Title
Essays on the Real Effects of Financial Market Fluctuations

Permalink
https://escholarship.org/uc/item/33p55579

Author
Giuliano, Fernando Mauro

Publication Date
2015

Peer reviewed|Thesis/dissertation
Essays on the Real Effects of Financial Market Fluctuations

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Economics

by

Fernando Mauro Giuliano

2015
Abstract of the Dissertation

Essays on the Real Effects of Financial Market Fluctuations

by

Fernando Mauro Giuliano

Doctor of Philosophy in Economics

University of California, Los Angeles, 2015

Professor Ariel Tomas Burstein, Chair

In the following essays I study the effects of disruptions in financial markets on aggregate outcomes.

In the first two chapters, I study the transmission mechanisms from financial crises to the real economy in emerging countries, in environments where firms set heterogeneous markups. The introduction of heterogeneous markups is backed by data: I document that there is evidence of firms setting heterogeneous markups using microdata for Argentina and Colombia. As an endogenous source of resource misallocation across firms, markups can potentially be an important driver of aggregate productivity and output dynamics during large financial crises.

The opening chapter is my first attempt to address the role of heterogeneous markups during financial crises. To investigate the extent to which this has a significant quantitative role, I adapt a model of imperfect competition where markups are a function of within-sector market shares. Using microdata from Argentina’s annual manufacturing survey, I document that market shares become more disperse during the Argentine 2001-02 crisis. Through the lens of the model this results in increased variability of markups, which decreases aggregate productivity. I perform an accounting exercise and find that markup-induced misallocation can explain between 6.4% and 15.6% of the fall in aggregate productivity during the Argentine crisis, or up to one third of the overall effect of resource misallocation.

In Chapter 2, joint with Gabriel Zaourak, we explicitly introduce financial frictions to
analyze the interaction between credit constraints and variable markups during a credit crunch. Financial frictions take the form of a collateral constraint on working capital. A financial crisis in this framework is modeled as an exogenous shock to the maximum amount of working capital that can be financed externally. Using microdata from financial statements and manufacturing surveys, we calibrate the model to match salient features of the Colombian economy for the 1998-99 financial crisis, and evaluate the transition dynamics of aggregate variables. The model replicates the fall and subsequent recovery of aggregate output and productivity, as well as the concentration patterns observed in the data. We find that in this case variable markups partially offset the resource misallocation triggered by a credit crunch, dampening the response of aggregate variables. The reason is that under variable markups firms try not to change their price (hence quantities) as much as they would under constant markups. This is an example of the ambiguous effect of distortions in a second best world.

The last chapter is an early empirical exploration of the link between price fluctuations in financial markets and aggregate labor market outcomes, using data from the United Kingdom. I build a quarterly wealth index from stock market prices and real estate prices for the 1971-2012 period. Using a VECM, I find a robust co-integrating relationship between the unemployment rate and the wealth index. Specifically, fluctuations in wealth Granger-cause the unemployment rate, but not the opposite. This relationship is true for both components of the wealth index individually, and is stable over time. This is consistent with a model where output is demand determined and fluctuations in asset prices affect the unemployment rate through changes in aggregate consumption.
The dissertation of Fernando Mauro Giuliano is approved.

Nico Voigtländer

Hugo Andres Hopenhayn

Francisco Buera

Andrew Granger Atkeson

Ariel Tomas Burstein, Committee Chair

University of California, Los Angeles
2015
To my wife Cecilia.

To my family.

This is theirs as much as mine.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>vi</td>
</tr>
<tr>
<td>List of Figures</td>
<td>viii</td>
</tr>
<tr>
<td>List of Tables</td>
<td>x</td>
</tr>
<tr>
<td>1 Variable Markups and Missallocation in the Argentine Crisis.</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.2 The Model</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Data</td>
<td>9</td>
</tr>
<tr>
<td>1.4 Calibration and Results</td>
<td>12</td>
</tr>
<tr>
<td>1.5 Conclusion</td>
<td>15</td>
</tr>
<tr>
<td>1.6 Tables and Figures</td>
<td>17</td>
</tr>
<tr>
<td>1.7 Appendices</td>
<td>22</td>
</tr>
<tr>
<td>1.7.1 Appendix 1: Estimation of Number of workers per firm</td>
<td>22</td>
</tr>
<tr>
<td>1.7.2 Appendix 2: Distribution of Wedges</td>
<td>23</td>
</tr>
<tr>
<td>1.7.3 Appendix 3: Sectors in AMS and Simulated Subsectors</td>
<td>24</td>
</tr>
<tr>
<td>2 Market Power and Aggregate Efficiency in Financial Crises.</td>
<td>26</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>26</td>
</tr>
<tr>
<td>2.2 Motivating Facts</td>
<td>31</td>
</tr>
<tr>
<td>2.2.1 Increase in Concentration</td>
<td>31</td>
</tr>
<tr>
<td>2.2.2 Heterogeneous Markups</td>
<td>32</td>
</tr>
<tr>
<td>2.3 A Model of Financial Frictions with Variable Markups</td>
<td>33</td>
</tr>
<tr>
<td>2.4 Data and Calibration</td>
<td>39</td>
</tr>
<tr>
<td>2.5 Quantitative Exercise</td>
<td>42</td>
</tr>
<tr>
<td>2.6 Conclusion</td>
<td>44</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>2.7</td>
<td>Tables and Figures</td>
</tr>
<tr>
<td>2.8</td>
<td>Appendices</td>
</tr>
<tr>
<td>2.8.1</td>
<td>Appendix 1: Kimball Aggregator and Demand Approximation</td>
</tr>
<tr>
<td>2.8.2</td>
<td>Appendix 2: Pricing Decisions for the Constrained Firms</td>
</tr>
<tr>
<td>2.8.3</td>
<td>Appendix 3: Solution Algorithm</td>
</tr>
<tr>
<td>2.8.4</td>
<td>Appendix 4: Markups from Microdata</td>
</tr>
<tr>
<td>3</td>
<td>Stock Market Fluctuations, Unemployment and Consumption in the United Kingdom.</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td>3.2</td>
<td>Detrending Time-Series</td>
</tr>
<tr>
<td>3.2.1</td>
<td>Variables in wage units</td>
</tr>
<tr>
<td>3.3</td>
<td>Stock Prices and Unemployment in the UK</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Data</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Results</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Robustness checks</td>
</tr>
<tr>
<td>3.3.4</td>
<td>A stylized model</td>
</tr>
<tr>
<td>3.4</td>
<td>Stock Prices and Consumption in the United Kingdom</td>
</tr>
<tr>
<td>3.4.1</td>
<td>A Framework for the Cointegrating Relationship Between Stock Prices and Consumption</td>
</tr>
<tr>
<td>3.4.2</td>
<td>The empirics</td>
</tr>
<tr>
<td>3.5</td>
<td>Conclusion</td>
</tr>
<tr>
<td>3.6</td>
<td>Tables and Figures</td>
</tr>
<tr>
<td>3.7</td>
<td>Appendices</td>
</tr>
<tr>
<td>3.7.1</td>
<td>Appendix 1: Data Sources</td>
</tr>
<tr>
<td>3.7.2</td>
<td>Appendix 2: Granger Non-causality Test</td>
</tr>
</tbody>
</table>

References .................................................................................. 87
LIST OF FIGURES

1.1 Real Value Added Index .................................................. 17
1.2 Market Share Distribution (logs), Largest Sectors .................... 17
1.3 Inverse Herfindahl Index .................................................. 18
1.4 Market Shares and Markups, 2001 ....................................... 18
1.5 Market Shares and Markups, 1997 ....................................... 19
1.6 Capital Wedge Distribution .............................................. 23
1.7 Scale Wedge Distribution ................................................. 23

2.1 Inverse Herfindahl Index (1997=100) .................................. 46
2.2 Revenue Productivity and Size .......................................... 46
2.3 GVP-to-Materials Ratio and Size ....................................... 47
2.4 Demand for Different Superelasticities ................................ 47
2.5 Partial Equilibrium Example ............................................. 48
2.6 Real Gross Domestic Product (1997=100) .............................. 48
2.7 Real Net Credit (1997=100) ............................................. 49
2.8 Credit as a Share of GDP ................................................. 49
2.9 Real Net Credit, Deviations from Trend (1997=100) ................. 50
2.10 Model Implied Leverage vs. Size ....................................... 50
2.11 Different Measures of Credit Crunch (1997=1) ....................... 51
2.12 Value of Production by Top 10% Firms ............................... 51
2.13 Aggregate Output .......................................................... 52
2.14 Aggregate Productivity .................................................... 52
2.15 Dispersion of Capital Across Firms .................................... 53
2.16 Cumulative Distribution of Entrepreneurs. Constant Markups .... 53
2.17 Cumulative Distribution of Entrepreneurs Variable Markups ...... 54
3.1 Real GDP (deviations from HP trend) ........................................ 78
3.2 Real GDP in Wage Units ......................................................... 78
3.3 FTSE100 vs. FTSEAS ............................................................ 79
3.4 Different Measures of Employment ........................................... 79
3.5 Unemployment Rate and the Stock Market ................................. 80
3.6 Actual and Forecasted Unemployment Rate ................................. 80
3.7 FTSEAS in Wage and GDP Units .............................................. 81
3.8 FTSEAS and Housing Wealth (Wage Units) ............................... 81
3.9 FTSEAS and Consumption ..................................................... 82
3.10 Response of Consumption to a 10% Shock in Stock Prices ............. 82
3.11 Response of Stock Prices to a 10% Shock in Stock Prices ............. 82
3.12 Wealth Index and Consumption ............................................. 83
LIST OF TABLES

1.1 Number of Active Firms per Year in AMS ........................................... 19
1.2 Concentration Measures ................................................................. 20
1.3 Percentage of the Fall in TFP explained by Variability of Markups ....... 20
1.4 Percentage of the Fall in TFP explained by Variability of Markups ....... 21
1.5 Sectors in the AMS ................................................................. 24
1.6 Artificial Subsectors Within Each Sector ........................................ 25
2.1 Calibration 1 ................................................................. 55
2.2 Calibration 2 ................................................................. 55
I am deeply indebted to Francisco Buera and Ariel Burstein for their patience and guidance in all stages of the project; and to Hugo Hopenhayn and Andrew Atkeson for their invaluable feedback. They all also helped me attend UCLA, for which I am very grateful. I would also like to thank Arnold Harberger for his support and his always insightful remarks. I have benefited from comments and suggestions from Pablo Fajgelbaum, Roger Farmer, Federico Grinberg, Dennis Kuo, Gabriel Zaourak, and seminar participants at UCLA and Universidad de Los Andes. Finally, the empirical analysis in Chapter 2 was completed with the help of technicians at DANE, the Colombian Statistical office.
VITA

2006 B.A. (Economics), Universidad Nacional de Córdoba, Argentina.

2009 Graduate Degree (Economics), Universidad Torcuato Di Tella, Argentina.

2011 M.A. (Economics), University of California, Los Angeles.

2011 C.Phil. (Economics), University of California, Los Angeles.

2003-2006 Teacher Assistant, Universidad Nacional de Córdoba, Argentina.

2007-2008 Teacher Assistant, Universidad Torcuato Di Tella, Argentina.

2009 Research Assistant, Universidad Torcuato Di Tella, Argentina.

2011-2015 Teaching Fellow, Department of Economics, University of California, Los Angeles.
CHAPTER 1

Variable Markups and Missallocation in the Argentine Crisis.

1.1 Introduction

Recent research has stressed the role of resource misallocation on the level of aggregate Total Factor Productivity (TFP). Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), for example, show how firm-specific distortions can result in important cross-country differences in TFP. Jeong and Townsend (2007), Buera et al. (2011), and Midrigan and Xu (2014) argue that financial frictions could be responsible for a considerable fraction of measured TFP, through their effects on the allocation of entrepreneurial talent and capital.

This approach could be relevant for the study of economic cycles, since TFP accounts for most of the measured fall in aggregate output during crises. Kehoe (2003), for example, documents that over 90% of the fall in output during the Argentine crisis of 2001/02 can be accounted for by TFP. Meza and Quintin (2005) find a similar result for the Mexican crisis in 1995. Ohanian (2001) claims that the contribution of TFP to the fall in output observed during the Great Depression is about 50%.

In a recent paper, Sandleris and Wright (2014) follow this path and claim that about half of the decline in TFP during the Argentine 2001/02 crisis can be attributed to resource misallocation. They measure misallocation as an increase in the dispersion of ”wedges” during the crisis. Wedges are recovered from the optimality conditions of the firm, under the assumption that markets for goods and factors of production are perfectly competitive. The authors are agnostic about the sources of wedges and the reason for the increased dispersion observed during the crisis.

What drives the increase in misallocation during crisis? Is it exogenous firm-specific
distortions, or endogenous decisions by the firms? In the former case, firms face exogenous wedges that distort relative prices of inputs and output across firms. These distortions could be interpreted as implicit taxes and subsidies and result in a socially inefficient use of resources. If those distortions were absent and firms faced the same relative prices, resources would reallocate across productive units to achieve efficiency. But inefficiency could also be the result of deliberate decisions by the firms even if those implicit taxes and subsidies were absent. For example, a firm could be producing too little output not because it is facing an implicit tax on its sales, but because it chooses to do so in order to exploit its monopolistic power.

In this paper I explore the role of variable markups, a particular endogenous distortion, in the fall of TFP during the Argentine 2001/02 crisis. The model is an adaptation of Atkeson and Burstein (2008), where firms engage in imperfect competition and markups are a function of the market share of a firm within a sector: the higher the market share, the higher the markup that the firm is able to charge. It delivers increased misallocation in the event of a crisis as long as the distribution of market shares becomes more disperse and industries become more concentrated, which I will show is the case during the Argentine 2001/02 crisis. This misallocation is endogenous in the sense that it is not driven by external factors that prevent firms from producing at the social optimum level, but by their unwillingness to do so given their market power. I also allow for exogenous wedges to output and capital/labor, to capture distortions other than markups present in the data. The model is close to Peters (2013), extended to account for wedges and applied to study the role of markups during crises.

