From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2017

Cyclicality and Heterogeneity of Price Markup

Hong, Sungki

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01736667118

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Cyclicality and Heterogeneity of Price Markup

Sung Ki Hong

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Economics
Adviser: Oleg Itskhoki

June 2017

© Copyright by Sung Ki Hong, 2017.

All rights reserved.

Abstract

This dissertation studies the importance of firm-level price markup dynamics for business cycle fluctuations. In chapter one, I use recent IO techniques to measure the behavior of markups over the business cycle at the firm level. I find that markups are countercyclical with an average elasticity of -0.9 with respect to real GDP, in line with the earlier industry-level evidence. Importantly, I find substantial heterogeneity in markup cyclicality across firms, with small firms having significantly more countercyclical markups than large firms.

In chapter two, I examine if two prominent models in the literature are consistent with the empirical findings. First, I explore the Atkeson and Burstein (2008) model of oligopolistic competition. Coupled with an exogenous second-moment shock to firm productivities, this model results in a countercyclical average markup, but predicts that smaller firms reduce their markups in recessions. Second, I calibrate both Calvo and menu cost models of price stickiness to match the empirical heterogeneity in price durations as in Goldberg and Hellerstein (2011). I find that both models can match the average counter-cyclicality of markups in response to monetary shocks. Quantitatively, however, only the menu cost model, through its selection effect, can match the extent of the empirical heterogeneity in markup cyclicality. In addition, both sticky price models imply pro-cyclical markup behavior in response to productivity shocks.

In chapter three, I develop a new general equilibrium model that embeds *customer capital* into a standard firm dynamics model with entry and exit. A key feature of the model is that a firm's decision about markups becomes dynamic – firms accumulate customer capital in the periods of fast growth by charging low markups, and choose to exploit it by charging high markups in the downturns. In particular, during recessions, the endogenous higher exit probability for smaller firms implies that they place lower weight on future profits, leading them to charge higher markups. This mechanism serves to endogenously increase the dispersion of firm sales and employment in recessions. Also, the resulting input misallocation

amplifies both the volatility and persistence of the exogenous productivity shocks driving the business cycle.

Acknowledgements

I am grateful to my adviser Oleg Itskhoki. This dissertation would not have been possible without his continual guidance and support. He is a generous and inspiring adviser. We often held long and productive discussions, in his office, over meals and coffee, or in the basement of Fisher Hall. His deep understandings and knowledge of economics makes him an excellent scholar. He could always explain the core of complicated economic models and their main mechanisms with the most intuitive words. He is and will be the role model who I aspire to be.

Moreover, I am indebted to Richard Rogerson and Jakub Kastl for their critical feedbacks. They teach me how to think about research questions in terms of "big picture", and effectively communicate with the audience in presentations. And I much appreciate constructive suggestions and comments from Mark Aguiar, Nobuhiro Kiyotaki, Jan De Loecker, Ben Moll, Edurado Morales, Ezra Oberfield, Esteban Rossi-Hansberg, Chris Sims, and all seminar participants at Princeton University, Federal Reserve Bank of St Louis, USC Marshall School of Business, Queen's University, and 2016 Asia Meeting of the Econometric Society. Also, Laura Hedden and Jennifer Bello provided great assistance during the job market season (I will miss your emails.).

This journey of PhD in Economics would not have started without influence of my undergraduate advisers at MIT: Jerry Hausman, and Whitney Newey. And this journey would not have been accomplished without the friends I have met in Princeton. Zhuozhi Yao has been an excellent roommate. I shared a lot of memorable moments with Stephen Lin even after he graduated. John Kim was my adviser to Korean culture and Whisky tasting. A longer-than-necessary coffee break with Se Hyoun Ahn, Kevin Lim, Terry Moon, Dmitry Mukhin, and CJ Verbeck always fits nicely into my busy research schedule. With Jung Eun Yoon and Chen Zhao, I experienced the job market madness in a cold Chicago winter. And I feel grateful to Ki-Soon Choi, Sooji Han, and Injee Jeong for being there at my lowest points in Princeton.

Lastly, I am greatly indebted to my father, Un Pyo Hong, my mother, Young Rea Oh, and my sister, Kyung Lim Hong. Their love and support helped me reach the end of the journey. All of these would have been meaningless without them, and I would like to share all the achievements with them.

To my family.

Contents

Abstract iii

Acknowledgements v

List of Tables xi

List of Figures xii

1 Empirical Evidence **1**

1.1 Introduction 1

1.2 Empirical Study of Markup Cyclicality 5

1.2.1 Estimation of Markup 5

1.2.2 Data 9

1.2.3 Empirical Modeling Strategy 10

1.3 Empirical Results 11

1.4 Choice of Input 14

1.4.1 Labor Adjustment Cost 15

1.4.2 Overhead Labor Cost 19

1.5 Potential Mechanism for Counter-cyclical Markup 20

1.5.1 Rotemberg Adjustment Cost for Price 20

1.6 Conclusion 25

1.7 Tables 27

2 A Tale of Two Models **34**

2.1 Introduction . . . 34
2.2 General Oligopolistic Competition Model . . . 36
2.2.1 General Framework . . . 36
2.3 Quantitative Analysis: Atkeson-Burstein . . . 41
2.3.1 Calibration and Simulation . . . 48
2.3.2 Impulse Response . . . 50
2.4 Sticky Price Model . . . 51
2.4.1 Household . . . 52
2.4.2 Firms . . . 53
2.4.3 Recursive Competitive Equilibrium . . . 54
2.4.4 The CalvoPlusPlus Model . . . 55
2.4.5 Calibration . . . 57
2.4.6 Simulation Results . . . 57
2.4.7 Robustness . . . 59
2.5 Conclusion . . . 60

3 Customer Capital, Markup Cyclicality, and Amplification 73

3.1 Introduction . . . 73
3.2 Model . . . 77
3.2.1 Household . . . 79
3.2.2 Firm . . . 81
3.2.3 Recursive Competitive Equilibrium . . . 88
3.2.4 Equilibrium Computation . . . 89
3.2.5 Empirically Comparable Variables . . . 91
3.3 Calibration . . . 92
3.3.1 Fixed Parameters . . . 93
3.3.2 Fitted Parameters . . . 93
3.4 Numerical Results and Analysis . . . 94

3.4.1 Steady State Results 95
3.4.2 Results with Aggregate Fluctuations 98
3.5 Alternative Functional Form Specifications 102
3.5.1 GHH Preference 102
3.5.2 Quasi-Difference Deep Habits 104
3.6 Conclusion 105
3.7 Graphs 109

A **116**

A.1 Production Function Estimation 116

B **119**

B.1 Oligopolistic Competition Model 119
B.2 Numerical Solution for CalvoPlusPlus Model 123

C **125**

C.1 FOCs for Incumbent's Problem 125
C.1.1 First Order Conditions for Incumbents 125
C.1.2 Symmetric Equilibrium 126
C.2 Numerical Solution and Simulation 129
C.2.1 Transformation of Firm's Problem 129
C.2.2 Solution Algorithm 132
C.2.3 Internal Accuracy Statistics 136
C.2.4 Impulse Response Calculation 138
C.2.5 A More Realistic Frisch Elasticity of Labor Supply 138
C.2.6 GHH Preference 140

Bibliography **141**

List of Tables

1.1 Heterogeneity in Markup Cyclicality ϕ: Cobb-Douglas 27

1.2 Markup Cyclicality (Material Costs): Cobb-Douglas 28

1.3 Heterogeneity in Markup Cyclicality ϕ: Translog 29

1.4 Heterogeneity in Markup Cyclicality ϕ: Cobb-Douglas 30

1.5 Age and Financial Variables 31

1.6 Markup Cyclicality (Labor Costs): Cobb-Douglas 32

1.7 Markup Cyclicality: Labor Adjustment Cost and Overhead Labor 33

1.8 Markup Cyclicality: Rotemberg 33

2.1 Parameter Values for Calibration 61

2.2 Summary Statistics: Goldberg and Hellerstein (2011) 62

2.3 Parameter Values for Simulation: Calvo Model 63

2.4 Summary Statistics: Calvo Model 64

2.5 Parameter Values for Simulation: Menu Cost Model 65

2.6 Summary Statistics: Menu Cost Model 66

3.1 Heterogeneity in Markup Cyclicality: Exit Indicator 107

3.2 Fixed Parameter 108

3.3 Fitted Paramters 108

3.4 Targeted Moments in Data 109

List of Figures

2.1 Second Moment Shock 67

2.2 Response of Small Firms VS Large Firms 68

2.3 First Moment Shock 69

2.4 Comparison of markup cyclicality ϕ between data, Calvo Model, and Menu Cost model. 71

2.5 Comparison of markup cyclicality ϕ between data and Menu Cost model. Data includes both Cobb-Douglas and Translog cases. 72

3.1 Timing of Firm's Decision in Period t 84

3.2 Stationary Distribution of Firms 95

3.3 Exit Probability of Firms: $1 - G(\psi^*)$ 97

3.4 Revenue Growth Rate $\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}$ 97

3.5 Markup Cyclicality by Firm Size: Model vs Data 110

3.6 Impulse Response to TFP A_t 111

3.7 Value-Added Cyclicality by Firm Size: Model vs Data 112

3.8 Dispersion of Firm Size and Productivity 113

3.9 Comparison of Aggregate Productivity A_t and Measured TFP Z_t^D 114

3.10 Comparison: Homogeneous vs Heterogeneous Firm Model 115

Chapter 1

Empirical Evidence

1.1 Introduction

Understanding dynamics of price-cost markups is an important issue in business cycle analysis. Chari, Kehoe, and McGrattan (2007) find that the movements of the *labor wedge*[1] account for much of aggregate fluctuations, and Bils, Klenow, and Malin (2016) find that variation in markups explains at least half of the variation in the labor wedge. Also, endogenous markup variation is an important propagation mechanism for business cycle models to generate a more volatile employment response. Moreover, markup cyclicality has important implications for policymakers. The standard New Keynesian Philips Curve (NKPC) implies that the dual main objectives of the central bank do not conflict: stabilizing inflation also stabilizes output at desired level. But this "divine coincidence" breaks down with cyclical movements in desired markups, as stabilizing inflation no longer ensures zero output gap. Blanchard and Gali (2007) stress that understanding the interactions between the NKPC and the variations in desired markups "should be high on macroeconomists' research agendas."

The existing literature has focused on measuring and modeling the behavior of the aggregate (or sectoral) markup. In contrast, in this chapter I measure the markup behaviors

[1]Labor wedge is defined as the ratio between the marginal rate of substitution of consumption for leisure and the marginal product of labor.

at the firm level. And in chapter two and three, I study the role of markups by explicitly modeling firm heterogeneity, and deriving its implications for aggregate economic outcomes when subjected to business cycle shocks. A key advantage of this approach is that it allows one to assess whether the firm level choices of markups in the model are consistent with the data. Aggregate measures abstract from heterogeneity in markups and potentially confound changes in markups at the level of the individual firm with changes in the composition of firms.

My dissertation makes three contributions. Key to this approach is the availability of reliable estimates of markup variability at the firm level. Because such estimates do not exist, the first chapter documents the contribution of using the recent IO technique to measure markup behavior at the firm level. In the second chapter, I assess whether the two prominent models in the literature are consistent with the empirical findings of micro-level markup behavior, but they turn out not match key moments in the data. In the third chapter, I develop a general equilibrium model with heterogeneous firm that can account for the salient features of markups in the micro data. The key features of the model is the interaction between customer capital and firm entry and exit decisions. Also, I explore the implications of this model when subjected to aggregate shocks.

In this chapter, I first conduct the empirical analysis. In particular, I use firm level data of four large European countries (France, Germany, Italy, and Spain) from Amadeus that covers manufacturing sectors[2]. The dataset contains both production data and financial balance sheet data. However, the dataset provides neither the price nor the marginal cost necessary for directly measuring markup. Instead, I follow insights from Hall (1986) and De Loecker and Warzynski (2012), relying on firm's optimality condition of cost minimization with respect to a static input. Therefore, I use material input for markup estimation, because material does not suffer much from adjustment costs and other dynamic considerations (see Levinsohn and Petrin 2003). Previous studies measure markup with labor inputs, but it

[2]I use France as the benchmark country since it has longer time period (2003 – 2013) available for online download from WRDS than other three countries.

is well known that measurement of true costs of labor inputs suffers from several frictions, including but not limited to hiring and firing cost, overhead costs, and etc (as discussed by Rotemberg and Woodford 1999). Also, I check robustness for functional forms for production function when estimating the marginal cost. My result suggests strong countercyclicality of markups. I estimate that markups are countercyclical with an average elasticity of -1.2 with respect to real GDP with Cobb-Douglas functional form, and -0.7 with Translog. In other words, when real GDP decreases by 1%, firms increase their markups by 0.7% - 1.2%. This is in line with the recent findings by Bils, Klenow, and Malin (2015) using US industry-level data. Furthermore, taking full advantage of firm level data, I find substantial heterogeneity of markup cyclicality among firms. In particular, if I split firms based on their revenue share within own industries, I find that small firms' markups are 50% more countercyclical than large firms'.

Furthermore, exploiting the granularity of my firm-level data, I find that the empirical results above are robust along several dimensions. First, I check if the same results hold in other three countries, namely, Germany, Italy, and Spain. Although these countries cover shorter time periods than France, the main results stay the same. Next, I check whether the firm size actually approximates other firm's characteristics that matters for heterogeneity. It is possible that the firm size actually reflects firm age, since young firms tend to be small, and old firms tend to be large. I find that the size effect on markup cyclicality remains significant, controlling for the age effect. The other possibility is that firm size represents firm's financial health, because large firms probably have less collateral constraints and better relationships with lending institutes. A recent study by Gilchrist, Schoenle, Sim, and Zakrajsek (2015), using samples from Compustat matched with PPI data by BLS, documents that low liquidity firms raised prices, while high liquidity firms decreased prices during the recent financial crisis. However, the firm size still matters after firms' financial balance sheet variables are included in the regression. Notice that the heterogeneity is not driven by sector, since I define small and large firms based on their revenue share within 4-digit sectors.

In addition, I measure the markup with labor cost as in the literature. I find that the labor cost-based markup is strongly pro-cyclical. When real GDP increases by 1%, labor cost-based markup increases by 2%. This is in contrast to material cost-based markup. I calibrate a partial equilibrium model that features labor adjustment cost and overhead labor cost, but I find that these frictions or biases are not sufficient to generate pro-cyclical markups.

Lastly, I briefly discuss potential mechanisms that could generate counter-cyclical markups. I find that a partial equilibrium model with Rotemberg (1982) price adjustment costs could match the average cyclicality of markup as in the data.

Related Literature This chapter adds to the extensive literature estimating the cyclicality of markup. Bils (1987), Rotemberg and Woodford (1999), and Nekarda and Ramey (2013) use labor input to test the markup cyclicality at the aggregate level. Other studies infer the cyclicality using inventories (Kryvtsov and Midrigan 2012), retail prices (Stroebel and Vavra 2014), and advertising expenses (Hall 2014). I follow Bils, Klenow, and Malin (2013) to use material inputs, but I estimate the markup at the firm level using IO techniques by De Loecker and Warzynski (2012). To the best of my knowledge, this paper is the first one to use this technique to address the price-cost cyclicality at the firm level. And the paper is also related to the literature estimating heterogeneity in markup cyclicality. Bils, Klenow, and Malin (2012) study how markup cyclicality changes based on durability of goods. Chevalier and Scharfstein (1996) examine the pricing behavior of supermarkets with leveraged buyout (LBO). Gilchrist, Schoenle, Sim, and Zakrajsek (2015) document that more financially constrained firms raise prices more during the recent financial crisis. My work exclusively explores the heterogeneity of markup cyclicality due to firm size, which endogenously affects firm's exit decision over business cycles. I show that the result is robust against alternative economic mechanisms in these earlier works.

1.2 Empirical Study of Markup Cyclicality

In this section, I describe the estimation strategy for firm-level markup and the data I use for empirical analysis. Then, I show the empirical findings about markup dynamics, and discuss potential economic mechanisms.

1.2.1 Estimation of Markup

For firm i at time t, I define the markup as price over marginal cost:

$$\mu_{it} = \frac{P_{it}}{MC_{it}}. \tag{1.1}$$

In practice, we seldom observe neither price nor marginal cost in data. To overcome this problem, I follow the insights from Hall (1986) and De Loecker and Warzynski's (2012) and procedures in the literature that relies on firm's cost minimization condition with respect to static inputs that do not have adjustment frictions. I closely follow De Loecker and Warzynski's (2012) approach, which involves estimations of both production function and markup at the firm level. However, this methodology still suffers from the lack of information in the data, especially the firm-level price data. I will discuss the potential issues for the estimation caused by this missing information.

In particular, consider a firm i at time t that produces gross output Y_{it} with production function:

$$Y_{it} = F_{it}(M_{it}, L_{it}, K_{it}, A_{it}), \tag{1.2}$$

where M_{it} is material input, L_{it} is labor input, K_{it} is capital input, and A_{it} is firm productivity. For a given level of output Y_{it}, its associated Lagrange function with cost minimization is

$$\mathcal{L}(M_{it}, L_{it}, K_{it}, \lambda_{it}) = P_{it}^{M} M_{it} + W_{it} L_{it} + R_{it} K_{it} + \lambda_{it}(Y_{it} - F_{it}(M_{it}, L_{it}, K_{it}, A_{it})), \tag{1.3}$$

where P_t^M, W_{it}, and K_{it} are input prices for material, labor, and capital, respectively, and λ_{it} is the Lagrange Multiplier associated with the output constraint. Economically, it stands for marginal cost of production for a given level of output Y_{it}. In large, material is categorized as static input, while labor and capital are categorized as dynamic inputs. A static input is free of any adjustment costs and firm's dynamic decisions. The firm's optimality condition for the material input M_{it} is

$$P_{it}^M = \lambda_{it} \frac{\partial F_{it}}{\partial M_{it}}. \tag{1.4}$$

Since markup is equal to $\mu_{it} = \frac{P_{it}}{\lambda_{it}}$ as defined above, rearranging the optimality condition yields the following:

$$\mu_{it} = \theta_{it}^M \left(\alpha_{it}^M\right)^{-1}, \tag{1.5}$$

where $\theta_{it}^M = \frac{\partial \log F_{it}}{\partial \log M_{it}}$ is output elasticity with respect to material input, and $\alpha_{it}^M = \frac{P_{it}^M M_{it}}{P_{it} Y_{it}}$ is material cost share of revenue. This states that variation in markup is driven by the output elasticity and the input share. In data, we can easily observe the input share, but we need to estimate the production function to recover the output elasticity. And the output elasticity depends on the assumption of production function. For empirical analysis, I use two functional forms popularly used in the literature, and discuss their pros and cons below.

In contrast, firm's decision about dynamic inputs involves solving the full dynamic problem. For example, if there is adjustment costs for labors, such as hiring and firing costs, a firm has to consider how labor decision affects not only today's outcome, but also future's expected profits. Hence, the wedge between the output elasticity and labor share captures an additional wedge caused by the dynamic decisions. In the previous macroeconomics literature, researchers rely on labor input for markup estimation and conclude that markup is procyclical. However, the results might be biased due to the wedge caused by the dynamic nature of labor inputs. Rotemberg and Woodford (1999) show that, controlling for the adjustment costs, the markup is actually countercyclical.

Due to the reasons stated above, I use material input to examine markup cyclicality. Furthermore, I run my analysis at the firm level, which is different from previous analysis conducted at the industry level. Industry-level markup is a revenue-weighted average markup, and thus the results includes the compositional effect of firm sizes. However, as I discuss below, it is important to examine markup cyclicality at the firm-level, because it helps us understand and find the underlying economic mechanisms that generate the empirical patterns. Also, finding the correct mechanism is important to understand its welfare implications.

Production Function

In general, the output elasticity θ_{it} depends on the form of production function. For the estimation of production function, I restrict to two functional forms of production function commonly used in the literature: Cobb-Douglas and Translog. For the case of Cobb-Douglas form, the production function has the following form:

$$y_{it} = \theta_m m_{it} + \theta_l l_{it} + \theta_k l_{it} + a_{it}, \tag{1.6}$$

where lower case letters denote the natural logarithm of variables. Note that all the output elasticity terms θ are constants, and hence there is no need for estimation since I am mostly interested in the dynamics of markup, not level. However, this choice of functional form is restrictive. It does not allow any nonlinearity of or interactions among inputs. Hence, I consider the case of Translog form, which allows for more flexibility in terms of production function estimation. Specifically, the Translog production function is:

$$\begin{aligned} y_{it} =& \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + \theta_{mm} m_{it}^2 + \theta_{ll} l_{it}^2 + \theta_{kk} k_{it}^2 \\ &+ \theta_{ml} m_{it} l_{it} + \theta_{mk} m_{it} k_{it} + \theta_{lk} l_{it} k_{it} + a_{it}. \end{aligned} \tag{1.7}$$

In this case, the output elasticity is a function of inputs[3], hence ignoring the nature of time varying elasticity might bias the result of markup cyclicality. However, a simple OLS regression estimate of production elasticities is biased, due to the unobserved productivities a_{it} in the error term. To overcome this issue, I closely follow the IO techniques developed by Olley and Pakes (1996), Levinsohn and Petrin (2003), and Ackerberg, Caves, and Frazer (2015). For details of implementation, please see Appendix A. Notice that the flexibility of Translog comes with a price. In my data set, like most other data set, I do not have firm-specific price deflator, rather I deflate the firm revenue with industry-level price index. Hence, the firm specific price deflator enters the error term, and potentially causes bias in the estimation[4]. Nevertheless, the Translog function provides a robustness check of my result.

1.2.2 Data

The firm-level data in manufacturing sectors for four large European countries comes from Bureau Van Dijck's (BvD) Amadeus dataset[5]. This dataset has recently been used in many studies[6]. I focus on manufacturing sector for it suits the estimation of production function. I only use the firm labeled as "unconsolidated" to avoid double-counting the sales with sales in its parent company. The dataset contains 12,392 firm observations on average each year from 2003 to 2013[7].

[3]For example, output elasticity with respect to material is

$$\frac{\partial f_{it}}{\partial m_{it}} = \theta_m + 2\theta_{mm} m_{it} + \theta_{ml} l_{it} + \theta_{mk} k_{it},$$

which is a function of material, labor, and capital.

[4]De Loecker and Goldberg (2014) has a thorough discussion of potential issues with not controlling firm specific output and input prices.

[5]The four countries include France, Germany, Italy, and Spain. I use France as my baseline case, since other countries have shorter time period available for online download.

[6]For example, Gopinath, Kalemli-Ozcan, Karabarbounis, and Villegas-Sanchez (2015) use the same dataset to study capital misallocation in European countries.

[7]I understand that the time period is relatively short for business cycle study. However, this is the best dataset I can get for firm-level study at this stage. Currently, I am applying for the access to US Census dataset, which covers longer time period, and I would do the same analysis on it as soon as I get the access.

I now define the variables used in the empirical analysis. An *industry* is defined as a 4-digit NACE Rev. 2 code. I measure firm's *revenue* $P_{it}Y_{it}$ with operating revenue (gross output). Since the dataset does not have information about firm-level price P_{it}, I instead use the 2-digit industry-level price deflator P_t to measure the *output* Y_{it}. I measure the *material input* M_{it} with total material costs. I use the number of employees as measure of *labor input* L_{it}. I measure the *capital input* K_{it} with tangible fixed assets. To control for the input quality across firms in production function estimation, I have included the firm-specific average wage, defined as the ratio of total labor costs and total number of employees. Finally, I estimate the production function coefficients at 2-digit industry level[8].

I take business cycle measures from Eurostat, and Federal Reserve website. We use four measures of business cycle: (1) real GDP, (2) total employment, (3) total working hours, and (4) recession period defined by NBER. All business cycle variables are quadratically detrended.

1.2.3 Empirical Modeling Strategy

To empirically test the markup cyclicality, I bring the following specification into the data

$$\log \mu_{it} = \alpha_i + \phi_0 \log Y_{it} + \log Y_{it} \times \Omega'_{it}\phi_1 + \chi'_{it}\beta + \epsilon_{it}, \tag{1.8}$$

where μ_{it} is firm level markup, α_i is firm's fixed effect, Y_{it} is a business cycle indicator, Ω'_{it} is a vector of firm's characteristic variables, and χ_{it} is a vector of control variables. ϕ_0 captures the average markup cyclicality for all firms. ϕ_1 captures the heterogeneity in markup cyclicality across firms.

However, the fixed-effect (FE) regression above is not efficient if ϵ_{it} has serial auto-correlation, which is highly likely for time-series data. Hence, to control for the potential

[8]Details of data cleaning process will be available in online appendix.

auto-correlation in ϵ_{it}, I run the following first-difference (FD) regression

$$\Delta \log \mu_{it} = \phi_0 \Delta \log Y_{it} + \log \Delta Y_{it} \times \Omega'_{it}\phi_1 + \Delta \chi'_{it}\beta + \Delta \epsilon_{it}. \tag{1.9}$$

Note that I take first difference in all terms except for the term that captures the heterogeneity in markup cyclicality Ω'_{it}.

1.3 Empirical Results

Cobb-Douglas

I begin my empirical analysis with the case of Cobb-Douglas production. With this specification, the movement in markup is proportional to the inverse of material share. I evaluate the markup cyclicality with fixed-effect regression (1.11). I do not estimate the production function with this specification. Instead, I control the potential differences in output elasticity with firm fixed effects. Table 1.1 reports the results for this specification. The column (1) shows that markup is countercyclical and statistically significant for all firms. The elasticity of markup with respect to aggregate output is -1.2. This means that one-percent increase of output from its non-linear trend causes the markup to decline by 1.2 percent.

Moreover, with micro-level data, I find that there is substantial heterogeneity in markup cyclicality. One interesting and important aspect to look at is whether the markup cyclicality depends on firm size. Markups of small and large firms potentially move differently for several reasons: market power, pricing frictions, and etc. To explore the heterogeneity, I split firms into small and large firms based on their sizes. I define a large firm as one with more than 1% of market share within 4-digit industry. Column (2) and (3) report the results for subsample of large and small firms, respectively. I find that large firms markup elasticity to be small, equal to -0.9, and small firms have larger elasticity of -1.3. This implies that small firms' markups fluctuate 50% more than large firms' along business cycle. Recently, Bils, Klenow,

and Malin (2016) use intermediate input to estimate markup cyclicality at industry level and find that elasticity with respect to real GDP to be equal to -0.9. My finding with the large firms is largely comparable to this number, since large firms drive most of the industry-level markup, by its definition. Another way to identify differential effects between small and large firms is by pooling all firms and interacting aggregate output Y_t with a dummy variable Large_i, as in column (4). The result shows that the difference is significant. Alternatively, instead of using a dummy variable, I interact aggregate output with market share $s_{it} = \frac{p_{it}y_{it}}{\sum_i p_{it}y_{it}}$. The results in column (5) indicate the same findings.

In a panel regression, the error term ϵ_{it} in the fixed-effect specification is likely to follow an AR(1) process, and the resulting serial auto-correlation could bias the results above. Hence, I use first-difference regression for robustness check. Another advantage of using the first-difference regression is that the constant output elasticity term under Cobb-Douglas specification drops out by first-differencing the markup. The results are reported in Table 1.2. I find the same results as with fixed-effect regression: markup is countercyclical, and small firms' markups fluctuate more.

Translog

Although Cobb-Douglas is an convenient assumption for estimation, it has a strong restriction in terms of specification that output elasticities to be independent of input usages. And one could wrongly attribute the change in production technology to the change in markup. To depart from Cobb-Douglas, I assume Translog production function, which allows for higher order and interaction terms of inputs. However, as discussed above, since I do not have firm-specific output deflator, the estimations of Translog production function is biased. Nevertheless, I use this alternative specification for robustness check.

I report the fixed-effect regression results in Table 1.3. The results are consistent with main findings under the Cobb-Douglas specification, but there are two differences in terms of magnitude. First, the markup is less countercyclical under Translog. It implies that

firms actually substitute among the inputs along business cycle, possibly due to different factor prices, adjustment costs and etc. Second, the cyclicality difference between small and large firms are larger. This means that firms adjust its input for production differently. For example, a firm will be more reluctant to change its labor inputs if it faces greater adjustment frictions in labor than other firms.

Robustness Check

Here, I investigate whether the results above are robust to alternative samples and. First, I check whether the results still hold in other countries. In the Amadeus dataset, there are other countries available for download, but they have shorter time periods available for online download from WRDS. Nevertheless, I choose Germany, Italy, and Spain for robustness check. Results are reported in Table 1.4. All the results are robust.

Next, I investigate if the firm size actually captures other characteristics of firms. Young firms tend to be small, and old firms tend to be large. Hence, the size effect on markup cyclicality could be due to age effect. I check this by adding an additional interaction term between aggregate output and age. The results are in column (1) of Table 1.5. The coefficient of the interaction term with market share remains strongly positive and statistically significant, after controlling for age. Hence, the size effect is important for firm's pricing behavior.

Also, a recent study by Gilchrist, Schoenle, Sim, and Zakrajsek (2015) finds that financially constrained firms raise markups more to avoid using costly external financing to pay fixed payment of operation during the Great Recession[9]. And firm size might be proxying the financial conditions of the firms. Therefore, I use the following four financial variables to control for financial constraints, which are *cash ratio* – a ratio of cash and cash equivalents to current liabilities, *current ratio* – a ratio of current assets to current liabilities, *liquidity ratio* – a ratio of liquid assets to current liabilities, and *solvency ratio* – a ratio of cashflow

[9] An earlier study by Chevalier and Scharfstein (1996) find thats more financially constrained supermarkets have more countercyclical markups.

to total liabilities. The results are reported in column (2) to (5). First, we see that the size effect remains strongly positive. Second, the coefficients of interaction term with all financial variables are small and statistically insignificant.

Overall, these results suggest that firm size plays a prominent role in determining the markup cyclicality. Therefore, any theory or model of markup cyclicality has to be consistent with the documented size effect.

1.4 Choice of Input

In principle, any input that is free of adjustment cost can be used to measure markup and should yield the same value. For my benchmark analysis, I use materials input for inference. Materials is believed to have less adjustment costs, and less overhead components than other inputs (See Levinsohn and Petrin (2003), and Bils, Klenow, and Malin (2016).). In the literature, the conventional way of estimating markup is using labor cost (See Nekarda and Ramey (2013)). However, if labor input incurs additional costs $AC_{it}(\cdot)$, its optimality condition is

$$\mu_{it} = \underbrace{\theta_{it}^{L}\left(\alpha_{it}^{L}\right)^{-1}}_{=\mu_{it}^{L}}\left[1+\frac{\partial(AC_{it}/W_{it})}{\partial L_{it}}\right]^{-1}, \tag{1.10}$$

where $\alpha_{it}^{L} = \frac{W_{it}L_{it}}{P_{it}Y_{it}}$ is labor share of revenue, and μ_{it}^{L} is the mis-measured markup if we wrongly assume that L_{it} is a static input. The distortion comes from $\partial_L(AC_{it}/W_{it})$ in the second term, and I refer to it as the real marginal cost bias[10]. The effect of the real marginal cost bias is twofold. First, it affects the level of the measured markup $\hat{\mu}_{it}^{L}$. If the real marginal cost bias is positive, $\hat{\mu}_{it}^{L}$ upward biases the true markup μ_{it}. Second, it affects the cyclicality of the markup, if the real marginal cost bias is time varying. Hence, we can not recover the true cyclicality of markup using labor input. Nevertheless, if the real marginal cost

[10]Please see Rotemberg and Woodford (1999) for a thorough discussion of possible sources of this bias and potential remedies for them.

bias affects all the firms the same way, then we can still infer the heterogeneity of markup cyclicality among firms.

To see how much bias the additional costs bring to the cyclicality of the markup, I run the following fixed-effect regression:

$$\log \mu_{it}^{L} = \alpha_i + \phi_0 \log Y_{it} + \log Y_{it} \times \Omega_{it}' \phi_1 + \chi_{it}' \beta + \epsilon_{it}. \tag{1.11}$$

The regression is identical to the main regression, except that the markup is now measured with labor share of revenue. The results are reported in table 1.6. Although large firms have more pro-cyclical markups than small firms as with the case of material input, the firms have strongly pro-cyclical markup. The average elasticity of markup with respect to real GDP is 2, in stark contrast to -1 in the baseline case.

To see whether the omission of the additional cost AC_{it} in measurement could bring such a huge bias, I calibrate and simulate two models in the following.

1.4.1 Labor Adjustment Cost

The first model I consider is a partial equilibrium with labor adjustment cost. Assume that a firm produces output Y_{it} with labor L_{it}:

$$Y_{it} = L_{it}^{\alpha}. \tag{1.12}$$

Assume that the firm faces the following demand function:

$$C_{it} = P_{it}^{-\theta_t} C_t, \tag{1.13}$$

where P_{it} is the price the firm charges, and C_t is aggregate consumption at time t. θ_t is the elasticity of demand at time t, and is time varying. Hence, the firm's desired markup changes over the time. I specify its law of motion below.

