 0.701 |
| Mean Dep. Var | 0.001 | 0.001 | 0.001 | -0.011 | 0.007 | -0.011 |
| Mean Gov. Shock (A) | 0.067 | 0.067 | 0.067 | 0.391 | 0.332 | 0.065 |
| Mean Own Task Share (B) | 0.001 | 0.001 | 0.005 | 0.007 | 0.027 | 0.001 |
| Mean Other Task Share (C) | 0.049 | 0.047 | 0.049 | 0.284 | 0.305 | 0.034 |

Notes: The sample includes annual data for 141 industries and 605 commuting zones from 2001-2014. Standard errors are twoway clustered at the state and industry level. Regression includes rolling 5 year changes and tasks are defined using Intermediate Work Activities (IWA). Regressions includes state-by-year fixed effects, industry-by-state fixed effect, and parent-industry-by-time fixed effects.

Table B6: Robustness: Relationship Between Local Industry Average Wages and Government Spending Shocks

| Dependent Variable: $\Delta$ ln(Wages) | | |
|---|---|---|
| | 1-year Change | Employment Weights |
| Gov. Spending | -0.001* | -0.006** |
| | (0.001) | (0.003) |
| 1-Own Task Share (Y) | 0.000* | -0.000 |
| | (0.000) | (0.000) |
| Other Task share (Z) | -0.005 | -0.014 |
| | (0.007) | (0.012) |
| State x Year Dummies | parent | parent |
| Industry x State Dummies | yes | yes |
| Industry x Year Dummies | yes | yes |
| No. Observations | 858586 | 740046 |
| $R^2$ | 0.137 | 0.462 |
| Mean Dep. Var | 0.028 | 0.141 |
| Mean Gov. Shock (A) | 0.067 | 0.332 |
| Mean Own Task Share (B) | 0.202 | 3.343 |
| Mean Other Task Share (C) | 0.056 | 0.287 |

Notes: The sample includes annual data for 141 industries and 605 commuting zones from 2001-2014. Standard errors are twoway clustered at the state and industry level. Regression includes rolling 5 year changes and tasks are defined using Intermediate Work Activities (IWA). Regressions includes state-by-year fixed effects, industry-by-state fixed effect, and parent-industry-by-time fixed effects.