The theoretical framework, as we will see below, requires many sectors with relatively few firms each for the imperfect competition assumption to be reasonable.\footnote{This will become apparent once the model is introduced in Section 2} In the Argentine microdata, however, firms are classified in few sectors, each of which has relatively many firms. To address this issue I perform two different exercises. In the first one, I quantify the role of variable markups using the sectoral classification provided in the microdata. In a second exercise I randomly divide each sector into subsectors, quantify the role of variable markups given those subsectors, and repeat the procedure 10,000 times to report the average role of markups across simulations. The first exercise has
the virtue of relying in the industry classification provided in the data, even though that classification does not map very well to the model. The second exercise better resembles the assumptions in the model, but relies on artificially generated subsectors.

I find that variable markups account for 6.4% to 15.6% of the fall in TFP during the Argentine crisis, or up to one-third of the portion of the fall in TFP explained by misallocation. This results from the comparison of the variable markup model with a constant markup benchmark. The constant markup benchmark is a specific parameterization of the general model.

The paper is structured as follows. Section 2 presents the model and finds the expression for aggregate TFP used in the empirical analysis. Section 3 discusses the manufacturing microdata for Argentina and presents relevant statistics. Section 4 discusses the calibration of the model and performs the empirical analysis. Section 5 concludes.

1.2 The Model

I adapt the model in Atkeson and Burstein (2008) to quantify the role of variable markups in the fall in aggregate TFP during crises. This model is appealing for several reasons. First, it allows for variable markups and strategic behavior in a relatively simple setting. Second, it nests the standard CES demand system with constant markups as a particular case. Third, it has proven to match relevant stylized facts on international relative prices. Finally, it can be readily calibrated from available microdata on market shares.

Aggregation of Goods from Individual Producers

There is a final good $Y$ produced using as input the output of a continuum of sectors $Y_s$, where $s \in [0, 1]$. The production function is CES with elasticity of substitution $\eta$

$$Y = \left[ \int_0^1 Y_s^{\frac{\eta-1}{\eta}} ds \right]^{\frac{\eta}{\eta-1}} \quad (1.1)$$

The implied demand for output from sector $s$ is
\[ Y_s = \left( \frac{P}{P_s} \right)^\eta Y \]  \hspace{1cm} (1.2)

where \( P \) is the standard aggregate price index

\[ P = \left[ \int_0^1 P_s^{1-\eta} ds \right] ^{\frac{1}{1-\eta}} \]  \hspace{1cm} (1.3)

Intermediate good \( Y_s \) is produced combining the output of the \( N_s \) firms in sector \( s \), through a discrete CES production function with elasticity of substitution \( \rho \)

\[ Y_s = \left[ \sum_{i=1}^{N_s} y_{si}^{1-\rho} \right] ^{\frac{1}{1-\rho}} \]  \hspace{1cm} (1.4)

The implied demand for output from firm \( i \) in sector \( s \) is:

\[ y_{si} = \left( \frac{P_s}{p_{si}} \right)^\eta Y_s \]  \hspace{1cm} (1.5)

where \( P_s \) is the price index for sector \( s \), i.e.:

\[ P_s = \left[ \sum_{i=1}^{N_s} p_{si}^{1-\rho} \right] ^{\frac{1}{1-\rho}} \]  \hspace{1cm} (1.6)

where \( N_s \) is the number of firms in sector \( s \). \( N_s \) is a number relatively small so that the assumption on the sectoral market structure holds, as shown below.

**Market Structure and the Problem of the Firm**

I assume that individual firms produce output combining capital and labor with a constant returns to scale Cobb-Douglas technology

\[ y_{si} = A_{si} k_{si}^{\alpha} l_{si}^{1-\alpha} \]  \hspace{1cm} (1.7)

where \( A_{si} \) is an idiosyncratic productivity factor.

As in Atkeson and Burstein, there is imperfect competition at the firm level. Specifically, firms engage in quantity competition taking the output from other firms in their
sector as given. That is, firms internalize the fact that their own output choice will affect sectoral output $Y_s$, hence sectoral prices $P_s$. Firms take the aggregate price $P$ and output $Y$ as given. This seems like a reasonable assumption given that within the model firms operate in one of infinitely many sectors, each one of which consists of a discrete number of firms. The capital rental and labor markets are assumed perfectly competitive, with prices $r$ and $w$ respectively.

I assume that $\rho > \eta > 1$, i.e. goods within a sector are more substitutable than goods across sectors. This assumption coupled with the imperfect competition market structure will be key to generate heterogeneous markups across firms.

I allow for firm-specific wedges in the output market, $(1 - \tau_{y_{si}})$ and in the capital rental market $(1 + \tau_{k_{si}})$. I introduce wedges to account for allocations that differ from the model’s predictions in the microdata. These exogenous wedges could capture many things from measurement errors, adjustment costs, miscalibration of the model, etc. The scale wedge captures all factors (other than markups) that result in a sub-optimal scale for a firm. The capital wedge, all factors that result in a suboptimal capital-labor ratio. A capital wedge higher than one is related to a capital-labor ratio lower than optimal, while a capital wedge lower than one reflects a capital-labor ratio higher than optimal. The capital wedge is, thus, isomorphic to a labor wedge.

The firms solves:

$$\max_{y_{si}, l_{si}, k_{si}} (1 - \tau_{y_{si}}) p_{si} y_{si} - w l_{si} - (1 + \tau_{k_{si}}) r k_{si}$$

subject to:

$$\frac{p_{si}}{P} = \left( \frac{y_{si}}{Y} \right)^{-1/\rho} \left( \frac{Y_s}{Y} \right)^{-1/\eta}$$

$$Y_s = \left[ \sum_{i=1}^{N_s} \frac{y_{si}^{\rho}}{\rho} \right]^{\frac{1}{\rho - 1}}$$

$$y_{si} = a_{si} k_{si}^{\alpha} l_{si}^{1-\alpha}$$

where (1.13) is obtained combining (1.2) and (1.5). Equation (1.13) reflects the imperfect competition assumption: i.e. firms are not price-takers but internalize the effect
of their output choice on sectoral output, hence prices.

To solve the model analytically I proceed in two stages, first minimize costs and then maximize profits. The cost minimization problem yields a cost function of the form:

\[ C_{si}(w, r, y_{si}) = MC_{si}y_{st} \]  

where

\[ MC_{si} = \frac{(1 + \tau_{k,si})^{\alpha}}{A_{si}} \left( \frac{w^{1-\alpha}r^\alpha}{\alpha^\alpha(1-\alpha)^{1-\alpha}} \right) = \frac{(1 + \tau_{k,si})^{\alpha}}{A_{si}} \Gamma \]  

I use this information to maximize profits:

\[ \max_{y_{si}} (1 - \tau_{y,si})p_{si}y_{si} - MC_{si}y_{si} \]

subject to

\[ \frac{p_{si}}{F} = \left( \frac{y_{si}}{Y_S} \right)^{-1/\rho} \left( \frac{Y_S}{Y} \right)^{-1/\eta} \]

\[ Y_s = \left[ \sum_{i=1}^{N_s} y_{si}^{\rho} \right]^{\frac{1}{\rho-1}} \]

The optimal pricing rule derived from the first order conditions of the problem is:

\[ p_{si} = \frac{\varepsilon(\omega_{si})}{\varepsilon(\omega_{si}) - 1} \frac{MC_{si}}{(1 - \tau_{y,si})} \]  

where

\[ \varepsilon(\omega_{si}) = \frac{1}{\rho} (1 - \omega_{si}) + \frac{1}{\eta} \omega_{si} \]  

and \( \omega_{si} \) is the market share of firm \( i \) in sector \( s \), i.e. \( \frac{p_{si}y_{si}}{F} \)

The pricing rule (1.13) states that firms set prices as a markup over their marginal costs (as in a standard CES demands system), in this case adjusting for wedges. In this model, however, markups are not a parameter but a function of the within-sector market
share. The higher the market share within a sector, the higher the markup. To illustrate, assume that there are infinitely many firms within a sector, i.e. \( \omega_{si} = 0 \). In that case \( \varepsilon(\omega_{si}) = \rho \) and (1.13) becomes:

\[
p_{si} = \frac{\rho}{\rho - 1} \frac{MC_{si}}{1 - \tau_{ys_{i}}}
\]

At the other extreme, if there is only one firm in a given sector, i.e. \( \omega_{si} = 1 \), then \( \varepsilon(\omega_{si}) = \eta \) and (1.13) becomes:

\[
p_{si} = \frac{\eta}{\eta - 1} \frac{MC_{si}}{1 - \tau_{ys_{i}}}
\]

Markups are greater in the latter case since \( \rho > \eta \). The intuition behind this result is simple. In the first case, an infinitesimally small firm mainly faces competition from within its sector, hence its relevant price elasticity is the within-sector elasticity \( \rho \). In the second case, a firm that dominates a sector mainly faces competition from other sectors, hence its relevant price elasticity is the across-sectors elasticity, \( \eta \). Between these two extremes the relationship is monotonic: the higher the within-sector market share, the higher the markup.

Note that the monotonic relationship between market shares and markups breaks down when \( \rho = \eta \). This is the standard CES case where markups are equal across firms and time.

In this framework, the distribution of markups changes over time as long as there are sizeable changes in the underlying distribution of market shares within sectors. As we will see below, this is the case during the Argentine crisis. Given that markups distort the efficient allocation of resources, this mechanism endogenously generates misallocation that could be quantitatively important in the aggregate.

**Aggregate TFP**

To assess the effect of variable markups on aggregate productivity I need to express aggregate TFP as a function of wedges and markups. I follow the approach in Edmond et al. (2015), adjusted to account for wedges.
Define a production function for sector $S$ as

$$Y_S = A_S K_S^{1-\alpha} L_S^{\alpha}$$  \hspace{1cm} (1.15)$$

where $A_S$ is sectoral TFP and $K_s$ and $L_s$ are sectoral capital and labor, respectively.

Using the first order conditions of the firm and the market clearing conditions for factor markets we can express sectoral TFP as

$$A_s = \frac{1}{\left[ \sum_i \frac{y_{si}}{Y_s} \left( \frac{1+\tau_{ksi}}{A_{si}} \right)^{\alpha-1} \right]^\alpha \left[ \sum_i \frac{y_{si}}{Y_s} \left( \frac{1+\tau_{ksi}}{A_{si}} \right)^{\alpha} \right]^{1-\alpha}}$$  \hspace{1cm} (1.16)$$

Define sectoral markup as

$$M_s = \frac{P_s A_s (1 - \tau_{SY})}{\Gamma (1 + \tau_{SK})^\alpha}$$  \hspace{1cm} (1.17)$$

where $(1 - \tau_{SY})$ and $(1 + \tau_{SK})$ are sectoral wedges for scale and capital, respectively, defined as revenue weighted averages of firm-specific wedges.

From (1.6)

$$1 = \left[ \sum_{i=1}^{N_s} \left( \frac{p_{si}}{P_s} \right)^{1-\rho} \right]^{\frac{1}{1-\rho}}$$

Replacing $p_{si}$ and $P_s$ using (1.13) and (1.17), and solving for $A_s$ we obtain:

$$A_s = \left[ \sum_{i=1}^{N_s} \left( \frac{M_s}{M_{si}} \left( \frac{(1 + \tau_{KS})}{(1 + \tau_{Ksi})} \right)^{\alpha} \frac{(1 - \tau_{YSi})}{(1 - \tau_{YS}) A_{si}} \right)^{\rho-1} \right]^{\frac{1}{\rho-1}}$$  \hspace{1cm} (1.18)$$

Equation (1.18) shows that dispersion in markups matters for sectoral TFP. In particular, higher dispersion in markups results in lower sectoral TFP, given the positive relationship between productivity and market share (hence markups). This is a particular result of the general notion that the dispersion of distortions reduces sectoral TFP, as in Hsieh and Klenow (2009). In this paper, however, distortions resulting from markups arise from conscious decisions by the firms, rather than exogenous distortions that prevent firms from producing optimally.
Following the same procedure as above, we can further aggregate sectoral TFP into aggregate TFP to obtain

\[
A = \left[ \int_0^1 \left( \frac{M}{M_s} \left( \frac{(1 + \tau_K)}{(1 + \tau_{KS})} \right)^{\alpha} \frac{(1 - \tau_{YS})}{(1 - \tau_Y)} A_s \right)^{\eta-1} ds \right]^{\frac{1}{\eta-1}} \quad (1.19)
\]

I will use the equation above to assess the role of variable markups in the observed fall in TFP during crises.

1.3 Data

I test the role of variable markups using Argentine manufacturing microdata for the 1997-2002 period. I believe the Argentine crisis is well suited for this experiment due to its magnitude, and due to the fact that previous studies have argued that misallocation played a significant role in the decline of aggregate TFP (see Sandleris and Wright).

Background: the Argentine Crisis

After years of relative growth and stability Argentina entered a protracted and deep recession in mid 1998. The height of the crisis started in late December 2001, following a bank run that resulted in banking holidays and in the government limiting the availability of demand deposits to the public. In January 2002, in the midst of political turmoil, the government declared default on its debt and devalued the Argentine peso after a decade-long peg to the US dollar. The fall in economic activity in the following months had no precedent in the country’s troubled economic history. Figure 1.1 shows the evolution of a real value added index for the economy as a whole and for the manufacturing sector.