The price of material is equal to the economy's aggregate price index. The real wage is a power function of aggregate consumption:

$$w_t = \omega C_t^{\sigma}, \tag{1.14}$$

where ω is the marginal disutility of labor, and σ measures the relative risk aversion[11]. Also, the firm has to pay adjustment costs for labor in units of labor as in Cooper and Willis (2009):

$$\frac{\gamma_l}{2}\left(\frac{L_{it}}{L_{it-1}} - 1\right)^2 L_{it-1} + F_{L_{it} \neq L_{it-1}}, \tag{1.15}$$

where γ_l governs the labor adjustment cost[12].

I assume that the aggregate output follows the following AR(1) process:

$$\log Y_t = (1 - \rho_y)\log \bar{Y} + \rho_y \log Y_{t-1} + \varepsilon_t^y \tag{1.16}$$

where $\bar{Y}$ is the steady state aggregate output, ρ_y governs the persistence degree of the process, and $\varepsilon_t^y \sim N(0, \sigma_y)$. Furthermore, I assume the following relation between the firm's desired markup and the aggregate output:

$$\log \frac{\theta_t}{\theta_t - 1} = \log \frac{\bar{\theta}}{\bar{\theta} - 1} + \phi(\log Y_t - \log \bar{Y}), \tag{1.17}$$

where $\frac{\bar{\theta}}{\bar{\theta}-1}$ is the steady state desired markup, and ϕ governs the cyclicality of the desired markup with respect to aggregate output. $\phi < 0$ implies that markup is counter-cyclical, while $\phi > 0$ pro-cyclical. And I impose that the aggregate consumption C_t is equal to the aggregate output Y_t.

[11] This equation comes from the assumption that a representative household has the following utility function in aggregate consumption and labor:

$$\frac{C_t^{1-\sigma}}{1-\sigma} - \omega L_t.$$

[12] I assume that the hiring and firing costs are symmetric

The firm's real profit at time t is the following:

$$\Pi_{it} = Y_{it}^{\frac{\theta_t - 1}{\theta_t}} Y_t^{\frac{1}{\theta_t}} - w_t L_{it} - w_t \left[\frac{\gamma_l}{2} \left(\frac{L_{it}}{L_{it-1}} - 1 \right)^2 L_{it-1} - F_{L_{it} \neq L_{it-1}} \right]. \tag{1.18}$$

The firm discounts the future real profit with the following discount factor

$$q_{t,t+1} \equiv \beta \left(\frac{Y_{t+1}}{Y_t} \right)^{-\sigma}. \tag{1.19}$$

Then the firm's value function can be written as:

$$V(L_{-1}, Y) = \max_L \Pi + \mathrm{E}(qV(L, Y')). \tag{1.20}$$

The value function has two state variables. One is the labor choice in the last period L_{-1}, and the other is current period's aggregate output level Y. The firm chooses the current labor level L to maximize its sum of profits.

For model calibration, I assume that one period is equal to one quarter, and the time discount factor is $\beta = 0.99$. The steady state elasticity of demand $\bar{\theta}$ is set to be equal to 6, such that the firm's average markup is 1.2. I set the elasticity of markup with respect to real GDP as $\phi = -1$, as observed in the data. The production function has a decreasing returns to scale of 0.8 ($\alpha = 0.8$). The constant relative risk aversion parameter is equal to $\sigma = 1$, which means that the "household" has log utility in consumption. Finally, I choose ω such that the steady state labor supply is 1/3 ($\hat{L} = 1/3$) [13].

[13] It follows that $\bar{Y} = (1/3)^{\alpha}$.

Measurement of Markup

The first order condition with respect to labor yields the following:

$$\begin{aligned}\frac{\theta_t}{\theta_t - 1} =& \alpha \left\{ \frac{w_t L_{it} + w_t \gamma_l \left(\frac{L_{it}}{L_{it-1}} - 1 \right) L_{it} + w_t \mathrm{E}_t \left[q_{t,t+1} \frac{\gamma_l}{2} \left(1 - \left(\frac{L_{it+1}}{L_{it}} \right)^2 \right) \right] L_{it}}{P_{it} Y_{it}} \right\}^{-1} \\ =& \mu_{it}^L \left\{ \frac{1 + \gamma_l \left(\frac{L_{it}}{L_{it-1}} - 1 \right) + \mathrm{E}_t \left[q_{t,t+1} \frac{\gamma_l}{2} \left(1 - \left(\frac{L_{it+1}}{L_{it}} \right)^2 \right) \right]}{P_{it} Y_{it}} \right\}^{-1}. \end{aligned} \tag{1.21}$$

Hence, if the choice of laborL_{it} is likely to be pro-cyclical, then the second terms in the expression above is counter-cyclical, and it would add pro-cyclical bias to the cyclicality of markup based on μ_{it}^L.

Simulation Results

To quantify the pro-cyclical bias in markup cyclicality, I run the following regression:

$$\log \mu_{it}^L = \phi_0 + \phi_1 \log Y_t + \epsilon_t. \tag{1.22}$$

I choose three different values of γ_l such that the standard deviation of log labor is equal to 80%, 60%, and 40% of flexible case ($\gamma_l = 0$). They correspond to $\gamma_l = 0.5$, $\gamma_l = 2.5$, and $\gamma_l = 9$, respectively. The results are reported in top panel of table 1.7. From the results, we see that the labor adjustment cost does bring pro-cyclical bias. In the flexible choice case, the markup cyclicality should be $\phi = -1$. We see that ϕ reduces to around -0.9 with labor adjustment costs. However, the pro-cyclical bias is not strong enough to generate the elasticity of 2 as observed in the data.

1.4.2 Overhead Labor Cost

Next, I consider a partial equilibrium model with overhead labor cost. Assume that a firm produces output Y_{it} with labor L_{it} but with overhead labor:

$$Y_{it} = (L_{it} - L^o)^{\alpha}, \tag{1.23}$$

where L^o denotes the amount of overhead labor[14].

Now, the firm's first-order condition becomes the following:

$$\begin{aligned} \frac{\theta_t}{\theta_t - 1} &= \alpha \left\{ \frac{w_t(L_{it} - L^o)}{P_{it}Y_{it}} \right\}^{-1} \\ &= \mu_{it}^L \frac{L_{it}}{L_{it} - L^o}. \end{aligned} \tag{1.24}$$

Since the choice of labor L_{it} is pro-cyclical, $\frac{L_{it}}{L_{it}-L^o}$ contributes pro-cyclical bias. In the simulation, I choose three different values of overhead labor cost L^o, which are 30%, 50%, and 70% of the steady state labor, respectively. The results are in the bottom panel of table 1.8. One can see that as the overhead labor cost increases, the "markup" becomes more and more pro-cyclical. But even with 70% of overhead labor cost, the elasticity of markup is only -0.08, which is still far from the estimated elasticity 2 in the data.

The results above suggest that other mechanisms are needed to explain the strong pro-cyclical gap between material-based and labor-based markups.

1.5 Potential Mechanism for Counter-cyclical Markup

Now I present a model that can generate counter-cyclical markup and see if it can match the markup cyclicality quantitatively through simulation. The model I am going to analyze is a New Keynesian model. The reason that a New Keynesian model could generate countercyclical markup is the following. With price stickiness, a procyclical marginal cost

[14]The firm's optimal choice of labor L_{it}^* is always above the overhead labor $\bar{L}$ in the simulation.

implies that in a boom, the gap between the price and the marginal cost shrinks, and hence decrease in the markup. In particular, I use a partial equilibrium model with Rotemberg price adjustment costs.

1.5.1 Rotemberg Adjustment Cost for Price

Assume that each firm produces output Y_{it} with a technology in labor L_{it}:

$$Y_{it} = L_{it}^{\alpha}. \tag{1.25}$$

Assume that the firm faces the following demand:

$$C_{it} = \left(\frac{P_{it}}{P_t}\right)^{-\theta} C_t, \tag{1.26}$$

where C_{it} denotes firm i's quantity of demand at time t, P_{it} denotes the nominal price of firm i, P_t denotes the economy's aggregate price deflator, and C_t denotes the aggregate consumption.

I assume that the logarithm of the aggregate price deflator follows a random walk with a trend:

$$\log P_t = \mu_p + \log P_{t-1} + \varepsilon_t^p, \tag{1.27}$$

where $\varepsilon_t^p \sim N(0, \sigma_p)$. Furthermore, I assume that the aggregate output Y_t is determined by the following New Keynesian Philips Curve:

$$\pi_t = \beta \mathrm{E}_t(\pi_{t+1}) + \kappa(\log Y_t - \log \bar{Y}), \tag{1.28}$$

where $\pi_t \equiv \log(P_t/P_{t-1})$ is the economy's inflation rate at time t, $\bar{Y}$ is the steady state aggregate output level, β is a constant discount rate, and κ is a reduced form parameter. Combined with equation (1.27), one can rewrite the equation above to get the following

expression for aggregate output:

$$\log Y_t = \log \bar{Y} + \frac{(1-\beta)\mu_p}{\kappa} + \frac{\varepsilon_t^p}{\kappa}. \tag{1.29}$$

I assume that the real wage is a power function of aggregate consumption:

$$\frac{W_t}{P_t} = \omega C_t^{\sigma}, \tag{1.30}$$

where ω measures the marginal disutility from labor, and σ is relative risk aversion parameter.

To adjust its prices relative to the constant inflation rate, the firm has to pay adjustment costs in terms of aggregate output:

$$\frac{\gamma_p}{2}\left(\frac{P_{it}}{(1+\mu_p)P_{it-1}} - 1\right)^2 Y_t. \tag{1.31}$$

In real terms, the firm's real flow profit Π_{it} is

$$\Pi_{it} = \frac{P_{it}}{P_t}Y_{it} - \frac{W_t}{P_t}L_{it} - \frac{\gamma_p}{2}\left(\frac{P_{it}}{(1+\mu_p)P_{it-1}} - 1\right)^2 Y_t. \tag{1.32}$$

Assume that the aggregate output is equal to aggregate consumption in this economy, $Y_t = C_t$ [15]. Combined with equations above, one can rewrite the real profit as the following:

$$\Pi_{it} = \left(\frac{P_{it}}{P_t} - \omega Y_t^{\sigma+\frac{1-\alpha}{\alpha}}\left(\frac{P_{it}}{P_t}\right)^{-\theta\frac{1-\alpha}{\alpha}}\right)\left(\frac{P_{it}}{P_t}\right)^{-\theta} Y_t - \frac{\gamma_p}{2}\left(\frac{1}{1+\mu_p}\frac{P_{it}}{P_t}\left(\frac{P_{it-1}}{P_{it}}\right)^{-1} - 1\right)^2 Y_t. \tag{1.33}$$

The firm discounts the future real profit with the following discount factor

$$q_{t,t+1} \equiv \beta\left(\frac{Y_{t+1}}{Y_t}\right)^{-\sigma}. \tag{1.34}$$

[15] Assume that the adjustment cost is small relative to the aggregate output.

Then the firm's value function can be written as the following:

$$V(P_i/P, \varepsilon^p) = \max_{P_i'} \Pi + \mathrm{E}[qV(P_i'/P', \varepsilon^{p'})]. \tag{1.35}$$

There are two state variable for the firm. One is relative price of last period to current aggregate price level, P_{it-1}/P_t. The other is current period's inflation shock ε_t^p, which is sufficient to pin down the current period's aggregate output level Y_t. The firm chooses a new price P_{it} to maximize its sum of discounted profits.

One period is equal to one quarter. I assume that the time discount factor is $\beta = 0.99$, which implies an annual rate of 0.96. The constant relative risk aversion (CRRA) parameter is $\sigma = 1$. The elasticity of demand is $\theta = 6$ to match the steady state level of markup 1.2. I choose the marginal disutility of labor ω such that the steady state aggregate employment is equal to 1/3, and so is the steady state aggregate output $\bar{Y} = 1/3$. The trend of the inflation rate is $\mu_p = 0.0021 \times 3$, and the volatility is $\sigma_p = 0.0032 \times \sqrt{3}$. For the parameter governing the price adjustment cost, I set it such that the New Keynesian Philips Curve are identical between the Rotemberg and the Calvo models up to first order approximation:

$$\frac{\theta - 1}{\gamma_p} = \frac{(1-\lambda)(1-\beta\lambda)}{\lambda}, \tag{1.36}$$

where λ is the probability that the firm keeps its current price. I choose $\lambda = 2/3$, which implies an average price duration of 3 quarters. Finally, I choose the NKPC reduced form parameter κ as follows:

$$\kappa = \frac{(1-\lambda)(1-\beta\lambda)}{\lambda} \frac{\alpha\sigma + 1 - \alpha}{\alpha + (1-\alpha)\theta}, \tag{1.37}$$

which is derived from the first order approximation of the Calvo model.

Measurement of Markup

The first-order condition with respect to price yields the following equation:

$$\begin{aligned}(\theta-1)\frac{P_{it}}{P_t} =& \frac{\theta}{\alpha}\frac{W_t}{P_t}Y_{it}^{\frac{1}{\alpha}-1} - \frac{\gamma_p}{1+\mu_p}\left(\frac{P_{it}}{(1+\mu_p)P_{it-1}}-1\right)\frac{P_{it}}{P_{it-1}Y_{it}}Y_t \\ &+ \mathrm{E}_t\left[q_{t,t+1}\frac{\gamma_p}{1+\mu_p}\left(\frac{P_{it+1}}{(1+\mu_p)P_{it}}-1\right)\frac{P_{it+1}}{P_{it}Y_{it}}Y_{t+1}\right]. \end{aligned} \tag{1.38}$$

Then the markup μ_{it}, which is defined as the price P_{it} over the marginal cost of production $\frac{1}{\alpha}W_tY_{it}^{\frac{1}{\alpha}-1}$, is the following:

$$\begin{aligned}\mu_{it} =& \frac{\theta}{\theta-1}\left\{1-\frac{\alpha\gamma_p}{\theta}\frac{P_{it}}{(1+\mu_p)P_{it-1}}\left(\frac{P_{it}}{(1+\mu_p)P_{it-1}}-1\right)\left(\frac{W_tL_{it}}{P_tY_t}\right)^{-1}\right. \qquad (1.39)\\ &\left.+\ \mathrm{E}_t\left[q_{t,t+1}\frac{\alpha\gamma_p}{\theta}\frac{P_{it+1}}{(1+\mu_p)P_{it}}\left(\frac{P_{it+1}}{(1+\mu_p)P_{it}}-1\right)\left(\frac{W_tL_{it}}{P_tY_{t+1}}\right)^{-1}\right]\right\}. \qquad (1.40)\end{aligned}$$

If there is no adjustment cost of price changes $\gamma_p = 0$, the firm always charges the desired markup $\frac{\theta}{\theta-1}$.

In the simulation, I estimate the markup cyclicality by running the following regression:

$$\log\mu_{it} = \phi_0 + \phi_1\log Y_t + \epsilon_t. \tag{1.41}$$

I run regressions for different values of α and σ. The reason that α and σ would entail different elasticities of markup is the following. Assume that the firm's relative price is held constant at $\bar{P}_i/P$, then the markup μ_{it} becomes

$$\begin{aligned}\mu_{it} =& \frac{\bar{P}_i/P}{\frac{1}{\alpha}\frac{W_t}{P}Y_{it}^{\frac{1}{\alpha}-1}} \\ =& \frac{\bar{P}_i/P}{\frac{1}{\alpha}\omega Y_t^{\sigma}Y_{it}^{\frac{1}{\alpha}-1}} \\ =& \frac{(\bar{P}_i/P)^{1-\theta(1/\alpha-1)}}{\frac{1}{\alpha}\omega Y_t^{\sigma+\frac{1}{\alpha}-1}},\end{aligned} \tag{1.42}$$

where the second line uses the real wage equation, and the third line uses the demand equation. Therefore, the elasticity of markup with respect to aggregate output is

$$\phi_1 \equiv \frac{\partial \log \mu_{it}}{\partial \log Y_t} = -\sigma - \frac{1}{\alpha} - 1. \tag{1.43}$$

From the equation above, one can see that as the marginal cost curve becomes steeper (α becomes smaller), the markup becomes more counter-cyclical. Also, as real marginal cost becomes more pro-cyclical (σ becomes larger), the markup becomes more counter-cyclical.

The simulation results are reported in table 1.8. With $\sigma = 1$, as α decreases, the markup becomes more counter-cyclical. However, one can see that the change is small, as the elasticity ranges from -0.89 to -1.08. In the next panel, with α held constant at 0.8, the counter-cyclicality increases with σ. The change is also greater than with the case of α, as the elasticity ranges from -0.20 to -1.75. Overall, the elasticity is very close to the empirical estimate -1.1.

1.6 Conclusion

This chapter documented the firm-level markup behavior over the business cycle. The next step is to investigate which model can successfully replicate the two empirical findings that not only firms have countercyclical markups, but also small firms' markups fluctuate more.

In chapter two, I test the markup behavior of the following two prominent models in the literature: (i) an oligopolistic competition model, and (ii) a New Keynesian model with heterogeneous price stickiness. First, I explore the Atkeson and Burstein (2008) model of oligopolistic competition, in which markups are an increasing function of firm market shares. Coupled with an exogenous *uncertainty* shock as in Bloom (2009), i.e. a second-moment shock to firm productivities in recessions, this model results in a countercyclical average markup, as in the data. However, in contrast with the data, this model predicts that smaller firms reduce their markups. Second, I calibrate both Calvo and menu cost models of price

stickiness to match the empirical heterogeneity in price durations across small and large firms, as in Goldberg and Hellerstein (2011). I find that both models can match the average counter-cyclicality of markups in response to monetary shocks. Furthermore, since small firms adjust prices less frequently, they exhibit greater markup counter-cyclicality, consistent with the empirical patterns. Quantitatively, however, only the menu cost model, through endogenous selection of firms adjusting prices, matches the extent of the heterogeneity in markup cyclicality in the data. Also, the markup cyclicality depends on the source of shock. In response to productivity shocks, both sticky price models imply pro-cyclical markup behavior.

Therefore, in third chapter, I propose a new general equilibrium model that features firm's endogenous accumulation process of customer capital, and entry and exit decision. In the model, firms' markup decisions become dynamics, and the decisions of firms closer to the extensive margin of exit are rather different from others. I describe the main intuition and mechanism of the model below.

1.7 Tables

Table 1.1: Heterogeneity in Markup Cyclicality ϕ: Cobb-Douglas

Dep. Variable	(1)	(2)	(3)	(4)	(5)
$\log \mu_{it}$	All	Large	Small	All	All
$\log Y_t$	-1.13***	-0.76***	-1.21***	-1.21***	-1.22***
	(0.17)	(0.21)	(0.16)	(0.16)	(0.18)
$\log Y_t \times \text{Large}_i$				0.45***	
				(0.12)	
$\log Y_t \times s_{it}$					6.13***
					(1.79)
N	96507	16048	80459	96507	96507
adj. R^2	0.89	0.87	0.89	0.89	0.89

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: $\log Y_t$ is log aggregate output quadratically detrended. $Large_i$ is an indicator for a firm with more than 1% market share within a 4-digit industry. s_{it} stands for market share in a 4-digit industry. Standard errors are clustered at time level.

Table 1.2: Markup Cyclicality (Material Costs): Cobb-Douglas

Dep. Variable	(1)	(2)	(3)	(4)	(5)
$\Delta \log \mu_{it}$	All	Large	Small	All	All
$\Delta \log Y_t$	-1.19***	-1.02***	-1.23***	-1.23***	-1.24***
	(0.25)	(0.22)	(0.25)	(0.25)	(0.26)
$\Delta \log Y_t \times \text{Large}_i$				0.21*	
				(0.11)	
$\Delta \log Y_t \times s_{it}$					3.98***
					(1.08)
N	66696	12452	54244	66696	66696
adj. R^2	0.01	0.01	0.01	0.01	0.01

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: $\log Y_t$ is log aggregate output. $Large_i$ is an indicator for a firm with more than 1% market share within a 4-digit industry. s_{it} stands for market share in a 4-digit industry. Standard errors are clustered at time level.

Table 1.3: Heterogeneity in Markup Cyclicality ϕ: Translog

Dep. Variable	(1)	(2)	(3)	(4)	(5)
$\log \mu_{it}$	All	Large	Small	All	All
$\log Y_t$	-0.69***	-0.32*	-0.80***	-0.80***	-0.79***
	(0.15)	(0.19)	(0.17)	(0.17)	(0.16)
$\log Y_t \times \text{Large}_i$				0.47**	
				(0.22)	
$\log Y_t \times s_{it}$					6.09***
					(1.63)
N	60701	13352	47349	60701	60701
adj. R^2	0.86	0.86	0.86	0.86	0.86

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: $\log Y_t$ is log aggregate output quadratically detrended. $Large_i$ is an indicator for a firm with more than 1% market share within a 4-digit industry. s_{it} stands for market share in a 4-digit industry. Standard errors are clustered at time level.

Table 1.4: Heterogeneity in Markup Cyclicality ϕ: Cobb-Douglas

Dep. Variable $\Delta\mu_{it}$	(1) France	(2) Germany	(3) Italy	(4) Spain
$\log Y_t$	-1.21***	-1.34***	-1.88***	-0.49
	(0.16)	(0.11)	(0.39)	(0.38)
$\log Y_t \times \text{Large}_i$	0.45***	0.45***	0.54*	0.351***
	(0.12)	(0.09)	(0.28)	(0.05)
N	96507	62844	409837	269211
adj. R^2	0.89	0.83	0.86	0.77

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: $\log Y_t$ is log aggregate output. $Large_i$ is an indicator for a firm with more than 1% market share within a 4-digit industry. Standard errors are clustered at time level.

Table 1.5: Age and Financial Variables

Dep. Variable $\log \mu_{it}$	(1)	(2)	(3)	(4)	(5)	(6)
$\log Y_t$	-1.21***	-1.23***	-0.85***	-1.21***	-0.82***	-0.87***
	(0.18)	(0.43)	(0.18)	(0.16)	(0.19)	(0.21)
$\log Y_t \times s_{it}$	6.13***	6.12***	5.79***	5.99***	5.79***	5.96***
	(1.79)	(1.62)	(1.50)	(1.13)	(1.51)	(1.61)
$\log Y_t \times \text{Age}_{it}$		0.01				
		(0.11)				
$\log Y_t \times \text{LIQD}_{it}$			-0.28			
			(0.20)			
$\log Y_t \times \text{CASH}_{it}$				0.33		
				(4.26)		
$\log Y_t \times \text{CURR}_{it}$					-0.22	
					(0.14)	
$\log Y_t \times \text{SOLV}_{it}$						-0.78
						(0.69)
N	96507	96507	94727	94625	94637	94428
adj. R^2	0.89	0.89	0.90	0.90	0.90	0.90

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: $\log Y_t$ is log aggregate output quadratically detrended. s_{it} stands for market share in a 4-digit industry. Age_{it} stands for firm's age. $LIQD_{it}$ for liquidity ratio, $CASH_{it}$ stands for cash ratio, $CURR_{it}$ for current ratio, and $SOLV_{it}$ for solvency ratio. Standard errors are clustered at time level.

Table 1.6: Markup Cyclicality (Labor Costs): Cobb-Douglas

Dep. Variable	(1)	(2)	(3)	(4)	(5)
$\log \mu_{it}^L$	All	Large	Small	All	All
$\log Y_t$	2.06***	2.33***	2.00***	2.00***	2.02***
	(0.31)	(0.43)	(0.28)	(0.28)	(0.30)
$\log Y_t \times \text{Large}_i$				0.33*	
				(0.17)	
$\log Y_t \times s_{it}$					3.10*
					(1.71)
N	96507	16048	80459	96507	96507
adj. R^2	0.92	0.94	0.90	0.92	0.92

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: $\log \mu_{it}^L$ is log markup measure based on labor share of revenue. $\log Y_t$ is log aggregate output quadratically detrended. $Large_i$ is an indicator for a firm with more than 1% market share within a 4-digit industry. s_{it} stands for market share in a 4-digit industry. Standard errors are clustered at time level.

Table 1.7: Markup Cyclicality: Labor Adjustment Cost and Overhead Labor

	$\gamma_l = 0.5$	$\gamma_l = 2.5$	$\gamma_l = 9$
ϕ_1	-0.96	-0.92	-0.88
	$L^o/\bar{L} = 0.3$	$L^o/\bar{L} = 0.5$	$L^o/\bar{L} = 0.7$
ϕ_1	-0.71	-0.45	-0.08

Table 1.8: Markup Cyclicality: Rotemberg

$\sigma = 1$	$\alpha = 1$	$\alpha = 0.8$	$\alpha = 0.6$
ϕ_1	-0.89	-0.97	-1.08
$\alpha = 0.8$	$\sigma = 0$	$\sigma = 1$	$\sigma = 2$
ϕ_1	-0.20	-0.97	-1.75

Chapter 2

A Tale of Two Models

2.1 Introduction

There is a long line of empirical studies regarding price markup fluctuations over business cycles. One could view the markup as the ratio of price over marginal cost, and it measures the distortion in the output market. A countercyclical markup contributes to the amplification of aggregate fluctuations. However, the determinants of markup movements are not well understood. Many models in the business cycle literature generate cyclical price markups. One approach is to assume that firm's markup follows an exogenous process and varies over time in the literature (as in Smets and Wouters 2003, 2007, and Steinsson 2003). In contrast, other models rationalize the variable markups with micro-founded models.

However, which model is the right one to consider? The aim of this chapter is to choose one that is consistent with the empirical findings at the micro level. Previous chapter finds that markups are countercyclical on average, and small firms' markups are more countercyclical than large firms'. In particular, I test the markup behavior of the following two models: (i) an oligopolistic competition model, and (ii) a New Keynesian model with heterogeneous price stickiness. First, I study the competition model in a general setting. I find that one needs varying second moment shock in firm's idiosyncratic productivities over time to

generate variable markups (as in Bloom 2009). In a special case with Atkeson and Burstein (2008), firm's pricing function is increasing and convex in its own market share. And due to Jensen's inequality, the changes in dispersion of market shares generates countercyclical markup at the aggregate level. However, since convexity is stronger for large firms than small firms, large firms' markups tend to be more countercyclical, which is not consistent with the data. Second, I calibrate the New Keynesian model with heterogeneous adjustment costs. With sticky price, firms adjust price more slowly compared to changes in marginal cost. Hence, with a procyclical marginal cost, markup is countercyclical. A recent empirical study by Goldberg and Hellerstein (2011) finds that small firms adjust less frequently as large firms. I calibrate both Calvo and menu cost models of price stickiness to match the empirical heterogeneity in price durations across small and large firms as in their study. The model is subject to nominal aggregate demand shock. I find that the model could successfully generate both countercyclical markup and that small firms' markup more countercyclical than large firms'. Quantitatively, however, only the menu cost model, through its selection effect, can match the extent of the empirical heterogeneity in markup cyclicality. In addition, both sticky price models imply pro-cyclical markup behavior in response to productivity shocks, since marginal costs become countercyclical.

The rest of the chapter proceeds as follows: section 2 derives theoretical results for oligopolistic competition in a general setup. Section 3 discusses the quantitative analysis of the oligopolistic competition in a specific setup, namely the Atkeson-Burstein (2008) model. Section 4 introduces and discusses the results of a New Keynesian model with new extensions. Section 5 concludes.

2.2 General Oligopolistic Competition Model

To think about markup cyclicality along business cycle, a natural first step is oligopolistic competition model. I start with a general imperfect competition framework as described

in Burstein and Gopinath (2013) to study how nature of firm competition and underlying marginal cost process affect cyclicality of markup.

2.2.1 General Framework

Consider an economy consisting of n firms each indexed by $i = \{1, ..., n\}$. Each firm has a constant-returns-to-scale production technology. Firm i's optimal pricing rule is markup over marginal cost

$$p_i = \mu_i + mc_i, \tag{2.1}$$

where $p_i \equiv \log P_i$ is the log price of firm, $\mu_i \equiv \log \mathcal{M}_i$ is the log markup of firm, and $mc_i \equiv \log MC_i$ is the log marginal cost of firm. Markup depends on both firm's log price p_i, and log industry price index $p \equiv \log P$. In particular, log of markup takes the form of $\mu_i = \mu(p_i - p)$. Many models generate this relationship between markup and relative price. The functional form of markup $\mu(\cdot)$ and industry price index p depend on the model.

Firms compete in product market and interact with each other through industry price index p. We can see this competition framework as a special case of interaction networks as in Acemoglu, Ozdaglar, and Tahbaz-Salehi (2015). To study cyclicality of markup, I focus on small marginal cost shocks to firms so I can use first several terms of Taylor expansions around initial states. In particular, variable of main interest is first difference in log industry markup defined as

$$\Delta\mu(\mu_1, ..., \mu_n) = \sum_{i=1}^{n} S_i \Delta\mu_i, \tag{2.2}$$

where S_i is firm i's market share in revenue. Intuitively, log industry markup change is revenue-weighted average of individual markup change. We will see that this definition is consistent with welfare-relevant measure in the following subsection, in which I introduce a specific imperfect competition model for calibration.

First Order Approximation

First, I start with a first-order Taylor approximation of change in individual markup with respect to change in marginal costs

$$\Delta\mu_i = -\Gamma_i \left[\sum_{k=1}^{n} \frac{\partial(p_i - p)}{\partial mc_k} \Delta mc_k \right], \tag{2.3}$$

where $\Gamma_i \equiv -\frac{\partial \mu_i}{\partial(p_i - p)}$ is the elasticity of markup with respect to the relative price. If desired markup is decreasing in relative price, $\Gamma_i > 0$. Also, Γ_i measures strength of strategic complementarities in pricing. To see this, take a first-order approximation of equation (2.1):

$$\Delta p_i = -\Gamma_i(\Delta p_i - \Delta p) + \Delta mc_i,$$

which leads to

$$\Delta p_i = \frac{\Gamma_i}{1+\Gamma_i}\Delta p + \frac{1}{1+\Gamma_i}\Delta mc_i.$$

Hence, price of a firm with higher Γ_i responds more to industry price index than its own marginal cost shock, and *vice versa*. Note that two coefficients sum to one.

I use a first order approximation for change in industry price index:

$$\Delta p = \sum_{j=1}^{n} S_j \Delta p_j, \tag{2.4}$$

which is revenue-weighted average of individual price change. This equation holds exactly in many models, including the one I use in the following subsection. Combine equation (2.4) with partial differentiation of equation (2.1), we get the following equation for each i

$$\frac{\partial p_i}{\partial mc_k} = -\Gamma_i \left(\frac{\partial p_i}{\partial mc_k} - \sum_{j=1}^{n} S_j \frac{\partial p_j}{\partial mc_k} \right) + \mathbf{1}\{i = k\},$$

where $\mathbf{1}\{i = k\}$ is an indicator function whether $i = k$. And it is straightforward to show that:

$$\frac{\partial p_i}{\partial mc_k} = \frac{\Gamma_i}{1+\Gamma_i}\left(\sum_{j=1}^{n}\frac{S_j}{1+\Gamma_j}\right)^{-1}\left(\frac{S_k}{1+\Gamma_k}\right) + \frac{\mathbf{1}\{i=k\}}{1+\Gamma_i}. \tag{2.5}$$

Thus how much marginal cost shock to firm k impacts firm i depends on either if firm i has strong strategic complementarities in pricing ($\frac{\Gamma_i}{1+\Gamma_i}$), or if firm k is relatively important in the industry ($\frac{S_k}{1+\Gamma_k}$). Additionally, if the marginal cost shock hits firm i itself, it responds through its own marginal cost channel.

Putting definition of the industry markup change (2.2), and equation (2.3) & (2.5) leads to the following linear approximation of industry markup change as a function of underlying marginal cost change:

Theorem 1 *The first-order approximation to the industry markup change is given by*

$$\Delta\mu^{(1)} = -\left(\sum_{j=1}^{n}\frac{S_j}{1+\Gamma_j}\right)^{-1} Cov_S\left(\frac{\Gamma_i}{\Gamma_i+1}, \Delta mc_i\right).^{1} \tag{2.6}$$

This result shows that industry markup change is proportional to negative covariance between strategic complementarities and marginal cost shock. Hence if firms with stronger complementarities are hit with greater marginal cost shock, industry markup decreases. Also, this result implies that if all firms are hit with identical shock ($\Delta mc_i = \Delta mc$, $\forall i$), industry markup stays the same. This is easy to understand since each firm's desired markup depends on relative price difference, hence to lead to aggregate effect, we need some heterogeneities in marginal cost shocks.

However, even if marginal cost shocks are independently and identically distributed, and have mean zero and variance σ^2, we have the following corollary for the expectation of industry markup change:

[1] I define $\mathrm{Cov}_S(X_i, Y_i)$ as the weighted covariance $\mathrm{Cov}_S(X_i, Y_i) \equiv \sum_i S_i X_i Y_i - (\sum S_i X_i)(\sum S_i Y_i)$, where weights sum to 1: $\sum S_i = 1$.

Corollary 1 $E[\Delta\mu^{(1)}] = 0$.

This corollary shows that first-order expansion is not informative about interaction between the competition network and the underlying marginal cost process. Therefore, it is natural to use second-order expansion in the following.

Second Order Approximation

I start with a second-order approximation for individual markup change:

$$\Delta\mu_i = -\Gamma_i \left[\sum_{k=1}^{n} \frac{\partial(p_i - p)}{\partial mc_k} \Delta mc_k\right] + \frac{1}{2}\sum_{k=1}^{n}\sum_{r=1}^{n} \frac{\partial^2 p_i}{\partial mc_k \partial mc_r} \Delta mc_k \Delta mc_r, \tag{2.7}$$

where first term is the same as first-order approximation, and second term comes from the fact that

$$\frac{\partial^2 \mu_i}{\partial mc_k \partial mc_r} = \frac{\partial^2 p_i}{\partial mc_k \partial mc_r}.$$

To derive this Hessian matrix for prices, I take second partial derivative of equation (2.1) to get

$$\begin{aligned}\frac{\partial^2 p_i}{\partial mc_k \partial mc_r} = &-\Gamma_i \left(\frac{\partial^2 p_i}{\partial mc_k \partial mc_r} - \sum_{j=1}^{n} S_j \frac{\partial^2 p_j}{\partial mc_k \partial mc_r} - \sum_{j=1}^{n}\sum_{j'=1}^{n} \frac{\partial^2 p}{\partial p_j \partial p_{j'}} \left(\frac{\partial p_{j'}}{\partial mc_k}\right)\left(\frac{\partial p_j}{\partial mc_r}\right)\right) \\ &+ \Gamma_{ii} \left(\frac{\partial(p_i - p)}{\partial mc_k}\right)\left(\frac{\partial(p_i - p)}{\partial mc_r}\right).\end{aligned} \tag{2.8}$$

where $\Gamma_{ii} \equiv -\frac{\partial \Gamma_i}{\partial p_i - p}$ is superelasticity of markup, which captures convexity (or concavity) of markup. If $\Gamma_{ii} > 0$, firms with lower relative price have more strength of strategic complementarities, and *vice versa*. Furthermore, I show the following result (see the Appendix for proof):

Proposition 1 *If market share S_j is a function of relative price $S_j = S\left(\frac{P_j}{P}\right)$, then the elasticity of market share with respect to relative price $-\frac{\partial \log S_j}{\partial(p_j - p)}$ is a constant for all j. And*

the Hessian matrix for industry price equals:

$$\frac{\partial^2 p}{\partial p_j \partial p_j'} = \Lambda S_j (S_{j'} - \mathbf{1}\{j = j'\}), \tag{2.9}$$

where Λ *denotes the market share elasticity* $-\frac{\partial \log S_j}{\partial (p_j - p)}$.

This proposition leads to simplification of equation (2.8) (see Appendix for derivation):

$$\frac{\partial^2 p_i}{\partial mc_k \partial mc_r} = \frac{\Gamma_i}{1 + \Gamma_i} \left(X_i^{kr} + \sum_{j=1}^{n} S_j \frac{\partial^2 p_j}{\partial mc_k \partial mc_r} \right), \tag{2.10}$$

where

$$X_i^{kr} \equiv \frac{\Gamma_{ii}}{\Gamma_i} \left(\frac{\partial (p_i - p)}{\partial mc_k} \right) \left(\frac{\partial (p_i - p)}{\partial mc_r} \right) - \Lambda \sum_j S_j \left(\frac{\partial (p_j - p)}{\partial mc_k} \right) \left(\frac{\partial (p_j - p)}{\partial mc_r} \right).$$

Note that $\frac{\Gamma_{ii}}{\Gamma_i} = -\frac{\mu_i''}{\mu_i'}$ measures the convexity of markup. Combining equation (2.10) with equation (2.7) leads to the following result:

Theorem 2 *The second-order approximation to the total markup change is given by*

$$\Delta\mu^{(2)} = \Delta\mu^{(1)} + \frac{1}{2} \sum_{k=1}^{n} \sum_{r=1}^{n} \left\{ \left(\sum_j \frac{S_j}{1 + \Gamma_j} \right)^{-1} \left(\sum_j S_j \frac{\Gamma_j}{1 + \Gamma_j} X_j^{kr} \right) \Delta mc_k \Delta mc_r \right\}, \tag{2.11}$$

where $\Delta\mu^{(1)}$ *is first-order approximation as in Theorem 1.*

To understand the intuition of this result, I take the expectation, and assume that all firms' initial states are the same to get the following (see Appendix for proof):

Corollary 2 *If all firms have the same initial states such that* $S_j = \frac{1}{n}$, $\Gamma_j = \Gamma'$, $\Gamma_{jj} = \Gamma''$, *then*

$$E[\Delta\mu^{(2)}] = \frac{1}{2} \sigma^2 \frac{n-1}{n} \frac{\Gamma'}{(1 + \Gamma')^2} \left(\frac{\Gamma''}{\Gamma'} - \Lambda \right). \tag{2.12}$$

This result implies that if the convexity of markup $\frac{\Gamma''}{\Gamma'}$ is greater than the elasticity of market share Λ, change in industry markup is an increasing function of variance σ^2.

2.3 Quantitative Analysis: Atkeson-Burstein

In this section, I use the oligopolistic competition framework introduced by Atkeson and Burstein (2008) for quantitative simulation.

Household

The representative household has an additively separable preference over consumption and labor

$$U(C, L) = \frac{C^{1-\sigma}}{1-\sigma} - \omega \frac{L^{1+\psi}}{1+\psi}, \tag{2.13}$$

where $\frac{1}{\sigma}$ is the intertemporal elasticity of substitution (IES), ω is the disutility parameter from labor, and $\frac{1}{\psi}$ is the Frisch elasticity of labor supply. Total consumption C consists of consumption from a continuum of sectors j:

$$C = \left(\int_0^1 C_j^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}}, \tag{2.14}$$

where C_j is consumption for sector j's good, and η is the elasticity of substitution between any two different sectoral goods. Within each sector j, there are n_j firms producing differentiated goods. The household has a CES type preference over finite number of differentiated goods for each sector j:

$$C_j = \left(\sum_{i=1}^{n_j} C_{ij}^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}, \tag{2.15}$$

where C_{ij} is consumption of good i in sector j, and ρ is the elasticity of substitution between any two differentiated goods within sector. It is assumed that the elasticity of substitution within sector is higher than the elasticity of substitution across sector, $\rho > \eta$.

The household chooses consumption $\{C_{ij}\}$ and labor L to maximize the utility function (2.13) subject to the following budget constraint

$$\int_0^1 \left(\sum_{i=1}^{n_j} P_{ij} C_{ij} \right) dj \leq WL, \tag{2.16}$$

where P_{ij} is the price of good i in sector j, and W is the nominal wage. The solution to the household's problem gives the demand function for C_{ij}:

$$C_{ij} = \left(\frac{P_{ij}}{P_j} \right)^{-\rho} \left(\frac{P_j}{P} \right)^{-\eta} C, \tag{2.17}$$

where P_j is sector j's price index defined as

$$P_j \equiv \left(\sum_{i=1}^{n_j} P_{ij}^{1-\rho} \right)^{\frac{1}{1-\rho}}, \tag{2.18}$$

and P is total economy price index defined as

$$P \equiv \left(\int_0^1 P_j^{1-\eta} \right)^{\frac{1}{1-\eta}}. \tag{2.19}$$

And the consumption and labor optimality condition is the following

$$\omega \frac{L^{\psi}}{C^{-\sigma}} = \frac{W}{P}. \tag{2.20}$$

Firm

Firm i in sector j produces output using labor

$$Y_{ijt} = a_{ijt} l_{ijt}, \tag{2.21}$$

where a_{ijt} is producer-level productivity and I discuss its composition and evolution in the next sub subsection. Firms engage in Cournot competition within sector.[2] Taking wage W and demand equation (3.4) as given, a firm i in sector j chooses its output Y_{ijt} to maximize its profit

$$\pi_{ijt} = \max_{Y_{ijt}} \left[\left(P_{ijt} - \frac{W}{a_{ijt}} \right) Y_{ijt} - W\phi \right] \mathbf{1}\{Y_{ijt} > 0\}, \tag{2.22}$$

where ϕ is fixed cost of production and is denominated in units of labor. A firm can choose not to produce to avoid paying the fixed cost ϕ. Hence ϕ captures the extensive margin of the oligopolistic competition.

The solution to the firm's profit maximization problem is a markup over marginal cost

$$P_{ijt} = \frac{\varepsilon(S_{ijt})}{\varepsilon(S_{ijt}) - 1} \frac{W}{a_{ijt}}, \tag{2.23}$$

where firm-specific demand elasticity $\varepsilon(S_{ijt})$ is a harmonic weighted average of elasticities of substitution ρ and η

$$\varepsilon(S_{ijt}) = \left(S_{ijt} \frac{1}{\eta} + (1 - S_{ijt}) \frac{1}{\rho} \right)^{-1}, \tag{2.24}$$

where S_{ijt} is firm's market share in sector j,

$$S_{ijt} = \frac{P_{ijt} Y_{ijt}}{\sum_{i=1}^{n_j} P_{ijt} Y_{ijt}} = \left(\frac{P_{ijt}}{P_{jt}} \right)^{1-\rho}. \tag{2.25}$$

Since there are finite number of firms in each sector, the firms are large enough ($S_{ijt} > 0$) to affect industry price index P_{jt}.

Also, firm's markup $\mathcal{M}_{ijt}$ can be expressed as

$$\frac{1}{\mathcal{M}_{ijt}} = \frac{\rho - 1}{\rho} - \left(\frac{1}{\eta} - \frac{1}{\rho} \right) S_{ijt}, \tag{2.26}$$

[2]Bertrand competition generates qualitatively the same results.

and the elasticity of markup with respect to relative price are:

$$\Gamma_i = -\frac{\partial \log \mathcal{M}_{ijt}}{\partial(\log P_{ijt} - \log P_{jt})} = (\rho - 1)\left(\frac{1}{\eta} - \frac{1}{\rho}\right) S_{ijt}\mathcal{M}_{ijt}. \tag{2.27}$$

Since $\rho > \eta$, markup is an increasing and convex function of market share. Respectively, the elasticity and super-elasticity of markup with respect to relative price are:

$$\Gamma_i = -\frac{\partial \log \mathcal{M}_{ijt}}{\partial(\log P_{ijt} - \log P_{jt})} = (\rho - 1)\left(\frac{1}{\eta} - \frac{1}{\rho}\right) S_{ijt}\mathcal{M}_{ijt} \tag{2.28}$$

$$\Gamma_{ii} = -\frac{\partial \Gamma_i}{\partial(\log P_{ijt} - \log P_{jt})} = \Gamma_i(\rho - 1 + \Gamma_i). \tag{2.29}$$

The market share elasticity with respect to relative price is:

$$\Lambda = -\frac{\partial \log S_{ijt}}{\partial(\log P_{ijt} - \log P_{jt})} = \rho - 1. \tag{2.30}$$

Hence $\frac{\Gamma_{ii}}{\Gamma_i} - \Lambda = \Gamma_i > 0$, and according to Corollary 2, change of industry markup is an increasing function of marginal cost shock variance in expectation.

Market Clearing

Denote L_t^* the optimal labor supply by the representative household, and l_{ijt}^* the labor demand of firm i in sector j. The labor market clearing condition is then

$$\int_0^1 \left(\sum_{i=1}^{n_j}(l_{ijt}^*) + \phi\right) = L_t^*. \tag{2.31}$$

And the good market clearing condition is

$$C_{ijt} = Y_{ijt} \quad \forall i, j, t \tag{2.32}$$

Aggregate Productivity and Markup

Define aggregate productivity as the following:

$$A_t \equiv \frac{Y_t}{\tilde{L}_t^*}, \tag{2.33}$$

where Y_t is the quantity of final output, and $\tilde{L}_t^*$ is the aggregate labor supply net of production fixed costs. From the labor market clearing condition (2.31), the aggregate productivity can be expressed as the quantity weighted harmonic average of individual productivity:

$$A_t = \left[\int_0^1 \left(\sum_{i=1}^{n_j} \frac{Y_{ijt}}{Y_t} \frac{1}{a_{ijt}} \right) dj \right]^{-1} \tag{2.34}$$

Define aggregate markup as the following:

$$\mathcal{M}_t \equiv P_t \left(\frac{W_t}{A_t} \right)^{-1}, \tag{2.35}$$

where P_t is the aggregate price index as defined in (2.19), and $\frac{W_t}{A_t}$ is the aggregate marginal cost. From equation (2.34), it is easy to see that the aggregate markup can be expressed as the market share weighted harmonic average of individual markup:

$$\mathcal{M}_t = \left[\int_0^1 S_{jt} \left(\sum_{i=1}^{n_j} \frac{S_{ijt}}{\mathcal{M}_{ijt}} \right) dj \right]^{-1}, \tag{2.36}$$

where $S_{jt} \equiv \frac{P_{jt} Y_{jt}}{P_t Y_t}$ is sector j's total revenue share of the economy.

Note that the aggregate productivity can be rewritten as

$$A_t = \left[\int_0^1 \left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t} \right)^{-\eta} a_{jt}^{\eta-1} \right]^{\frac{1}{\eta-1}}, \tag{2.37}$$

where $\mathcal{M}_{jt} \equiv P_{jt}\left(\frac{W_t}{a_{jt}}\right)^{-1}$ is the sectoral markup and a_{jt} is the sectoral productivity defined as

$$a_{jt} \equiv \left[\sum_{i=1}^{n_j}\left(\frac{\mathcal{M}_{ijt}}{\mathcal{M}_{jt}}\right)^{-\rho} a_{ijt}^{\rho-1}\right]^{\frac{1}{\rho-1}}. \tag{2.38}$$

We can compare this to the first best (FB) aggregate productivity attained by a social planner:

$$A_t^{\mathrm{FB}} = \left(\int_0^1 a_{jt}^{\mathrm{FB}\eta-1}\right)^{\frac{1}{\eta-1}}, \tag{2.39}$$

where the first best sectoral productivity is

$$a_{jt}^{\mathrm{FB}} \equiv \left(\sum_{i=1}^{n_j} a_{ijt}^{\rho-1}\right)^{\frac{1}{\rho-1}}. \tag{2.40}$$

We see that the markup dispersion in the product market distorts the resource allocation and hence causes TFP loss in the economy. Hsieh and Klenow (2009), Restuccia and Rogerson (2008), and Edmond, Midrigan, and Xu (2015) analyze this misallocation effect in cross-section.

However, it might be a different picture if we think in terms of business cycle. Along business cycle, standard deviation of idiosyncratic productivities is countercyclical. Even though the aggregate TFP is lower than the level could be attained by FB, but the aggregate TFP might be countercyclical due to the well-known Oi-Hartman-Abel effect. I illustrate that it is indeed the case in the simulation.

Implications for Aggregate Output

In this sub subsection, I discuss how the imperfect firm competition affects the total output along the business cycles. Change in log total output can be written as

$$\Delta \log Y_t = \Delta \log A_t + \Delta \log L_t. \tag{2.41}$$

For the simplicity of illustration, I ignore the fixed cost for production in the analysis. From the representative household's consumption and labor optimality condition (2.20), I can express the labor supply as a function of the aggregate productivity and the aggregate markup

$$L_t = \left(\frac{A_t^{1-\sigma}}{\omega \mathcal{M}_t}\right)^{\frac{1}{\psi+\sigma}}. \tag{2.42}$$

Then change in log total output becomes

$$\Delta \log Y_t = \frac{\psi+1}{\psi+\sigma}\Delta \log A_t - \frac{1}{\psi+\sigma}\Delta \log \mathcal{M}_t. \tag{2.43}$$

Hence, countercyclical aggregate markup amplifies the fluctuation of output along business cycle.

And for change in log aggregate productivity, I show the following result (see the Appendix for proof)

Proposition 2 *Change in aggregate productivity can be decomposed into three parts:*

$$\Delta \log A_t = \Delta \log \tilde{A}_t - \frac{\eta}{\eta-1}\Delta \log \tilde{\mathcal{M}}_{st} - \frac{\rho}{\rho-1}\Delta \log \tilde{\mathcal{M}}_{wt}. \tag{2.44}$$

First, $\Delta \log \tilde{A}_t$ *is TFP loss due to misallocation*

$$\Delta \log \tilde{A}_t \equiv \int_0^1 \left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)^{-1} S_{jt} \left(\sum_{i=1}^{n_j} \left(\frac{\mathcal{M}_{ijt}}{\mathcal{M}_{jt}}\right)^{-1} S_{ijt}\Delta \log a_{ijt}\right) dj. \tag{2.45}$$

Second term $\Delta \log \tilde{\mathcal{M}}_{st}$ *is TFP loss due to sectoral markup cyclicality*

$$\Delta \log \tilde{\mathcal{M}}_{st} \equiv \int_0^1 \left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)^{-1} S_{jt}(\Delta \log \mathcal{M}_{jt} - \Delta \log \mathcal{M}_t) dj. \tag{2.46}$$

Third term $\Delta \log \tilde{\mathcal{M}}_{wt}$ is TFP loss due to within-sector markup cyclicality

$$\Delta \log \tilde{\mathcal{M}}_{wt} \equiv \int_0^1 \left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)^{-1} S_{jt} \left(\sum_{i=1}^{n_j} \left(\frac{\mathcal{M}_{ijt}}{\mathcal{M}_t}\right)^{-1} S_{ijt}(\Delta \log \mathcal{M}_{ijt} - \Delta \log \mathcal{M}_{jt})\right) dj. \quad (2.47)$$

2.3.1 Calibration and Simulation

Household Preference Parameters

Household has a log utility in consumption ($\sigma = 1$). I set Frisch elasticity of labor supply $1/\psi = 1$, as suggested by Chang, Kim, Kwon, and Rogerson (2014). Then from equation (2.42), movement in labor supply is simply driven by only movement in aggregate markup: $L_t = (\omega \mathcal{M}_t)^{1/2}$. There is no effect of aggregate productivity on labor supply, since income and substitution effects cancel out perfectly due to unit intertemporal elasticity. Finally, I set disutility from labor supply parameter such that labor supply in the steady state equal to one third.

Elasticities of Substitution

I infer the within-sector elasticity of substitution ρ and the across-sector elasticity of substitution η by running a regression of firm's markup on firm's market share as in (2.48). Note that a firms's optimal pricing rule is the markup over the marginal cost, hence the markup can be expressed as:

$$\mathcal{M}_{ijt} = \frac{P_{ijt}}{W_t / a_{ijt}} = \frac{P_{ijt} Y_{ijt}}{W_t l_{ijt}},$$

where the second equality results from multiplying the denominator and the numerator by output Y_{ijt}. Hence, I can replace the dependent variable of equation (2.48) with the labor cost share:

$$\frac{W_t l_{ijt}}{P_{ijt} Y_{ijt}} = \gamma_0 + \gamma_1 S_{ijt}.^3 \quad (2.48)$$

I can infer the values of ρ and η from the ratio of the coefficient estimates γ_0/γ_1:

$$\eta = \left(\frac{1}{\rho} - \frac{\gamma_1}{\gamma_0}\left(\frac{\rho-1}{\rho}\right)\right)^{-1}$$

The estimate of the ratio $\frac{\gamma_1}{\gamma_0}$ is -0.973. I choose $\rho = 11$ such that firms' markup equal to 1.1 under perfect competition, and hence $\eta = 1.026$.

Firm Parameters

Each firm's TFP a_{ijt} consists of common TFP A_t^M and firm specific TFP a_{ijt}^F: $a_{ijt} = A_t^M \times a_{ijt}^F$. $\log A_t^M$ and $\log a_{ijt}^F$ follow AR(1) processes respectively:

$$\log A_t^M = \rho_m \log A_{t-1}^M + \nu_m \xi_t^m, \quad \xi_t^m \sim N(0,1) \tag{2.49}$$

$$\log a_{ijt}^F = (1-\rho_f)\ln \alpha_{ij} + \rho_f \log a_{ijt-1}^F + d_t \xi_{ijt}^f, \quad \xi_{ijt}^F \sim N(0,1). \tag{2.50}$$

Note that the variance of the firm-level shock d_t is itself time-varying. In the normal period, I set $d_L = 0.05$, and it spikes to $d_H = 0.15$ during the recession period.

Finally, I set the number of firms in each sector to be 30, which is close to the mean number of firms in the data.

2.3.2 Impulse Response

I analyze several business cycle moments with impulse response analysis. Specifically, I test with two scenarios: (i) a spike in variance of firm specific productivity d_t, and (ii) a drop in common TFP A_t^M.

[3] If firms have labor production elasticity β_l different from unity, equation (2.48) can be extended to $\frac{W_t l_{ijt}}{P_{ijt} Y_{ijt}} = \gamma_{0j} + \gamma_1 S_{ijt}$, where γ_{0j} is a dummy variable for sector j to capture heterogeneous labor production elasticities across sectors. In this case, I cannot identify $\hat{\rho}$ from $\hat{\gamma}_0$.

Second Moment Shock

In this experiment, I set the variance of firm specific productivity $d_t = 0.15$ at period 0 for one period, which is three times as high as the normal period value $d_L = 0.05$. The impulse response results are in Figure 2.1. With increased dispersion in idiosyncratic productivities and the result of *corollary 2*, the aggregate markup increases by around 2.5%. And labor supply decreases by around 1.2% accordingly. However, due to Oi-Hartman-Abel effect, the aggregate TFP actually increases in the recession. Bloom (2009) discusses this undesired effect, but since there are adjustment costs for both labor and capital usage in his model, misallocation effect dominates and aggregate TFP decreases. Finally, since increase in aggregate TFP dominates decrease in labor supply, aggregate output turns out to increase during recession.

The model also has a wrong prediction for response of small and large firms. On average, small firms have smaller markups while large firms have larger ones, the model predicts that small firm's markup is procyclical while large firm's is countercyclical (as in Figure 2.2). But in empirical analysis of markup cyclicality, I actually find that small firm's markup is more countercyclical than large firm's.

Moreover, with the same second moment shock, I now assume that firms have to pay operating cost to produce in the economy. Specifically, I assume that firms have to pay 4% of mean profit in the steady state. Now in the recession, the number of operating firms decrease by around 12%. Jaimovich and Floetotto (2008) emphasize this extensive margin effect on markup cyclicality. However, as seen in Figure 2.1, we see that this effect is almost negligible. The reason is only small firms drop out of the market and they have marginal effect for large firms remaining in the market.

First Moment Shock

In this experiment, I set the common TFP A_t^M drops by 3% at period 0. From *Theorem 1*, It is not surprising to see that it has no effect on aggregate markup. And since movement

of labor supply is only determined by markup in our parameter specification, labor supply stays constant. Hence all firms profit stay constant across the time period and hence no firm exits the market even though they have to pay operating cost.

2.4 Sticky Price Model

In the previous section, we have seen that the oligopolistic competition successfully generates countercyclical markup at the aggregate level, but is inconsistent with micro-level evidence. Now I examine another model that could generate countercyclical markup - sticky price model.

The reason that a standard New Keynesian model could generate countercyclical markup is the following. Under monopolistic competition and constant consumer price elasticity θ, a firm's optimal pricing strategy is a constant markup $\frac{\theta}{\theta-1}$ over marginal cost. However, with price stickiness, a procyclical marginal cost implies that in a boom, the gap between the price and the marginal cost shrinks, and hence decrease in the markup.

To match the cross-sectional markup cyclicality in the data, small firms should exhibit more price stickiness than large firms. Goldberg and Hellerstein (2011) find that it is indeed the case. They categorize firms into three equal bins by their size, and they find that the largest firms have a frequency of price adjustment 18.20%, while the smallest firms have a frequency of price adjustment 10.50%[4]. We see that large firms adjust prices almost twice as frequently as small firms. Hence, the sticky price model implies markup cyclicality that is consistent with my empirical finding qualitatively. To investigate if heterogeneity in price stickiness is large enough to generate heterogeneity in markup cyclicality, I examine the following New Keynesian model in general equilibrium. The innovation of my model is that cost and probability of price adjustment depends on firm's size.

[4]Please see Table 2.2.

2.4.1 Household

The representative household has an additively separable preference over consumption and labor and maximizes the following

$$\max \mathrm{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\sigma}}{1-\sigma} - \omega \frac{L_t^{1+\psi}}{1+\psi} \right) \right\}, \tag{2.51}$$

where $\frac{1}{\sigma}$ is the inter temporal elasticity of substitution (IES), ω is the disutility parameter from labor, and $\frac{1}{\psi}$ is the Frisch elasticity of labor supply. And C_t is Dixit-Stiglitz aggregator of differentiated goods consumption over varieties i,

$$C_t = \left(\int_0^1 c_{it}^{\frac{\theta-1}{\theta}} di \right)^{\frac{\theta}{\theta-1}}.$$

The budget constraint for household is

$$\int_0^1 p_{it} c_{it} di + \mathrm{E}_t[Q_{t,t+1} B_{t+1}] \leq B_t + W_t L_t + \int_0^1 \pi_{it} di.$$

A complete set of Arrow-Debreu state-contingent assets is traded, so that B_{t+1} is a random variable that delivers payoffs in period $t+1$. $Q_{t,t+1}$ is the stochastic discount factor used to price them.

The first-order conditions of the household's maximization problem is

$$\begin{aligned} \frac{W_t}{P_t} &= \omega \frac{L_t^{\psi}}{C_t^{-\sigma}} \\ Q_{t,t+1} &= \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \frac{P_t}{P_{t+1}} \end{aligned}$$

Finally, I assume that the aggregate nominal value-added $S_t \equiv P_t C_t$ follows an exogenous random walk:

$$\log S_t = \log S_{t-1} + \mu_S + \eta_t, \quad \eta_t \sim N(0, \sigma_S). \tag{2.52}$$

We can think of this as the central bank has a targeted path of nominal value-added, and it does so by adjusting interest rate accordingly.

2.4.2 Firms

Each firm produces output c_{it} using a technology in labor l_{it}:

$$c_{it} = a_{it} l_{it}, \tag{2.53}$$

where a_{it} is firm-specific idiosyncratic productivity, which follows an AR(1) process

$$\log a_{it} = \rho_a \log a_{it-1} + \epsilon_{it}, \quad \epsilon_t \sim N(0, \sigma_a).$$

And each firm faces the following demand:

$$c_{it} = \left(\frac{p_{it}}{P_t}\right)^{-\theta} C_t, \tag{2.54}$$

where p_{it} is price of good i, P_t is the aggregate price level, and C_t is the aggregate consumption.

To change its price, a firm must pay a fixed cost κ_{it} in units of labor. Structure of κ_{it} will be specified below. Hence, a firm's nominal profit equals to

$$\pi_{it} = \left(p_{it} - \frac{W_t}{a_{it}}\right)\left(\frac{p_{it}}{P_t}\right)^{-\theta} C_t - \kappa_{it} W_t I_{p_{it} \neq p_{it-1}}.$$

Krusell-Smith Forecast Rule

To solve the model in general equilibrium, it is necessary to keep track of distribution of firms over idiosyncratic productivities and prices, and thus determines the aggregate price level. Here, I assume that the aggregate price level itself is self predictable. In particular, I

assume that each firm perceives a Krusell-Smith type law of motion for S_t/P_t

$$\log \frac{S_t}{P_t} = \gamma_0 + \gamma_1 \log \frac{S_t}{P_{t-1}}.$$

Given this conjecture, a firm's state variables are: (i) last period's individual price over the nominal value-added $\frac{p_{it-1}}{S_t}$, (ii) idiosyncratic productivity a_{it}, (iii) ratio of nominal value-added over aggregate price level $\frac{S_t}{P_t}$, and (iv) size of adjustment cost κ_{it}. And firm's problem can be written recursively in real term as

$$V\left(\frac{p_{it-1}}{S_t}, a_{it}, \frac{S_t}{P_t}, \kappa_{it}\right) = \max_{p_{it}} \left\{ \frac{\pi_{it}}{P_t} + \mathrm{E}_t \left[Q_{t,t+1} V\left(\frac{p_{it}}{S_{t+1}}, a_{it+1}, \frac{S_{t+1}}{P_{t+1}}, \kappa_{it+1}\right)\right]\right\}.$$

Please see appendix for numerical solution outline.

2.4.3 Recursive Competitive Equilibrium

A recursive competitive equilibrium is a law of motion (γ_0, γ_1), a set of price level path $\{P_t\}$, and a set of wage path $\{W_t\}$ that are consistent with

1. Household utility maximization problem
2. Firm profit maximization problem
3. Goods market clearing
4. Arrow-Debreu market clearing
5. Evolution of nominal aggregate demand S_t and idiosyncratic productivity a_{it}

2.4.4 The CalvoPlusPlus Model

To match the heterogeneity in price stickiness, there are two ways to implement it. First, the cost of price adjustment (menu cost) depends on the firm size. Second, the Calvo probability of price adjustment depends on the firm size.

Nakamura and Steinsson (2010) introduces the *CalvoPlus* model, where a firm has a probability $1-\lambda$ to face an infinite menu cost, and a probability λ to face a small menu cost, but large enough that it makes some of the firms unwilling to adjust their prices still. The last assumption is different from the usual Calvo model, in which all firms adjust their prices with probability λ. In my model, both the size and probability of menu cost depend on the firm's size, hence I call this new extension *CalvoPlusPlus Model*.

Menu Cost Model

To adjust its price, a firm has to pay the following menu cost

$$\kappa_{it} = \kappa_0 \left(\frac{p_{it} c_{it}}{P_t} \right)^{\kappa_1}.$$

The value of the cost depends on its revenue, as in Gertler and Leahy (2008). Note that $\kappa_1 = 0$ corresponds to the case of a constant menu cost.

Calvo Model

A firm has a certain probability of not paying any cost to adjust its price

$$\kappa_{it} = \begin{cases} 0 & \text{w.p. } \lambda_{it} \\ \bar{\kappa} & \text{otherwise,} \end{cases}$$

where probability of zero menu cost λ_{it} depends on last period's revenue

$$\lambda_{it} = \lambda_0 \left(\frac{p_{it-1} c_{it-1}}{P_{t-1}} \right)^{\lambda_1}.$$

$\bar{\kappa}$ is set such that firms almost never pay $\bar{\kappa}$ to adjust prices.

Interpretation

How should we understand these heterogeneous adjustment costs? I do not see them as the literal cost of changing the menu. Instead, I see them as a general way of capturing the cost associated with adjusting the listed prices, which includes survey cost of current market condition, paying a manager to collect information, and etc. And this cost could weigh large or small relative to a firm's total revenue. Midrigan (2011), and Bhattarai and Schoenle (2014) find that multi-product firms tend to change prices more frequently than single-product firms. They construct a model where firms can pay one cost to change prices of all the underlying products, and it matches their empirical finding. Gertler and Leahy (2008) introduce a size-dependent menu cost to keep price adjustment decision of the firm homogeneous of its size. Carvalho (2006) introduces exogenous heterogeneity in price stickiness across sectors, and find that monetary shocks tend to have larger effects in the heterogeneous model, compared to an identical price stickiness model. My model is an addition to this heterogeneity in price stickiness, which depends on the firm size in particular. I leave it to the future research to study the microstructure underlying the heterogeneous adjustment costs I introduce here.

2.4.5 Calibration

In the model, one period equals to one month in the data. The monthly discount factor is $\beta = 0.997$. For the representative household, I assume log utility in consumption $\sigma = 1$, and infinite Frisch elasticity of labor supply $\psi = 0$ as in Hansen (1985) and Rogerson (1988). Hence, the real wage is a linear function of the aggregate consumption $W_t/P_t = \omega C_t$, this means that we do not need to keep the aggregate labor supply as a state variable.

For elasticity of substitution, I set $\theta = 5$, which is aligned with most of empirical findings. The growth rate and standard deviation of value-added S_t are taken from Nakamura and Steinsson (2010). The values I find in France data are quite close to these values. Firm's idiosyncratic productivity has persistence $\rho_a = 0.9$, and standard deviation $\sigma_a = 0.03$.

For the parameters of price adjustment cost, I set them such that the model matches top and bottom firms' price adjustment frequency. Please see Table 2.3 and Table 2.5 for parameter specifications for the Calvo model, and menu cost model, respectively.

2.4.6 Simulation Results

I present and discuss the simulation results of the CalvoPlusPlus model under two alternative assumptions about adjustment costs, (i) Calvo model, and (ii) menu cost model.

Calvo Model

The main statistics from the model is summarized in Table 2.4. Compared to Hellerstein and Goldberg's (2011) finding in Table 2.2, I find that firms increase prices more frequently in the model, and the size of price adjustment is smaller in the model, too. For example, the size of adjustment for middle is 6% in the data, while 5.38% in the model. However, most of the values are in the same magnitude as in the data. This is surprising since the only moments that I target in the calibration is price adjustment frequency of top and bottom firms.

Furthermore, I compare the markup cyclicality in the model to my empirical finding. In the simulation, I run the same regression as I run in the data: regress change in log markup $\Delta \log \mathcal{M}_{it}$ on change in aggregate output $\Delta \log Y_t$. In Figure 2.4, I present both markup cyclicality from the data and the model. Number 1 on the vertical axis stands for the smallest firms in terms of market share, number 2 for firms with middle market share, and number 3 for firms with largest market share. I find that the model generates the same magnitude of markup cyclicality as in the data, and it captures the heterogeneity in markup cyclicality qualitatively. Small firms adjust prices less frequently, hence more firms are unable to adjust prices while the underlying marginal cost fluctuates procyclically with the aggregate output. Therefore small firms' markup are more countercyclical relative to

large firms'. However, we can see that the model does not capture the heterogeneity of markup cyclicality closely as in the data.