Real value added for the economy as a whole fell 19% from peak to trough. The performance of the manufacturing sector was worse, falling over 30% between mid 1998 and mid 2002. The manufacturing sector represented about 20% of total value added during the 1990s.

In the empirical implementation I compare 1997 (the last year of positive growth in the sample) with 2002 (the last year with negative growth in the sample). When I talk about ”the fall in TFP during the crisis”, or the change in certain statistic or distribution
during the crisis, I refer to the contrast between 1997 and 2002. Results are robust to using 1998 as benchmark.

The Annual Manufacturing Survey

My data source is the Annual Manufacturing Survey (henceforth AMS), conducted by Argentina’s federal statistical agency, INDEC. It surveys manufacturing establishments with at least 10 employees, which represent about 90% of aggregate value added in manufacturing. The database is an unbalanced panel. Establishments are selected randomly, and once they are surveyed they don’t leave the sample unless they stop producing. They report data on revenues, investments, and expenses on energy, labor, intermediate inputs, taxes, etc.

There are an average of 3731 active establishments per-year in the sample (see Table 1.1), categorized in 22 two-digit manufacturing industries (equivalent to the three-digit disaggregation in the NAICS code). For confidentiality reasons, all variables are expressed as a ratio to workers in payroll. The number of workers in payroll per establishment is not provided, but its growth rate is. Moreover, establishments are classified as small, medium or large according to the number of workers in payroll: between 10 and 80, between 81 and 200, and over 200, respectively. Following the methodology in Sandleris and Wright (2014), I use information on the size bin and the growth rate of workers in payroll to recover the variables in levels.\(^2\)

I use sectoral producer price indexes to deflate nominal variables, since I do not observe establishment-specific or product-specific prices. My measure of output per establishment is value added, which I compute as the difference between the gross value of production and the cost of intermediate inputs. To deflate the cost of intermediate inputs, I construct sector-specific intermediate input deflators using information from the 1993 input-output matrix.

\(^2\)See Appendix 1 for further details.
Relevant Statistics

In the light of the theoretical model, variable markups could explain a portion of the fall in TFP during the crisis as long as market shares become more disperse and sectors become more concentrated. We observe both features during the Argentine crisis.

Figure 1.2 shows that during the crisis market shares became more disperse. It plots the distribution of within-sector market shares for the nine largest manufacturing sectors for 1997 and 2002, the last year before the recession started and the year of the height of the crisis, respectively. This pattern holds for all other sectors. The overall standard deviation in market shares increases from 2.6% to 4.5%.

Figure 2.1 shows that sectors became more concentrated. It plots the inverse Herfindahl Index by industry for 1997 and 2002. The inverse Herfindahl Index is a measure of the number of “effective competitors” in an industry. The lower the Index, the more concentrated a sector. Industry concentration increased in 19 out of the 22 sectors in the AMS (i.e. dots below the 45-degree line).

Even though market shares become more disperse and industries become more concentrated with the crisis, market shares and concentration indexes are very low (see Table 2.12). The average of the within-sector median market share, for example, is around 0.5%. The average of the within-sector percentile 95 reaches 6.9% in 2002. Even the average maximum market share across sectors does not seem particularly high, 30.4%. For comparative purposes, the figures for Chile during the 2009 financial crisis are 2%, 30% and 35%, respectively. Regarding concentration, the mean Inverse Herfindahl Index in Argentina is 12.3 at the height of the crisis, compared to 7.8 for Chile during the 2009 crisis. This does not necessarily mean that the level of concentration of the manufacturing sector in Argentina is relatively low, but could be the result of the high level of aggregation in Argentine data.

These facts represent a challenge for the exercise in two fronts. First, the theoretical model requires many sectors with relatively few firms within each sector for the assumptions that lead to imperfect competition to hold. Second, the larger the sectors, the less

---

3 An Inverse Herfindahl Index of (say) 5, indicates that a sector has a concentration similar to having 5 firms that equally split the market.

4 Data from Encuesta Industrial Nacional Anual (ENIA), 2009, the Chilean equivalent of Argentina’s AMS. ENIA classifies plants into three-digit industries.
likely the variable markup mechanism will have a relevant contribution to the fall in TFP. Recall from Section 2 that in the limit, when there are infinitely many firms within a sector, markups are the same for every firm and constant across time.

To address this issue, I perform two exercises. In the first one, I compute the contribution of variable markups to the fall in TFP during the crisis using the 22 two-digit sector classification provided in the AMS. In the second one, I randomly divide each large sector in the database into several subsectors. I do this 10,000 times. I then compute the contribution of variable markups to the fall in TFP during the crisis for each of the 10,000 simulations. Finally, I report the average contribution of variable markups across simulations.\(^5\)

### 1.4 Calibration and Results

I need to set values for the production function parameter \(\alpha\), the within-sector elasticity of substitution \(\rho\), and the across-sector elasticity of substitution \(\eta\). Since labor shares are usually underestimated in emerging countries’ national accounts (see Gollin, 2002), I do not use Argentine data to estimate \(\alpha\). Rather, I set \((1 - \alpha) = 2/3\) as the usual approximation for the labor share in US data. Following Edmond et al. (2015), who use a similar model to address gains from trade, I set \(\eta = 1.25\) and \(\rho = 8.45\). Given these values for the elasticities of substitution I can recover firm markups from observed market shares using (1.13). I believe this parametrization to be reasonable since it yields an aggregate revenue weighted markup of about 1.31 (or 31%), in line with previous studies (see for example Nekarda and Ramey (2013) and De Loecker and Warzynski (2012)).

The blue line in Figures 1.4 and 1.5 displays the mapping between market share and markups given the parameterization above. The vertical lines show the market share percentiles shown in Table 2.12. The distribution of market shares, hence markups, has fatter tails during the crisis. This effect plus the increased variability in market shares during the crisis results in a lower TFP.

\(^5\)In each of the 10,000 simulations we generate 120 subsectors from the existing 22 sectors. This is to match the number of subsectors in a comparable database with higher level of disaggregation (3-digit) such as Chile’s ENIA.
I can recover the capital wedge from the capital-labor ratio optimality condition:

\[
\frac{k_{si}}{l_{si}} = \frac{\alpha}{1 - \alpha} \frac{w}{(1 + \tau_{k_{si}}) r}
\]

\[
(1 + \tau_{k_{si}}) = \frac{\alpha}{1 - \alpha} \frac{w l_{si}}{r k_{si}}
\]

The wage bill \(wl_{si}\) is obtained from the microdata. Although we do not have information on a plant’s capital stock, I approximate expenditures on capital services with expenditures on electricity and gas used in production. This is consistent with the assumption that every unit of capital needs a fixed proportion of energy to produce the capital services needed for production (see Sandleris and Wright (2014)). I set \(r = 0.1\) as in Hsieh and Klenow (2009), consistent with a 5% real return on capital and 5% depreciation rate. Any idiosyncratic difference in these numbers across firms will be captured by the capital wedge. Notice that the value of the capital wedge is not influenced to any extent by the fact that variables in the microdata are expressed as a ratio to workers in payroll.

The scale wedge is obtained rearranging the pricing function (1.13)

\[
(1 - \tau_{ys_i}) = \frac{\varepsilon(\omega_{si})}{\varepsilon(\omega_{si}) - 1 (1 - \alpha_s) p_{si}} \frac{MC_{si}}{y_{si}}
\]

Multiplying and dividing by \(y_{si}\) in the right-hand side and noting that \(C_{si} = MC_{si} y_{si} = \frac{w l_{si}}{(1 - \alpha_s)}\) yields

\[
(1 - \tau_{ys_i}) = \frac{\varepsilon(\omega_{si})}{\varepsilon(\omega_{si}) - 1 (1 - \alpha_s) y_{si} P_{si}} \frac{w l_{si}}{y_{si} P_{si}}
\]

It is worth noting that markups and wedges recovered from the data are uncorrelated. Specifically, the correlation between markups and scale wedges is -0.01 and between markups and capital wedges is 0.01. This is consistent with markups being an independent source of distortions from wedges.

As in Hsieh and Klenow I recover (physical) productivity from the assumptions on
the demand structure and profit maximization. Specifically

\[ A_{si} = \kappa \left( \frac{p_{si}y_{si}}{k_{si}^{1-\alpha}} \right)^{\frac{1}{\rho-1}} \]

with \( \kappa = \left( \frac{P_s}{Y_s} \right)^{\frac{1}{\rho-1}} \). I can recover \( l_{si} \) from micro data, as well as (an energy proxy for) \( k_{si} \). As for \( P_s \), I use the sectoral deflators described in the Data section.

### Quantitative Exercises

To disentangle the contribution of variable markups to the fall in TFP during the Argentine crisis I proceed in two steps. First, I compute the fall in TFP between 1997 and 2002 using the calibration from the previous section. In the second step, I shut down the variable markup channel setting \( \eta = \rho \). Recall that under this assumption markups are equal to \( \frac{\rho}{\rho-1} \) for all firms. The difference between the fall in TFP computed with variable markups, and the fall in TFP when the variation in markups is zero is attributed to variable markups. In the constant markup case I set \( \eta = \rho = 4.2 \) to match the revenue-weighted aggregate markup that results from the variable markup calibration.

As explained in the previous section, I perform this exercise under two settings. In the benchmark setting I use AMS’ 22-sector classification to compute within-sector market shares, hence markups. In the second one, I simulate subsectors within each large sector, for a total of 120 subsectors. In this specification, the market share of a firm is computed with respect to the subsector I generated. It is important to note that in this second setting, two firms that are originally in different sectors in the AMS classification cannot end up in the same artificially created subsector. More details on AMS’ sectors and artificial subdivisions can be found in Appendix 2.

Since I do not explicitly model the entry-exit decision of a firm, I report results for the full sample of firms and for continuing firms (i.e. firms present in both 1997 and 2002). As is usual in the wedges literature, I trim the top and bottom 1% of wedges before running the exercises.

---

6We impose our functional assumptions to recover physical productivity since the use of sectoral price indexes to recover firm productivity from nominal variables results in a measure of "revenue productivity", i.e: \( p_{si}A_{si} \) (see Foster et al. (2008)).
The results for the exercise using AMS’ sector classification are shown in Table 1.3. If we look at continuing firms, variable markups account for 6.4% of the fall in TFP between 1997 and 2002. When I include firms that are either present in both years or in any one of the two, variable markups account for 7.9% of the fall in TFP. For the exercise with simulated subsectors (Table 1.4), if we look at continuing firms, variable markups account for 15.6% of the fall in TFP between 1997 and 2002. When I include firms that are either present in both years or in any one of the two, variable markups account for 8.3% of the fall in TFP.

Taking into account that in Sandleris and Wright up to half of the fall in TFP is explained by misallocation, these results suggest that variable markups could explain up to one-third of the fall in TFP attributed to misallocation.

1.5 Conclusion

Distortions at the firm level conspire against the efficient allocation of resources and result in a lower aggregate TFP. They have been proven to be quantitatively relevant to explain TFP differences across countries, and lately have been used to explain the cyclical variations of aggregate productivity in a given country.

This paper focuses on a particular source of misallocation: markups. If they depend on market shares, and the distribution of market shares becomes more volatile and concentrated with a crisis, then they contribute to the fall in TFP observed during crises. My exercises suggest that variable markups account for 7.8% to 15.6% of the fall in TFP in the Argentine 2001/02 crisis.

Although these exercises suggest that a fraction of misallocation could be explained by endogenous factors, this paper is silent about the underlying causes behind the change in the distribution of market shares. Why do sectors become more concentrated with the crisis? Is it exit? Is it financial constraints that prevent some firms from operating at their desired level?

I think it is important to identify and model the source of the shock, since it could interact with variable markups in a non-trivial way. That is, it is likely that the source of the shock and the type of response by firms are not independent events. In Chapter 2
of this dissertation I explicitly address these questions.
1.6 Tables and Figures

Figure 1.1: Real Value Added Index

Figure 1.2: Market Share Distribution (logs), Largest Sectors

Source: INDEC.

Source: Own computations from Annual Manufacturing Survey.
Figure 1.3: Inverse Herfindahl Index

Source: Own computations from Annual Manufacturing Survey.