Menu Cost Model

The main statistics about the menu cost model is summarized in Table 2.6. The result is surprising, since the model captures all the moments astonishingly well, including size of price adjustment and etc. Furthermore, I compare the markup cyclicality in the model to empirical counterparts as I do in the Calvo model, and I find that the model captures both the magnitude and heterogeneity quite well.

The reason that the menu cost model generates more heterogeneity in markup cyclicality is the following: In a menu cost model, only a firm that has its markup substantially far away from its optimal markup $\mu^* \equiv \frac{\theta}{\theta-1}$ would adjust its price to obtain optimal profit. Upon a positive demand shock, firms that are close to the optimal mark do not adjust their price, which contributes countercyclicality to the aggregate markup. While firms that are far from the optimal markup are willing to pay the adjustment cost, and increase their price with respect to the increased nominal marginal cost, which contributes procyclicality to the aggregate markup. In contrast, in a Calvo model, the selection of which firms adjusting their prices is independent of how far they are from optimal markups; the firms chosen by a random probability λ_{it} are allowed to adjust their prices. Hence, the strong selection effect in the menu cost model generates large heterogeneity in the markup cyclicality.

2.4.7 Robustness

The business cycle of the benchmark model is driven by the nominal value-added shock. To check the robustness of my result, I investigate a New Keynesian model with an aggregate TFP shock in partial equilibrium. I find that markup becomes procyclical, in contrast to countercyclical markup with nominal value-added shock. The reason is that upon a positive TFP shock, the nominal marginal cost shifts downward, instead of upward upon a positive

demand shock, hence with a sticky price, markup increases during a boom. The result of the model with TFP shock is not presented here, but is available upon request.

2.5 Conclusion

Markup cyclicality is an important magnification mechanism in the business cycle models. Previous literatures either assume an exogenous process for markup cyclicality, or use models that generate markup cyclicality without examining their validities at the micro level. In this chapter, I examine two representative models, an oligopolistic competition model, and a New Keynesian model. First, I find that the oligopolistic competition model can generate the countercyclical aggregate markup, but fails to capture markup cyclicality at the firm level. Second, I introduce heterogeneous price adjustment costs into a standard New Keynesian model, and discipline the parameters to match heterogeneity in price adjustment frequencies. The resulting model successfully captures all the important moments, in the data, and in particular, the magnitude and heterogeneity in markup cyclicality in the Cobb-Douglas production function case. However, both sticky price models imply procyclical markup behavior in response to productivity shocks.

Tables and Figures

Table 2.1: Parameter Values for Calibration

Parameter	Value	Rationale
Intertemporal Elasticity of Substitution (IES)	$1/\sigma = 1$	Unit IES
Frisch Elasticity of Labor Supply	$1/\psi = 1$	Chang, Kim, Kwon, and Rogerson (2014)
Disutility Parameter from Labor	$\omega = 7$	
Across-sector Elasticity of Substitution	$\eta = 1.026$	Labor Cost Share and Market Share
Within-sector Elasticity of Substitution	$\rho = 11$	Labor Cost Share and Market Share
Number of Firms	$n_j = 30$	Moment in the data
Fixed Cost of Production	$\phi = 4\%$	
Persistence of Firm Productivity	$\rho_f = 0.95$	
SD of Firm Productivity (Low)	$\nu_f = 0.05$	
SD of Firm Productivity (High)	$\nu_f = 0.15$	

Table 2.2: Summary Statistics: Goldberg and Hellerstein (2011)

Weighted Median	Top	Middle	Bottom
Frequency of Adjustment	18.20%	12.20%	10.50%
Frequency of Increases	13.60%	10.30%	8.20%
Frequency of Decreases	5.50%	1.60%	1.50%
Adjustment Size Change	5.60%	6.00%	6.00%
Upward Size Change	5.70%	5.40%	5.60%
Downward Size Change	5.60%	5.90%	6.70%

Note: Top, Middle, and Bottom refers to terciles in terms of firms revenues. Large firms adjust prices more frequently, and adjust less than small firms.

Table 2.3: Parameter Values for Simulation: Calvo Model

Parameter	Value	Rationale
Monthly Discount Factor	$\beta = 0.997$	Nakamura and Steinsson (2010)
Elasticity of Substitution	$\theta = 5$	
Inverse of Intertemporal Elasticity of Substitution	$1/\sigma = 1$	
Inverse of Frisch Elasticity of Labor Supply	$\psi = 0$	
Steady State Labor Supply	$L_{ss} = 1/3$	
Nominal Aggregate Demand Growth Rate	$\mu_S = 0.0028$	Nakamura and Steinsson (2010)
Nominal Aggregate Demand Std. Deviation	$\sigma_S = 0.0065$	Nakamura and Steinsson (2010)
Idiosyncratic Productivity Persistence	$\rho_a = 0.9$	
Idiosyncratic Productivity Std.Deviation	$\sigma_a = 0.03$	moments in the data
Calvo Constant	$\lambda_0 = 3.1200$	moments in the data
Calvo Curvature	$\lambda_1 = 3.0169$	moments in the data

Table 2.4: Summary Statistics: Calvo Model

Weighted Median	Top	Middle	Bottom
Frequency of Adjustment	18.20%	13.90%	10.60%
Frequency of Increases	11.20%	8.80%	6.80%
Frequency of Decreases	6.90%	5.10%	3.60%
Adjustment Size Change	4.87%	5.38%	5.93%
Upward Size Change	5.12%	5.75%	6.45%
Downward Size Change	4.46%	4.75%	4.96%
$\text{Corr}(\Delta \ln \mathcal{M}_{it}, \Delta \ln Y_t)$	-0.1304	-0.1351	-0.1392
ϕ	-0.9643	-1.0155	-1.0519
Mean of Markup	1.2544	1.2591	1.2683
Std of Δ log Markup	0.0403	0.0410	0.0412

Table 2.5: Parameter Values for Simulation: Menu Cost Model

Parameter	Value	Rationale
Monthly Discount Factor	$\beta = 0.997$	Nakamura and Steinsson (2010)
Elasticity of Substitution	$\theta = 5$	
Inverse of Intertemporal Elasticity of Substitution	$1/\sigma = 1$	
Inverse of Frisch Elasticity of Labor Supply	$\psi = 0$	
Steady State Labor Supply	$L_{ss} = 1/3$	
Nominal Aggregate Demand Growth Rate	$\mu_S = 0.0028$	Nakamura and Steinsson (2010)
Nominal Aggregate Demand Std. Deviation	$\sigma_S = 0.0065$	Nakamura and Steinsson (2010)
Idiosyncratic Productivity Persistence	$\rho_a = 0.9$	
Idiosyncratic Productivity Std. Deviation	$\sigma_a = 0.03$	moments in the data
Menu Costs Constant	$\kappa_0 = 0.00043\%$	moments in the data
Menu Costs Curvature	$\kappa_1 = -7$	moments in the data

Table 2.6: Summary Statistics: Menu Cost Model

Weighted Median	Top	Middle	Bottom
Frequency of Adjustment	18.30%	14.00%	10.40%
Frequency of Increases	12.70%	10.10%	8.10%
Frequency of Decreases	5.60%	3.80%	2.20%
Adjustment Size Change	5.37%	5.83%	6.24%
Upward Size Change	5.01%	5.40%	5.77%
Downward Size Change	6.16%	6.96%	7.86%
$\text{Corr}(\Delta \ln \mathcal{M}_{it}, \Delta \ln Y_t)$	-0.1278	-0.1394	-0.1514
ϕ	-0.9407	-1.0774	-1.2071
Mean of Markup	1.2399	1.2364	1.2326
Std of Δ log Markup	0.0276	0.0289	0.0299

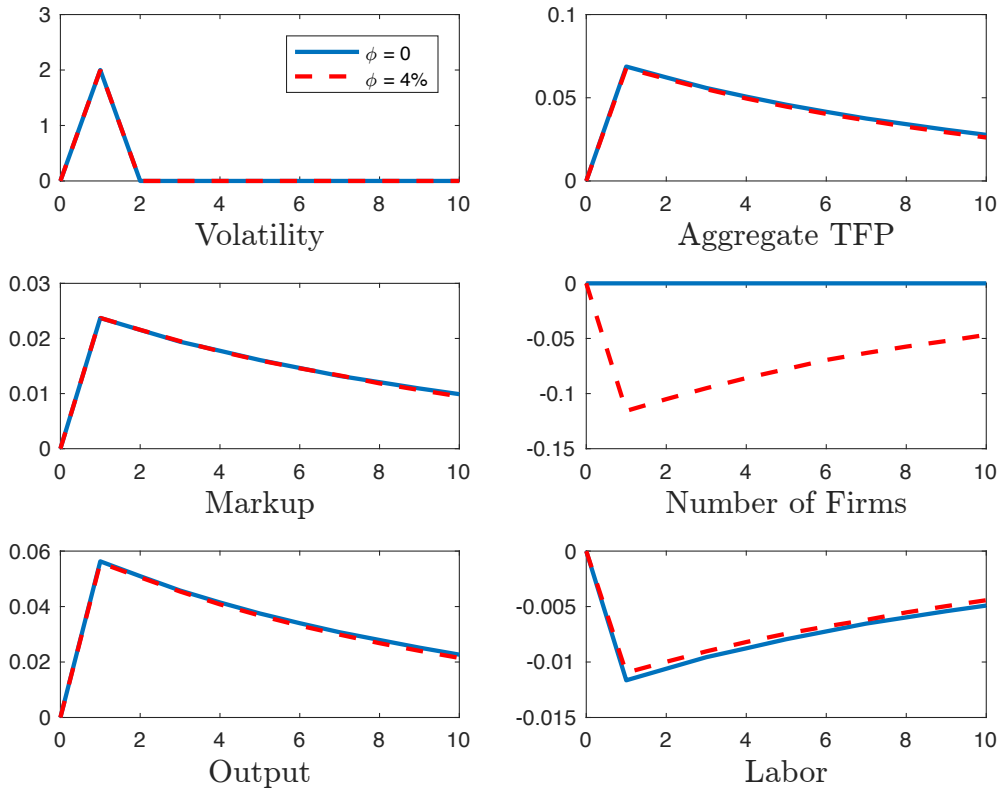


Figure 2.1: Second Moment Shock

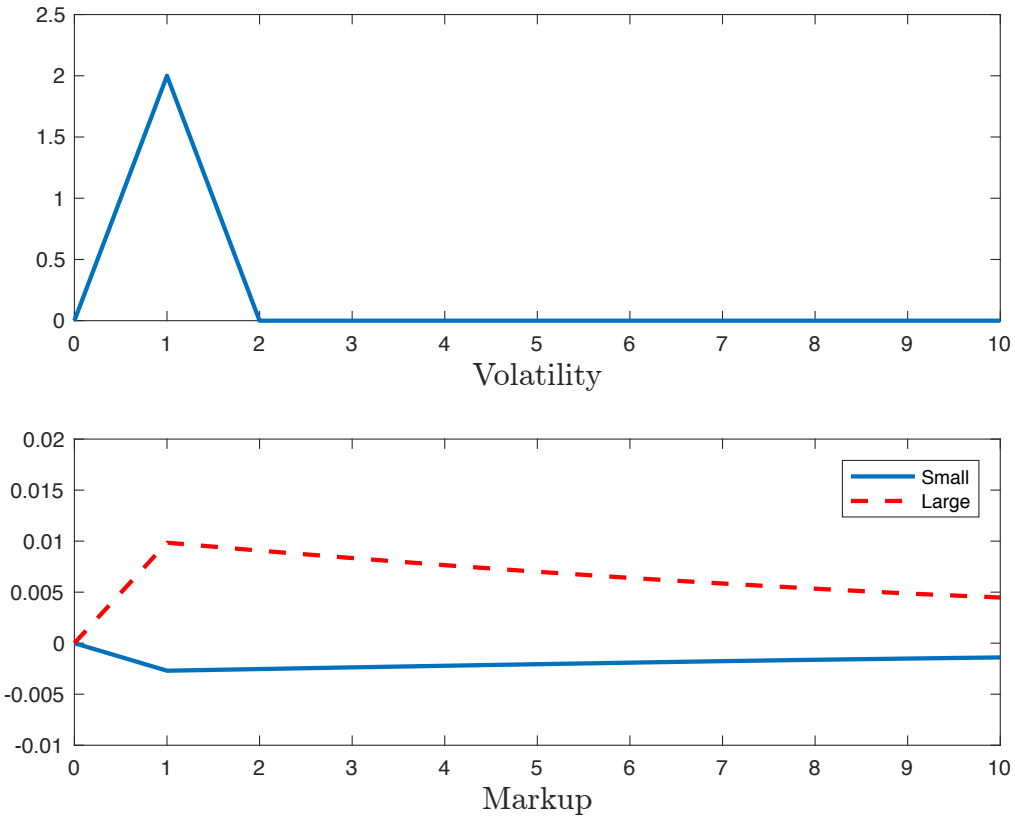


Figure 2.2: Response of Small Firms VS Large Firms

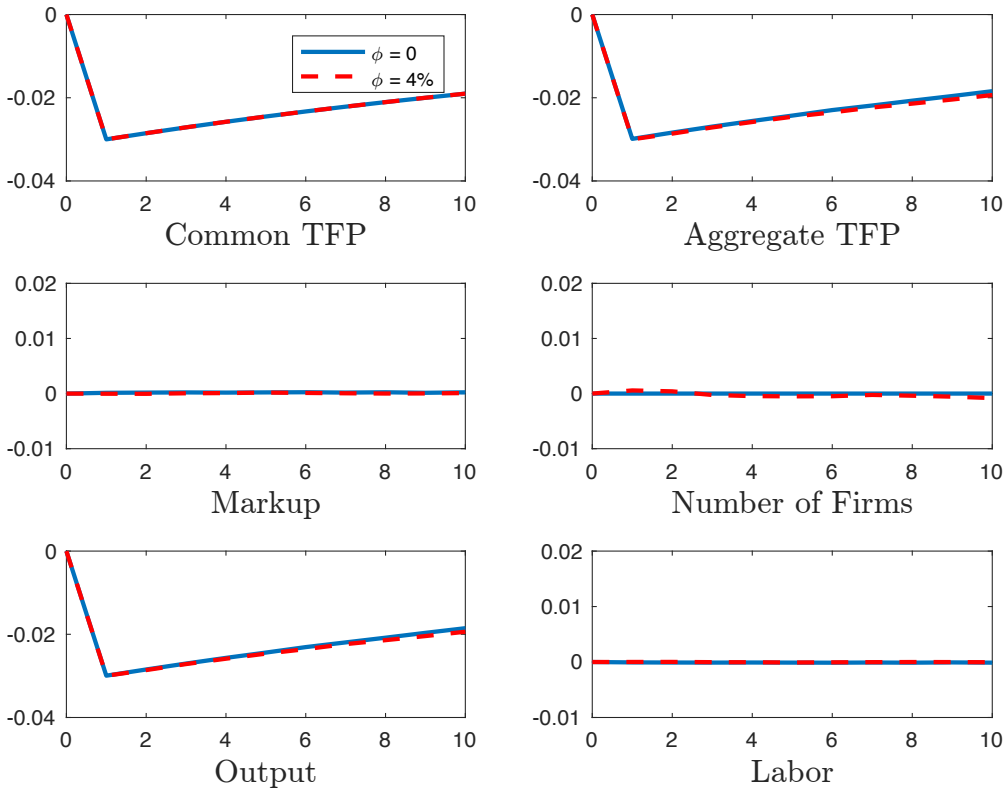
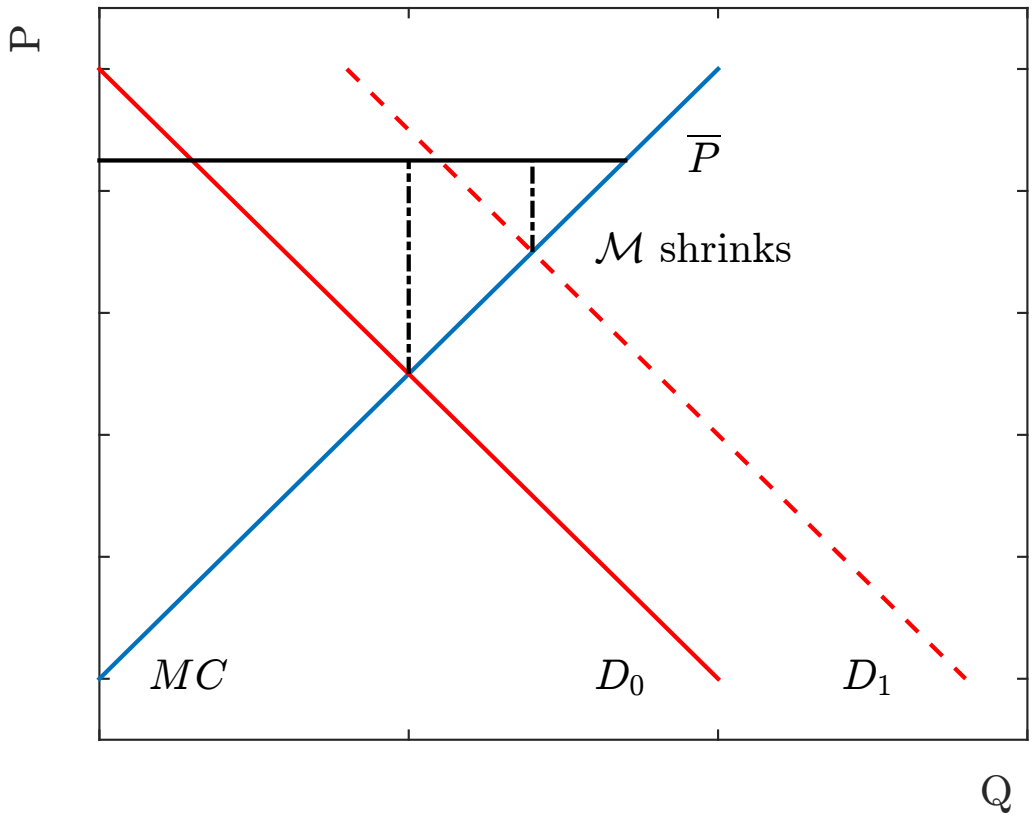


Figure 2.3: First Moment Shock



- Initially, firm sets price at $\overline{P}$. When demand curve shifts from D_0 to D_1, marginal cost MC increases. Due to price stickiness, price stays at $\overline{P}$, hence markup $\mathcal{M}$ shrinks.
- To match the cross-sectional markup cyclicality in the data, small firms should exhibit more price stickiness than large firms.

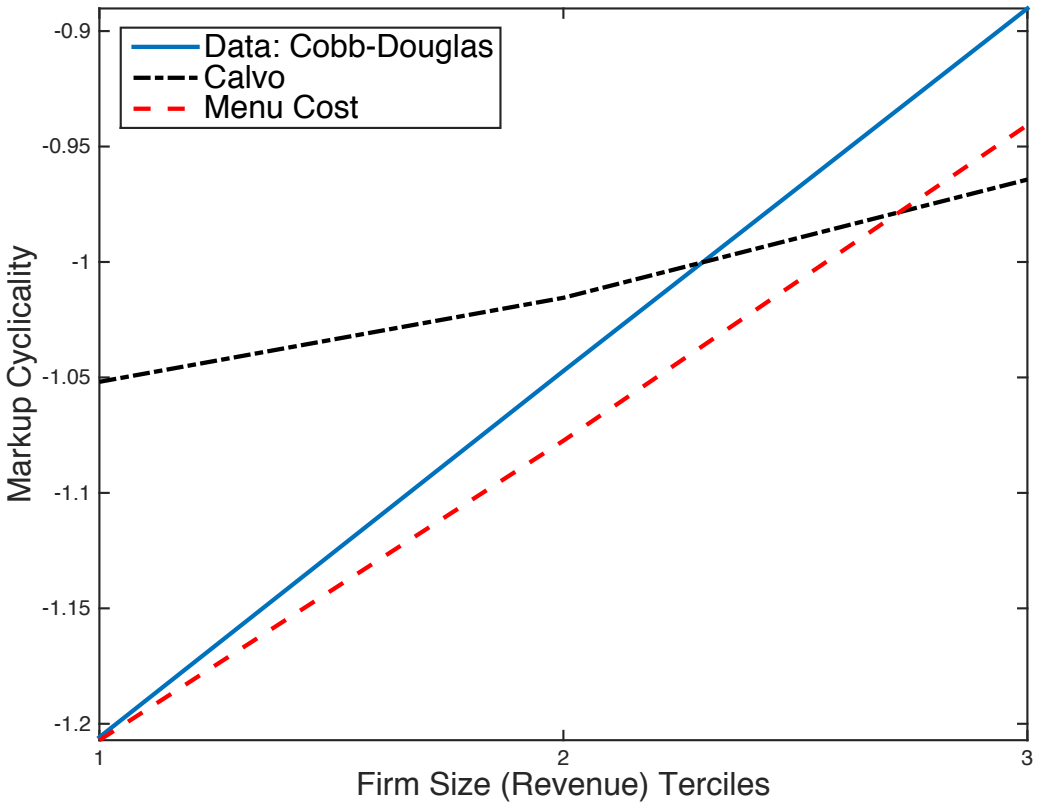


Figure 2.4: Comparison of markup cyclicality ϕ between data, Calvo Model, and Menu Cost model.

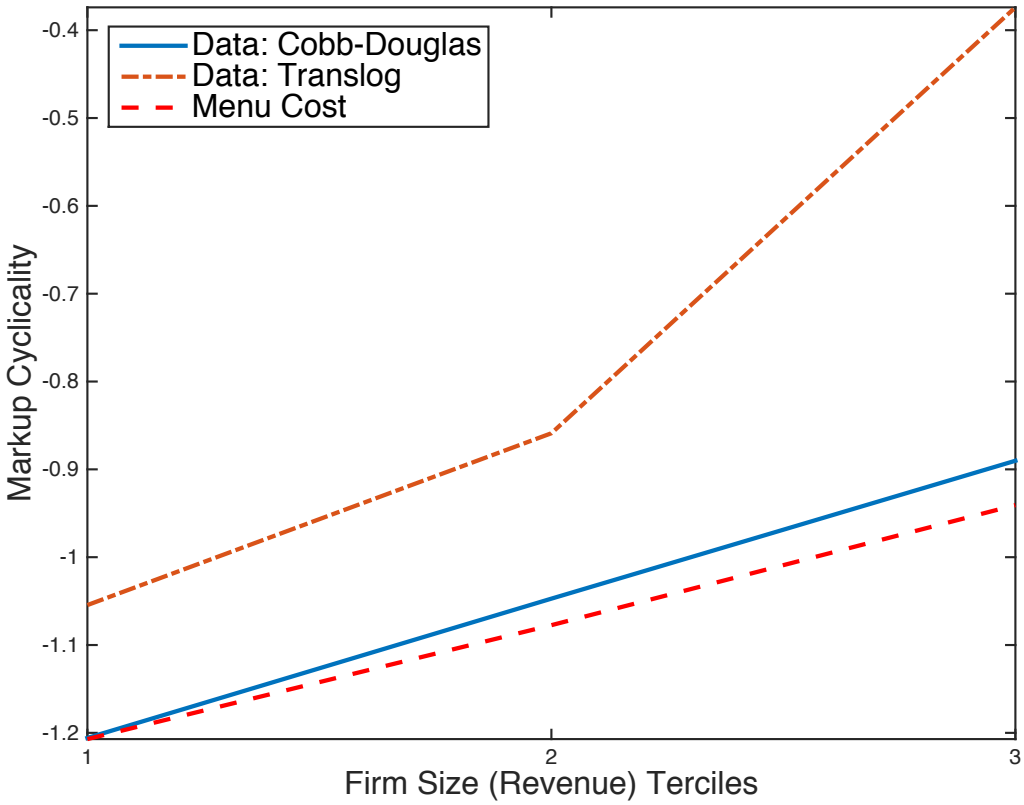


Figure 2.5: Comparison of markup cyclicality ϕ between data and Menu Cost model. Data includes both Cobb-Douglas and Translog cases.

Chapter 3

Customer Capital, Markup Cyclicality, and Amplification

3.1 Introduction

In this chapter, I explore the economic mechanism that gives rise to two empirical patterns in chapter one by building a general equilibrium model. In particular, I embed *customer capital* (due to deep habits as in Ravn, Schmitt-Grohe, and Uribe 2006) into a standard Hopenhayn (1992) model of firm dynamics with entry and exit. The driving force of business cycle in the economy is the aggregate productivity shock. A key feature of the model with customer capital is that a firm's decision about markups becomes dynamic – higher markups today imply higher profits per unit sold today, but by decreasing the quantity sold lead to lower customer capital in the future. Firms calculate the present value of all future profits with procyclical stochastic discount factors. Thus the incentive for firms to *invest* in customer capital is high in the booms, and low in the recessions, resulting in countercyclical markups. Also, the endogenous firm entry and exit decision also affects firms pricing. During recessions, the endogenous probability of exit increases, and this implies that firms place lower weight on future profits, leading them to charge higher markups. Since the exit risk is larger for

small firms during downturns, the extensive margin of exit affects them more, resulting in more countercyclical markups.

Taken together, the model is able to replicate the two main empirical findings qualitatively, and quantitatively as well: 1) firms have countercyclical markups, and 2) small firms have more countercyclical markups than large firms. Furthermore, while the model can explain around one-third to two-thirds of the magnitude of markup elasticity with respect to aggregate output, it closely captures the heterogeneity of markup cyclicality between small and large firms as estimated in the data.

Interestingly, although the model was only disciplined to explain firm-level markup behaviors, it also captures other features over business cycles. First, the model is able to generate endogenous fluctuations of firm size distribution over business cycles caused by aggregate productivity fluctuations. Heterogenous markup dynamics of firms imply that, during recessions, small firms produce less and shrink relative to large firms by raising their markups, which results in increase in dispersion of firm size. This implication has support in the data. In the model, small firms' deflated sales are 13% more responsive to aggregate output than large firms', while the difference is 22% in the data. Notice that the model generates the countercyclical dispersion of firm size without relying on exogenous second moment shock in productivity growth as in Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2014). Indeed, the exogenous process that governs firms' idiosyncratic productivities is time-invariant in my model. What drives the heterogeneous responses of firms is the endogenous probability of staying in the market, which depends not only on the productivity, but also the customer capital that firms have accumulated.

Finally, the other business cycle implication of the model is that the resulting input misallocations amplifies the aggregate productivity. I define the measured Total Factor Productivity (TFP) as the ratio of total output to total input, and it has a non-trivial dynamic in the model. There are two opposing effects that drive the movements of measured TFP relative to the aggregate productivity. The first effect comes from the selection of firms'

productivities at the extensive margin (*cleansing effect*). During recessions, the productivity threshold for entrants and exiters increases, hence the mean for continuing firms' productivities increases. This cleansing effect tends to make the measured TFP procyclical. The second effect comes from the misallocation of inputs due to dispersion of firm size, which contributes to the countercyclicality of the measured TFP. In the simulation, upon a negative aggregate productivity shock, the misallocation effect dominates in the beginning. Later on, the selection effects dominates. Hence, the measured TFP initially drops, rebounds, and then overshoots compared to the aggregate productivity. I find that the misallocation effect, which is the main focus of this paper, amplified the standard deviation of the measured TFP by 25%.

Related Literature First, this chapter relates to models with variable markups. Some business cycle models assume an exogenous markup process, sometimes called "cosh-push" shock, without giving explicit explanations for the underlying economic mechanisms (e.g. Clarida, Gali, and Gertler 1999, Smets and Wouters 2003, 2007, Steinsson 2003). Other models try to ground the source of markup variation. In the New Keynesian model, sticky price with a procyclical marginal cost generates countercyclical markup. In Rotemberg and Saloner (1986), implicit collusion of oligopolistic firms gives rise to variation in markup. In Jaimovich and Floetotto (2008) and Bilbiie, Ghironi, and Melitz (2012), due to firm entry and exit, variation in product variety affects the goods elasticity of substitution, and thus markup. Edmond and Veldkamp (2009) analyze a model where the countercyclical earnings dispersion leads to countercyclical markup. These models explain the movements of aggregate markups, but are either silent to the heterogeneity of cyclicality, or have opposite predictions at the firm level. A closely related model to this paper is by Gilchrist, Schoenle, Sim, and Zakrajsek (2015). In their model, firms also have dynamic pricing problems in the customer market. During downturns, more financially constrained firms tend to finance their fixed payments (e.g. operating cost, debt payment) through internal financing by raising markups, but at the cost of future market shares. However, my model shows that, even without financial

frictions, if firms face entry and exit decisions at the extensive margin (as in Hopenhayn 1992, Clementi and Palazzo 2013, and Clementi, Khan, Palazzo, and Thomas 2014), heterogeneity of markup cyclicality could still emerge. Hence, my model serves as a complement to their financial friction model[1]. Another related literature on variable markups is the large body of research in international macroeconomics regarding exchange rate pass-through. These studies focus on responses of prices to exchange rate shocks, instead of shocks over business cycles. They find that, due to strategic complementarities in pricing, firms' prices respond sluggishly and heterogeneously to a common exchange rate shock (see, for instance, Atkeson and Burstein 2008, Amiti, Itskhoki, and Konings 2014, 2016). In contrast to these, firm's pricing in my model is driven by intertemporal decision of customer capital accumulation.

Second, the chapter speaks to literature on dispersion of firm size and input misallocation. Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2014) document countercyclical dispersion of innovations in firm productivity. Similarly, Kehrig (2015) finds that the *level* of firm productivity is countercyclical. The model in this paper generates countercyclical firm size dispersion without any underlying exogenous second moment shock in firm productivity. The dispersion occurs endogenously since small firms close to the entry-exit margin respond differently from large incumbents at the other end of the distribution. Previous studies (Restuccia and Rogerson 2008, Hsieh and Klenow 2009) also stress that the misallocation of inputs can lead to decrease in measured TFP. Instead of imperfection in the input markets, several papers emphasize how the distortion in the output market through markup variation could give rise to inputs misallocation. Peters (2013) studies the effect of endogenous misallocation due to market power on economic growth. Opp, Parlour, and Walden (2014) extend the collusion model of Rotemberg and Saloner (1986) under the general equilibrium

[1]Moreira (2015) builds a similar model to this paper to study effects of initial business conditions on entrants' size. However, there are several difference between her work and mine. First, the focus of my paper is markup cyclicality, while her paper focuses on size of new born businesses. Second, in terms of methodology, it is a partial equilibrium model in her model, and the business cycle is driven by an exogenous demand shock. In contrast, I build a general equilibrium model where the aggregate shock comes from aggregate productivity. This key difference allows for a stochastic discount factor, which is the main driver of the markup cyclicality. Also, I can quantify the welfare implication of the model.

framework, and analyze the resource misallocation across industries. Edmond, Midrigan, and Xu (2015) explore the pro-competitive gains from trade liberalization quantitatively. In contrast, my paper focuses on the inputs misallocation due to heterogeneity of markup cyclicality along business cycle, and I find that its effect on measured TFP is substantial in the quantitative exercises.

The remainder of the chapter is organized as follows. In the next section, I document the markup behavior at the firm level. In Section 3, I build a general equilibrium model of firm dynamics and describe its key mechanism. Then, I calibrate the model to match standard moments in the data in Section 4. In Section 5, I show that the model can match the empirical findings and its business cycle implications. In Section 6, I describe alternative preference specification. I then conclude in Section 7.

3.2 Model

In this section, I develop a general equilibrium model that is able to quantitatively match the empirical findings of markup dynamics in the previous section. As other standard RBC models, the business cycles are driven by exogenous shocks to aggregate productivities.

Specifically, I embed *customer capital* into a standard Hopenhayn's (1992) endogenous firm entry and exit model. In the model, the customer capital determines the level of demands of firms' outputs. To enlarge its customer capital, a firm can sell more of its products today to gain more market shares in the future. In other words, a firm sees the product sales as a form of investment into customer capital. Some earlier works in the literature capture this idea in the model. In Rotemberg and Woodford (1991), firms compete to gain market shares by lowering markups. In Klemperer (1995), customers purchasing from one firm have switching costs to a competitor firm. On one hand, firms have incentives to lock in customers by lowering its price. On the other hand, they have some degree of market power over their current customers. To capture this investment in customer capital idea in a tractable way,

I adopt the framework of *deep habits* model developed by Ravn, Schmitt-Grohe, and Uribe (2006), in which households have good-specific habit formations. However, even though the deep habits model has the advantage of tractability, it possibly has different welfare implications from other models, and needs modifications when compared to the data. In the quantitative analysis, I will discuss this welfare measurement bias caused by the good-specific habit formation. Moreover, firms face entry and exit problems as in Hopenhayn (1992).

This simple set-up is able to generate heterogenous markup cyclicality. In the model, a firm always faces a trade-off between *invest* and *harvest* motives. On one hand, a firm wants to lower its price to attract customers to invest in its own customer capital. On the other hand, since a firm has a certain degree of market power over the locked-in customers, it wants to raise its price to harvest the profit. During recessions, since small firms are more likely to exit the market, they put less weight on the future benefit of the customer capital, and increase the price to exploit the customers as much as possible before exiting the market.

Many recent papers have enriched the *deep habits* model to study firm dynamics. There are two papers closely related to my model. One paper is by Gilchrist, Schoenle, Sim, and Zakrajsek (2015), in which firms face financial constraint while making pricing decisions in customer-markets. The other paper is Moreira (2015), which uses the similar framework of customer capital and firm entry-exit dynamics. It studies the procyclical entrant size in a partial equilibrium model. However, my model differs in the following aspects: (i) My model is a general equilibrium model with heterogeneous firms, hence I can study the GE effects, the interaction among firms, and welfare implications over business cycle; (ii) When making decisions to maximize profits, no firms are subject to either nominal rigidity in price setting, or financial frictions; (iii) While Moreira (2015) studies long-run implication on entrant's size, my model focuses on markup dynamics and its implications in short run.

3.2.1 Household

Time is discrete and is indexed by $t = 1, 2, 3, ...$, and is of infinite horizon. A representative household consumes a variety of products produced by a continuum of firms indexed by $i \in [0, 1]$. The household maximizes the discounted expected utility

$$\mathrm{E}_t \sum_{\tau=0}^{\infty} \beta^{\tau} \left[\frac{{X_{t+\tau}}^{1-\sigma}}{1-\sigma} - \omega \frac{{L_{t+\tau}}^{1+\nu}}{1+\nu} \right], \tag{3.1}$$

where X_t is habit-adjusted consumption bundle, and L_t is labor. Household has constant relative risk aversion σ, disutility of labor ω, and inverse Frisch elasticity of labor supply ν.

Household consumes out of a continuum of firms, each producing a specific variety $i \in I$. For any given time, only a subset of varieties $I_t \subset I$ are available for consumption. The habit-adjusted Dixit-Stiglitz consumption aggregator is defined as

$$X_t = \left[\int_{i \in I_t} \left(c_{it} b_{it}^{\theta} \right)^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}}, \tag{3.2}$$

where c_{it} denotes household's consumption of variety i. b_{it} is variety i's habit stock. The household takes variety-specific habit stocks as given. Namely, the household follows the "Keeping up with the Joneses" behavior in consumption[2]. External habit stock b_{it} evolves according to

$$b_{it+1} = (1 - \delta) b_{it} + \delta c_{it}, \tag{3.3}$$

where $\delta \in [0, 1]$ governs both depreciation rate of past habit stock, and conversion rate of consumption into habit stock[3]. Finally, $\rho > 1$ governs the intratemporal elasticity of substitution across habit-adjusted consumption, and θ governs the degree of habit formation.

[2]Nakamura and Steinsson (2011) study the case when the household takes habit formation as internal, and find that firms face time-inconsistency problem in this case.

[3]Paciello, Pozzi, and Trachter (2015) microfound this law of motion for customer flow in a model, where customers constantly search for goods with search costs to buy from a firm which provides lower price. We can see (3.3) as a reduced form of this search process.

Household's demand for variety i is given by:

$$c_{it} = \left(\frac{p_{it}}{\tilde{p}_t}\right)^{-\rho} b_{it}^{\theta(\rho-1)} X_t, \tag{3.4}$$

where p_{it} is product i's price, and $\tilde{p}_t$ is habit-adjusted price index:

$$\tilde{p}_t = \left[\int_{i\in I_t} \left(\frac{p_{it}}{b_{it}^{\theta}}\right)^{1-\rho} di\right]^{\frac{1}{1-\rho}}. \tag{3.5}$$

The household is subject to the following budget constraint:

$$\tilde{p}_t X_t + \mathrm{E}_t(q_{t,t+1} D_{t+1}) = D_t + w_t L_t + \Phi_t, \tag{3.6}$$

where D_{t+1} is Arrow-Debreu securities, w_t is nominal wage, and Φ_t is household's share of firms' profits.

With all the specifications and constraints above, household's optimization problem yields the following two FOCs:

$$\frac{w_t}{\tilde{p}_t} = \omega L_t^{\nu} X_t^{\sigma}, \tag{3.7}$$

$$q_{t,t+\tau} = \beta^{\tau} \left(\frac{X_{t+\tau}}{X_t}\right)^{-\sigma} \frac{\tilde{p}_t}{\tilde{p}_{t+\tau}}, \tag{3.8}$$

which are household's intratemporal trade-off between consumption and leisure, and intertemporal deicision, respectively.

3.2.2 Firm

Incumbent

Incumbent i produces output y_{it} with the following production technology

$$y_{it} = A_t a_{it} l_{it}^{\alpha}, \tag{3.9}$$

where A_t is aggregate productivity to all firms, a_{it} is firm-specific productivity, and l_{it} is labor input for production. α controls the degree of returns-to-scale in production technology[4]. A_t and a_{it} follow the AR(1) processes respectively:

$$\log A_t = \rho_A \log A_{t-1} + \varepsilon_t^A, \tag{3.10}$$

$$\log a_{it} = \rho_a \log a_{it-1} + \varepsilon_{it}^a, \tag{3.11}$$

where both innovations ε_t^A and ε_{it}^a are drawn from normal distribution with standard deviation σ_A and σ_a, respectively. Denote $\Pi_A(A_{t+1}|A_t)$ and $\Pi_a(a_{t+1}|a_t)$ as conditional distribution of aggregate and idiosyncratic productivity, respectively.

At all $t \geq 0$, the distribution of incumbents over the two dimensions of idiosyncratic productivities and customer capitals is denoted by $\Gamma_t(b, a)$. Note that the mass of incumbents N_t is the integral of the distribution, $N_t = \int\int d\Gamma_t(b, a)$. And let Λ_t denote the vector of aggregate state variables, and its transition operator is $B(\Lambda_{t+1}|\Lambda_t)$. In the later section, I show that $\Lambda_t = \{A_t, \Gamma_t\}$.

Each period t, incumbent i faces demand (3.4), and takes into account that the habit stock b_{it} evolves as (3.3). After production, firm draws an operating cost ψ from distribution G, and decides whether to pay the cost to continue operating. If the firm decides to exit the market, it scraps the exit value, and can not reenter the market in the future. For ease of calculation, ψ is measured in terms of habit-adjusted aggregate consumption. Firms discount each period's profit with a one-period stochastic discount factor $q(\Lambda, \Lambda')$, which depends on the current aggregate state Λ, and the future aggregate state Λ'. Also, firms take the habit-adjuted price index $\tilde{p}(\Lambda)$, and the wage $w(\Lambda)$ as given, where both variables depend on the current aggregate state Λ.

[4]Aggregate productivity A_t is the source of aggregate fluctuation in my model. But note that the result of my model does not depend on whether it is supply-side shock or demand-side shock. I have constructed a partial equilibrium model with an exogenous aggregate demand, and all the results in this paper hold. This is in contrast to a New Keynesian model, in which markup is counter-cyclical against demand-side shock, while pro-cyclical against supply-side shock.

Given customer capital b, idiosyncratic productivity a, and aggregate state Λ, an incumbent's dynamic problem is

$$V(b,a,\Lambda) = \max_{y,l,p,b'} \; py - w(\Lambda)l + \int_{\psi} \max\left\{0, -\tilde{p}(\Lambda)\psi + \mathrm{E}_{a,\Lambda}\left[q(\Lambda,\Lambda')V(b',a',\Lambda')\right]\right\} dG(\psi)$$
$$\text{s.t. constraints (3.3), (3.4), and (3.9) satisfy.} \tag{3.12}$$

Without loss of generality, I normalize the exit value to be zero. An incumbent decides to stay in the economy if and only if $\psi \leq \psi^*$, where ψ^* is firm-specific threshold value of operating cost implicitly defined by

$$\tilde{p}(\Lambda)\psi^*(b',a,\Lambda) = \mathrm{E}_{a,\Lambda}\left[q(\Lambda,\Lambda')V(b',a',\Lambda')\right]. \tag{3.13}$$

Since optimal choice of next period's customer capital b' is based on state variables $\{b,a,\Lambda\}$, so the threshold value can also be written as $\psi^*(b,a,\Lambda)$.

The mass of exiting firms at time t is

$$N_t^x = \int\int \Pr\left[\psi \geq \psi^*(b,a,\Lambda)\right] d\Gamma_t(b,a), \tag{3.14}$$

where x denotes *exit*. And the mass of incumbents continue to next period is

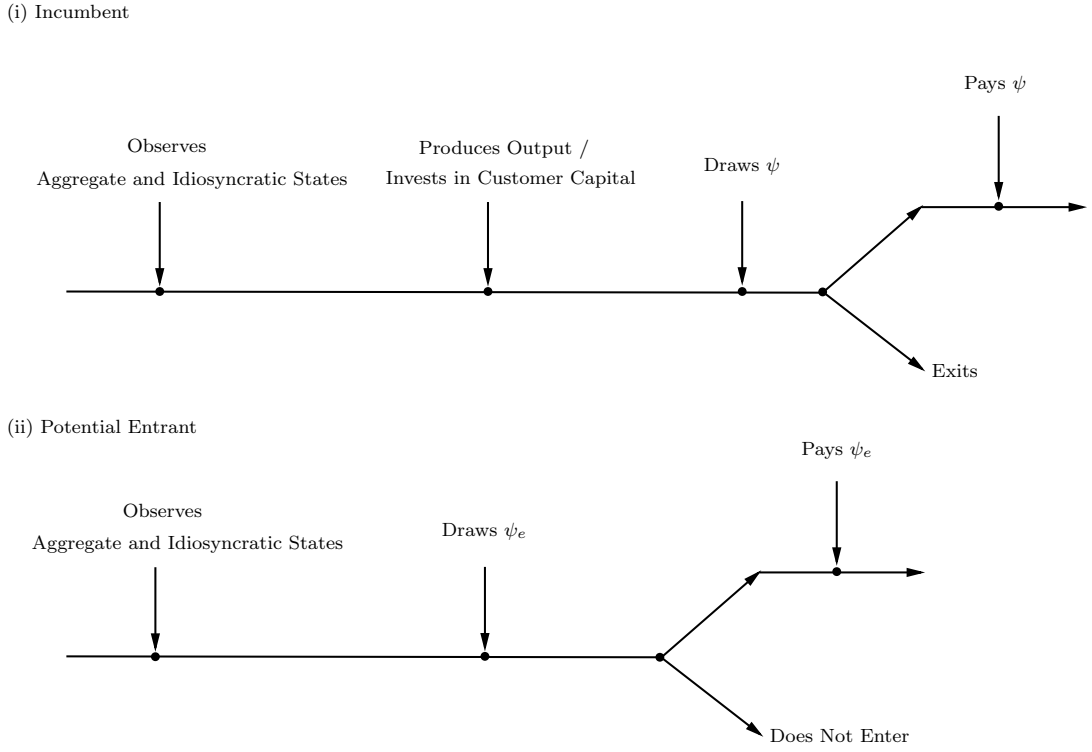
$$N_t^c = N_t - N_t^x, \tag{3.15}$$

where c denotes *continuation*.

Markup Dynamics

To help understand the intuition of the model, I write down the equation that summarizes the markup dynamics below (For full derivation of FOCs for incumbent's problem, please

Figure 3.1: Timing of Firm's Decision in Period t



see Appendix C):

$$\mu_{it}^{-1} - \bar{\mu}^{-1} = G(\psi_{it}^*)\mathrm{E}_t \left\{ q_{t,t+1} \left[(1-\delta)\frac{p_{it+1}}{p_{it}} \left(\mu_{it+1}^{-1} - \bar{\mu}^{-1} \right) + \delta \frac{\theta(\rho-1)}{\rho} \frac{p_{it+1} y_{it+1}}{p_{it} b_{it+1}} \right] \right\}, \quad (3.16)$$

where μ_{it} is firm i's price markup over marginal cost, $\bar{\mu} \equiv \frac{\rho}{\rho-1}$ is monopolistic firm's constant price markup, $G(\psi_{it}^*)$ is firm i's probability of staying in the market in the next period $t+1$, $q_{t,t+1}$ is the stochastic discount factor. Notice that when customer capital stays at a constant ($\delta = 0$), or there is no good-specific habit formation ($\theta = 0$), then the firm always charges a monopolistic markup in a standard economy $\mu_{it} = \bar{\mu}$. In contrast, when a firm faces a customer market, its markup is always below the monopolistic markup $\mu_{it} < \bar{\mu}$[5]. This is

[5]One can easily see this by iterating (3.16), and get the following:

$$\mu_{it}^{-1} - \bar{\mu}^{-1} = \frac{\delta}{1-\delta} \frac{\theta(\rho-1)}{\rho} \mathrm{E}_t \left\{ \sum_{j=1}^{\infty} q_{t,t+j} (1-\delta)^j \left[\Pi_{j'=0}^{j} G(\psi_{it+j'}^*) \right] \frac{p_{it+j} y_{it+j}}{p_{it} b_{it+j}} \right\}$$

Since all the right-hand-side terms are positive, we have $\mu_{it} < \bar{\mu}$.

because a firm has an incentive to *invest* in its customer capital, hence it does not fully exploit its market power and sets a lower markup than $\bar{\mu}$. Hence, in the equilibrium, whenever a firm raises its markup, a firm always has an option to increase its markup close to the level of a monopolist to boost the short term profits.

To help pin down what factors affect the movements of markups, I consider a special case where the lagged customer capital fully depreciates $\delta = 1$[6]. In this case, the customer capital in the next period is equal to the current output $b_{it+1} = y_{it}$ according to (3.3), and (3.16) simplifies to:

$$\mu_{it} = \bar{\mu} \left\{ 1 + \theta G(\psi_{it}^*) \mathrm{E}_t \left(q_{t,t+1} \frac{p_{it+1} y_{it+1}}{p_{it} y_{it}} \right) \right\}^{-1}. \tag{3.17}$$

Notice that there are three factors that come into play: (i) the probability of staying in the market in the next period $G(\psi_{it}^*)$; (ii) the stochastic discount factor $q_{t,t+1}$; (iii) the revenue growth rate of the firm $\frac{p_{it+1} y_{it+1}}{p_{it} y_{it}}$. In the original *deep habits* model by Ravn, Schmitt-Grohe, and Uribe (2006), the second factor is the main driver of markup cyclicality. Given a procyclical stochastic discount factor, a firm values the benefit of extra future customer capital more in booms rather than in recessions, thus resulting in countercyclical markups[7]. Since the stochastic discount factor is common across all firms, it affects all firms by the same magnitude[8].

However, in my model, I want to emphasize two other channels that bring heterogeneity into markup dynamics. First, the movement of $G(\psi_{it}^*)$ is heterogeneous among firms. In downturns, small firms are less likely to stay in the market than large firms. Hence, even

[6] I choose $\delta = 0.19$ in the quantitative exercise.

[7] This result can be overturned in other models. Opp, Parlour, and Walden (2014) study a theoretical model of price wars among oligopolistic firms as in Rotemberg and Saloner (1986). They extend the analysis in a general equilibrium, and find that if stochastic discount factor is highly pro-cyclical, present value of future cooperation could be higher than current profit from deviation, thus resulting in procyclical markup.

[8] With a representative household with access to complete market, the stochastic discount factor is the same for all firms. This assumption can be relaxed to assume heterogenous households with incomplete market, then, depending on the ownership structure of the firm, the stochastic discount factor will be different across firms. This might be an interesting extension future work, but it is not the main focus of this paper.

discounting the future profits with the same stochastic discount factors, small firms have greater incentives to exploit their current customers with high markups. Maintaining a large customer capital is an attractive option only if firms continue operations in the market. Benefit of maintaining their own customer base is meaningless if they do not stay in the market any more. To test this channel empirically in the data, I include an interaction term of output and a dummy variable for exit in the fixed-effect regression (1.11). Exit is defined as the year of the firm's final appearance in the data. For this empirical analysis, I restrict my sample to year 2005 – 2010 for France. The reasons of not using the whole sample period are twofold. First, the data tends to drop a firm that does not report for a certain period, which creates potential survival bias. Second, any online download of the data (WRDS) caps the number of firms to be downloaded, which makes exit dummy a noisy measure. I choose year 2005 to 2010 because this time period has an average exit rate close to the estimates in Bellone, Musso, Nesta, and Quere (2006). The results are reported in Table 3.1. Column (1) is the baseline results from Section 2. Column (2) shows the results with interaction of output and exit dummy, and we see that the coefficient is negative and significant at 10%-level. The reason that it is not significant at 5% is probably due to two reasons. First, censoring of samples in the data makes the measure of exit noisy. Second, firm's exit decision might contain more randomness in normal periods, hence the exit dummy does not fully reflect the firm's optimal decision. Hence, I interact the aggregate output with both exit dummy and recession dummy, equal to one if year is either 2008 or 2009. The results are in column (3). We see that the coefficient of exit dummy during recessions is negative and significant. This result suggests that exit decision during recessions are more endogenous.

Second, the revenue growth rate $\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}$ is different for firms in general. Cross-sectionally, firms with high expected growth rates charge low markups to quickly attain optimal size of customer capital. Over business cycles, small firms tend to grow faster after recessions, and *vice versa*, contributing to more procyclical markups for them. However, the

effect of this channel is quantitatively small. I have calibrated a model features heterogenous firms but without entry and exit, and I find that the difference of markup cyclicality among firms is small.

Entrant

Every period, there is a mass of prospective entrants M_t

$$M_t = 1 - N_t. \tag{3.18}$$

The total number of varieties in the economy is fixed, and is normalized to one here. Only one firm can produce good of variety i in the economy. A potential entrant can only enter into the production line of variety i absent of an incumbent [9].

Before entering the market, a potential entrant draws a signal $\tilde{a}$ about its idiosyncratic productivity if it becomes an incumbent. $\tilde{a}$ is drawn from the distribution $h(\tilde{a})$. Foster, Haltiwanger, and Syverson (2008) find that the mean of entrants' physical productivities are not different from incumbents. Hence, I assume that $h(\cdot)$ is the stationary distribution of incumbent's idiosyncratic productivity AR(1) process, $\log \tilde{a} \sim N(0, \sigma_a/(1-\rho_a^2))$[10]. And its actual idiosyncratic productivity upon entry a' follows the same process as incumbent's AR(1) process: $\log a' = \rho_a \log \tilde{a} + \varepsilon^a$. To enter the economy in the next period, it has to pay an entry cost ψ_e, which is drawn from a distribution denoted as $G^e(\psi_e)$[11]. All entrants are endowed with the same level of initial customer capital b_0[12]. Note that even though the

[9] There is a one-to-one mapping between variety and firm. Clementi, Khan, Palazzo, and Thomas (2014) use the same assumption. They assume that there is a fixed stock of blueprints in the economy, and only a blueprint not used by an incumbent can be used by an entrant. This assumption allows depletion effects of incumbents.

[10] There are other choices of potential entrants' productivities distribution in the literature. They include, but not limited to, uniform distribution in Hopenhayn and Rogerson (1993), exponential distribution in Lee and Mukoyama (2013), and pareto distribution in Clementi and Palazzo (2013).

[11] Given the discrete nature of my solution technique, randomness of entry cost helps smooth the response of number of entrants along business cycle, and hence facilitates the search for general equilibrium solution. Clementi, Khan, Palazzo, and Thomas (2014) use the same assumption.

[12] This assumption can be relaxed by allowing the firms to endogenously choose the level of initial customer capital by paying some sunk costs. We can think of the sunk costs as advertising costs of its product.

entrants draw their productivities from the same distribution as incumbents, they are still smaller because they have less initial product demand b_0.

Given a draw of entry cost ψ_e from $G^e(\cdot)$, initial customer capital b_0, signal $\tilde{a}$, and aggregate A potential entrant's problem can be written as

$$V_e(\psi_e, b_0, \tilde{a}, \Lambda) = \max\left\{0, -\tilde{p}(\Lambda)\psi_e + \mathrm{E}_{\tilde{a},\Lambda}\left[q(\Lambda, \Lambda')V(b_0, a', \Lambda')\right]\right\}, \tag{3.19}$$

A potential entrant enters the economy if and only if the entry cost ψ_e is less than or equal to a threshold $\psi_e^*(b_0, \tilde{a}, \Lambda)$, which is implicitly given by the following:

$$\tilde{p}(\Lambda)\psi_e^*(b_0, \tilde{a}, \Lambda) = \mathrm{E}_{\tilde{a},\Lambda}\left[q(\Lambda, \Lambda')V(b_0, a', \Lambda')\right]. \tag{3.20}$$

And since all entrants have the same initial customer capital b_0, for any $t \geq 0$, denote entrant's distribution as $\Gamma_t^e(a')$. And the mass of actual entrants into the next period is

$$N_{t+1}^e = M_t \int \Pr[\psi_e \leq \psi_e^*(b_0, \tilde{a}, \Lambda)] dG^e(\psi_e). \tag{3.21}$$

3.2.3 Recursive Competitive Equilibrium

A ***Recursive Competitive Equilibrium*** of the economy is a list of functions, (i) household policy functions $\mathbf{c}(\Lambda)$, $L_s(\Lambda)$, and $\mathbf{D}(\Lambda)$, (ii) firm value functions and aggregate profits $V(b, a, \Lambda)$, $V_e(b_0, \tilde{a}, \Lambda)$, and $\Phi(\Lambda)$, (iii) firm policy functions $y(b, a, \Lambda)$, $l(b, a, \Lambda)$, $p(b, a, \Lambda)$, and $b'(b, a, \Lambda)$, (iv) wage and prices $w(\Lambda)$, $\tilde{p}(\Lambda)$ and $q(\Lambda, \Lambda')$, and (v) entrant's distribution Γ^e, and incumbent's distribution Γ such that

1. Taking $\mathbf{p}$, $w(\Lambda)$, $q(\Lambda, \Lambda')$, and $\Phi(\Lambda)$ as given, $\mathbf{c}(\Lambda)$, $L_s(\Lambda)$, and $\mathbf{D}(\Lambda)$ solve household's problem, where $\mathbf{p}$ is a vector of available goods' prices.

2. Taking $\tilde{p}(\Lambda)$, $w(\Lambda)$, and $q(\Lambda, \Lambda')$ as given, $V(b, a, \Lambda)$, $V_e(b_0, \tilde{a}, \Lambda)$, $y(b, a, \Lambda)$, $l(b, a, \Lambda)$, $p(b, a, \Lambda)$, and $b'(b, a, \Lambda)$ solve firm's problem.

3. Aggregate profit is given by

$$\Phi(\Lambda) = \int [p(b,a,\Lambda)y(b,a,\Lambda) - w(\Lambda)l(b,a,\Lambda)]\, d\Gamma(b,a). \tag{3.22}$$

4. The budgets constraint (3.6) satisfies.

5. $w(\Lambda)$ is given by (3.7).

6. $\tilde{p}$ is given by (3.5).

7. $q(\Lambda, \Lambda')$ is given by (3.8).

8. For all measurable sets of entrants' idiosyncratic productivities $\mathcal{A}$,

$$\Gamma^e(\mathcal{A}) = M \int\int\int \mathbb{1}\{a' \in \mathcal{A}\} \times \mathbb{1}\{\psi_e \leq \psi_e^*(b_0, \tilde{a}, \Lambda)\} \times d\Pi_a(a'|\tilde{a}) \times dh(\tilde{a}) \times dG^e(\psi_e), \tag{3.23}$$

where M and ψ_e^* are given by (3.18) and (3.20), respectively.

9. For all measurable sets of customer capital and idiosyncratic productivities $\mathcal{B} \times \mathcal{A}$,

$$\begin{aligned}\Gamma'(\mathcal{B} \times \mathcal{A}) = \int\int\int & \mathbb{1}\{b'(b,a,\Lambda) \in \mathcal{B}\} \times \mathbb{1}\{a' \in \mathcal{A}\} \times \mathbb{1}\{\psi \leq \psi^*(b,a,\Lambda)\} \\ & \times d\Gamma(b,a) \times d\Pi_a(a'|a) \times dG(\psi) + \mathbb{1}\{b_0 \in \mathcal{B}\} \times \Gamma^e(\mathcal{A}),\end{aligned} \tag{3.24}$$

where ψ^* is given by (3.13).

3.2.4 Equilibrium Computation

In this section, I sketch the outline of the equilibrium computation. The solution algorithm I employ is closely related to the ones in Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2014), and Clementi and Palazzo (2013). For readers who are interested in the detail of the implementation, please refer to Appendix C.

First, firm's problem can be expressed in terms of the habit-adjusted price index. In particular, incumbent and entrant's value function can be normalized by the habit-adjusted price index as $\tilde{V} \equiv V/\tilde{p}$, and $\tilde{V}_e \equiv V_e/\tilde{p}$, respectively. Hence, the normalized value functions for an incumbent becomes

$$\tilde{V}(b,a,\Lambda) = \max_{y,l,p,b'} \frac{py - w(\Lambda)l}{\tilde{p}(\Lambda)} + \int_{\psi} \max\left\{0, -\psi + \mathrm{E}_{a,\Lambda}\left[\frac{\tilde{p}(\Lambda')}{\tilde{p}(\Lambda)} q(\Lambda,\Lambda')\tilde{V}(b',a',\Lambda')\right]\right\} dG(\psi),$$

and the normalized value function for an entrant becomes:

$$V_e(\psi_e, b_0, \tilde{a}, \Lambda) = \max\left\{0, -\psi_e + \mathrm{E}_{\tilde{a},\Lambda}\left[\frac{\tilde{p}(\Lambda')}{\tilde{p}(\Lambda)} q(\Lambda,\Lambda')V(b_0,a',\Lambda')\right]\right\}.$$

The advantage of the transformation above is that the expression of the real wage $w(\Lambda)/\tilde{p}(\Lambda)$ is given by the household's FOC (3.7), and the real stochastic discount factor is simply

$$\frac{\tilde{p}(\Lambda')}{\tilde{p}(\Lambda)} q(\Lambda,\Lambda') = \beta\left(\frac{X(\Lambda')}{X(\Lambda)}\right)^{-\sigma}, \tag{3.25}$$

which only depends on the ratio of the habit-adjusted aggregate consumption X.

In the Appendix C, I show that the vector of aggregate state variables Λ actually consists of the aggregate productivity A, and the distribution of firm $\Gamma(b,a)$ over the customer capital b and the idiosyncratic productivity a, denotes as $\Lambda = \{A, \Gamma\}$. In general, the firm distribution Γ is an infinitely-dimensional object, hence it is computational impossible to perfectly keep track of it. Instead, following the strategy as suggested in Krusell and Smith (1998), and its application in Clementi and Palazzo (2013), in which firms face similar entry and exit problems as in mine, I conjecture that the law of motion for habit-adjusted aggregate consumption is given by the following forecast rule:

$$\log X_{t+1} = \beta_{X0} + \beta_{X1}\log X_t + \beta_{X2}\log A_{t+1} + \beta_{X3}\log A_t. \tag{3.26}$$

As in Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2014), I assume that the inverse of Frisch elasticity of labor supply $\nu = 0$, which implies that real wage is a linear function of habit-adjusted aggregate consumption $w_t/\tilde{p}_t = \omega X_t^{\sigma}$. This means that I do not need an extra forecast rule for the aggregate labor (or the real wage). Therefore, the state variables become $\{b, a, A, X\}$, and a firm uses the forecast rule (3.26) when calculating the continuation value in the future. In the solution algorithm, I keep updating the forecast rule coefficient $\beta^X = \{\beta_{X0}, \beta_{X1}, \beta_{X2}, \beta_{X3}\}$, until some accuracy level is reached. The definition of the accuracy statistics and its assessment is described in the Appendix C.

3.2.5 Empirically Comparable Variables

For the calibration, I need to define several variables that is consistent with the measurement in the data. The reasons are twofold. First, the construction of consumer price index (CPI) in the real world suffers from the well-known new product bias, since it does not put much weight on utility gain from product variety. Even if it adjusts for love-of-variety effect, it does so at a lower frequency than in my model. Second, as seen from (3.5), the welfare-consistent price index is adjusted by habit stocks, which is different from the standard composite price index. Hence, I need to adjust for these two biases for moment matching of aggregate output.

From now on, for any variable J_t in the model, I denote J_t^D for data-consistent counterpart of it. Then, the CPI $\tilde{p}_t^D$ is given by $\tilde{p}_t^D \equiv N_t^{\frac{1}{\rho-1}} \left(\int_{i \in I_t} p_{it}^{1-\rho} di \right)^{1/(1-\rho)}$, and data-consistent counterpart for aggregate output is $X_t^D \equiv (\tilde{p}_t X_t)/p_t^D$[13]. With some algebra, it can be

[13]This can be seen by considering a more general case of Dixit-Stiglitz preference:

$$X_t = N_t^{\xi - \frac{1}{\rho-1}} \left[\int_{i \in I_t} \left(c_{it} b_{it}^{\theta} \right)^{\frac{\rho-1}{\rho}} di \right]^{\frac{\rho}{\rho-1}},$$

where ξ measures the degree of love-of-variety as discussed in Benassy (1996). Its corresponding aggregate price indicator is $\tilde{p}_t = N_t^{-\xi + \frac{1}{\rho-1}} \left(\int_{i \in I_t} p_{it}^{1-\rho} di \right)^{1/(1-\rho)}$. $\xi = \frac{1}{\rho-1}$ corresponds to the usual Dixit-Stiglitz case, while $\xi = 0$ if there is no love-of-variety.

written as

$$X_t^D = \underbrace{N_t^{-\frac{1}{\rho-1}}}_{\xi\text{-bias}} \times \underbrace{\left[\frac{\int_{i\in I_t}\left(\mu_{it} l_{it}^{1-\alpha} a_{it}^{-1} b_{it}^{-\theta}\right)^{1-\rho} di}{\int_{i\in I_t}\left(\mu_{it} l_{it}^{1-\alpha} a_{it}^{-1}\right)^{1-\rho} di}\right]^{\frac{1}{1-\rho}}}_{\theta\text{-bias}} \times X_t, \tag{3.27}$$

which has two bias components, ξ-bias and θ-bias. These stand for bias due to love-of-variety, and deep habit formation, respectively. Bilbiie, Ghironi, and Melitz (2012) emphasize the ξ-bias in their quantitative analysis over business cycle. In the real world, a statistical agent adjusts for the new product bias in the CPI calculation, but at a slow pace[14]. θ-bias measures the welfare difference due to good-specific habit formation. And to see the relation of dynamics between X_t^D and X_t more clearly, I restrict to a symmetric equilibrium case[15], then the θ-bias can be simplified to b^θ, and growth in aggregate output is:

$$\Delta \log X_t^D = -\frac{1}{\rho-1}\Delta \log N_t - \theta \Delta \log b_t + \Delta \log X_t. \tag{3.28}$$

In the simulation, number of incumbents N_t is procyclical, and b_t is procyclical with a lag of one period, hence X_t^D tends to underestimate the true growth rate of output in absolute term.

3.3 Calibration

One main objective of the quantitative exercise is to see whether the model can match two key moments in the empirical findings–the magnitude and the heterogeneity of markup cyclicality. To do so, I choose common values used in literature for a subset of parameters, and calibrate the rest to match standard moments in the data.

[14]The UK added smartphone into the basket of good in 2011.

[15]Note that even if all the firms have the same productivity, they might still differ in sizes, due to the fact that entrant's customer capital is different from incumbent's, and it needs time to accumulate customer capital to catch up with the incumbent. Hence to guarantee a symmetric equilibrium, one needs a time-varying initial customer capital for entrant, which is equal to incumbent's.

3.3.1 Fixed Parameters

The externally fixed parameters are listed in Table 3.2. To be comparable with the data, I set one period in the model equal to one year. The annual discount factor equals to $\beta = 0.96$. I assume log utility in consumption, $\sigma = 1$. For labor supply, I assume inverse Frisch elasticity $\nu = 0$ as in Hansen (1985) and Rogerson (1988). And as explained above, this simplifies the equilibrium computation[16]. I set the production elasticity of labor input as $\alpha = 0.7$. I choose labor disutility parameter ω such that the steady state labor supply $L_{ss} = 1/3$.

For the parameters that govern the product demand dynamics, I use the results of structural estimation from Foster, Haltiwanger, and Syverson (2016). The elasticity of substitution is $\rho = 1.6$, the degree of habit formation is $\theta = 1.490$, and the depreciation rate of customer capital is $\delta = 0.19$. $\rho = 1.6$ implies that a simple monopolistic firm would charge a markup $\mu = 2.7$. However, due to strong incentives to invest in customer capital, the steady state markup is actually around $\mu = 1.2$, close to what most literature has found.

3.3.2 Fitted Parameters

For the remaining parameters that I choose to match the empirical moments are listed in Table 3.3. Since calibration of the full set of parameters in a general equilibrium requires significant amount of time, I split the calibration procedure into two steps. In the first step, I calibrate the parameters that govern the cross-sectional moments of firms. They are the parameters regarding the distribution of entry cost, the relative size of entrants, the distribution of operating cost, and idiosyncratic productivities. I calibrate theses parameters in the steady state equilibrium ($A_t = 1$). I assume that the entry cost ψ_e is drawn from a uniform distribution with bounds $[0, \bar{\psi}_e]$. To be consistent with the finding in Foster et al. (2008), I set $\bar{\psi}_e$ such that the entrants have the same average productivities as the incumbents. I set initial customer capital for entrants b_0 to match the relative size of entrants

[16]For robustness check, I have tried a more realistic Frisch elasticity $\nu = 2$ and laid out the detail of computation procedure in the appendix. My main result is robust to this alternative value.