Figure 1.4: Market Shares and Markups, 2001

Source: Own computations from Annual Manufacturing Survey.
Figure 1.5: Market Shares and Markups, 1997

Table 1.1: Number of Active Firms per Year in AMS

<table>
<thead>
<tr>
<th>year</th>
<th># active firms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>4126</td>
</tr>
<tr>
<td>1998</td>
<td>3855</td>
</tr>
<tr>
<td>1999</td>
<td>3580</td>
</tr>
<tr>
<td>2000</td>
<td>3265</td>
</tr>
<tr>
<td>2001</td>
<td>3026</td>
</tr>
<tr>
<td>2002</td>
<td>2742</td>
</tr>
</tbody>
</table>
Table 1.2: Concentration Measures

<table>
<thead>
<tr>
<th></th>
<th>1997</th>
<th>2002</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Market Shares</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Median</td>
<td>0.4%</td>
<td>0.5%</td>
</tr>
<tr>
<td>Average Percentile 95</td>
<td>4.3%</td>
<td>6.9%</td>
</tr>
<tr>
<td>Average Maximum</td>
<td>19.5%</td>
<td>30.4%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>2.6%</td>
<td>4.5%</td>
</tr>
<tr>
<td><strong>Inverse Herfindahl Index</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>19.2</td>
<td>12.3</td>
</tr>
<tr>
<td>Median</td>
<td>11.5</td>
<td>6.4</td>
</tr>
</tbody>
</table>

Table 1.3: Percentage of the Fall in TFP explained by Variability of Markups

<table>
<thead>
<tr>
<th>Universe</th>
<th>Contribution to Fall in TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuing Firms</td>
<td>6.4%</td>
</tr>
<tr>
<td>Full Sample</td>
<td>7.9%</td>
</tr>
</tbody>
</table>
Table 1.4: Percentage of the Fall in TFP explained by Variability of Markups

<table>
<thead>
<tr>
<th>Universe</th>
<th>Contribution to Fall in TFP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuing Firms</td>
<td></td>
</tr>
<tr>
<td>Average across simulations</td>
<td>15.6%</td>
</tr>
<tr>
<td>Full Sample</td>
<td></td>
</tr>
<tr>
<td>Average across simulations</td>
<td>8.3%</td>
</tr>
</tbody>
</table>
1.7 Appendices

1.7.1 Appendix 1: Estimation of Number of workers per firm

All variables in the database are in "per workers in payroll" terms. The level of workers in payroll has not being provided, but we have two variables that can be used to try to recover it. The first one is the annual rate of change of workers in payroll. The second is size, which is defined in terms of workers in payroll. A firm is classified as small if it has between 10 and 80 workers in payroll, medium if it has between 81 and 200 workers in payroll, and large if it has over 200 workers in payroll. I follow Sandleris and Wright assuming that firms small or medium had a number of workers in payroll corresponding to the midpoint of their respective bin for the first year of the sample (1996). I then compute the number of workers in payroll for subsequent years with the provided rate of growth. If in any year the estimate of workers in payroll lies outside the size bin variable, I adjust the initial estimate so that the number of workers in payroll for that conflicting year lies in the border of the size bin. For large establishments that do not change size bins, I set the number of workers to match employment in aggregated data.
1.7.2 Appendix 2: Distribution of Wedges

The distribution of wedges for 1997 and 2002 is displayed in Figures 1.6 and 1.7. Capital wedges do not seem to vary much with the crisis. Scale wedges become more volatile (the standard deviation almost tripled) and tails became more fat with the crisis.

Figure 1.6: Capital Wedge Distribution

```
Distribution of Capital Wedge (in logs)

Source: Own computations from Annual Manufacturing Survey.
```

Figure 1.7: Scale Wedge Distribution

```
Distribution of Scale Wedge (in logs)

Source: Own computations from Annual Manufacturing Survey.
```
1.7.3 Appendix 3: Sectors in AMS and Simulated Subsectors

Argentina’s AMS classifies establishments in one of 22 two-digit industries. Their description and the total number of active firms in AMS (pooling figures for 1997 and 2002) are shown on the following Table.

Table 1.5: Sectors in the AMS

<table>
<thead>
<tr>
<th>Industries</th>
<th>#Active</th>
</tr>
</thead>
<tbody>
<tr>
<td>15 Food Products and Beverage</td>
<td>1635</td>
</tr>
<tr>
<td>16 Tobacco Products</td>
<td>15</td>
</tr>
<tr>
<td>17 Textile Products</td>
<td>494</td>
</tr>
<tr>
<td>18 Clothing Products</td>
<td>289</td>
</tr>
<tr>
<td>19 Leather Products</td>
<td>186</td>
</tr>
<tr>
<td>20 Wood and Cork Products (exc. Furnitures)</td>
<td>430</td>
</tr>
<tr>
<td>21 Pulp and Paper</td>
<td>178</td>
</tr>
<tr>
<td>22 Printing, Editing and Recording Activities</td>
<td>418</td>
</tr>
<tr>
<td>23 Petroleum and Coke (fuel) Products</td>
<td>11</td>
</tr>
<tr>
<td>24 Chemicals Products</td>
<td>415</td>
</tr>
<tr>
<td>25 Rubber and Plastic Products</td>
<td>402</td>
</tr>
<tr>
<td>26 Non-Metallic Mineral Products</td>
<td>401</td>
</tr>
<tr>
<td>27 Basic Metals</td>
<td>183</td>
</tr>
<tr>
<td>28 Fabricated Metal Products (exc. Machinery)</td>
<td>585</td>
</tr>
<tr>
<td>29 Mechanical Machinery and Equipment</td>
<td>593</td>
</tr>
<tr>
<td>30 Office Machinery</td>
<td>19</td>
</tr>
<tr>
<td>31 Electrical Machinery and Components</td>
<td>248</td>
</tr>
<tr>
<td>32 Radio, TV and Communication Devices</td>
<td>47</td>
</tr>
<tr>
<td>33 Medical, Optical and Precision Instruments</td>
<td>90</td>
</tr>
<tr>
<td>34 Motor Vehicles and Trailers</td>
<td>344</td>
</tr>
<tr>
<td>35 Other Type of Transportation Vehicles</td>
<td>93</td>
</tr>
<tr>
<td>36 Furnitures</td>
<td>352</td>
</tr>
</tbody>
</table>

In the paper, I randomly generate subsectors within each sector to resemble a 3-digit classification. Table 1.6 details the number of artificial subsectors within each sector.
Table 1.6: Artificial Subsectors Within Each Sector

<table>
<thead>
<tr>
<th>AMS Original Industry</th>
<th># Artificial Subsectors</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>28</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>24</td>
<td>9</td>
</tr>
<tr>
<td>25</td>
<td>7</td>
</tr>
<tr>
<td>26</td>
<td>8</td>
</tr>
<tr>
<td>27</td>
<td>3</td>
</tr>
<tr>
<td>28</td>
<td>9</td>
</tr>
<tr>
<td>29</td>
<td>10</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>34</td>
<td>5</td>
</tr>
<tr>
<td>35</td>
<td>1</td>
</tr>
<tr>
<td>36</td>
<td>6</td>
</tr>
</tbody>
</table>
CHAPTER 2

Market Power and Aggregate Efficiency in Financial Crises.

2.1 Introduction

Large financial crises in emerging economies are very costly in terms of output and productivity. To a large extent, their scope and magnitude as well as the mechanisms that shape aggregate dynamics, are not well understood.

One feature of financial crises whose implications have not been studied in detail is the resulting increase in market concentration. Using a detailed manufacturing database we document that during the Colombian 1998/99 financial crisis, large firms become relatively larger and small firms become relatively smaller. This is true if we look at concentration within narrowly defined sectors, with measures such as the Herfindahl Index, and is also true for the economy as a whole, as measured by the evolution of the share of production by the largest firms.

Our hypothesis is that the impact on concentration is the result of financial factors that affect firms asymmetrically. A vast number of studies have shown that smaller, younger firms are more likely to face liquidity constraints during downturns (see for example Gertler and Gilchrist (1994) or Perez-Quiroz and Timmermann (2000)), for reasons that may not be directly related to productivity. Small firms, for example, usually lack the collateral and/or reputation needed to obtain a loan; they typically cannot access international credit markets; they don’t have the lobbying power to access discretionary public credit; they have less access to formal sources of external finance; etc.

In this paper we analyze the aggregate consequences of the change in market power across firms that results from a credit crunch. If firms translate their market power into higher markups, more concentration could, in principle, result in higher markup
dispersion, which affects aggregate productivity negatively. On the other hand, the ability to change markups in response to a credit shock may act as a buffer that counteracts the distortions generated by the credit crunch. The overall quantitative effect will depend upon the calibration of the model.

We address this question quantitatively using a model where firms are subject to a collateral constraint on the amount of capital that they can rent. Based on Buera and Shin (2011), agents are heterogeneous with respect to their ability and wealth. Those that decide to be entrepreneurs are forced to self-finance at least a portion of their working capital needs. A financial crisis in this framework is modeled as an exogenous, unexpected tightening of the collateral constraint, or, alternatively, an increase in the fraction of working capital that needs to be self-financed by entrepreneurs.

To allow for heterogeneous market power across firms we depart from the canonical model of constant markups. In our framework, large firms set higher markups than small firms.\footnote{In what follows, unless otherwise noted, we use quantities produced as the measure of a firms size. In the model this is equivalent to hiring more workers or using more intermediate inputs.} This results from a demand system of the Kimball (1995) family, that allows for price elasticity to decrease with quantities. Large firms face a more inelastic demand for their differentiated product and thereby charge higher markups.

We discipline the model to match salient features of the Colombian economy before the 1998/99 financial crisis. To that end we use balance sheet data from Supersociedades, and data from the annual manufacturing survey by Colombia’s statistical agency, DANE. A convenient feature of the variable markups model that we adopt is that it nests the standard model of constant markups as a special case. We then calibrate two versions of the model: one with variable markups and another one where we restrict markups to be constant, across firms and time.

We use the model to measure the extent to which changes in market power that result from a financial crisis matter at the aggregate level. To do so we shock the economy with an unexpected tightening of financial conditions, intended to reproduce the actual credit crunch faced by firms, which can be inferred—with the help of the model— from financial data. To identify the role of market power quantitatively, we contrast the transition dynamics of the two versions of the model, variable vs. constant markups.
We find that heterogeneous markups dampen the negative response of output and aggregate productivity that follows a credit crunch. They act as a buffer that partially offsets the misallocation forced by the financial crisis. Intuitively, the credit crunch forces previously constrained firms (and firms that were close to the threshold before the credit crunch) to downsize. Large, unconstrained firms take the slack by increasing production in the constant markup case, and by (partially) increasing their markup in the variable markup case. The response in quantities is hence smaller in the variable markup case than in the constant markup case. Market power thus acts as a real rigidity, preventing firms from changing quantities (and prices) too much with respect to the initial equilibrium.

There are two sources of distortions in this model: financial frictions and heterogeneous markups. Financial frictions are exogenous to the firm, which tries to circumvent it through the accumulation of assets for self-financing purposes. Since the collateral constraint is imposed over capital, it affects the optimal capital-labor allocation, and is equivalent to a wedge on the capital-labor ratio. Heterogeneous markups are endogenous, in the sense that they can be interpreted as the deliberate decision of a firm to raise or lower their price in response to the competitive environment. It is equivalent to an endogenous wedge on size, or, as has been recently called by Bils et al. (2014), a product market wedge.

The assumption of heterogeneous markups, with larger firms charging higher markups, is informed by empirical evidence. First, revenue productivity, which should be equalized across firms if firms have the same markup, is increasing in size in the data. While markups are not the only reason for this pattern to hold, as any type of correlated distortion would generate such a pattern, it is certainly a possibility that this is the result of markups. Second, the ratio between the gross value of production and the value of intermediate inputs, which is positively related to markups, is increasing in size. In the next section we will further discuss the empirical evidence on heterogeneous markups.

There is yet another reason to include heterogeneous markups. One of the results of our quantitative exercise shows that a model of constant markups calibrated to the Colombian economy fails to reproduce the increase in market concentration triggered by a credit crunch. Variable markups are required to reproduce the pattern.

The contribution of this paper is twofold. First, it documents the increase in concen-
tration that occurs during a financial crisis, a feature that, to our knowledge, has not been previously discussed. Second, we contribute to the literature on resource misallocation over the business cycle, identifying specific sources of misallocation and analyzing the aggregate implication of their interaction under disruptive events.

This paper relates to many strands of literature. First, it relates to the literature that studies the misallocation of resources as an important determinant of aggregate productivity. Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) focus on distortions in the form of wedges that act as implicit taxes/subsidies that distort the allocation of capital and labor across firms. This group of papers, typically known as the ”indirect” approach, are often silent about the underlying channels through which misallocation takes place and consider the overall effect of all the potential sources of distortions in the economy. On the opposite side, those papers where misallocation arises as a result of a particular policy or institution are referred to as the ”direct” approach. This branch of the literature typically picks a particular friction and using quantitative models tries to assess its contribution to the misallocation of resources and its impact on aggregate TFP.

Whereas those seminal articles study cross-country income differences at a point in time, our paper focuses on time-series dynamics. We share this interest with Oberfield (2013) and Sandleris and Wright (2014), who document that misallocation of resources across firms accounts for a large portion of TFP losses during crises. We differentiate from these papers in two ways. First, they do not point to a particular source of misallocation (i.e. they take an indirect approach). We take a stance on the specific sources of misallocation (namely: financial frictions and heterogeneous markups) and evaluate their interaction in response to a financial crisis. Recently, Bils et al. (2014) use US data to argue that distortions in the form of product market wedges (markups) deserve a central place in macroeconomic research. Second and most important, while their approach is closer to Hsieh and Klenow (2009) in the sense that they perform an accounting exer-

---

2Bartelsman et al. (2013) is part of this set of papers. They followed a different method to measure cross country differences in the allocation of resources. However, the qualitative results are similar to the ones found in the aforementioned papers: the correlation between size and productivity is weaker in developing countries.

3The name of these approaches was first introduced by Restuccia and Rogerson (2013).

4For a more exhaustive description of the resource misallocation literature see Hopenhayn (2014), Jones (2011), Restuccia and Rogerson (2013) and references therein.
cise, we propose a model that accounts for optimal responses based on forward looking behavior over a transition path.

Our paper also relates to the vast literature that studies quantitatively the interaction between financial frictions and aggregate TFP (Amaral and Quintin (2010); Buera et al. (2011); Buera and Shin (2011); Midrigan and Xu (2014); Moll (2014)). Following those papers, we model financial frictions as a collateral constraint that represents a limited commitment problem between borrowers and lenders. The main difference between these papers and ours is the existence of competitive effects in product markets.\(^5\)

We relate to the many quantitative models that explore the macro and micro implications of heterogeneous markups in different scenarios. Atkeson and Burstein (2008), for example, use a model where variable markups arise from imperfect competition and strategic interaction to explain deviations from relative purchasing power parity. For simplicity, our model abstracts from strategic interactions, but still captures the main prediction of that model: larger firms set higher markups. Edmond et al. (2015) studied the role of pro-competitive effects in reducing product market distortions in a trade liberalization event. Opening the economy to international trade could generate a reduction in markup variability as the result of more competition in the domestic market. They provide conditions under which a trade reform could push the economy closer to the pareto optimum allocation. Instead of focusing in trade reform, our paper studies the interaction of these forces with a tightening in credit conditions faced by domestic firms. Peters (2013) explores the interaction between innovation, entry-exit decisions of firms and markup adjustment when entry costs are modified. He finds that reductions in entry costs reduce misallocation derived from the markups, and that fosters the growth rate of the economy through innovation. In our model, the occupational choice of agents will determine the entry and exit of the economy, and will also have a key role in the intensive margin of competition. However, by introducing a financial shock that affects incumbents in a heterogeneous way, our model will also have implications for the intensive margin of competition through changes in market shares.