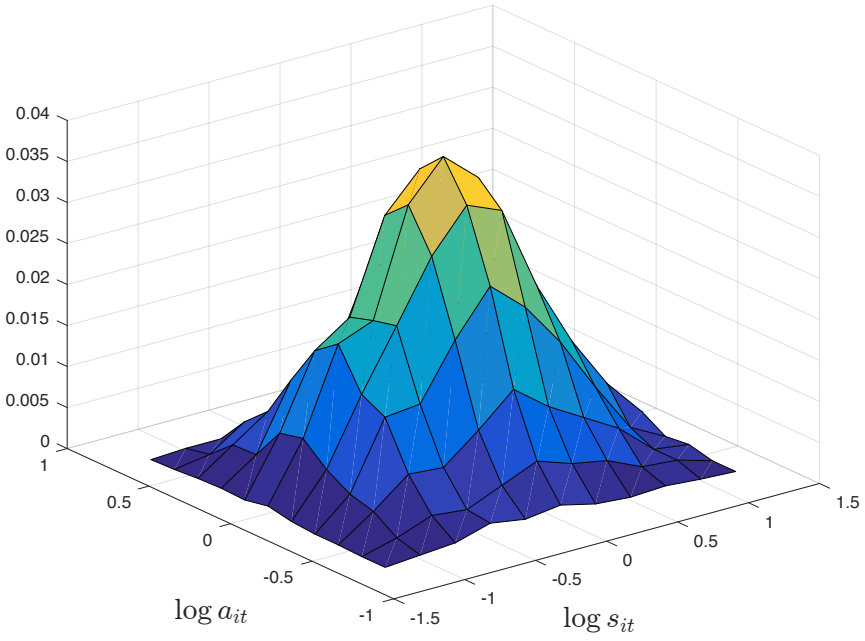
to incumbents. I assume that the operating cost ψ is drawn from a log normal distribution with mean m_ψ and standard deviation σ_ψ. I set $\{m_\psi, \sigma_\psi\}$ to match the average exit rate, and the relative size of exiters. And for idiosyncratic productivities, I set ρ_a and σ_a to match the persistence and the standard deviation of market share. Note that I only need $\rho_a = 0.83$ to match the high persistence of market share (0.96). This is because the customer capital b_{it} is also autocorrelated across periods. And this value is very close to the estimate in Foster et al. (2008)[17].

In the second step, I calibrate the persistence ρ_A and standard deviation σ_A of aggregate productivity in the general equilibrium. I set these two parameters such that the persistence and the standard deviation of X_t^D match the data.

3.4 Numerical Results and Analysis

In this section, I first describe the model's implication for firm dynamics in the steady state. Then, I discuss the numerical results of the model with aggregate fluctuations. In particular, I do impulse response analysis to analyze variables of interest. I show that my model is able to generate heterogeneity in markup cyclicality between small and large firm, as observed in the data. Moreover, in contrast to a standard business cycle model, my model predicts countercyclical firm size dispersion with only a first moment shock in aggregate productivity. And due to the increased misallocation of inputs caused by the firm size dispersion in recessions, the fluctuations of the measured TFP is amplified. Finally, I compare the baseline model to the one with homogeneous firms to see the difference.

Figure 3.2: Stationary Distribution of Firms



Note: a_{it} is idiosyncratic productivity, and s_{it} is market share. s_{it} is demeaned for illustration purpose.

3.4.1 Steady State Results

Figure 3.2 displays the stationary distribution of firms over idiosyncratic productivity a_{it} and market share in revenue $s_{it} = \frac{p_{it}y_{it}}{\sum_i p_{it}y_{it}}$[18]. In contrast to Hopenhayn (1992), where there is a one-to-one mapping between idiosyncratic productivity and size (Unless stated otherwise, size refers to market share in this paper.), customer capital b_{it} jointly affects firms' size. The correlation between idiosyncratic productivity and size is high $\rho_{\log a_{it},\log s_{it}} = 0.73$, but not perfect. Previous papers generate this pattern with firms facing adjustment costs for inputs (e.g. capital). However, there is no such adjustment costs in this model, and firms are simply constrained by the level of current demand (b_{it}). A firm with high productivity but low demand can only grow in size slowly by accumulating its customer capital.

[17]With firm-level output data, they find that the persistence of physical idiosyncratic productivity to be 0.814

[18]The market share in the figure is already demeaned.

Figure 3.3: Exit Probability of Firms: $1 - G(\psi^*)$

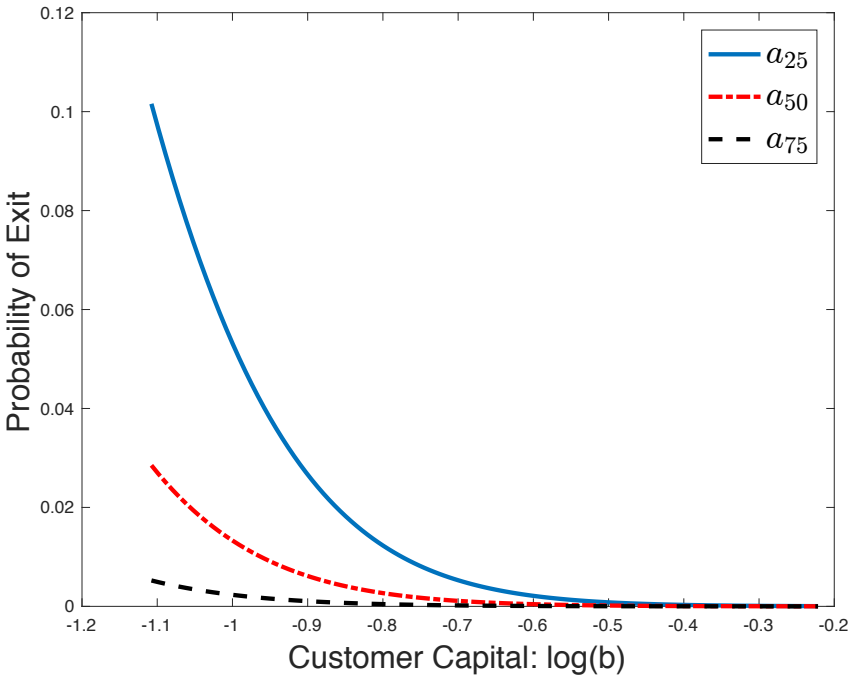
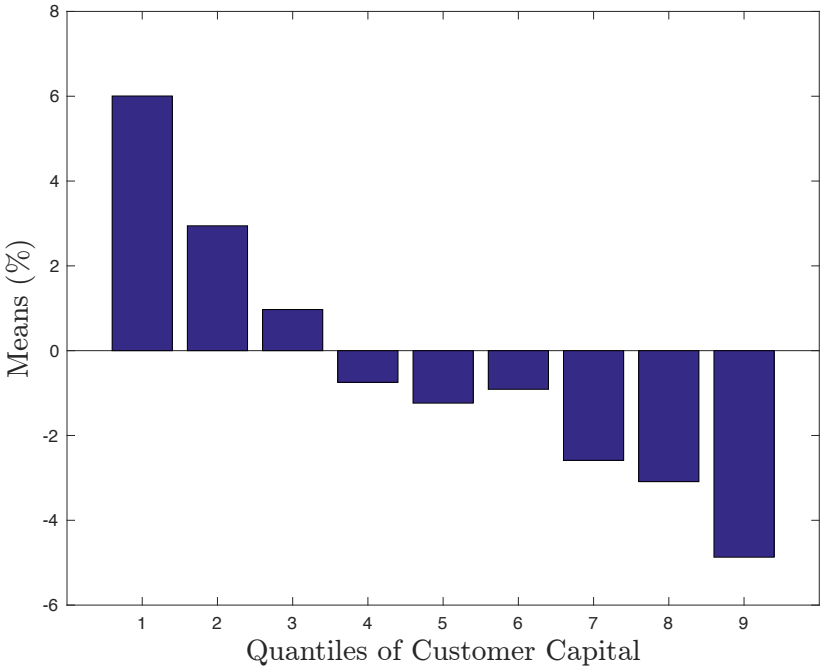


Figure 3.3 displays the steady-state exit probability of firms. I plot exit probability for firms with 25th percentile, 50th percentile, and 75th percentile of idiosyncratic productivities in the steady state. First, we note that, conditional on customer capital, the exit rate drops rapidly as productivity increases. This is due to high persistence of productivity as calibrated in the model ($\rho_a = 0.83$). A firm knows that it can keep its high productivity once drawn, hence its discounted expected profit increases substantially, and lowers the exit probability. On the other hand, for a given productivity level, exit probability smoothly increases as its customer capital shrinks. Hence, a firm with high productivity can still possibly exit due to low demand. Also, the increase of exit probability is smooth because the operating cost ψ is drawn from a continuous distribution[19].

Lastly, Figure 3.4 displays the mean of revenue growth rate conditional on customer capital. Overall, firms with low customer capital are growing, while firms with high customer capital are shrinking. This is due to two forces. The first force is that low b firms are

[19]There will be discrete jump if the operating cost is instead a singleton.

Figure 3.4: Revenue Growth Rate $\frac{p_{it+1}y_{it+1}}{p_{it}y_{it}}$



expanding its output to catch up with optimal size implied by its high a, while high b firms are shrinking to adjust to its optimal size implied by its new low a. The second force is that the productivity process is mean-reverting, and there is positive, though low[20], correlation between customer capital and productivity $\rho_{\log a_{it},\log b_{it}} = 0.35$. Overall, the revenue growth rate is quite small, since the model is calibrated to match the high persistence of market share in the data (0.96).

3.4.2 Results with Aggregate Fluctuations

Heterogeneity in Markup Cyclicality

I now move to the case with aggregate fluctuations of A_t. First, I show that my model is able to match the heterogeneity of markup cyclicality, which is the key moment to target in this paper. To compare my model to the data, for each period t in both data and simulation,

[20]It is low because the model is calibrated such that the entrants have the same average productivities as the incumbents, which are relatively larger in size.

I first split firms into five categories based on their market share s_{it}: $\mathcal{C} \in \{1, 2, 3, 4, 5\}$. Each category $\mathcal{C}$ consists of a quintile of firms[21], with 1 consisting of the smallest firms and 5 the largest firms. And then, for each category, I separately run the following first difference regression[22]:

$$\Delta \log \mu_{it} = \phi_{\mathcal{C}} \Delta \log X_t^D + \varepsilon_{it}. \tag{3.29}$$

My results are summarized in Figure 3.5. First, as discussed in the empirical section, small firms have more countercyclical markups than large firms in the data. And the result is robust against alternative specifications of the production function (Cobb-Douglas and Translog). My model generates the same magnitude of heterogeneity in markup cyclicality. For the markup cyclicality of large firms, the model matches the counterpart for Translog production function case. The result is encouraging, since the only moments that I target in the calibration are static moments, but the model successfully captures the heterogenous dynamics of markups.

Second, the model captures one half of average cyclicality when compared to Cobb-Douglas case (-0.5 vs -1.1), and 60% for Translog (-0.7 vs -1.1). This possibly implies that other models of markup cyclicality are still important. For example, one element that the model lacks is sticky price. It will be an important extension to add this feature into the model for future research to understand the interactions between the sticky price and the cyclical variations in desired markups. Nevertheless, there are possible options within the current framework to match the magnitude. I discuss the alternative functional form specifications form that would help towards the right direction below.

[21] Since the first difference regression involves the variable at time t and $t-1$, I categorize firms based on their average market share at time t and $t-1$, $\frac{s_{it}+s_{it-1}}{2}$.

[22] Note that to be consistent with the empirical work, I use the data-consistent variable X_t^D as the regressor, instead of the welfare-consistent variable X_t.

Countercyclical Firm Size Dispersion

To analyze the firm dynamics upon a negative aggregate productivity shock, I do the following Impulse Response analysis. Specifically, I run a simulation of 2000 independent economies, each with T periods. At time $t = T_0 < T$, I impose a one-and-a-half standard deviation shock to the aggregate productivity A_t, and all economies evolve normally afterwards. For each period t, I calculate the average value of interested variables and normalized to give the percentage deviation from their pre-shock values.

Figure 3.6 depicts the heterogeneous responses of firms. The middle panel depicts the markup responses of small and large firms, where small and large are defined as above. Again, a small firm's markups rise more than a large firm's upon a negative aggregate productivity shock. An interesting and important implication of the markup dynamics is its transmission into firm's output dynamics. The bottom panel shows the output responses of small and large firms. Since a small firm raises markup more than a large firm, its output declines relatively more. Hence, the dispersion of firm output increases endogenously[23]. And note that the output gap between two groups remains for a long period after the negative shock. The reason is that the law of motion for customer capital (3.3) is a persistent process, so it takes time for a small firm to catch up with a large firm.

Unfortunately, we don't observe output in the data. However, we can test the dispersion by looking at deflated value-added[24]. In my simulation, to be consistent with the data, I deflate the sales with the CPI p_t^D. For both simulation and data, I run the following regressions:

$$\Delta \log \left(\frac{p_{it} y_{it}}{p_t^D} \right) = \phi_{\mathcal{C}}^{py} \Delta \log X_t^D + \varepsilon_{it}. \tag{3.30}$$

[23]The same result holds with market share and labor.

[24]Since my model does not incorporate intermediate inputs, hence it is more natural to compare the sales in the model to value-added in the data.

where $\mathcal{C} \in \{1, 2, 3, 4, 5\}$ as above. The results are in Figure 3.7. We can see that large firms tend to have less volatile sales, both in the model and data. The model explains 62% of heterogeneity in output cyclicality.

In Figure 3.8, I plot the time profile of labor, output, and idiosyncratic productivity dispersion. Again, similar to output, the dispersion of labor goes up upon the shock. And due to decreasing returns-to-scale for labor input, the dispersion of labor is greater than the dispersion of output. In the literature, much interest has been placed on the second moment of productivity. Using the US census data, Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2014) find that second moment of productivity growth is counter-cyclical, and Kehrig (2015) find that the second moment of level of productivity is also counter-cyclical. Hence, the time-varying distribution of productivity might be driving the result of size dispersion. But interestingly, the dispersion of idiosyncratic productivities actually goes down in the model. The reasoning is that there is a strong selection effect at the entry and exit. During recessions, only firms with higher than average productivity survive, so the distribution of productivity compresses. Overall, if anything, the dispersion of productivity works against the dispersion of firm size.

Dynamics of Measured TFP

Next, I analyze the model's implications for the measured TFP. Following the standard business cycle accounting framework, I define the measured TFP Z_t as a ratio of total output divided by total inputse:

$$Z_t \equiv \frac{X_t}{L_t^{\alpha}}, \tag{3.31}$$

where L_t is total labor supply $L_t \equiv \int l(b, a, \Lambda_t) d\Gamma_t(b, a)$. With some algebra, it can be shown that:

$$Z_t = A_t \times \underbrace{\mathrm{E}(a)_t}_{\textit{Compositional Effect}} \times \underbrace{\left[\int \left(\frac{a}{\mathrm{E}(a)_t} b^{\theta} \right)^{\frac{\rho-1}{\rho}} \left(\frac{l(b, a, \Lambda_t)}{L_t} \right)^{\alpha \frac{\rho-1}{\rho}} d\Gamma_t(b, a) \right]^{\frac{\rho}{\rho-1}}}_{\textit{Misallocation Effect}}. \tag{3.32}$$

As seen above, the difference between Z_t and A_t are due to two terms associated with compositional and misallocation effect. First, $\mathrm{E}(a)_t$ is the mean of incumbents' idiosyncratic productivity. This variable is time varying due to composition of firms over business cycles. High-productivity firms survive during recessions, and low-productivity firms enter the economy during booms. So $\mathrm{E}(a)_t$ is counter-cyclical. The second term measures the change in TFP due to misallocation of labor input. The second term above $(\cdot)^{\alpha\frac{\rho}{\rho-1}}$ is a concave function, and if labor input l is more dispersed, the measured TFP Z_t is more downward biased relative to A_t due to Jensen's inequality. Finally, I denote data-consistent counterpart for measured TFP as Z_t^D, and define it as $Z_t^D \equiv X_t^D/L_t^\alpha$ accordingly.

To see the amplification of measured TFP due to misallocation, I depict the time profile of measured TFP and correlation of productivity and labor in Figure 3.9. Upon the negative shock, the measured TFP Z_t^D drops by 10% more than the true aggregate productivity A_t. However, Z_t^D surpasses A_t at $t = 3$ and remains above for 10 periods after the shock. This *rebound* phenomenon is due to the strong selection effect at the extensive margin. During recessions, the surviving firms have higher idiosyncratic productivities, which contributes to the *rebound*. Hence, to examine the amplification effect due to misallocation, I also plot the measured TFP net of the compositional effect, $Z_t^D/\mathrm{E}(a)_t$. We see that the alternative measure is always below the true productivity A_t, and converges more slowly. In fact, the persistence and the standard deviation of measured TFP (net of compositional effect) has increased by 14% and 25%, respectively. In the bottom panel, I plot the labor allocative efficiency measured by correlation between idiosyncratic productivity a_{it} and labor l_{it}. Clearly, we see that the allocation efficiency drops at the beginning, and slowly recovers to pre-shock level.

Lastly, to show the importance of heterogeneity of firm dynamics in the model, I compare the baseline model to a model that features homogeneous firms without entry and exit. In Figure 3.10, without heterogeneity, we see that the measured TFP perfectly overlaps with

the true aggregate productivity. Hence, all the amplified dynamics of the measured TFP comes from firm heterogeneity.

3.5 Alternative Functional Form Specifications

As shown in the quantitative analysis, the model successfully matches the heterogeneity of markup cyclicality, but only a half of the magnitude. In this section, I discuss two potential ways to improve on this dimension.

3.5.1 GHH Preference

The volatility of stochastic discount factor $q_{t,t+1}$ determines the magnitude of markup cyclicality. One way to increase its volatility is to increase the value of relative risk aversion from the value used in the baseline calibration, $\sigma = 1$. However, a high σ implies a strong wealth effect on labor supply. In fact, the aggregate labor becomes counter-cyclical in the simulation when $\sigma = 2$, an upper bound suggested by Chetty (2006). Therefore, to reduce the wealth effect on labor supply, one option would be to use the non-separable utility function introduced by Greenwood, Hurcowitz, and Huffman (1988)[25] as below:

$$U(X_t, L_t) = \frac{1}{1-\sigma}\left(X_t - \omega\frac{L_t^{1+\nu}}{1+\nu}\right)^{1-\sigma}. \tag{3.33}$$

Under GHH, the household's intratemporal and intertemporal conditions become the following:

$$\frac{w(\Lambda)}{\tilde{p}(\Lambda)} = \omega L^{\nu}(\Lambda), \tag{3.34}$$

$$q(\Lambda, \Lambda') = \beta\left(\frac{X(\Lambda') - \omega\frac{L^{1+\nu}(\Lambda')}{1+\nu}}{X(\Lambda) - \omega\frac{L^{1+\nu}(\Lambda)}{1+\nu}}\right)^{-\sigma}\frac{\tilde{p}(\Lambda)}{\tilde{p}(\Lambda')}. \tag{3.35}$$

[25]GHH preference is one special case of a more general preference introduced by Jaimovich and Rebelo (2009). The other special case is the one proposed by King, Plosser, and Rebelo (1988).

Hence, wealth effect on labor supply is entirely eliminated. In contrast to the baseline case, now we need two forecast rules for habit-adjusted aggregate consumption X_t and aggregate labor L_t respectively:

$$\log X_{t+1} = \beta_{X0} + \beta_{X1} \log X_t + \beta_{X2} \log A_{t+1} + \beta_{X3} \log A_t, \tag{3.36}$$

$$\log L_t = \beta_{L0} + \beta_{L1} \log X_t + \beta_{L2} \log A_t. \tag{3.37}$$

This procedure takes significantly more time for computation, and I find that the increase in markup cyclicality is modest at this stage[26].

3.5.2 Quasi-Difference Deep Habits

In the baseline model, the household's aggregate consumption depends on the quasi-product of product-specific consumption and habit stock. Instead, we can consider one that depends on the quasi-difference of consumption and habit stock, as introduced as the baseline case in Ravn et al. (2006). Specifically, the household has the following preference over consumption:

$$X_t = \left[\int_{i \in I_t} (c_{it} - \theta b_{it})^{\frac{\rho-1}{\rho}} \, di \right]^{\frac{\rho}{\rho-1}}. \tag{3.38}$$

Then household's demand for consumption c_{it} is given by

$$c_{it} = \left(\frac{p_{it}}{p_t} \right)^{-\rho} X_t + \theta b_{it}. \tag{3.39}$$

As before, the past habit stock b_{it} affects today's demand. So firms still take into account that current price decision affects future demand when making decisions. If the present value of future profit is high, firms have incentive to invest in customer capital, and vice versa.

[26]For this model, I assume that the constant relative risk aversion $\sigma = 2$, and the inverse of Frisch elasticity of labor supply $\nu = 2$, as suggested by Chetty, Guren, Manoli, and Weber (2011). The gain in average markup cyclicality is not large, probably due to strong cyclicality of labor compared to consumption, which reduces the volatility of stochastic discount factors.

And the cyclicality of stochastic discount factor determines the present value. This is called *intertemporal effect*. However, now the habit stock b_{it} enters as an additive term into the equation. This generates an extra effect called *price-elasticity effect*. The first term in the equation above is the usual price-elastic term with price elasticity ρ, and the second term is price inelastic. The total price elasticity is the weighted average between ρ and 0. An increase in X_t in boom induces the price elasticity to increase, and hence the markup goes down.

However, before using the quasi-difference form of deep habits to increase the magnitude of markup cyclicality, one needs to solve one issue. Due to the perfectly inelastic term, a monopolist actually has an incentive to set an infinite price. Schmitt-Grohe and Uribe (2007) show that any finite deviation from the symmetric equilibrium price is suboptimal, but this does not need to hold in general. One simple but not desirable solution is to enforce a price cap for firms. Finding a good remedy for this issue is beyond the scope of this paper and is left for future research.

3.6 Conclusion

This chapter established a general equilibrium model with heterogeneous firms consistent with the empirical findings. Among manufacturing firms from Amadeus dataset, firms display strong countercyclical markups, and interestingly, small firms' markups are more countercyclical than large firms'. I then build a quantitative model that features customer capital and endogenous firm entry and exit. In the model, a negative aggregate productivity shock increases small firms' exit probabilities more than large firms', and thus small firms increase their short-term profits by raising markups and put less weight on maximizing long-term profits. Also, the model has two further implications. First, because small firms have more volatile exit risks over the business cycles than large firms, their pricing, and thus output responses are more volatile. Hence, during recessions, small firms output decline more than

large firms, which gives rise to a more spread-out firm size dispersion. In addition to that, the resulting input misallocation amplifies the standard deviation of the measured TFP.

In spite of the already rich structure of the model, I did not incorporate two other features that are important for business cycle and policy analysis. First, I abstracted from the adjustment frictions for the labor input. Incorporating this element into the model is important for two reasons. The first reason is that the current model could not explain the seemingly contradictive evidence of markup cyclicality while using static inputs (material) and dynamics inputs (labor). The second one is that the incorporation would allow for a decomposition analysis of labor wedge into product market wedge (markup) and labor market wedge, but the model in this paper attributes all the labor wedge to the markup. The second feature absent from the model is nominal rigidity in price setting. With price stickiness, we can study the dynamics between inflation rate and output fluctuations, and provide guidance for policymakers to obtain a better monetary policy. Adding these two features will be important extensions for my future research agenda.

Table 3.1: Heterogeneity in Markup Cyclicality: Exit Indicator

Dep. Variable $\log \mu_{it}$	(1)	(2)
$\log Y_t$	-1.39***	-1.36***
	(0.17)	(0.17)
$\log Y_t \times s_{it}$	3.84*	3.81*
	(1.81)	(1.81)
$\log Y_t \times \text{EXIT}_{it+1}$		-0.37***
		(0.09)
N	44343	44343
adj. R^2	0.93	0.93

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note: $\log Y_t$ is log aggregate output quadratically detrended. s_{it} stands for market share in a 4-digit industry. Age_{it} stands dummy for firm exit, which is equal to one if firms exits in the next period. Standard errors are clustered at time level.

Table 3.2: Fixed Parameter

Parameter	Symbol	Value
Annual Discount Factor	β	0.96
Coefficient of Relative Risk Aversion	σ	1
Inverse of Frisch Elasticity of Labor Supply	ν	0
Degree of Production Returns-to-Scale	α	0.7
Steady-state Labor Supply	L_{ss}	1/3
Elasticity of Substitution	ρ	1.6
Degree of Habit Formation	θ	1.490
Depreciation Rate of Customer Capital	δ	0.19

Table 3.3: Fitted Paramters

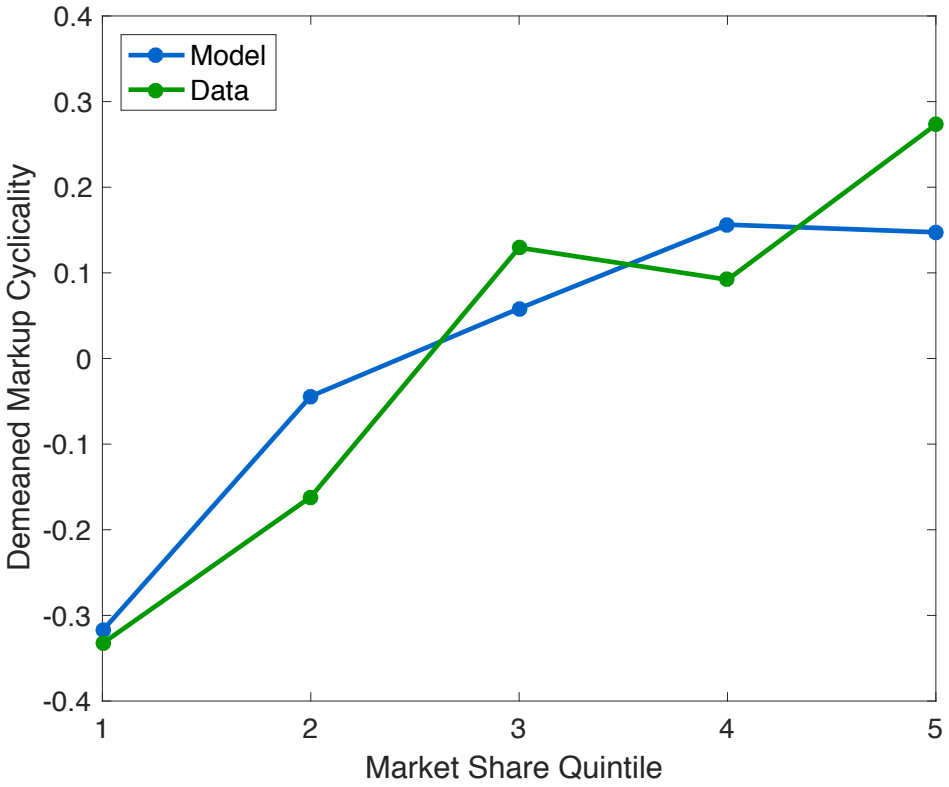
Paramter	Symbol	Value
Entry Cost	ψ_e	0.12
Initial Customer Capital	b_0	$0.58 \times b_{ss}$
Mean Parameter of Operating Cost	m_ψ	-2.73
Std Deviation Parameter of Operating Cost	σ_ψ	0.48
Persistence of Idiosyncratic Productivity	ρ_a	0.83
Standard Deviation of Idiosyncratic Productivity	σ_a	0.16
Persistence of Aggregate Productivity	ρ_A	0.61
Standard Deviation of Aggregate Productivity	σ_A	0.0138

Table 3.4: Targeted Moments in Data

Target	Data	Model
Average Exit (Entry) Rate	7.5%	8.5%
Entrant's Relative Size	51%	54%
Exiter's Relative Size	41%	43%
$corr(s_{it}, s_{it-1})$	0.96	0.94
IQR of $std(\Delta \log s_{it})$	0.22	0.22
Autocorr. of Aggregate Output	0.609	0.608
Std. Dev. of Aggregate Output	0.027	0.027

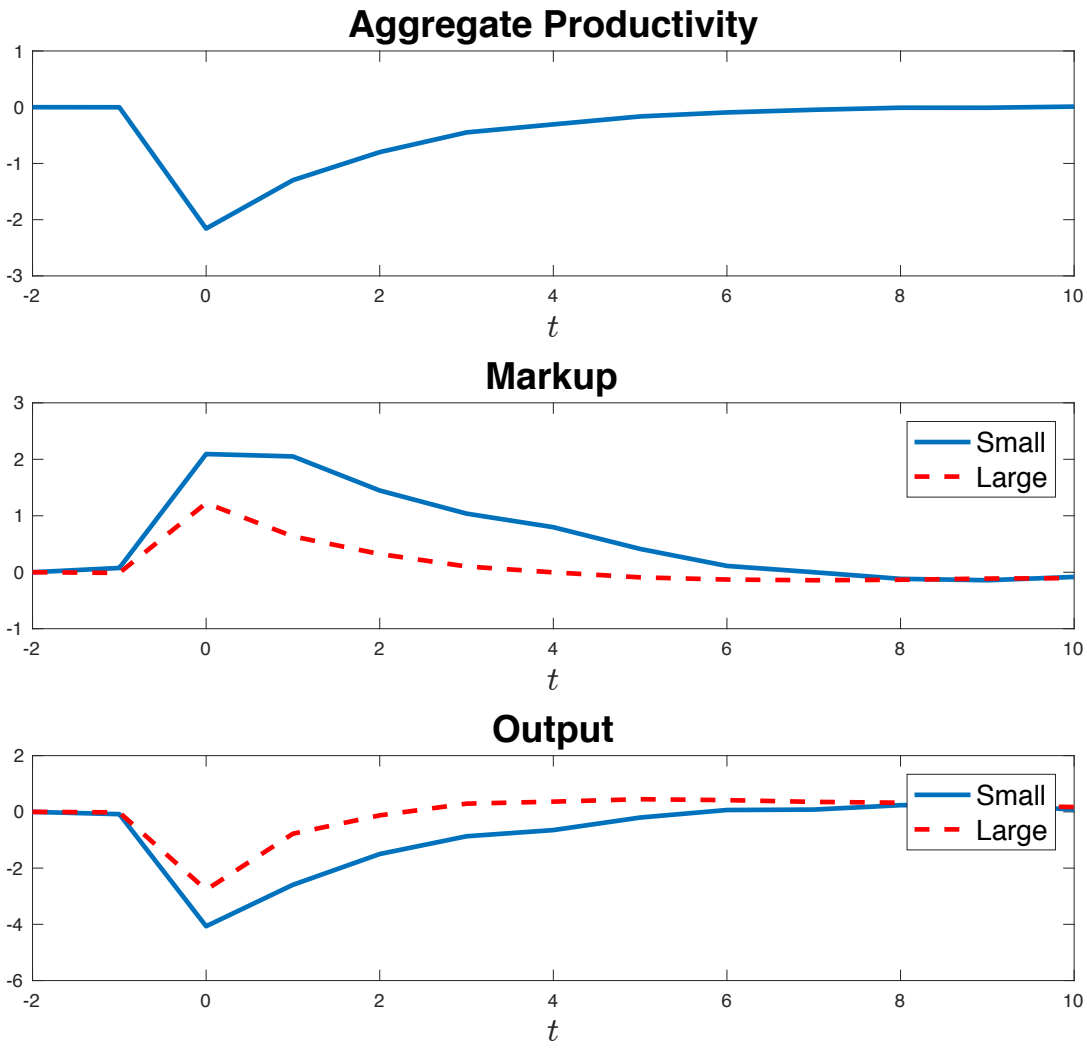
3.7 Graphs

Figure 3.5: Markup Cyclicality by Firm Size: Model vs Data



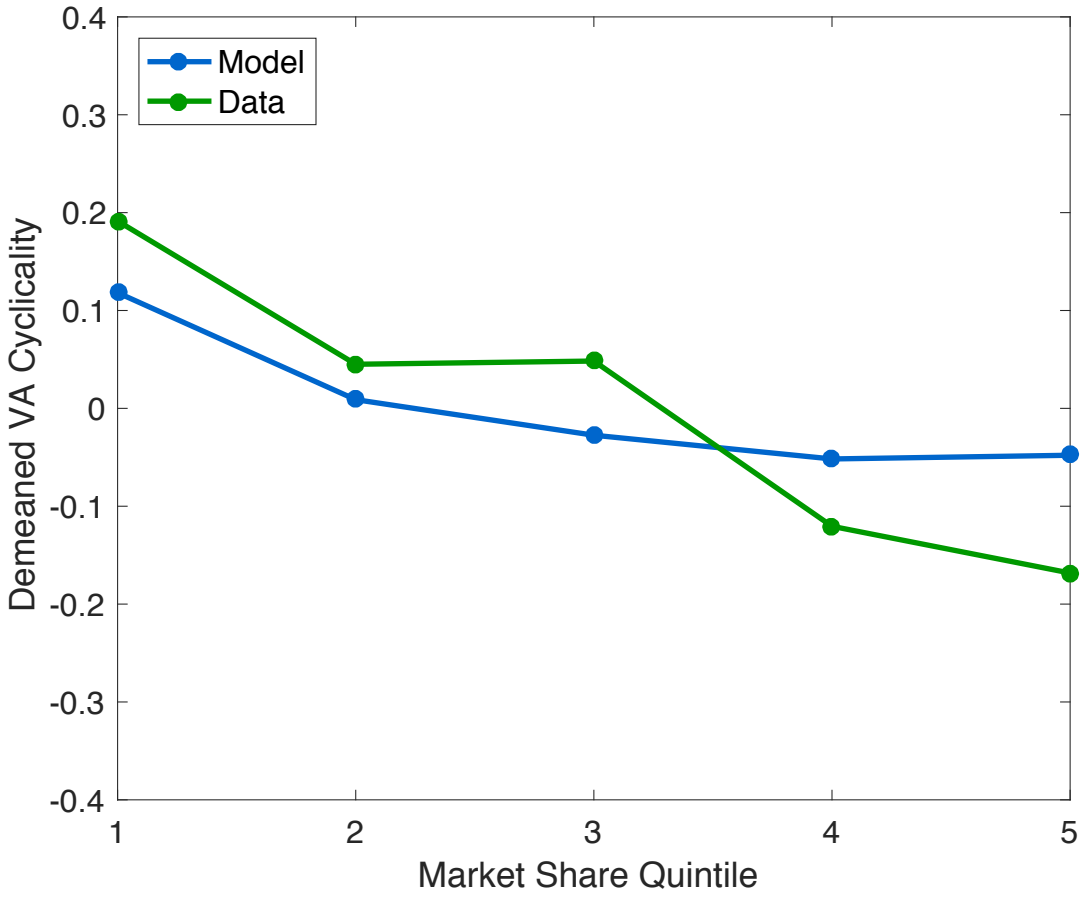
Note: Markup cyclicality with respect to aggregate consumption X_t^D. Firms in regression samples are split into five categories by market share.

Figure 3.6: Impulse Response to TFP A_t



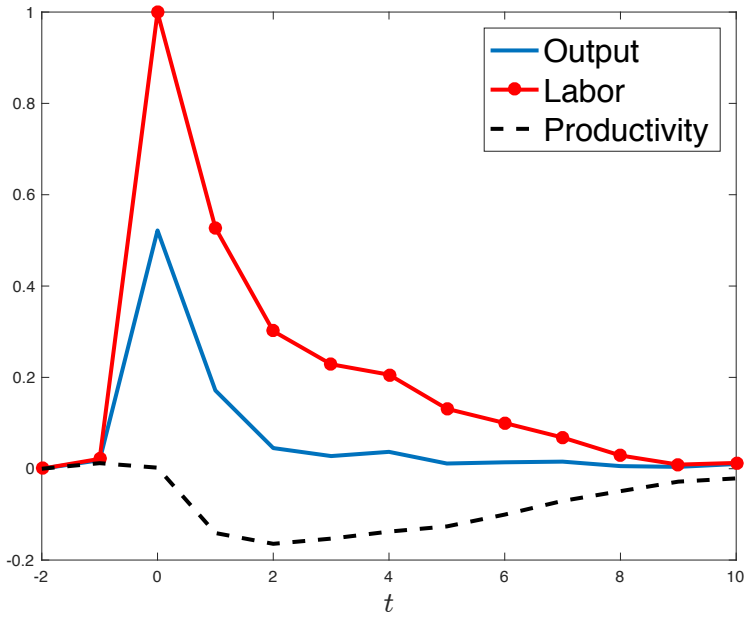
Note: Heterogeneous Responses of Small and Large Firms.

Figure 3.7: Value-Added Cyclicality by Firm Size: Model vs Data



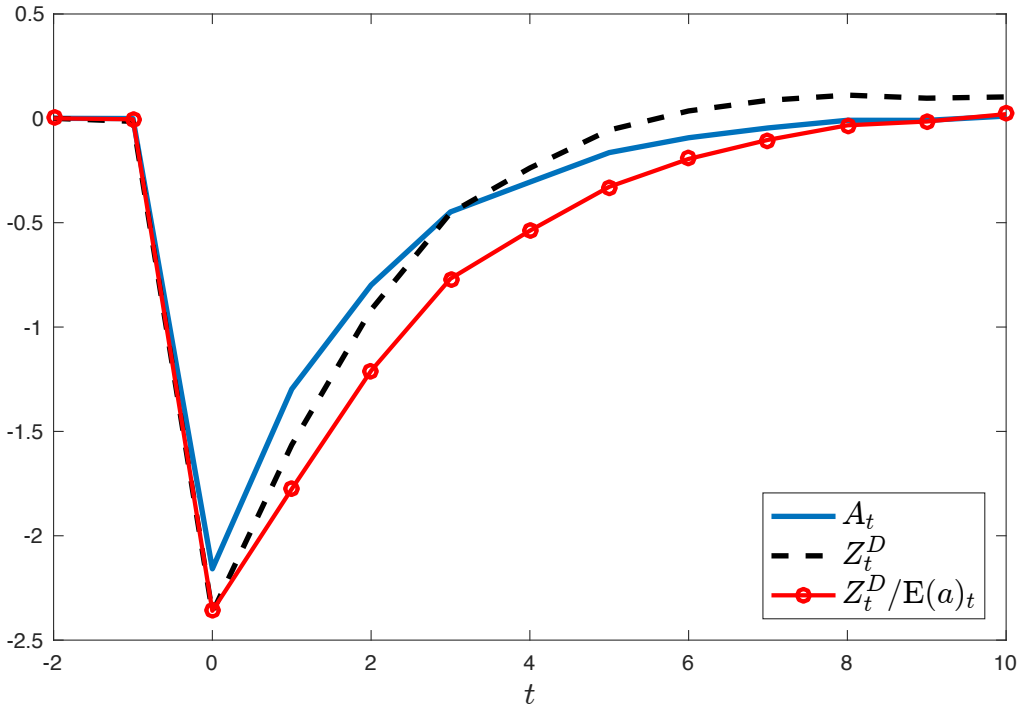
Note: Value-Added cyclicality with respect to aggregate consumption X_t^D. Firms in regression samples are split into five categories by market share.

Figure 3.8: Dispersion of Firm Size and Productivity



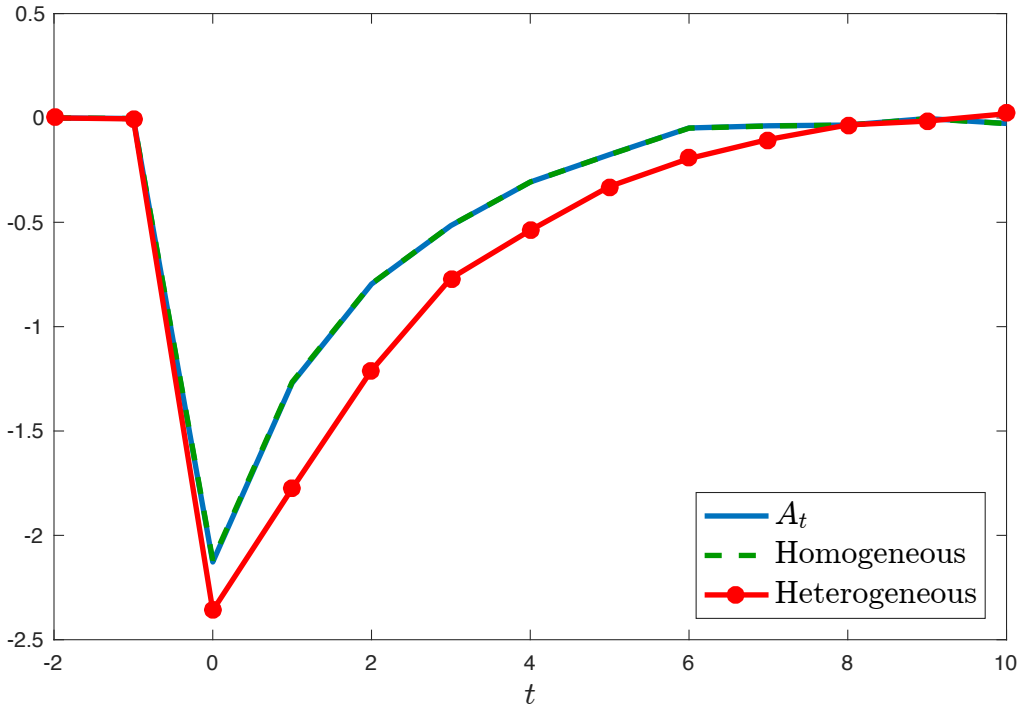
Note: The solid line depicts the standard deviation of output y_{it}. The dotted-solid line depicts the standard deviation of labor l_{it}. The dashed line depicts the standard deviation of idiosyncratic productivity a_{it}.

Figure 3.9: Comparison of Aggregate Productivity A_t and Measured TFP Z_t^D



Note: The figure depicts the time profile of aggregate productivity A_t, the measured TFP Z_t^D, and the measured TFP net of compositional effect $Z_t^D/\mathrm{E}(a)_t$.

Figure 3.10: Comparison: Homogeneous vs Heterogeneous Firm Model



Note: The Heterogenous case stands for the baseline case with heterogenous agents with entry and exit. The Homogenous case stands for the case with homogeneous firms without entry and exit.

Appendix A

A.1 Production Function Estimation

In this appendix, I present the estimation procedure for a Translog production function. For a Cobb-Douglas production function, all the steps are the same except for setting higher order terms for inputs to be zero. In particular, I estimate the following gross output production function:

$$\begin{aligned} y_{it} =& \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + \theta_{mm} m_{it}^2 + \theta_{ll} l_{it}^2 + \theta_{kk} k_{it}^2 \\ &+ \theta_{ml} m_{it} l_{it} + \theta_{mk} m_{it} k_{it} + \theta_{lk} l_{it} k_{it} + a_{it} + \epsilon_{it}, \end{aligned}$$

where m_{it} denotes material, l_{it} denotes labor, k_{it} denotes capital, a_{it} is unobserved productivity, and ϵ_{it} is some unanticipated production shock to the firm or classical measurement error. The main challenge of the estimation is controlling for the unobserved productivity a_{it}, which is in general correlated with the inputs. Olley and Pakes (1996), and Levinsohn and Petrin (2003) suggest a "control function" method to solve the simultaneity issue.

I closely follow Ackerberg, Caves, and Frazer's (2015; hereafter ACF) assumption and approach for estimation, and my procedure is the following:

1. Given the timing assumption in ACF[1], firm's demand function for material is

$$m_{it} = m(a_{it}, k_{it}, l_{it}, \mathbf{x}_{it}),$$

where $\mathbf{x}_{it}$ is a set of control variables, which include wage and time fixed effects. Assume that the demand function for material can be inverted for a_{it}, and substitute it into the production function to get

$$\begin{aligned} y_{it} =& \theta_m m_{it} + \theta_l l_{it} + \theta_k k_{it} + \theta_{mm} m_{it}^2 + \theta_{ll} l_{it}^2 + \theta_{kk} k_{it}^2 \\ &+ \theta_{ml} m_{it} l_{it} + \theta_{mk} m_{it} k_{it} + \theta_{lk} l_{it} k_{it} + m^{-1}(m_{it}, k_{it}, l_{it}, \mathbf{x}_{it}) + \epsilon_{it}. \end{aligned}$$

No parameter is identified at this first stage. The purpose of this stage is to get an estimate of the "anticipated" output $\hat{h}_{it}$ free of the error $\hat{\epsilon}_{it}$

$$y_{it} = \hat{h}(m_{it}, l_{it}, k_{it}, \mathbf{x}_{it}) + \hat{\epsilon}_{it}.$$

Since $h(\cdot)$ is a highly nonlinear function in general, I estimate it with a high-order polynomial by running an OLS regression.

2. After the first stage, for any set of parameters $\boldsymbol{\theta} = (\theta_m, \theta_l, \theta_k, \theta_{mm}, \theta_{ll}, \theta_{kk}, \theta_{ml}, \theta_{mk}, \theta_{lk})$, we can compute the implied productivity:

$$\begin{aligned} a(\boldsymbol{\theta})_{it} =& \hat{h}(m_{it}, l_{it}, k_{it}, \mathbf{x}_{it}) - \theta_m m_{it} - \theta_l l_{it} + \theta_k k_{it} - \theta_{mm} m_{it}^2 - \theta_{ll} l_{it}^2 - \theta_{kk} k_{it}^2 \\ &- \theta_{ml} m_{it} l_{it} - \theta_{mk} m_{it} k_{it} - \theta_{lk} l_{it} k_{it}. \end{aligned}$$

And by running an AR(1) regression on $a(\boldsymbol{\theta})_{it}$, we can recover the innovation shock to the productivity $\hat{\xi}_{it}(\boldsymbol{\theta})$:

$$a(\boldsymbol{\theta})_{it} = \rho_a a(\boldsymbol{\theta})_{it-1} + \xi_{it}(\boldsymbol{\theta}).$$

[1] ACF assume that labor choice is decided some time between $t-1$ and t, after k_{it} is chosen at $t-1$, but before m_{it} is chosen at t.

3. The main identification assumption is that the innovation to productivity $\xi_{it}(\boldsymbol{\theta})$ is independent of a set of lagged variables (m_{it-1}, l_{it-1}, k_{it}, m_{it-1}^2, l_{it-1}^2, k_{it}^2, $m_{it-1}l_{it-1}$,$m_{it-1}k_{it}$, $l_{it-1}k_{it}$). I use a standard GMM technique to obtain the estimation of $\boldsymbol{\theta}$. These moment conditions are similar to those in ACF.

After the estimation of the production function, I can easily recover the firm-specific markup. In particular, say for material input, the estimated markup $\hat{\mu}_{it}$ is equal to

$$\hat{\mu}_{it} = (\hat{\theta}_m + 2\hat{\theta}_{mm}m_{it} + \hat{\theta}_{ml}l_{it} + \hat{\theta}_{mk}k_{it}) \left(\frac{P_{it}^M M_{it}}{P_{it} \frac{Y_{it}}{\exp \hat{\epsilon}_{it}}} \right)^{-1},$$

where $\hat{\epsilon}_{it}$ is the estimated error term from the first-stage regression. This is to capture the timing assumption that a firm makes input decision before realizing the "unanticipated" shock ϵ_{it}.

Appendix B

B.1 Oligopolistic Competition Model

Proof of Proposition 1

If industry price index p is a continuous function of individual firm's price p_j, the symmetry of second partial derivatives holds

$$\frac{\partial^2 p}{\partial p_j \partial p_{j'}} = \frac{\partial^2 p}{\partial p_{j'} \partial p_j}.$$

Since $\frac{\partial p}{\partial p_j} = S_j$, it leads to

$$\begin{aligned}
\frac{\partial S_j}{\partial p_{j'}} &= \frac{\partial S_{j'}}{\partial p_j} \\
S_j \frac{\partial \log S_j}{\partial p_j - p}(\mathbf{1}\{j' = j\} - S_{j'}) &= S_{j'} \frac{\partial \log S_{j'}}{\partial p_{j'} - p}(\mathbf{1}\{j = j'\} - S_j) \\
\Rightarrow \frac{\partial \log S_j}{\partial p_j - p} &= \frac{\partial \log S_{j'}}{\partial p_{j'} - p} \quad \forall j, j'.
\end{aligned}$$

Derivation of Equation (11)

From Proposition 1, we have that $\frac{\partial^2 p}{\partial p_j \partial p'_j} = \Lambda S_j (S_{j'} - \mathbf{1}\{j = j'\})$, hence

$$
\begin{aligned}
&\sum_{j=1}^{n}\sum_{j'=1}^{n} \frac{\partial^2 p}{\partial p_j \partial p_{j'}} \left(\frac{\partial p_{j'}}{\partial mc_k}\right)\left(\frac{\partial p_j}{\partial mc_r}\right) \\
=& -\Lambda \sum_{j=1}^{n} S_j \left(\frac{\partial p_j}{\partial mc_k} - \sum_{j'=1}^{n} S_{j'} \frac{\partial p_{j'}}{\partial mc_k}\right) \frac{\partial p_j}{\partial mc_r} \\
=& -\Lambda \sum_{j=1}^{n} S_j \left(\frac{\partial p_j}{\partial mc_k} - \sum_{j'=1}^{n} S_{j'} \frac{\partial p_{j'}}{\partial mc_k}\right)\left(\frac{\partial p_j}{\partial mc_r} - \sum_{j''=1}^{n} S_{j''} \frac{\partial p_{j''}}{\partial mc_r} + \sum_{j''=1}^{n} S_{j''} \frac{\partial p_{j''}}{\partial mc_r}\right) \\
=& -\Lambda \sum_{j=1}^{n} S_j \left(\frac{\partial p_j}{\partial mc_k} - \sum_{j'=1}^{n} S_{j'} \frac{\partial p_{j'}}{\partial mc_k}\right)\left(\frac{\partial p_j}{\partial mc_r} - \sum_{j''=1}^{n} S_{j''} \frac{\partial p_{j''}}{\partial mc_r}\right) \\
&-\Lambda \left(\sum_{j''=1}^{n} S_{j''} \frac{\partial p_{j''}}{\partial mc_r}\right) \underbrace{\left(\sum_{j=1}^{n} \frac{\partial p_j}{\partial mc_k} - \sum_{j'=1}^{n} \frac{\partial p_{j'}}{\partial mc_k}\right)}_{=0} \\
=& -\Lambda \sum_{j=1}^{n} S_j \left(\frac{\partial p_j - p}{\partial mc_k}\right)\left(\frac{\partial p_j - p}{\partial mc_r}\right),
\end{aligned}
$$

and the rest follows.

Proof of Corollary 2

Since marginal cost shock Δmc_i are independently and identically distributed with mean zero, only $(\Delta mc_i)^2$ terms matter in expectation. Hence

$$
\mathrm{E}[\Delta\mu^{(2)}] = \frac{1}{2}\sigma^2 \left(\sum_{j=1}^{n} \frac{S_j}{1+\Gamma_j}\right)^{-1} \left[\sum_{j=1}^{n} S_j \frac{\Gamma_j}{1+\Gamma_j} \left(\sum_{k=1}^{n} X_j^{kk}\right)\right].
$$

Since I assume that all firms have the same initial states, $S_j = \frac{1}{n}$, $\Gamma_j = \Gamma'$, $\Gamma_{jj} = \Gamma''$, and $X_j^{kk} = X^{kk}$. Putting X^{kk} with equation (2.5) leads to

$$\begin{aligned}
&\sum_{k=1}^{n} X^{kk} \\
=&\frac{\Gamma''}{\Gamma'}\left(\frac{1}{1+\Gamma'}\right)^2\left[\left(\sum\frac{S}{1+\Gamma'}\right)^{-2}\left(\sum\left(\frac{S}{1+\Gamma'}\right)^2\right)-2\left(\sum\frac{S}{1+\Gamma'}\right)^{-1}\left(\frac{S}{1+\Gamma'}\right)+1\right] \\
&-\Lambda\left(\sum S\left(\frac{1}{1+\Gamma'}\right)^2\right)\left(\sum\frac{S}{1+\Gamma'}\right)^{-2}\left(\sum\left(\frac{S}{1+\Gamma'}\right)^2\right) \\
&-\Lambda S\left(\frac{1}{1+\Gamma'}\right)^2\left[n-2\left(\sum\frac{S}{1+\Gamma'}\right)^{-1}\left(\sum\frac{S}{1+\Gamma'}\right)\right] \\
=&\frac{\Gamma''}{\Gamma'}\left(\frac{1}{1+\Gamma'}\right)^2\frac{n-1}{n} \\
&-\Lambda\left(\frac{1}{1+\Gamma'}\right)^2\frac{1}{n}-\Lambda\left(\frac{1}{1+\Gamma'}\right)^2\frac{n-2}{n} \\
=&\frac{n-1}{n}\left(\frac{1}{1+\Gamma'}\right)^2\left(\frac{\Gamma''}{\Gamma'}-\Lambda\right),
\end{aligned}$$

and the rest follows.

Proof of Proposition 2

Take full log differentiation of $\log A$, and we have

$$\begin{aligned}
d\log A =&\frac{1}{\eta-1}\left[\int_0^1(\eta-1)\left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)^{-\eta}\left(\frac{a_{jt}}{A_t}\right)^{\eta-1}\frac{da_{jt}}{a_{jt}}dj\right. \\
&\left.+\int_0^1-\eta\left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)^{-\eta-1}\left(\frac{a_{jt}}{A_t}\right)^{\eta-1}d\left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)dj\right] \\
=&\int_0^1\left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)^{-\eta}\left(\frac{a_{jt}}{A_t}\right)^{\eta-1}d\log A_{jt}dj-\frac{\eta}{\eta-1}\int_0^1\left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)^{-\eta}\left(\frac{a_{jt}}{A_t}\right)^{\eta-1}d\log\left(\frac{\mathcal{M}_{jt}}{\mathcal{M}_t}\right)dj.
\end{aligned}$$

For log sectoral productivity change $d \log a_{jt}$, we have

$$d \log a_{jt} = \sum_{i=1}^{n_j} \left[\left(\frac{\mathcal{M}_{ijt}}{\mathcal{M}_{jt}} \right)^{-\rho} \left(\frac{a_{ijt}}{a_{jt}} \right)^{\rho-1} d \log a_{ijt} - \frac{\rho}{\rho - 1} \left(\frac{\mathcal{M}_{ijt}}{\mathcal{M}_{jt}} \right)^{-\rho} \left(\frac{a_{ijt}}{a_{jt}} \right)^{\rho-1} d \log \left(\frac{\mathcal{M}_{ijt}}{\mathcal{M}_{jt}} \right) \right]$$

Also, respectively, firm market share, and sectoral market share can be expressed as

$$\begin{aligned} S_{ijt} &= \left(\frac{P_{ijt}}{P_{jt}} \right)^{1-\rho} \\ &= \left(\frac{\mathcal{M}_{ijt}/a_{ijt}}{\mathcal{M}_{jt}/a_{jt}} \right)^{1-\rho}, \end{aligned}$$

and

$$\begin{aligned} S_{jt} &= \left(\frac{P_{jt}}{P_t} \right)^{1-\eta} \\ &= \left(\frac{\mathcal{M}_{jt}/a_{jt}}{\mathcal{M}_t/A_t} \right)^{1-\eta}. \end{aligned}$$

Substitute these into the equations above and the result follows.

B.2 Numerical Solution for CalvoPlusPlus Model

The firm's real profit of posting price P_{it} in period t is

$$\begin{aligned}\Pi_{it}^{R}(p_{it}) &= \left(\frac{p_{it}}{P_t}\right)^{1-\theta} C_t - \frac{W_t}{P_t} L_{it} \\ &= \left(\frac{p_{it}}{S_t} - \frac{\omega}{a_{it}}\right)\left(\frac{p_{it}}{S_t}\right)^{-\theta}\left(\frac{S_t}{P_t}\right)^{2-\theta},\end{aligned}$$

where I have used the identity real wage $W_t/P_t = \omega C_t$, and $C_t = S_t/P_t$ in the second line.

Hence, I can define the state space for firm i as $\mathcal{S}_{it} = \left\{\frac{p_{it-1}}{S_t}, a_{it}, \frac{S_t}{P_t}\right\}$, and rewrite the firm's value function in real term:

$$V(\mathcal{S}_{it}) = \max\{V_N(\mathcal{S}_{it}), V_A(\mathcal{S}_{it})\},$$

where the value of not adjusting price $V_N(\mathcal{S}_{it})$ and adjusting price $V_A(\mathcal{S}_{it})$ are respectively given by

$$V_N(\mathcal{S}_{it}) = \Pi_{it}^{R}(p_{it-1}) + \beta \mathrm{E}_t\left[\frac{S_t/P_t}{S_{t+1}P_{t+1}} V(\mathcal{S}_{it+1})\right],$$

and

$$V_A(\mathcal{S}_{it}) = \max_{p_{it}} \Pi_{it}^{R}(p_{it}) - \kappa_{it}\omega\left(\frac{S_t}{P_t}\right) + \beta \mathrm{E}\left[\frac{S_t/P_t}{S_{t+1}/P_{t+1}} V(\mathcal{S}_{it+1})\right].$$

Specific form of the adjustment cost κ_{it} depends on the nature of the adjustment cost. Under a menu cost model, the adjustment cost is

$$\kappa_{it} = \kappa_0 \left(\frac{p_{it}}{S_t}\right)^{\kappa_1(1-\theta)} \left(\frac{S_t}{P_t}\right)^{\kappa_1(2-\theta)}.$$

And under a calvo model, the adjustment cost is

$$\kappa_{it} = \begin{cases} 0 & \text{w.p. } \lambda_{it} \\ \bar{\kappa} & \text{otherwise,} \end{cases}$$

where the Calvo probability can be written as

$$\lambda_{it} = \lambda_0 \left(\frac{p_{it-1}}{S_{t-1}} \right)^{\lambda_1(1-\theta)} \left(\frac{S_{t-1}}{P_{t-1}} \right)^{\lambda_1(2-\theta)}.$$

In the simulation with the calibrated parameters, λ_{it} is always below one.

Appendix C

C.1 FOCs for Incumbent's Problem

C.1.1 First Order Conditions for Incumbents

Incumbent's problem can be written in terms of the habit-adjusted price index $\tilde{V} \equiv V/\tilde{p}$. And for simplicity of the derivation, I omit firm's idiosyncratic productivity a from the state variables. Then, incumbent's transformed value function becomes the following:

$$\begin{aligned}\tilde{V}(b; A, X) = \max_{y,l,p,b'} \frac{py - wl}{\tilde{p}} + \kappa \left[Al^{\alpha} - y\right] \\ + \lambda \left[(1-\delta)b + \delta y - b'\right] + \eta \left[\left(\frac{p}{\tilde{p}}\right)^{-\rho} b^{\theta(\rho-1)} X - y\right] \\ + \int \max\left\{0, -\psi + \beta \mathrm{E}\left[\left(\frac{X'}{X}\right)^{-\sigma} \tilde{V}(b'; A', X')\right]\right\} dG(\psi),\end{aligned}$$

where κ is the Lagrange multiplier for the production of output constraint, λ for law of motion for customer capital constraint, and η for firm product demand constraint, respectively. The FOCs with respect to the choice of output y, labor l, price p, and next period's customer

capital b' are the following:

$$\begin{aligned}
y: &\quad \eta = \frac{p}{\tilde{p}} - \kappa + \delta\lambda \\
l: &\quad \kappa\alpha A l^{\alpha-1} = \frac{w}{\tilde{p}} \\
p: &\quad \frac{p}{\tilde{p}} = \rho\eta \\
b': &\quad \lambda = \beta G(\psi^*)\mathrm{E}\left\{\left(\frac{X'}{X}\right)^{-\sigma}\left[(1-\delta)\lambda' + \theta(\rho-1)\eta'\frac{y'}{b'}\right]\right\}.
\end{aligned}$$

Note that normalization of the exit value to be 0 helps simplify the last FOC with Leibniz integral rule by a lot.

Combining first three FOCs yields the following

$$\frac{p}{\tilde{p}}\left(\mu^{-1} - \bar{\mu}^{-1}\right) = \delta\lambda,$$

where $\bar{\mu} \equiv \frac{\rho}{\rho-1}$ is the optimal markup in an economy without habit (δ or θ equals to zero). Note that firm's markup is always below $\bar{\mu}$. And the firm's intertemporal FOC becomes

$$\mu^{-1} - \bar{\mu}^{-1} = G(\psi^*)\mathrm{E}\left\{q\left[(1-\delta)\frac{p'}{p}\left(\mu'^{-1} - \bar{\mu}^{-1}\right) + \delta\frac{\theta(\rho-1)}{\rho}\frac{p'y'}{pb'}\right]\right\},$$

where $q = \beta\left(\frac{X'}{X}\right)^{-\sigma}\frac{\tilde{p}}{\tilde{p}'}$.

C.1.2 Symmetric Equilibrium

It follows that to have a symmetric equilibrium, only exclusion of idiosyncratic productivity is not sufficient. For any $t > 0$, all entrants need to have the same customer capital as the incumbent's, $b_0 = b_t$. Under symmetric equilibrium, habit-adjusted aggregate consumption

and price index equal to:

$$X = N^{\frac{\rho}{\rho-1}} y b^{\theta}$$
$$\tilde{p} = N^{-\frac{1}{\rho-1}} \frac{p}{b^{\theta}},$$

where N is the measure of incumbents. Then the one-period stochastic discount factor is

$$q = \beta \frac{p}{p'} \left(\frac{y}{y'}\right)^{\sigma} \left(\frac{b}{b'}\right)^{\theta(\sigma-1)} \left(\frac{N}{N'}\right)^{\frac{\rho\sigma-1}{\rho-1}}.$$

Baseline Case: $\sigma = 1$

If household has log utility function in consumption $\sigma = 1$ as in the baseline calibration, then the stochastic discount factor simplifies to

$$q = \beta \frac{Npy}{N'p'y'}.$$

And firm's intertemporal FOC becomes:

$$\mu^{-1} - \bar{\mu}^{-1} = \beta G(\psi^*) \mathrm{E} \left\{ (1-\delta) \frac{Ny}{N'y'} \left(\mu'^{-1} - \bar{\mu}^{-1}\right) + \delta \frac{\theta(\rho-1)}{\rho} \frac{Ny}{N'b'} \right\}.$$

By iterating forward, one can obtain:

$$\mu_t^{-1} - \bar{\mu}^{-1} = \frac{\delta}{1-\delta} \frac{\theta(\rho-1)}{\rho} \mathrm{E}_t \left\{ \sum_{j=1}^{\infty} \left[\beta(1-\delta)\right]^j \left[\Pi_{j'=0}^{j} G(\psi^*_{t+j'})\right] \frac{N_t y_t}{N_{t+j} b_{t+j}} \right\}.$$

One special case is that $\delta = 1$, then the equation becomes:

$$\mu_t^{-1} - \bar{\mu}^{-1} = \frac{\theta(\rho-1)}{\rho} \beta G(\psi_t^*) \frac{N_t}{N_{t+1}}.$$

Steady State without Entry and Exit

In the steady state, aggregate productivity is at a constant $A_t = A_{ss}$, $\forall t > 0$. Since there is neither firm entry nor exit, the number of incumbents stays constant and I normalize it to $N_t = 1$, $\forall t > 0$. Then a firm charges a steady state markup:

$$\mu_{ss} = \frac{\rho_{ss}}{\rho_{ss} - 1}, \tag{C.1}$$

where

$$\rho_{ss} = \rho \frac{1 - \beta(1 - \delta)}{1 - \beta(1 - \delta + \delta\theta(\rho - 1))} > \rho, \tag{C.2}$$

which means that firms' markup is below optimal markup without habit formation $\mu_{ss} < \frac{\rho}{\rho - 1}$.

In the steady state, habit stock is equal to the output $b_{ss} = y_{ss}$, habit-adjusted consumption index is $X_{ss} = y_{ss}^{1+\theta}$, and habit-adjusted price index is $\tilde{p}_{ss} = p_{ss} y_{ss}^{-\theta}$. Hence steady state output y_{ss} is

$$y_{ss} = \left[\frac{\alpha}{\mu_{ss}\omega} A_{ss}^{(1+\nu)/\alpha} \right]^{\frac{1}{\frac{1+\nu}{\alpha} + (\sigma-1)(1+\theta)}}. \tag{C.3}$$

And steady state aggregate labor is

$$L_{ss} = \left[\frac{\alpha}{\mu_{ss}\omega} A_{ss}^{-(\sigma-1)(1+\theta)} \right]^{\frac{1}{1+\nu+\alpha(\sigma-1)(1+\theta)}}. \tag{C.4}$$

C.2 Numerical Solution and Simulation

In this computational appendix, first I lay out the solution algorithm for the model. To solve the heterogeneous agents model in general equilibrium, I follow the linear forecast rule approach for firms as in Krusell and Smith (1998). Throughout the solution procedure, I provide the technical details turned out to be useful in the implementation. Also, I provide accuracy statistics regarding the forecast rule proposed below. Finally, I describe the calculation of the impulse response of the aggregate TFP shock in the model.