As we explain in detail below, the way we introduce variable markups follows the

\(^5\)For a comprehensive review of the literature on misallocation and aggregate productivity, see Hopenhayn (2014). For a review focused on financial frictions and entrepreneurship, see Buera et al. (2015)
model first developed by Kimball (1995), and implemented by Klenow and Willis (2006) and Gopinath and Itskhoki (2010). They propose a demand system that allows for non-constant price elasticity, but that nest the standard constant elasticity benchmark as a special case. This allows us to contrast the variable markup case with constant markup case parsimoniously, by changing the value of a single parameter.

This paper is structured as follows. In the following section, we present some motivating facts: the increase in concentration that occurred during the financial crisis and evidence on heterogeneous markups. We then present the model that we use as framework. In Section 4 we present the data and calibration strategy, with a brief description of the Colombian crisis. Section 5 shows our quantitative exercise. Section 6 concludes.

2.2 Motivating Facts

2.2.1 Increase in Concentration

In this section (and in the calibration of the model) we use data from Colombia’s Annual Manufacturing Survey (AMS). The AMS is conducted by DANE, Colombia’s statistical agency. It surveys all plants with at least 10 employees, or a minimum gross value of production that is updated periodically. It has about 8,000 observations per year, all in manufacturing. Surveyed firms produce over 95% of aggregate value added. It has information on sales, production, number of workers, wages, value of intermediate inputs, etc. It classifies plants in 4-digit ISIC sectors, and has information on prices and products at a 7-digit level.

The Colombian economy suffered its most severe economic crisis since the 1930s in 1998/99, with real GDP falling approximately 5%. As a result, the Colombian economy became more concentrated, as measured by several indicators. Figure 2.1, for example, shows the evolution of the average (and median) Inverse Herfindahl Index (IHI) across sectors in Colombian manufacturing. The IHI is defined as the inverse of the square root of an index of the squared market shares of producers, as a measure of market concentration. A higher IHI indicates a more concentrated market.

---

6 A different implementation is suggested in Dotsey and King (2005)
7 A version of the database built by Eslava et al. (2004) with data covering the 1982-1998 period is readily available online. Microdata for a different time period (or to analyze different variables) can be accessed upon request in-situ, in the DANE offices in Bogotá.
8 We leave a more detailed description of the Colombian crisis for later.
9 A sector is defined at the 7-digit NAICS classification level.
sum of market shares within an industry. It is interpreted as the number of "effective competitors" within an industry: the lower the IHI, the more concentration in a sector. The fall in the IHI that results from the crisis is approximately 13\%\textsuperscript{10}. This pattern is partially explained by firm exit, but it also holds for continuing firms. Concentration as measured by the maximum market share also increases, on average, from 66\% in 1997 to 70\% in 1999.

Another measure of concentration used in quantitative applications is the value of production produced by the largest firms, as a share of total output. This measure also shows a (milder) increase of concentration. Before the crisis, the largest 10\% of firms produced 79\% of total value of production. At the height of the crisis they produced 81\%.

A caveat on the results above is that we are not including imported varieties in our computations. We think that, if anything, we are underestimating the increase in concentration, since financial crises in developing economies usually involve local currency depreciation that increases the relative price of imported varieties, many of which drastically reduce their market share or even disappear in local markets.

### 2.2.2 Heterogeneous Markups

There is evidence in Colombian microdata that firms set heterogeneous markups; more specifically, that larger firms set higher markups. The first evidence is given by revenue productivity. The literature on resource misallocation uses revenue productivity ($tfpr$) as a summary measure of distortions in an economy. It is defined as the price set by a firm times its physical productivity (i.e. $tfpr_{it} = p_{it} z_{it}$). It is easy to show that if prices are set as a constant markup over marginal costs, absent other distortions $tfpr$ is constant across firms.\textsuperscript{11} This does not hold in the data. Figure 2.2 shows a kernel regression between revenue productivity and firm size for Colombian manufacturing firms. The relationship is positive, consistent with larger firms setting higher markups.

More direct evidence on heterogeneous markups can be inferred from the ratio of the

\textsuperscript{10}The fall is similar if we define sectors at the 4-digit NAICS level

\textsuperscript{11}Under constant returns to scale, and absent other distortions, marginal costs take the form $mc_{it} = \Gamma_i z_{it}$, where $\Gamma_i$ is constant across firms. Hence, under constant markups $tfpr_{it} = markup \Gamma_i z_{it} = tfpr$ i.e. constant across firms.
gross value of production to the value of intermediate inputs. De Loecker and Warzynski (2012) show that markups are positively related to this ratio. Specifically:

\[ \mu_i = \nu \frac{p_i y_i}{p_i m_i} \]

where, \( \nu \) is the elasticity of output with respect to intermediate inputs and \( \mu \). Assuming that this elasticity is constant across firms (or that there is no systematic correlation between elasticity and firm size), markups move one to one with changes in this ratio. Figure 2.3 shows that, controlling by sector, this ratio is positively correlated with size.\(^\text{12}\)

### 2.3 A Model of Financial Frictions with Variable Markups

Based on the evidence presented above, we build an heterogeneous agents model where larger firms set higher markups. Following Buera and Shin (2011), we model financial frictions as a collateral constraint on the amount of capital that a firm can rent. That is, firms are forced to self-finance a fraction of their capital rental. A financial crisis in this environment is modeled as an exogenous increase in the fraction of capital that needs to be self-financed.

**Agent Heterogeneity and Occupational Choice**

There is a continuum of mass 1 of infinitely-lived agents \( i \). They have standard CRRA preferences over the consumption of a final good, and are heterogeneous with respect to their entrepreneurial ability and financial wealth. Each period, with probability \( \gamma \), an agent loses her entrepreneurial ability \( z_i \in Z \) and has to re-draw from the invariant distribution \( G(z) \). To match some features of the data, in our calibration exercise we assume that \( G(z) \) is a Pareto distribution with shape parameter \( \phi \). Financial wealth \( a_i \) is determined endogenously from savings decisions. Savings take the form of risk-free claims on physical capital. As we will discuss below, savings will serve two purposes: as self-insurance against idiosyncratic shocks, and as collateral to finance working capital.

---

\(^{12}\)We use deviations from industry means, where industry is defined at the 4-digit level. We do this to control for potential differences in technology parameter \( \nu \) across firms, i.e., firms within narrowly-defined sectors are more likely to have similar technologies. Examples of 4-digit sectors are "Soap and Detergent", and "Refrigerators and Washing Machines"
requirements (if the agent is an entrepreneur). We implicitly assume throughout that agents cannot borrow (i.e. it has to be the case that \( a \geq 0 \), as in Aiyagari (1994)).

Every period, an agent makes an occupational choice: she decides to work for a wage \( w_t \) or run her own business for a profit \( \pi_{it} \). Formally, it chooses \( o(z_{it}, a_{it}) \in \{0,1\} \), where \( o(z_{it}, a_{it}) = 1 \) if \( w < \pi_{it} \), and \( o(z_{it}, a_{it}) = 0 \) otherwise. The problem of an agent in period \( t \) can be expressed recursively as:

\[
V_t(z_{it}, a_{it}) = \frac{C_{it}^{1-\sigma}}{1-\sigma} + \beta E \left[ V_{t+1}(z_{it+1}, a_{it+1}) \mid z_{it} \right]
\]

subject to

\[
C_{it} + a_{it+1} = \max \{ w, \pi_{it}(z_{it}, a_{it}) \} + (1 + r_t)a_{it}
\]

where \( C_{it} \) is the consumption of the final good, which bundles all varieties in the economy. Notice how profits depend on both the ability of the entrepreneur that runs the firm, and her wealth; the latter as a result of the financial frictions. We explain the problem of the firm below.

**Final Good**

The only final good in this economy is produced by a competitive firm that aggregates all intermediate varieties using a constant return to scale aggregator of the family first described in Kimball (1995). Kimball’s aggregators are implicitly defined as

\[
\frac{1}{\Omega_t} \int_{\Omega_t} \Upsilon \left( \frac{\Omega_y \Omega_t}{\Upsilon_t} \right) di = 1, \text{ where } \Upsilon(1) = 1, \Upsilon' > 0, \text{ and } \Upsilon'' < 0, \text{ and } \Omega_t \text{ is the measure of varieties.}
\]

We use a particular specification of the Kimball aggregator, proposed by Klenow and Willis (2006) and later used in Gopinath and Itskhoki (2010). Under this specification, the demand for intermediate varieties takes the following form:

\[13\text{This restriction can potentially be binding for workers only, not for entrepreneurs, as they need } a \geq 0 \text{ to run their business, as we will discuss below.}

\[14\text{For simplicity, we abstract from explicitly mentioning the aggregate states (price, output, and financial friction parameter -see below-).}

\[15\text{More details on the Klenow and Willis (2006) specification can be found in Appendix A1. See Dotsey and King (2005) for an alternative specification.}

\]
\[ y_{it} = \begin{cases} \left[ 1 - \eta \ln \left( \frac{p_{it}}{P_t} \right) \right]^{\frac{\sigma}{\eta}} Y_t & \text{if } 1 - \eta \ln \left( \frac{p_{it}}{P_t} \right) > 0 \\ 0 & \text{otherwise} \end{cases} \] (2.2)

It is easy to show that the price elasticity of demand faced by firm \( i \) at time \( t \), \( \varepsilon_{it} \) is given by:

\[ \varepsilon_{it} = \frac{\sigma}{1 - \eta \ln \left( \frac{p_{it}}{P_t} \right)} \] (2.3)

This illustrates two important features of this demand system. First, for a given level of \( \eta \) and \( \sigma \), demand is more elastic the higher the relative price. Second, this model nests the standard constant elasticity of substitution (CES) case, since when \( \eta \to 0 \), \( \varepsilon_{it} \to \sigma \). This will be useful to use as a constant markup benchmark in the quantitative exercises.

Parameter \( \eta \) pins down the rate at which demand elasticity changes along the demand curve, or superelasticity \( \varepsilon^s \):

\[ \varepsilon^s_{it} = \frac{\eta}{1 - \eta \ln \left( \frac{p_{it}}{P_t} \right)} \] (2.4)

Again, when \( \eta \to 0 \), \( \varepsilon^s_{it} \to 0 \); i.e. elasticity is constant along the demand curve. Figure 2.4 displays the resulting demand functions for different values of \( \eta \). As mentioned before, \( \eta = 0 \) corresponds to the CES case.

The Problem of the Entrepreneur

Agents that decide to become entrepreneurs in period \( t \) hire labor, materials and capital to produce a differentiated intermediate good \( y_{it} \). They have access to a constant returns to scale Cobb-Douglas technology indexed by their own ability:

\[ y_{it} = z_{it} \left( k_{it}^{\alpha} l_{it}^{1-\alpha} \right)^{\nu} m_{it}^{(1-\nu)}. \]

The amount of capital that an entrepreneur can rent in any period is subject to a collateral constraint of the form \( k_{it} \leq \lambda_t a_{it} \). The economy-wide parameter \( \lambda_t (\geq 1) \) summarizes in a parsimonious way the degree of financial frictions or credit constraints in the economy. A low \( \lambda_t \) is associated with a high degree of financial frictions in an economy. In the limit,
when $\lambda_t \to 1$, firms must self-finance all of their capital rental. At the other extreme, when $\lambda_t \to \infty$, the economy has perfect capital markets. In this case, saving decisions are independent of production decisions and have consumption smoothing as their sole motive.\textsuperscript{16}. Materials is a composite good, produced with the same technology as the final good, as in Basu (1995). Its price is thus equal to that of the final good in equilibrium. We introduce materials to the model to better back up markups from the data, as we explain below.

Let $mc_{it}$ stand for the marginal cost. In this setting the problem of the entrepreneur is:

$$
\pi_t(z_{it}) = \max_p p_{it}y_{it} - mc_{it}y_{it}
$$

subject to

$$
y_{it} = \left[1 - \eta \ln \left(\frac{p_{it}}{P_t}\right)\right]^\frac{\sigma}{\eta} Y_t
$$

$$
k_{it} \leq \lambda_t a_{it}
$$

In an unconstrained optimum (i.e. $k^*_{it} < \lambda_t a_{it}$) $mc_{it} = \frac{\Gamma_t}{z_{it}}$, where

$$
\Gamma_t = \left(\frac{w_t}{(1-\alpha)^\nu}\right)^\nu \left(\frac{m_t}{(1-\alpha)}\right)^\nu \left(\frac{p_m}{(1-\nu)}\right)^{(1-\nu)}
$$

That is, given factor prices and entrepreneurial ability, the marginal cost of production is constant for the unconstrained entrepreneur. In that case, the optimal pricing rule is:

$$
p_{it} = \frac{\varepsilon_{it}}{(\varepsilon_{it} - 1)} mc_{it} \tag{2.5}
$$

Substituting $\varepsilon_{it}$ by the expression in (2.3), the (gross) markup set by firm $i$ at time $t$, $\mu_{it}$, is given by:

$$
\mu_{it} = \frac{\sigma}{\sigma - 1 + \eta \ln \left(\frac{p_{it}}{P_t}\right)} \tag{2.6}
$$

Unsurprisingly, the markup charged by an unconstrained firm is a function of the

\textsuperscript{16}Financial constraints of this form capture in a simple way patterns that arise endogenously in limited commitment environments. See Buera and Shin (2011) for further details
demand elasticity it faces, which varies along the demand curve as long as \( \eta > 0 \). Specifically, markups are higher the more inelastic the demand. Once again, when \( \eta \to 0 \) markups are constant and equal to \( \left( \frac{\sigma}{\sigma - 1} \right) \).