C.2.1 Transformation of Firm's Problem

Bellman equations describing incumbent and entrant's problem in the main text are reproduced below. Firm's value functions are normalized by habit-adjusted price index, and hence incumbent's value function becomes $\tilde{V} \equiv V/\tilde{p}$:

$$\tilde{V}(b,a,\Lambda) = \max_{y,l,p,b'} \frac{py - w(\Lambda)l}{\tilde{p}(\Lambda)} + \int_{\psi} \max\left\{0, -\psi + \mathrm{E}_{a,\Lambda}\left[\frac{\tilde{p}(\Lambda')}{\tilde{p}(\Lambda)} q(\Lambda,\Lambda')\tilde{V}(b',a',\Lambda')\right]\right\} dG(\psi),$$

and entrant's value function becomes $\tilde{V}_e \equiv V_e/\tilde{p}$:

$$V_e(\psi_e, b_0, \tilde{a}, \Lambda) = \max\left\{0, -\psi_e + \mathrm{E}_{\tilde{a},\Lambda}\left[\frac{\tilde{p}(\Lambda')}{\tilde{p}(\Lambda)} q(\Lambda,\Lambda')V(b_0,a',\Lambda')\right]\right\}.$$

As outlined in the main text, household's optimization problem yields the following two equations for the real wage and the stochastic discount factor in the equilibrium:

$$\frac{w(\Lambda)}{\tilde{p}(\Lambda)} = \omega L(\Lambda)^{\nu} X(\Lambda)^{\sigma}$$
$$q(\Lambda,\Lambda') = \beta \left(\frac{X(\Lambda')}{X(\Lambda)}\right)^{-\sigma} \frac{\tilde{p}(\Lambda)}{\tilde{p}(\Lambda')}.$$

In the calibration, I choose the parameters such that the household has log utility in consumption ($\sigma = 1$), and an infinite Frisch elasticity of labor supply ($\nu = 0$). With these parameters, the two equations above simplify to

$$\frac{w(\Lambda)}{\tilde{p}(\Lambda)} = \omega X(\Lambda)$$
$$q(\Lambda, \Lambda') = \beta \frac{\tilde{p}(\Lambda) X(\Lambda)}{\tilde{p}(\Lambda') X(\Lambda')}.$$

First, the real wage is now a function of only habit-adjusted aggregate consumption $X(\Lambda)$. Second, in the firm's problem, we see that the normalized stochastic discount factor simplifies to $\frac{\tilde{p}(\Lambda')}{\tilde{p}(\Lambda)} q(\Lambda, \Lambda') = \beta \frac{X(\Lambda)}{X(\Lambda')}$. Hence, the only aggregate variable matters to firm's problem is the habit-adjusted aggregate consumption $X(\Lambda)$, and it can be expressed as the following in log

$$\log X(\Lambda) = \log A + \log \Omega(\Lambda),$$

where $\Omega(\Lambda) \equiv \left[\int \left(a l^{\alpha}(b, a, \Lambda) b^{\theta} \right)^{\frac{\rho-1}{\rho}} d\Gamma(b, a) \right]^{\frac{\rho}{\rho-1}}$, and $l(b, a, \Lambda)$ is the optimal labor demand decision for a firm with a set of state variables $\{b, a, \Lambda\}$ in the equilibirum. Hence, we see that the sufficient variables for the aggregate state Λ are the aggregate TFP A, and the distribution of firms $\Gamma(b, a)$ ($\Lambda = \{A, \Gamma\}$).

Steady State Computation

I outline the computation for the steady state of the model with $A_t = 1$ for all t, hence A_t is omitted from the state variable. Note that idiosyncratic productivities are not omitted. Overall, it is a root finding problem for the habit-adjusted aggregate consumption X^*: (1) Given a guess of X^*, solve for firm's optimization problem[1]; (2) Compute firm's policy function of out $y(b, a, X^*)$, and calculate the habit-adjusted aggregate consumption

[1]Note that the stochastic discount factor is a constant in the steady state $q_t = \beta \left(\frac{X^*}{X^*} \right)^{-\sigma} = \beta$.

from $X^* = \left[\int \left(y(b,a,X^*)b^\theta\right)^{\frac{\rho-1}{\rho}} d\Gamma(b,a,X^*)\right]^{\frac{\rho}{\rho-1}}$; (3) Repeat until the market-clearing X^* is found.

Krusell-Smith Algorithm

To describe the dynamics of the habit-adjusted aggregate consumption $X(\Lambda)$, I need to keep track of the evolution of the firm distribution Γ. Unfortunately, Γ is an infinitely-dimensional object, and it is impossible to keep track of it perfectly in practice. Hence, to lessen the complexity of the problem, I follow the Krusell and Smith (1998) approach, and conjecture that $\log \Omega'$ is a linear function of its past variable $\log \Omega$, and the current aggregate TFP $\log A'$. Then the conjecture rule for the habit-adjusted aggregate consumption is

$$\log \hat{X}' = \beta_{X0} + \beta_{X1} \log X + \beta_{X2} \log A' + \beta_{X3} \log A. \tag{C.5}$$

Below, I will test and discuss the internal accuracy of this forecast rule with several statistics widely used in the literature.

Incumbent's Problem

Now, given the linear forecast rule in X, the aggregate state variables for a firm is reduced to $\Lambda = \{A, X\}$. With constraints (3.3), (3.4), and (3.9) substituted into the value function, incumbent's normalized value function $\tilde{V}$ can be written as

$$\begin{aligned}\tilde{V}(b,a,A,X) = \max_{b'} \Bigg\{ & b^{\theta(\rho-1)/\rho} X^{1/\rho} \left(\frac{b' - (1-\delta)b}{\delta}\right)^{\frac{\rho-1}{\rho}} - \omega X^\sigma (Aa)^{-1/\alpha} \left(\frac{b' - (1-\delta)b}{\delta}\right)^{\frac{1}{\alpha}} \\ & - \Pr[\psi \le \psi^*(b,a,A,X)] \mathrm{E}\left[\psi | \psi \le \psi^*(b,a,A,X)\right] \\ & + \beta \Pr[\psi \le \psi^*(b,a,A,X)] \mathrm{E}_{a,A}\left[\left(\frac{\hat{X}'}{X}\right)^{-\sigma} \tilde{V}(b',a',A',\hat{X}')\right] \Bigg\}.\end{aligned}$$

The operating cost ψ is drawn from a log normal distribution $\log \psi \sim N(m_\psi, \sigma_\psi)$, then probability of less than ψ^* is equal to

$$\Pr(\psi \leq \psi^*) = \Phi\left(\frac{\log \psi^* - m_\psi}{\sigma_\psi}\right), \tag{C.6}$$

where $\Phi(\cdot)$ is cdf function for standard normal distribution. And the conditional expectation of operating cost is

$$\mathrm{E}(\psi|\psi \leq \psi^*) = \exp\left(m_\psi + \frac{\sigma_\psi^2}{2}\right)\Phi\left(\frac{\log \psi^* - m_\psi - \sigma_\psi^2}{\sigma_\psi}\right) \Big/ \Phi\left(\frac{\log \psi^* - m_\psi}{\sigma_\psi}\right). \tag{C.7}$$

Entrant's Problem

Similarly, an entrant's normalized value function $\tilde{V}_e$ can be written as

$$\tilde{V}_e(\psi_e, b_0, \tilde{a}, A, X) = \max\left\{0, -\psi_e + \beta \mathrm{E}_{\tilde{a},A}\left[\left(\frac{\hat{X}'}{X}\right)^{-\sigma}\tilde{V}(b_0, a', A', \hat{X}')\right]\right\}, \tag{C.8}$$

and the entrant enters if and only if

$$\psi_e \leq \beta \mathrm{E}_{\tilde{a},A}\left[\left(\frac{\hat{X}'}{X}\right)^{-\sigma}\tilde{V}(b_0, a', A', \hat{X}')\right]. \tag{C.9}$$

C.2.2 Solution Algorithm

With the transformed Bellman equations above, I now lay out the outline of the solution algorithm. First, guess initial values for the coefficients of the forecast rule $\left(\beta_{X0}^{(1)}, \beta_{X1}^{(1)}, \beta_{X2}^{(1)}, \beta_{X3}^{(1)}\right)$ to solve the model, and perform the following iterations $m = 1, 2, 3...$ of the solution algorithm

1. Given the forecast rule $\left(\beta_{X0}^{(m)}, \beta_{X1}^{(m)}, \beta_{X2}^{(m)}, \beta_{X3}^{(m)}\right)$ from the previous iteration, solve the value functions for incumbent $\tilde{V}^m$, and entrant $\tilde{V}_e^m$.

2. With the firm's policy function, simulate the economy for T periods with some arbitrary initial conditions (A_0, Γ_0).

3. Using the simulated variables obtained from the simulation, update the forecast rule $\left(\beta_{X0}^{(m+1)}, \beta_{X1}^{(m+1)}, \beta_{X2}^{(m+1)}, \beta_{X3}^{(m+1)}\right)$ accordingly.

4. Repeat from step (1) to (3) until some convergence criteria is attained.

Below, I explain the implementation of each step of the solution algorithm laid out above in greater detail.

Step 1: Firm Problem

I solve the incumbent problem using policy function iteration on grid points, also known as Howard's improvement algorithm. I use $n_b = 400$ grid points for the choice of customer capital in the next period b'. For the exogenous process of the aggregate TFP A, and the idiosyncratic productivity a, I discretize them following the method of Tauchen (1986), and use $n_A = 7$ and $n_a = 11$ grid points, respectively. I denote Π^A and Π^a as the transition matrix probability for the aggregate TFP and the idiosyncratic productivity, respectively. Finally, I choose $n_X = 7$ grid points for the habit-adjusted aggregate consumption X. And the incumbent uses the forecast rule $\beta_X^{(m)} = \left(\beta_{X0}^{(m)}, \beta_{X1}^{(m)}, \beta_{X2}^{(m)}, \beta_{X3}^{(m)}\right)$ to calculate the continuation value. However, given the discrete nature of the solution method, the forecasted habit-adjusted aggregate consumption in the next period $\hat{X}'$ does not fall on the given grid points in general. Hence, I compute the continuation value $\mathrm{E}\tilde{V}'$ off the grid points using linear interpolation. Within each loop for the policy function iteration, I iterate the value function with 100 steps forward. The policy function iteration procedure stops when the supremum norm of the percentage change of the value function is less than 10^{-6}.

For the entrant, the signal $\tilde{a}$ is drawn from the uniform distribution with bounds $[0, \bar{\psi}_e]$. Its draw of productivity for production follows the AR(1) process $\log a = \rho_a \log \tilde{a} + \varepsilon^a$. And each potential entrant also draws an entry cost ψ_e from distribution $G^e(\cdot)$. A potential

entrant decides to enter the market in the next period if and only if the entry cost is lower than the continuation value. Once again, I need to use linear interpolation to calculate the continuation value for entrant.

Step 2: Simulation of the Model

I simulate the economy with a period of $T = 700$. I generate a series of realizations for the aggregate TFP $\{A_t\}_{t=1,...,T}$, which follows the discrete Markov chain process as discussed above. Also, I initiate the economy with an arbitrary distribution $\Gamma_0(b, a)$.

Throughout the simulation, I follow the histogram-based approach to track the cross-sectional distribution as proposed by Young (2010). This method avoids the sampling error from the Monto Carlo approach. In particular, I keep track of the distribution over the customer capital and idiosyncratic productivity $\Gamma_t(b, a)$. Within each period t, given the firm's policy function $b'(b, a, A, x)$, the next period's distribution is determined by

$$\begin{aligned}\Gamma_{t+1}(b,a) = &\sum_{j}^{n_a}\sum_{i}^{n_b} \mathbb{1}\left\{b'(b_i,a_j,A_t,X_t)=b\right\} \times \Pi^a(a|a_j) \\ &\times (1-\gamma) \times \Pr\left(\psi \leq \psi^*(b_i,a_j,A_t,X_t)\right) \times \Gamma_t(b_i,a_j) \\ &+ \sum_{i}^{n_{\tilde{a}}} \mathbb{1} \times \Pr\left(\psi_e \leq \psi_e^*(b_0,\tilde{a},A_t,X_t)\right) \times (1-\gamma) \times \Pi_a(a|\tilde{a}_i) \times h(\tilde{a}_i) \times M_t,\end{aligned}$$

where M_t is the mass of potential entrants at time t.

Within each period, the incumbent's decision must be consistent with the market clearing condition for output X_t. In particular, given a predetermined distribution of firm $\Gamma_t(b, a)$, for any guess of current habit-adjusted aggregate consumption $\dot{X}$, an incumbent with customer capital b and idiosyncratic productivity a uses the forecast rule to calculate the continuation

value, and solves the following problem

$$\tilde{V}(b,a,A,\dot{X}) = \max_{b'} \Bigg\{ b^{\theta(\rho-1)/\rho}\dot{X}^{1/\rho}\left(\frac{b'-(1-\delta)b}{\delta}\right)^{\frac{\rho-1}{\rho}} - \omega\dot{X}(Aa)^{-1/\alpha}\left(\frac{b'-(1-\delta)b}{\delta}\right)^{\frac{1}{\alpha}}$$
$$- \Pr[\psi \leq \psi^*(b,a,A,\dot{X})]\mathrm{E}\left[\psi|\psi \leq \psi^*(b,a,A,\dot{X})\right]$$
$$+\beta(1-\gamma)\Pr[\psi \leq \psi^*(b,a,A,\dot{X})]\mathrm{E}_{a,A}\left[\left(\frac{\hat{X}'}{\dot{X}}\right)^{-1}\tilde{V}(b',a',A',\hat{X}')\right]\Bigg\}.$$

The habit-adjusted aggregated consumption implied by the firm's policy function $b'(b,a,A,\dot{X})$ is

$$\ddot{X} = \left[\sum_{b,a}\left(\frac{b'(b,a,A,\dot{X})-(1-\delta)b}{\delta}b^{\theta}\right)^{\frac{\rho-1}{\rho}}\Gamma_t(b,a)\right]^{\frac{\rho}{\rho-1}}.$$

The market clearing condition for output is satisfied if and only if $\ddot{X} = \dot{X}$. In practice, I use golden section search method to search for the solution.

Lastly, there is one point worth mentioning to speed up the algorithm in the implementation. In practice, it is costly to search for the maximizer of the customer capital b' over the grid for a given $\dot{X}$ every time. Hence, I instead do a linear interpolation of the policy function obtained in step (1) along the grid for X to approximate the solution. The result turns out to be quite close to the one obtained with search for the maximizer, and the speed gain is substantial at the same time.

Step 3: Forecast Rule Update

After running the simulation for $T = 700$ periods, we have obtained a series of aggregate TFP and habit-adjusted aggregate consumption $\{A_t, X_t\}_{t=1,\ldots,T}$ [2]. To update the forecast rule, discard first T_0-1 periods of variables. This is called the "burning" process, and we try to insulate the effects of initial conditions on the equilibrium outcome as much as possible.

[2]Note that x_t is obtained from the equilibrium search within each period, not from the forecast rule.

I set $T_0 = 200$ in the implementation. Now with the simulated variables $\{A_t, X_t\}_{t=T_0,\ldots,T}$, I run the following OLS regression:

$$\log X_{t+1} = \beta_{X0} + \beta_{X1} \log X_t + \beta_{X2} \log A_{t+1} + \beta_{X3} \log A_t.$$

Denote the set of coefficients obtained from the regression above as $\hat{\beta}_X = \left(\hat{\beta}_{X0}, \hat{\beta}_{X1}.\hat{\beta}_{X2}, \hat{\beta}_{X3}\right)$. I update the forecast rule as the weighted sum of $\hat{\beta}_X$, and the forecast rule from the previous iteration $\beta_X^{(m)}$:

$$\beta_X^{(m+1)} = w_{KS}\hat{\beta}_X + (1 - w_{KS})\beta_X^{(m)},$$

where I choose the weight $w_{KS} = 0.8$ in the implementation.

Step 4: Convergence Criteria

There are several convergence metrics widely used in the literature. The most commonly used metric is the change in the forecast rule regression coefficients. Here, I use the maximum error statistics as proposed by Den Haan (2010). The Den Haan statistics is the maximum difference in log between the X_t^{KS} generated by repeatedly applying the forecast rule β_X only, and the X_t generated from the equilibrium search within each period in the simulation. The convergence criteria is that the Den Haan statistics is less than equal to 10^{-4}. The detail of the maximum Den Haan statistics is described in the subsection below.

C.2.3 Internal Accuracy Statistics

The prediction rule used by firm to solve the value function is the following:

$$\log \hat{X}' = \beta_{X0} + \beta_{X1} \log X + \beta_{X2} \log A' + \beta_{X3} \log A.$$

The first metric commonly used in the literature is the R^2 obtained from the regression of the equation above. In my model, the R^2 is greater than 99% in all of the alternative

specifications. However, Den Haan (2010) points out three main problems associated with the use of R^2: (i) R^2 assesses the goodness of fit conditional on the variables generated by the true law of motion. In other words, it only checks one-period ahead forecast error. (ii) The R^2 is an average statistics, and might hide some large errors. In particular, it will be problematic if an large error occurs when we are interested in the model's behavior in response to a large shock. (iii) The R^2 scales the forecast error by the variance of the dependent variable.

In response to this, Den Haan (2010) proposes to assess the accuracy of the model by calculating the maximum error between the actual habit-adjusted aggregate consumption, and the habit-adjusted aggregate consumption generated by iteration of the forecast rule. This method allows the error of the forecast rule to be accumulated over the time. Denote X_t as the market-clearing consumption within each period. And denote X_t^{DH} as the outcome of the iteration of the forecast rule β_X,

$$\log X_{t+1}^{DH} = \beta_{X0} + \beta_{X1} \log X_t^{DH} + \beta_{X2} \log A_{t+1} + \beta_{x3} \log A_t,$$

where the initial point of the iteration is equal to the actual habit-adjusted aggregate consumption $X_1^{DH} = X_1$. Then the maximum Den Haan statistics for a given forecast rule β_X is defined as

$$DH(\beta_X) \equiv \max_t |\log X_t - \log X_t^{DH}|.$$

Hence, we can interpret the Den Haan statistics as the maximum percentage error of the forecasted variable. In my model, the Den Haan statistics over 500 periods is 1%, which means that the error of the firm's prediction of X_t is at most 1% over 500 years. The average absolute error defined in the same fashion is much smaller (0.43%).

C.2.4 Impulse Response Calculation

In this subsection, I describe the calculation of impulse response of interested variables to an exogenous shock.

With the firm's policy function obtained from the numerical solution above, I simulate $M = 2000$ economies each with length of periods $T_M = 200$. At $T_{\text{shock}} = 190$, I impose a one-standard deviation negative shock to the aggregate TFP A_t. And after the shock period T_{shock}, the aggregate TFP A_t evolves normally according to the Markov chain.

For any interested variable Y, I denote Y_{mt} as the simulated value from economy m at period t, and define the impulse response of Y as

$$\text{IR}_Y = 100 \times \left(\frac{\bar{Y}_t}{\bar{Y}_{T_{\text{shock}}-2}} - 1 \right),$$

where $\bar{Y}_t = \frac{1}{M} \sum_m Y_{mt}$ is the average of Y_{mt} across all the simulated economies at time t. For the presented figures, I set $T_{\text{shock}} = 0$.

C.2.5 A More Realistic Frisch Elasticity of Labor Supply

In my benchmark model, I have assumed that the inverse of Frisch elasticity of labor supply to be zero, $\nu = 0$. However, this choice of parameter is based on complexity of computation, not economic findings. Chang et. al. (2011) suggest a Frisch elasticity of 0.75 in a representative model, Chang et. al. (2014) suggest a value slightly larger than unity, Chetty et al. (2011) suggest 2 that reconciles the micro and macro elasticity of labor supply. In this extension exercise, I set $\nu = 2$.

Steady State Computation

The computation for the steady state is similar to the one when $\nu = 0$, except that the real wage now becomes

$$\frac{w}{\tilde{p}} = \omega L^{\nu} X^{\sigma},$$

which depends on both aggregate labor supply l, and habit-adjusted aggregate consumption X. Hence, I need to find roots for both X^* and L^* in the following way: (1) Given a guess of (X^*, L^*), solve for firm's optimization problem; (2) Compute firm's policy function for output $y(b, a, X^*, L^*)$, and labor demand $l(b, a, X^*, L^*)$, and calculate $X^* = \left[\int \left(y(b, a, X^*, L^*)b^\theta\right)^{\frac{\rho-1}{\rho}} d\Gamma(b, a, X^*, L^*)\right]^{\frac{\rho}{\rho-1}}$, and $L^* = \int l(b, a, X^*, L^*)d\Gamma(b, a, X^*, L^*)$; (3) Repeat until the market-clearing X^* and L^* are found.

Augmented Krusell-Smith Algorithm

To compute the numerical solution for the general equilibrium model, the main challenge now is that, for a finite Frisch elasticity,we need to consider the additional labor supplied in each period. Aggregate labor supplied $L(\Lambda)$ depends on the aggregate state variable Λ, and this determines the real wage jointly with the habit-adjusted aggregate consumption $X(\Lambda)$. Firms take the real wage as given, and solve for the optimization problem. The solution of the optimization problem leads to the policy function for labor demand, $L(b, a, L(\Lambda), X(\Lambda))$. At the equilibrium, the labor supplied $L(\Lambda)$ must be equal to the labor demand integrated over all incumbents $L(\Lambda) = \int L(b, a, X(\Lambda), L(\Lambda))d\Gamma(b, a, X(\Lambda), L(\Lambda))$.

From the equilibrium condition above, we see that the aggregate labor only depends on the current aggregate state Λ, hence I conjecture the following log-linear function for aggregate labor

$$\log L_t = \beta_{L0} + \beta_{L1} \log X_t + \beta_{L2} \log A_t. \tag{C.10}$$

Hence a firm uses β_L to compute current profit, and β_X to calculate the continuation value for its optimization problem. In the simulation, I need to search for both X_t^* and L_t^* that clear the good and labor markets for each period t.

C.2.6 GHH Preference

Under GHH preference, the equilibrium computation is quite similar to the finite Frisch elasticity case ($\nu > 0$). One only needs to change the real wage and the stochastic discount factor into the following

$$\frac{w(\Lambda)}{\tilde{p}(\Lambda)} = \omega L^{\nu}(\Lambda), \tag{C.11}$$

$$q(\Lambda, \Lambda') = \beta \left(\frac{X(\Lambda') - \omega \frac{L^{1+\nu}(\Lambda')}{1+\nu}}{X(\Lambda) - \omega \frac{L^{1+\nu}(\Lambda)}{1+\nu}} \right)^{-\sigma} \frac{\tilde{p}(\Lambda)}{\tilde{p}(\Lambda')}. \tag{C.12}$$

And the rest of the implementation is exactly the same as the finite Frisch elasticity case.

Bibliography

[1] Daron Acemoglu, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. Networks, shocks, and systemic risk. Technical report, National Bureau of Economic Research, 2015.

[2] Daniel A Ackerberg, Kevin Caves, and Garth Frazer. Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451, 2015.

[3] Mary Amiti, Oleg Itskhoki, and Jozef Konings. Importers, exporters, and exchange rate disconnect. *The American Economic Review*, 104(7):1942–1978, 2014.

[4] Mary Amiti, Oleg Itskhoki, and Jozef Konings. International shocks and domestic prices: how large are strategic complementarities? Technical report, National Bureau of Economic Research, 2016.

[5] Andrew Atkeson and Ariel Burstein. Pricing-to-market, trade costs, and international relative prices. *The American Economic Review*, 98(5):1998–2031, 2008.

[6] Flora Bellone, Patrick Musso, Michel Quéré, and Lionel Nesta. Productivity and market selection of french manufacturing firms in the nineties. *Revue de l'OFCE*, (5):319–349, 2006.

[7] Jean-Pascal Benassy. Taste for variety and optimum production patterns in monopolistic competition. *Economics Letters*, 52(1):41–47, 1996.

[8] Saroj Bhattarai and Raphael Schoenle. Multiproduct firms and price-setting: Theory and evidence from us producer prices. *Journal of Monetary Economics*, 66:178–192, 2014.

[9] Florin Bilbiie, Fabio Ghironi, and Marc J Melitz. Endogenous entry, product variety, and business cycles. *Journal of Political Economy*, 120(2):304–345, 2012.

[10] Mark Bils. The cyclical behavior of marginal cost and price. *The American Economic Review*, pages 838–855, 1987.

[11] Mark Bils, Peter J Klenow, and Benjamin A Malin. Testing for keynesian labor demand. In *NBER Macroeconomics Annual 2012, Volume 27*, pages 311–349. University of Chicago Press, 2012.

[12] Mark Bils, Peter J Klenow, and Benjamin A Malin. Resurrecting the role of the product market wedge in recessions. Technical report, National Bureau of Economic Research, 2016.

[13] Olivier Blanchard and Jordi Galí. Real wage rigidities and the new keynesian model. *Journal of Money, credit and Banking*, 39(s1):35–65, 2007.

[14] Nicholas Bloom. The impact of uncertainty shocks. *Econometrica*, 77(3):623–685, 2009.

[15] Nicholas Bloom, Max Floetotto, Nir Jaimovich, Itay Saporta Eksten, and Stephen Terry. Really uncertain business cycles. *US Census Bureau Center for Economic Studies Paper No. CES-WP-14-18*, 2014.

[16] Ariel Burstein and Gita Gopinath. International prices and exchange rates. Technical report, National Bureau of Economic Research, 2013.

[17] Carlos Carvalho. Heterogeneity in price stickiness and the real effects of monetary shocks. *The B.E. Journal of Macroeconomics*, 6(3):1–58, 2006.

[18] Yongsung Chang, Sun-Bin Kim, Kyooho Kwon, and Richard Rogerson. Individual and aggregate labor supply in a heterogeneous agent economy with intensive and extensive margins. *Unpublished Manuscript*, 2014.

[19] Varadarajan V Chari, Patrick J Kehoe, and Ellen R McGrattan. Business cycle accounting. *Econometrica*, 75(3):781–836, 2007.

[20] Raj Chetty. A new method of estimating risk aversion. *The American Economic Review*, 96(5):1821–1834, 2006.

[21] Raj Chetty, Adam Guren, Day Manoli, and Andrea Weber. Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *The American Economic Review*, 101(3):471–475, 2011.

[22] Judith A Chevalier and David S Scharfstein. Capital-market imperfections and countercyclical markups: Theory and evidence. *The American Economic Review*, 86(4):703–725, 1996.

[23] Richard Clarida, Jordi Gali, and Mark Gertler. The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature*, 37:1661–1707, 1999.

[24] Gian Luca Clementi, Aubhik Khan, Berardino Palazzo, and Julia K Thomas. Entry, exit and the shape of aggregate fluctuations in a general equilibrium model with capital heterogeneity. *Unpublished Working Paper*, 2014.

[25] Gian Luca Clementi and Berardino Palazzo. Entry, exit, firm dynamics, and aggregate fluctuations. Technical report, National Bureau of Economic Research, 2013.

[26] Russell Cooper and Jonathan L Willis. The cost of labor adjustment: Inferences from the gap. *Review of Economic dynamics*, 12(4):632–647, 2009.

[27] Jan De Loecker and Frederic Warzynski. Markups and firm-level export status. *The American Economic Review*, 102(6):2437–2471, 2012.

[28] Wouter J Den Haan. Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. *Journal of Economic Dynamics and Control*, 34(1):79–99, 2010.

[29] Chris Edmond, Virgiliu Midrigan, and Daniel Yi Xu. Competition, markups, and the gains from international trade. *The American Economic Review*, 105(10):3183–3221, 2015.

[30] Chris Edmond and Laura Veldkamp. Income dispersion and counter-cyclical markups. *Journal of Monetary Economics*, 56(6):791–804, 2009.

[31] Lucia Foster, John Haltiwanger, and Chad Syverson. The slow growth of new plants: Learning about demand? *Economica*, 83(329):91–129, 2016.

[32] Mark Gertler and Simon Gilchrist. Monetary policy, business cycles, and the behaviors of small manufacturing firms. *The Quarterly Journal of Economics*, 109(2):309 – 340, 1994.

[33] Mark Gertler and John Leahy. A phillips curve with an ss foundation. *Journal of Political Economy*, 116(3):533–572, 2008.

[34] Simon Gilchrist, Raphael Schoenle, Jae Sim, and Egon Zakrajsek. Inflation dynamics during the financial crisis. *FEDS Working Paper*, 2015.

[35] Pinelopi Goldberg and Rebecca Hellerstein. How rigid are producer prices? *FRB of New York Staff Report*, (407), 2011.

[36] Gita Gopinath, Sebnem Kalemli-Ozcan, Loukas Karabarbounis, and Carolina Villegas-Sanchez. Capital allocation and productivity in south europe. Technical report, National Bureau of Economic Research, 2015.

[37] Jeremy Greenwood, Zvi Hercowitz, and Gregory W Huffman. Investment, capacity utilization, and the real business cycle. *The American Economic Review*, pages 402–417, 1988.

[38] Robert E Hall. Market structure and macroeconomic fluctuations. *Brookings papers on economic activity*, (2):285–322, 1986.

[39] Robert E Hall. What the cyclical response of advertising reveals about markups and other macroeconomic wedges. *Hoover Institution, Stanford University*, 2014.

[40] Gary D Hansen. Indivisible labor and the business cycle. *Journal of monetary Economics*, 16(3):309–327, 1985.

[41] Hugo Hopenhayn and Richard Rogerson. Job turnover and policy evaluation: A general equilibrium analysis. *Journal of political Economy*, pages 915–938, 1993.

[42] Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150, 1992.

[43] Chang-Tai Hsieh and Peter J Klenow. Misallocation and manufacturing tfp in china and india. *Quarterly Journal of Economics*, 124(4):1403–1448, 2009.

[44] Nir Jaimovich and Max Floetotto. Firm dynamics, markup variations, and the business cycle. *Journal of Monetary Economics*, 55(7):1238–1252, 2008.

[45] Nir Jaimovich and Sergio Rebelo. Can news about the future drive the business cycle? *The American economic review*, 99(4):1097–1118, 2009.

[46] Matthias Kehrig. The cyclical nature of the productivity distribution. *Earlier version: US Census Bureau Center for Economic Studies Paper No. CES-WP-11-15*, 2015.

[47] Robert G King, Charles I Plosser, and Sergio T Rebelo. Production, growth and business cycles: I. the basic neoclassical model. *Journal of monetary Economics*, 21(2-3):195–232, 1988.

[48] Paul Klemperer. Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *The review of economic studies*, 62(4):515–539, 1995.

[49] Per Krusell and Anthony A Smith, Jr. Income and wealth heterogeneity in the macroeconomy. *Journal of political Economy*, 106(5):867–896, 1998.

[50] Oleksiy Kryvtsov and Virgiliu Midrigan. Inventories, markups, and real rigidities in menu cost models. *The Review of Economic Studies*, page rds028, 2012.

[51] John Leahy. A survey of new keynesian theories of aggregate supply and their relation to industrial organization. *Journal of Money, Credit and Banking*, 43(s1):87–110, 2011.

[52] Yoonsoo Lee and Toshihiko Mukoyama. Entry, exit, and plant-level dynamics over the business cycle. *Working Paper*, 2013.

[53] James Levinsohn and Amil Petrin. Estimating production functions using inputs to control for unobservables. *The Review of Economic Studies*, 70(2):317–341, 2003.

[54] Virgiliu Midrigan. Menu costs, multiproduct firms, and aggregate fluctuations. *Econometrica*, 79(4):1139–1180, 2011.

[55] Sara Moreira. Firm dynamics, persistent effects of entry conditions, and business cycles. *Working Paper*, 2015.

[56] Emi Nakamura and Jon Steinsson. Monetary non-neutrality in a multisector menu cost model. *The Quarterly Journal of Economics*, 125(3):961–1013, 2010.

[57] Emi Nakamura and Jón Steinsson. Price setting in forward-looking customer markets. *Journal of Monetary Economics*, 58(3):220–233, 2011.

[58] Christopher J Nekarda and Valerie A Ramey. The cyclical behavior of the price-cost markup. Technical report, National Bureau of Economic Research, 2013.

[59] G Steven Olley and Ariel Pakes. The Dynamics of Productivity in the Telecommunications Equipment Industry. *Econometrica*, 64(6):1263–97, November 1996.

[60] Marcus M Opp, Christine A Parlour, and Johan Walden. Markup cycles, dynamic misallocation, and amplification. *Journal of Economic Theory*, 154:126–161, 2014.

[61] Luigi Paciello, Andrea Pozzi, and Nicholas Trachter. Price dynamics with customer markets. *Working Paper*, 2014.

[62] Michael Peters. Heterogeneous mark-ups, growth and endogenous misallocation. *Working Paper*, 2013.

[63] Morten Ravn, Stephanie Schmitt-Grohé, and Martin Uribe. Deep habits. *The Review of Economic Studies*, 73(1):195–218, 2006.

[64] Diego Restuccia and Richard Rogerson. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic dynamics*, 11(4):707–720, 2008.

[65] Richard Rogerson. Indivisible labor, lotteries and equilibrium. *Journal of monetary Economics*, 21(1):3–16, 1988.

[66] Julio J Rotemberg. Sticky prices in the united states. *Journal of Political Economy*, 90(6):1187–1211, 1982.

[67] Julio J Rotemberg and Garth Saloner. A supergame-theoretic model of price wars during booms. *The American Economic Review*, 76(3):390–407, 1986.

[68] Julio J Rotemberg and Michael Woodford. Markups and the business cycle. In *NBER Macroeconomics Annual 1991, Volume 6*, pages 63–140. MIT Press, 1991.

[69] Julio J Rotemberg and Michael Woodford. The cyclical behavior of prices and costs. *Handbook of macroeconomics*, 1:1051–1135, 1999.

[70] Stephanie Schmitt-Grohé and Martın Uribe. Why the monopolist does not deviate from the symmetric equilibrium in a model with good-specific subsistence points. *Note, Duke University*, 2007.

[71] Frank Smets and Raf Wouters. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, 1(5):1123–1175, 2003.

[72] Frank Smets and Rafael Wouters. Shocks and frictions in us business cycles: A bayesian dsge approach. *The American Economic Review*, 97(3):586–606, 2007.

[73] Jon Steinsson. Optimal monetary policy in an economy with inflation persistence. *Journal of Monetary Economics*, 50(7):1425–1456, 2003.

[74] Johannes Stroebel and Joseph Vavra. House prices, local demand, and retail prices. Technical report, National Bureau of Economic Research, 2014.

[75] George Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. *Economics letters*, 20(2):177–181, 1986.

[76] Eric R Young. Solving the incomplete markets model with aggregate uncertainty using the krusell–smith algorithm and non-stochastic simulations. *Journal of Economic Dynamics and Control*, 34(1):36–41, 2010.