This model does not deliver imperfect competition through any type of strategic interaction. It does, however, capture in a simple way the main predictions of such models, namely: the fact that larger firms set higher markups.\(^\text{17}\) We will refer to this fact, somewhat loosely, as large firms exploiting their market power.

In a constrained optimum (i.e. \( k^*_t > \lambda t a_t \)), the entrepreneur leverages as much as possible, sets \( k^*_t = \lambda t a_t \), and marginal cost become an increasing function of \( y_{it} \).\(^\text{18}\) Two firms with the same productivity but different net worth (i.e. different \( a_{it} \)) may charge different prices, if the collateral constraint is binding for the poorest one.

For illustration purposes, Figure 2.5 shows a simple partial equilibrium example of two such firms facing a linear demand. In a world with perfect credit markets (or with imperfect credit markets but where firms are wealthy enough), marginal costs are constant for both firms, and equal to \( mc_u \), they both produce at the unconstrained optimum \( q_u \), charge unconstrained prices \( p_u \), and set a markup given by the vertical segment CD. If instead there is a collateral constraint that is binding for one of the firms, the marginal cost curve for that firm becomes \( mc_c \), increasing in quantities from the point the firm becomes constrained. The constrained firm now produces at the point where constrained marginal costs are equal to marginal income, \( q_c < q_u \), and charges a price \( p_c > p_u \). Markups for the constrained firm are given by the vertical segment AB. The markup set by a constrained firm will then be equivalent to that of an unconstrained firm with marginal costs depicted as an horizontal line through point B.

\(^{17}\)See for example Atkeson and Burstein (2008)

\(^{18}\)More details on the algebraic expression for marginal costs of the constrained firm can be found on Appendix 2
Equilibrium

Given an initial joint distribution of wealth and entrepreneurial ability \( G_0(z, a) \), an equilibrium in this economy consists of allocations
\[
\{ C_s(z_t, a_t), a_s(z_t, a_t), k_s(z_t, a_t), l(z_t, a_t), m_s(z_t, a_t), oco_s(z_t, a_t) \}_{s=t}^{\infty}
\]
for all \( t \geq 0 \), sequences of joint distribution for wealth and entrepreneurial ability \( \{ G_t(z, a) \}_{t=1}^{\infty} \), individual prices \( \{ p_s(z_t, a_t) \}_{s=t}^{\infty} \), and aggregate prices \( \{ w_t, r_t, p_m^t, P_t \}_{t=1}^{\infty} \), such that

1. Given \( \{ w_t, r_t, p_m^t, P_t \} \), \( z_t \), and \( a_t \),
\[
\{ C_s(z_t, a_t), a_s(z_t, a_t), k_s(z_t, a_t), l(z_t, a_t), m_s(z_t, a_t), oco_s(z_t, a_t) \}_{s=t}^{\infty}
\]
solves the agent’s problem for all \( t \geq 0 \).

2. Markets for labor, capital, and final goods clear

\[
\int_{z \in Z} \int_{a(z_t)}^{\infty} l_t(z_t, a_t) G_t(da, dz) = \int_{z \in Z} G_t(a(z_t), dz)
\]

\[
\int_{z \in Z} \int_{a(z_t)}^{\infty} l_t(z_t, a_t) G_t(da, dz) = \int_{z \in Z} a G_t(a(z_t), dz)
\]

\[
\int_{z \in Z} \int_{a(z_t)}^{\infty} C_t(z_t, a_t) G_t(da, dz) + \int_{z \in Z} \int_{a(z_t)}^{\infty} m_t(z_t, a_t) G_t(da, dz) +
\]

\[
+ \int_{z \in Z} (\lambda_{t+1}(z_t, a_t) - (1 - \delta) a_t(z_{t-1}, a_{t-1})) G_t(a(z_t), dz) = Y_t
\]

3. The joint distribution of wealth and entrepreneurial ability
\( \{ G_t(z, a) \}_{t=1}^{\infty} \) evolves according to the following law of motion:

\[
G_{t+1}(a, z) = \gamma \int_{a_{t+1}(a, z) \leq a} G_t(da, z)
\]

\[
+ (1 - \gamma) J(z) \int_{a_{t+1}(a, z) \leq a} G_t(da, z)
\]
2.4 Data and Calibration

We calibrate the model to match salient features of the 1998-1999 financial crisis in Colombia. We chose Colombia for several reasons. First, it is a country where financial frictions are prevalent. According to the World Bank’s World Enterprise Survey, which surveys managers on a number of issues regarding the business environment, access to external finance is the biggest obstacle faced by firms. Second, it suffered a large financial crisis in 1998-99 (see subsection below). Third, it has wide availability of high-quality microdata that spans the crisis period.

The Colombian 1998/99 Financial Crisis

One of the most salient features of the Colombian economy relative to other Latin American countries is its macroeconomic stability in terms of business cycles. Even though there were some minor recessions in 1983, 1991 and 1996, the crisis of the end of the twentieth century in Colombia 1998/99 was the most severe since the 1930s, both in terms of depth and duration. To be more specific, Colombian GDP fell about 5% between 1997 and 1999, while the unemployment rate rose to a peak of 20.5% in the third quarter of 2000.

As we can see in Figure 2.6, the decrease in output was generalized in all sectors. Agriculture, Manufacturing and Services suffered an important decline in production during that period.

Carrasquilla et al. (2000), Garzón (2001), and Urrutia and Llano (2011) present evidence suggesting that the reduction in economic activity was due to a big drop in the supply of credit which is typically refer in the literature as credit crunch. The magnitude of the credit crunch was remarkable. Figure 2.7 and Figure 2.9 show the evolution of the amount of loans in Colombia (in real terms), and its deviations from trend, respectively. Total credit was 20% below trend at the height of the crisis, a fall of over 30% in absolute terms. The decline was common to all major lines of credit (see Figure

---

19 Specifically, 29% of the managers surveyed in Colombia identified access to finance as the biggest obstacle, (in Latin America the average is 15% and in emerging economies 17%). 41% see it as a major constraint, also the highest ranked constraint in Colombia (compare to 30% in Latin America, 29% in emerging economies).

20 See for example Lorenzoni and Guerrieri (2009) or Buera and Moll (2015)
2.8). Between January of 1998 and December of 1999, the stock of loans fell by 17%.

The natural question then is, what are the mechanism behind this reduction in the supply of credit? Two key factors have been pointed out in the literature:

1. A reduction in external supply of credit due to the financial crisis in Asia and Russia.

2. An increase in the perception of risk by banks due to a worsening in the quality of bank’s portfolios.

The decrease in the flow of funds from the rest of the world was a consequence of the increase perception of risk in the global economy due to the financial crisis in Asia in the first place, and then it was exacerbated with the Russian crisis in mid 1998.

In this paper, we do not take a stand on any particular mechanism leading to the credit crunch. Without loss of generality, we take a reduced form approach and parsimoniously model it as an exogenous tightening of the collateral constraint over the rent of capital.

**Data Sources**

We use two main data sources to discipline the model quantitatively: the Annual Manufacturing Survey (previously described), and balance sheet data.

Balance sheet data comes from *Supersociedades*, the government office that regulates corporations. The database includes all firms with assets or income over 20,000 monthly minimum wages, multinational corporations, or firms owned at least 20% by multinationals. Firms that do not comply with any of those requirements may still be included if the Superintendent decides to do so. Coverage varies across years, but the data set contains around 9,000 observations per year, starting in 1995. Manufacturing observations amount to about 2,500 per year.

**Calibration of the Steady State**

We calibrate a quantitative version of the model to the state of the Colombian economy before the 1998/99 financial crisis. To solve quantitatively for the steady state, we discretize the state space of entrepreneurial abilities and asset levels \((z, a)\), and use value
function iteration. We rely on parallelization methods for efficiency purposes.\textsuperscript{21} The model has 10 parameters, namely \{\(\beta, \delta, \alpha, \nu, \gamma, \phi, \lambda, \eta, \sigma\}\}. The first 5 parameters are relatively standard, and we borrow them from related literature. We set the time preference parameter \(\beta = 0.93\) and the utility preference parameter \(\rho = 1.5\). We set the depreciation rate \(\delta = 0.06\). We set \(\alpha\), the share of capital in value added equal to 0.33. The elasticity of output with respect to intermediate inputs is taken from Eslava et al. (2004), who estimate it using Colombia’s AMS for the pre-crisis period. We set \(\nu = 0.54\).

We choose the probability of losing ability \(z\) in any given period to match firm exit rate in the AMS. This results in \(\gamma = 0.1\).

We calibrate \(\lambda\) from financial data. In the model, constrained firms have \(\frac{k-a}{a} = \lambda - 1\), and unconstrained firms have \(\frac{k-a}{a} < \lambda - 1\). The numerator in the left hand side represents the intra-period loans that the firm uses to finance its working capital requirements. The denominator, the assets that the firm uses as collateral. Figure 2.10 shows the model implications for the relationship between this ratio and firm size, as measured by collateralizable assets. Smaller firms display the same ratio since they are, on average, constrained. We use this feature of the model to discipline \(\lambda\) in steady state. In particular, we use balance sheet data to compute the ratio of short-term loans to fixed assets (which we use as proxy to collateralizable assets) for the smallest 25\% of firms. This results in \(\lambda = 2.75\).\textsuperscript{22}

The remaining parameters are the price elasticity parameter \(\sigma\), the parameter driving the superelasticity, \(\eta\), and the shape of the Pareto distribution of abilities \(\phi\). In our quantitative exercise we contrast the response of a constant markup vs. a variable markup model, so these parameters may vary across calibrations.

In the variable markup specification, our strategy is to recover markups from AMS microdata. To do so we rely on the previously mentioned result by De Loecker and Warzynski (2012) (reproduced here for convenience):

\[
\mu_i = \nu \frac{p_i y_i}{p_i^m m_i}
\]

\textsuperscript{21}See Appendix 3 for further technical details on the computational procedures.  
\textsuperscript{22}We will show in our quantitative exercise that different potential measures of \(\lambda\) from financial data display the same evolution before, during, and after the crisis.
That is, given the production function parameter for intermediate inputs, we recover markups from the ratio of gross value of production to expenditure on intermediate inputs, data available in the AMS.\textsuperscript{23} We set $\sigma = 6$ to match the average markup of 1.19 observed in the data. We jointly calibrate $\eta = 0.5$ and $\phi = 2.6$ to match the variability of markups and the share of gross value of production produced by the largest 10% of firms.

For the constant markup model we set trivially $\eta = 0$. i.e. markups do not vary across firms. Again, we set $\sigma = 6$ such that the average markups matches the data and the variable markup specification. We calibrate $\phi = 2.3$ to match concentration at the top 10% of firms. Tables 2.1 and 2.2 summarize the calibration just described.

### 2.5 Quantitative Exercise

How do aggregate variables respond to a credit crunch if firms have market power? To answer this question, we compare the transition dynamics of aggregate variables in a variable markups model with those in a constant markups model.

To simulate a credit crunch, we assume a one-time, unexpected shock to the financial friction parameter $\lambda$. The future path of $\lambda$ is perfectly known by all agents after the initial shock. We back up the path of $\lambda$ from the median ratio of short-term loans to fixed assets for relatively small firms, as described in the Calibration section. Figure 2.11 displays the evolution of this ratio using different balance-sheet concepts for short-term loans and collateralizable assets. The solid line is our preferred choice. Their paths are similar, so we are confident that the shock that we feed into the model resembles the financial frictions actually faced by Colombian firms during that period. Figure 2.11 shows that it took approximately 6 years for the financial conditions prevailing before the crisis to fully recover. From then, we assume that $\lambda$ remains constant forever.

For the following transition exercise we assume a small open economy where the interest rate remains unchanged at its steady state level. We take this approach to reflect the fact that Colombia is a small country, with a relatively open capital account. This assumption also yields more computational precision, since it allows the program to run

\textsuperscript{23}See Appendix 4 for details on how the (De Loecker and Warzynski, 2012) framework map into our model with financial frictions.
more efficiently.

Figure 2.12 shows the evolution of concentration in economic activity that results from the financial crisis. We measure concentration as the gross value of production produced by the top 10% firms. The dashed line is actual data, the dotted line correspond to the constant markup case, and the solid line to the variable markup model. The constant markup model calibrated as explained above fails to reproduce the increase in concentration that follows the credit crunch. In fact, it results in less concentration. If instead we allow for market power, the model generates the concentration patterns observed in the data.

Regarding output, Figure 2.13 contrasts the evolution of yearly, hp-detrended aggregate output with those from the models. The first thing to note is how both the variable markup and constant markup models reproduce quite well the qualitative path of aggregate output: a large fall, followed by a slower but steady recovery that takes about 6 years. We think of this as a success of our calibration strategy for the financial shock, since even though it does not rely on targeting aggregate variables (but rather on looking at the model-implied financial conditions of small firms), it still delivers the same aggregate pattern.

The orders of magnitude are also similar, but the output losses predicted by the variable markup model are milder. The same is true for aggregate productivity, depicted on Figure 2.14: it displays roughly the same pattern in the model and in the data, but the fall is not as steep in the variable markup case. In particular, aggregate productivity explains half of the fall in output in the data, but only a third of aggregate aggregate output in the constant markup model, and less than that in the variable markup case.

Overall, heterogeneous markups act as a buffer of a credit crunch. What drives the differences across models? Part of the answer lies in the larger reallocation of capital that results from a credit crunch in a constant markup model. Figure 2.15 displays the evolution of capital dispersion across firms. Dispersion increases in both models, but the increase is larger in the constant markup case. That is, the reallocation of capital across firms triggered by the credit crunch is larger in the standard constant markup model. Intuitively, the credit crunch forces previously constrained firms (and firms close to the threshold before the credit crunch) to downsize. Large, unconstrained firms take the
slack by increasing production in the constant markup case, and by (partially) increasing their markup in the variable markup case. Market power thus acts as a real rigidity, preventing firms from changing prices too much with respect to the initial equilibrium.

Another reason is that the average size of firms in the constant markup case is smaller than in the variable markup case, which translates into larger TFP losses. This follows from differences at the intensive and extensive margins across models. At the intensive margin, incumbent entrepreneurs that are negatively affected by the credit crunch downsize more in the constant markup case, for the reasons explained above. At the extensive margin, the entry of infra-marginal entrepreneurs is higher in the constant markup case, which also reduces the average firm size.

The fact that there is net entry of entrepreneurs under both versions of the model may seem at odds with the intuition of higher exit rates of firms during crises. It is, however, consistent with existing evidence on the countercyclical nature of self-employment, as documented for example in Mondragón Vélez and Pena (2008) for Colombia, or in Earle and Sakova (2000) for eastern European countries. Figures 2.16 and 2.17 show the differences across models at the extensive margin. The model with constant displays a fatter left tail after the crisis than the model with heterogeneous markups, i.e. there’s more entry of low productivity entrepreneurs into self-finance.

2.6 Conclusion

In this paper we document the increase in market concentration that resulted from the 1998/99 Colombian financial crisis. We analyze the extent to which market power helps shape aggregate dynamics. We find that heterogeneous markups are needed in order to reproduce the concentration patterns that follow from a financial crisis. However, variable markups dampen the response of output and productivity in a credit crunch, since they partially offset the resource misallocation that follows. This does not mean that there is less resource misallocation in steady state in an environment with market power, but rather it illustrates the dynamics that follow a negative credit shock.

In line with recent research, we find that resource misallocation can help explain a non-negligible share of the large output losses during financial crises. But we bring attention
to the fact that more distortions do not necessarily result in larger losses. The interaction between financial frictions and heterogeneous markups during a negative credit shock is an example of the ambiguous effect of distortions in a second best world.
2.7 Tables and Figures

Figure 2.1: Inverse Herfindahl Index (1997=100)

Source: Annual Manufacturing Survey.

Figure 2.2: Revenue Productivity and Size

Source: Annual Manufacturing Survey.
Figure 2.3: GVP-to-Materials Ratio and Size

Source: Annual Manufacturing Survey.

Figure 2.4: Demand for Different Superelasticities
Figure 2.5: Partial Equilibrium Example

Figure 2.6: Real Gross Domestic Product (1997=100)

Source: DANE.
Figure 2.7: Real Net Credit (1997=100)

Source: Central Bank of Colombia.

Figure 2.8: Credit as a Share of GDP

Source: Central Bank of Colombia.
Figure 2.9: Real Net Credit, Deviations from Trend (1997=100)

Source: Central Bank of Colombia.

Figure 2.10: Model Implied Leverage vs. Size
Figure 2.11: Different Measures of Credit Crunch (1997=1)

Source: Own computations from Supersociedades.

Figure 2.12: Value of Production by Top 10% Firms

Source: Own computations and Annual Manufacturing Survey.
Figure 2.13: Aggregate Output

Source: Own computations and DANE.

---

Figure 2.14: Aggregate Productivity

Source: Own computations and DANE.
Figure 2.15: Dispersion of Capital Across Firms

Figure 2.16: Cumulative Distribution of Entrepreneurs. Constant Markups
Figure 2.17: Cumulative Distribution of Entrepreneurs Variable Markups
### Table 2.1: Calibration 1

**Invariant Across Calibrations**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Concept</th>
<th>Target/Source</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>Intertemporal preference</td>
<td>Literature</td>
<td>0.93</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Preference parameter</td>
<td>Literature</td>
<td>1.5</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Depreciation</td>
<td>Literature</td>
<td>0.06</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Capital share in value added</td>
<td>Literature</td>
<td>0.3</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Production function parameter</td>
<td>(Eslava et al., 2004)</td>
<td>0.57</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Ability loss probability</td>
<td>Exit Rate</td>
<td>0.1</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Financial friciton</td>
<td>Liab./F. Assets</td>
<td>2.75</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Elasticity</td>
<td>Average Markups</td>
<td>6</td>
</tr>
</tbody>
</table>

### Table 2.2: Calibration 2

**Vary Across Calibrations**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Concept</th>
<th>Target/Source</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Constant Markups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta$</td>
<td>Superelasticity</td>
<td>Exogenous</td>
<td>0</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Pareto Shape</td>
<td>Concentration top 10%</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td><strong>Variable Markups</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\eta$</td>
<td>Superelasticity</td>
<td>Markup Variability</td>
<td>0.5</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Pareto Shape</td>
<td>Concentration top 10%</td>
<td>2.3</td>
</tr>
</tbody>
</table>
2.8 Appendices

2.8.1 Appendix 1: Kimball Aggregator and Demand Approximation

As stated above, Kimball’s aggregator satisfies:

\[ \frac{1}{\Omega_t} \int_{\Omega_t} \Upsilon \left( \frac{\Omega_t y_{it}}{Y_t} \right) \, di = 1 \]

In the Klenow-Willis specification, function \( \Upsilon(x) \) takes the following form:

\[ \Upsilon(x) = 1 + (\epsilon - 1) \exp \left( \frac{1}{\eta} \right) \eta^{\frac{\eta}{\epsilon - 1}} \left( \Gamma \left( \frac{\epsilon}{\eta}, \frac{1}{\eta} \right) - \Gamma \left( \frac{\epsilon}{\eta}, \frac{x}{\eta} \right) \right) \]

where:

\[ \Gamma(u, z) = \int_z^\infty s^{u-1} e^{-s} \]

Demand for intermediate variety \( i \) is: \(^{24}\)

\[ y_{it} = \left[ 1 - \eta \ln \left( \frac{\sigma x_i}{\sigma - 1} \right) \right]^{\frac{\eta}{\sigma}} Y_t \]

where \( x_i = D_t \frac{P_t}{Y_t} \), \( \sigma > 1 \), \( \eta > 0 \), and

\[ D_t = \frac{\sigma - 1}{\sigma} \int \frac{y_{it}}{Y_t} \exp \left\{ \frac{1}{\eta} \left[ 1 - \left( \frac{\Omega y_{it}}{Y_t} \right)^\frac{\eta}{\epsilon} \right] \right\} \, di \]

Gopinath and Itskholi (2010) show that, up to a first order approximation, \( D_t \approx \frac{\sigma - 1}{\sigma} \).

Using this result, demand for variety \( i \) becomes: \(^{25}\)

\[ y_{it} = \left[ 1 - \eta \ln \left( \frac{p_{it}}{P_t} \right) \right]^{\frac{\eta}{\sigma}} Y_t \]

which is the expression presented in the body of the text.

---

\(^{24}\) See (Klenow and Willis, 2006) for further details on the derivation of the demand function.

\(^{25}\) In the quantitative exercises, we use this approximation to compute the results and then check that it holds in equilibrium.
2.8.2 Appendix 2: Pricing Decisions for the Constrained Firms

As we mentioned in the text, we can think of the profit maximization problem in two stages. First, firms solve the unconstrained problem (as in the no financial frictions case), where \( mc_{it} = \frac{\Gamma}{z_{it}} \). If unconstrained capital is \( k_{it}^u \leq \lambda_t a_{it} \) then the unconstrained solution holds and firms set their prices according to

\[
p_{it} = \frac{\varepsilon(x_i)}{(\varepsilon(x_i) - 1)} \frac{\Gamma_t}{z_{it}}
\]

If \( k_{it}^u > \lambda_t a_{it} \), then the entrepreneur will set \( k_{it} = \lambda_t a_{it} \). We can derive the pricing decision in the following way. From the first order conditions of the cost minimization problem of an unconstrained entrepreneur, we have

\[
m_{it} = (1 - \nu) \frac{\nu(1 - \alpha)}{\nu(1 - \alpha)} \frac{w_t}{p_t} \frac{1}{x_i} \frac{(1 - \nu)}{\nu(1 - \alpha)} \frac{1}{p_t}
\]

Plugging this in the production function of a constrained firm yields

\[
y_{it} = z_{it} (\lambda_t a_{it})^{\alpha \nu} \left[ \frac{(1 - \nu) \frac{w_t}{p_t} \frac{1}{x_i} (1 - \nu)}{\nu(1 - \alpha)} \frac{1}{p_t} \right]^{1 - \alpha \nu}
\]

so we can express the conditional labor demand as

\[
l_{it} = \frac{y_{it} z_{it} (\lambda_t a_{it})^{\alpha \nu} \left[ \frac{(1 - \nu) \frac{w_t}{p_t} \frac{1}{x_i} (1 - \nu)}{\nu(1 - \alpha)} \frac{1}{p_t} \right]^{1 - \alpha \nu}}{(1 - \nu) \frac{w_t}{p_t} \frac{1}{x_i} (1 - \nu)}
\]

and use this in the objective function to express the constrained problem as:

\[
\begin{align*}
\max_{p_{it}} p_{it} y_{it} - & \left[ 1 + \frac{(1 - \nu) w^{1 - \alpha \nu}}{\nu(1 - \alpha)} \right] \Delta^{\frac{1}{1 - \alpha \nu}} [y_{it}]^{\frac{1}{1 - \alpha \nu}} - r \lambda a_{it} \\
\text{st.} \quad & y_{it} = Y_t \left[ 1 - \eta \ln \left( \frac{p_{it}}{P_t} \right) \right]^{\frac{\eta}{\sigma}}
\end{align*}
\]

with

\[
\Delta = \left[ \frac{\nu(1 - \alpha)}{w^{1 - \alpha \nu} y_{it}} \right] \frac{(1 - \nu) \frac{w_t}{p_t} \frac{1}{x_i} (1 - \nu)}{\nu(1 - \alpha)} \frac{1}{p_t}
\]

For simplicity, let \( \Delta' = \left[ 1 + \frac{(1 - \nu) w^{1 - \alpha \nu}}{\nu(1 - \alpha)} \right] \Delta^{\frac{1}{1 - \alpha \nu}} \). Plugging in the demand into the objective function, and letting \( \Xi = 1 - \eta \ln \left( \frac{p_{it}}{P_t} \right) \)

\[
\begin{align*}
\max_{p_{it}} p_{it} y_{it} - & \left[ 1 + \frac{(1 - \nu) w^{1 - \alpha \nu}}{\nu(1 - \alpha)} \right] \Delta^{\frac{1}{1 - \alpha \nu}} [y_{it}]^{\frac{1}{1 - \alpha \nu}} - r \lambda a_{it} \\
\end{align*}
\]

\footnote{Notice that this is equivalent to having \( mc_{it}(y_{it}) = \left[ 1 + \frac{(1 - \nu) w^{1 - \alpha \nu}}{\nu(1 - \alpha)} \right] (1 - \alpha \nu) \Delta^{\frac{1}{1 - \alpha \nu}} [y_{it}]^{\frac{\alpha \nu}{1 - \alpha \nu}} \) (plus a fixed cost in the form of \( r \lambda_t a_{it} \)).}
\[ st. \quad y_{it} = Y_t \left[ 1 - \eta \ln \left( \frac{p_{it}}{P_t} \right) \right]^{\frac{\sigma}{\eta}} \]

with

\[ \Delta = \left[ \frac{w^{\nu(1-\alpha)} w^{\tau(1-\alpha)} y_{it}}{z_{it}(\lambda a_{it})^{\alpha \nu} \left[ (1-\nu) \frac{1}{\nu} \frac{1}{\eta^{\nu}} (1-\nu) \right]} \right] \]

For simplicity, let \( \Delta' \equiv \left[ 1 + \frac{(1-\nu)w^{1-\alpha}}{\nu(1-\alpha)} \right] \Delta \frac{1}{\tau-\alpha}, \)

Plugging in the demand into the objective function, and letting \( \Xi(p_{it}) \equiv 1 - \eta \ln \left( \frac{p_{it}}{P_t} \right) \)

\[ \max_{p_{it}} p_{it} Y_t \left[ \Xi(p_{it}) \right]^{\frac{\sigma}{\eta}} - \Delta' Y_t^{\frac{1}{\tau-\alpha}} \left[ \Xi(p_{it}) \right]^{\frac{\sigma}{\eta}} \frac{1}{\tau-\alpha} - r \lambda a_{it} \]

FOC

\[ \left[ \Xi(p_{it}) \right]^{\frac{\sigma}{\eta}} - \sigma \left[ \Xi(p_{it}) \right]^{\frac{\sigma-\eta}{\eta}} P_t + \Delta' Y_t^{\frac{1}{\tau-\alpha}} \frac{1}{\eta} \left\{ \left[ \Xi(p_{it}) \right]^{\frac{\sigma-\eta(1-\alpha)}{\eta(1-\alpha)}} \right\} \frac{\eta P_t}{p_{it}} = 0 \]

If \( P_t = 1 \)

\[ \left[ \Xi(p_{it}) \right]^{\frac{\sigma}{\eta}} - \sigma \left[ \Xi(p_{it}) \right]^{\frac{\sigma-\eta}{\eta}} + \Delta' Y_t^{\frac{1}{\tau-\alpha}} \frac{1}{\eta} \left\{ \left[ \Xi(p_{it}) \right]^{\frac{\sigma-\eta(1-\alpha)}{\eta(1-\alpha)}} \right\} \frac{1}{p_{it}} = 0 \]

Solving quantitatively for \( p_{it} \) in the equation above yields the pricing decision for the constrained firm.
2.8.3 Appendix 3: Solution Algorithm

We solve quantitatively for the steady state of the model as follows. We first discretize the space of abilities on a grid with 40 points. We set \( z_1 = 1 \), and choose \( z_{38} \) such that the cumulative probability up to \( z_{38} \) is 0.99 according to a pareto distribution (i.e. \( G(z_{38}) = 0.99 \)). The grid is equally spaced between \( z_1 \) and \( z_{38} \). We set \( z_{39} \) and \( z_{40} \) such that \( G(z_{39}) = 0.995 \) and \( G(z_{40}) = 0.999 \).

The grid for asset holdings consist of 1451 points. The first 400 points are relatively close to each other, and are equally spaced. The second part of the grid (the last 1051 points) is coarser. For interpolation purposes during the value function iteration routine, each bin in the grid is divided in 400 sub-bins, such that the value function is evaluated at 1451*400=580,400 points.

We normalize the price of the aggregate good to 1. Given aggregate output, wages, and the rental rate, we compute profits for constrained and unconstrained entrepreneurs, and check for the occupational choice condition. We then solve the recursive problem of the agents through value function iteration. We compute the invariant distribution of assets iterating over the joint distribution of assets and abilities. We then check the market clearing conditions for the rental rate, wages and aggregate output (the loop variables).

The algorithm is written in Fortran 90, and we use MPI parallelization methods to solve over 40 nodes, where each node is a point in the ability space.

To compute the transition dynamics of the system, we assume that agents suffer a completely unexpected shock, and then have perfect foresight over the future path of exogenous \( \lambda_t \). For computational reasons, we compute the transitions of the system assuming a small open economy, such that \( r \) is given.
2.8.4 Appendix 4: Markups from Microdata

As in De Loecker and Warzynski (2012), suppose we have a production function that produces using \( V \) variable inputs (e.g. labor, materials, energy), and one factor that is treaded as a dynamic input (capital). In general, the production function of firm \( i \) in period \( t \) is given by:

\[
Y_{it} = Y_{it} \left( X_{it}^1, ..., X_{it}^V, K_{it} \right)
\]

where \( X_{it}^v \) \( v = 1, ..., V \) are variable inputs with no adjustment cost, and \( K_{it} \) is the capital stock.

Assuming that producers in this economy are minimizing cost when producing, we have that the lagrangian in an economy with financial frictions is given by:

\[
L(X_{it}^1, ..., X_{it}^V, K_{it}) = \sum_{v=1}^{V} P_{it}^v X_{it}^v + r_{it} K_{it} + \Lambda_{it} \left[ Y_{it} - Y_{it} (\cdot) \right] + \Psi_{it} \left[ \lambda_{it} - K_{it} \right]
\]

where \( P_{it}^v \) is the price of input \( v \) and \( r_{it} \) is the rental rate of capital. In this case, \( \Lambda_{it} \) is the lagrange multiplier of the production function, and \( \Psi_{it} \) is the lagrange multiplier of the collateral constraint of the firm. The first order conditions using Kuhn Tucker for the factors are given by:

\[
[X_{it}] : \quad P_{it}^v = \Lambda_{it} \frac{\partial Y_{it} (\cdot)}{\partial X_{it}^v} \\
[K_{it}] : \quad r_{it} = \Lambda_{it} \frac{\partial Y_{it} (\cdot)}{\partial K_{it}} + \Psi_{it}
\]

Now, defining the markup as \( \mu_{it} = \frac{P_{it}}{mc_{it}} \) where \( mc_{it} \) is the marginal cost of the firm in the optimal solution, and using the fact that lagrange multiplier \( \Lambda_{it} \) is the marginal cost of the firm we have that

\[
\frac{\partial Y_{it} (\cdot)}{\partial X_{it}^v} \frac{X_{it}^v}{Y_{it} (\cdot)} = \mu_{it} \frac{p^m m}{p_{it} y_{it}}
\]

where \( \frac{\partial Y_{it} (\cdot)}{\partial X_{it}^v} \frac{X_{it}^v}{Y_{it} (\cdot)} \) is the elasticity of output with respect to materials, which in our model
is equal to \( (1 - v) \). That is, the derivation of markups from the ratio of the gross value of production to the value of materials presented in De Loecker and Warzynski (2012) holds in our framework.
CHAPTER 3

Stock Market Fluctuations, Unemployment and Consumption in the United Kingdom.

3.1 Introduction

The collapse of Lehman Brothers and subsequent crisis in 2008 have reignited interest in the link between stock market fluctuations and economic activity. In this article, I claim that in the UK there is evidence consistent with the hypothesis that stock market fluctuations drive business cycles. In particular, I show that stock market prices and the unemployment rate are cointegrated for UK data, and that the unemployment rate adjusts in response to movements in stock market prices. This result holds when I remove the same trend from each series, unlike usual practice, and is robust to different specifications and time periods.

This is consistent with a model where employment is demand determined and asset prices are driven by self-fulfilling beliefs. That is, pessimism about the economy drives asset prices down, which decreases aggregate demand, which in turn decreases employment and output, convalidating the initial pessimism. In such a model beliefs are a fundamental of the economy, just like technology or preferences, and high unemployment could in principle persist forever. This could have strong policy implications: if policy makers could credibly commit to stabilize asset prices, there would be no room for pessimism in the first place. Hence the self-fulfilling belief channel to recessions could be mitigated.

The relationship between stock market prices and unemployment is, however, weaker in the UK than in the US. Since in such a theoretical framework fluctuations in the stock market affect unemployment through changes in aggregate demand, I conjecture that less flexible labor markets in the UK relative to the US may prevent changes in aggregate demand to affect the unemployment rate immediately. If this story were true, one should
observe a close relationship between stock market and consumption fluctuations, and a weaker link between stock market and unemployment rate fluctuations. I argue that this holds in UK data. In particular, aggregate consumption and stock market prices are cointegrated, and the correlation between is strong. Furthermore, there is evidence that stock market prices Granger-cause consumption, and not the opposite.

The article is organized as follows. Section 2 describes briefly the detrending methodology I use following Farmer (2012). Section 3 describes the data and results, and performs robustness checks for different specifications. Section 4 investigates theoretically and empirically the relationship between stock market prices and aggregate consumption in the UK. Section 5 concludes.

3.2 Detrending Time-Series

Since the early 1980s, the standard approach in the empirical business cycle literature has been to remove the "cyclical component" of data using the Hodrick-Prescott filter (henceforth HP filter). Moments of the filtered series are then computed, and models to explain these facts are built. One major shortcoming of this approach is that the Hodrick-Prescott filter eliminates a different trend from each series. This could in principle hide important relationships among variables.

3.2.1 Variables in wage units

Farmer proposes an alternative method that extracts the same trend from any series. It consists on taking ratios of variables that are not stationary but cointegrated. If I use wages as the detrending variable, we say that we are expressing the series in "wage units". This idea can be traced back to the original work of Keynes (1936) and in practice extracts the nominal, productivity, and population trends. By detrending data series in this manner we can discover relationships that are not present when series are detrended using the HP filter.

Figures 3.1 and 3.2 illustrate this point for the UK. It shows the unemployment rate and the filtered series of GDP for the last decade using two different detrending techniques. Figure 3.1 uses an HP filter; Figure 3.2 expresses GDP in wage units. It’s hard to see
a relationship between unemployment and real GDP detrended with an HP filter, but
the relationship is clear if we express GDP in wage units. This is just an example of the
underlying principle that by automatically using the HP (or in fact, any) filter to detrend
series one could miss relationships of first order importance between series.

3.3 Stock Prices and Unemployment in the UK

3.3.1 Data

I restrict the time frame to the 1971:Q1-2012:Q1 period. I start our analysis in 1971 since
quarterly homogeneous coverage for some key variables start in that year. Specifically,
labour-related variables as reported by the Labour Force Statistics publication are not
available (or not comparable) before 1971. Compensation to Employees and Nominal and
Real GDP were taken out of the United Kingdom Economic Accounts. The source for
stock market data is Bloomberg. I use the so-called FTSE All Share Index (henceforth
FTSEAS), a market-capitalization weighted index of all companies in the London Stock
Exchange’s main market. It not only is more representative than the most widely used
FTSE 100, but it also has longer coverage, since the FTSE 100 was first computed in
1984. The two series move very closely together as Figure 3.3 shows. I also briefly assess
the relationship between unemployment and housing wealth. To construct an index of
housing wealth I use the Nationwide House Price Index, the longest unbroken series for
house prices in the UK.

I detrend variables using wages. I would ideally compute wages the following way:

\[ w = \frac{C^E}{L} \]

where \(C^E\) is Compensation to Employees and \(L\) is full-time equivalent employees.
However, full-time equivalent employees is not computed by the Office for National Statis-
tics. Moreover, quarterly data on full-time and part-time employees is available only since
1992. For these reasons I use jobs instead. To see whether jobs is a good approximation,
I construct a proxy for full-time equivalent employees starting in 1992 with data from
full-time and part-time employees, and compare its evolution with that of jobs. Figure 3.4 shows that jobs resembles quite closely the behavior of the full-time equivalent proxy, even better than persons in employment. To check for the robustness of our results, I will also use persons in employment and our proxy for full-time equivalent employees for the post 1992 period.

Every variable in "wage units" will be the ratio of that variable to wages, constructed as explained above. Further detail on the data used in this article can be found in the Data Appendix.

3.3.2 Results

Our benchmark exercise will study the relationship between FTSEAS (in wage units) and unemployment. It is important to note that the unemployment rate is bounded above and below, and that the FTSEAS is bounded below. Since standard cointegration analysis of bounded series is invalid (see for example Cavaliere (2005)), I need to transform the series so that they are unbounded. Following Farmer I take logarithm of FTSEAS and of a logistic transformation of the percent unemployment rate. The resulting series are presented in Figure 3.5.

Augmented Dickey-Fuller tests reveal it is not possible to reject the hypothesis that the transformed unemployment rate is non-stationary. The same is true for FTSEAS in wage units. Are they cointegrated? In our benchmark specification, where series have no discernible trends and in the long run they keep a fixed distance from each other there is evidence for cointegration. This is the so-called ”restricted constant” specification. There is no evidence of cointegration in the ”unrestricted constant” specification.

In our benchmark specification, the estimated Vector Error Correction Model (VECM) yields the following result (standard errors in parenthesis)

\footnote{The full-time equivalent proxy is constructed as full-time employees plus part-time employees divided by two. That is, we assume that two part-time employees are equivalent to one full-time employee.}
\[ \Delta u_t = 0.48 \Delta u_{t-1} + 0.02 \Delta p_{t-1} - 0.03 \left( u_{t-1} + 0.95 p_{t-1} - 7.67 \right) \]  \hspace{1cm} (3.1)

\[ \Delta p_t = 0.08 \Delta p_{t-1} + 0.27 \Delta u_{t-1} - 0.01 \left( u_{t-1} + 0.95 p_{t-1} - 7.67 \right) \]  \hspace{1cm} (3.2)

where \( u_t \) is the transformed unemployment rate and \( p_t \) is the logarithm of FTSEAS in wage units.

The first thing to notice is that there is a cointegrating relationship given by

\[ u_t = 7.67 - 0.95 p_t \]  \hspace{1cm} (3.3)

Note also that no coefficient is significantly different from zero in equation (3.2); the stock market in wage units is a random walk. The fact that the coefficient on the cointegrating vector is significant in equation (3.1) but not in equation (3.2) says that while the stock market price in wage units wanders randomly, the unemployment rate adjusts over time to the cointegrating relationship.

Is this relationship stable? Since there is no clear structural break candidate, I divide the sample in two sub-periods of equal length: from 1971Q1 to 1991Q3 and from 1991Q4 to 2012Q1. I use the first sub-period to estimate a VECM, and the second sub-period to evaluate the out-of-sample performance of the model. The estimated model for the first sub-period is

\[ \Delta u_t = 0.52 \Delta u_{t-1} + 0.01 \Delta p_{t-1} - 0.03 \left( u_{t-1} + 0.61 p_{t-1} - 5.67 \right) \]  \hspace{1cm} (3.4)

\[ \Delta p_t = 0.07 \Delta p_{t-1} + 0.39 \Delta u_{t-1} + 0.03 \left( u_{t-1} + 0.61 p_{t-1} - 5.67 \right) \]  \hspace{1cm} (3.5)

The estimates are not significantly different from those for the full-sample (equations (3.1) and (3.2)). Moreover, the one-step-ahead forecast of the unemployment rate matches very closely the actual series, as depicted in Figure 3.6. This highlights both the stability of the relationship between unemployment and stock prices, and the good fit of the model.
3.3.3 Robustness checks

Persons in employment

Given that I do not have a measure of full-time equivalent employees I used jobs. If I use persons in employment instead, results do not significantly change. To be specific, the estimated VECM for the entire sample is

\[
\Delta u_t = 0.59 \Delta u_{t-1} + 0.01 \Delta p_{t-1} - 0.03 \left( u_{t-1} + 2.06 p_{t-1} - 14.26 \right) \\
\Delta p_t = 0.11 \Delta p_{t-1} + 0.12 \Delta u_{t-1} - 0.01 \left( u_{t-1} + 2.06 p_{t-1} - 14.26 \right)
\]

(3.6) (3.7)

The estimates are not significantly different than when I use jobs. The estimation is robust to this alternative definition of wage rate.

Full-time equivalent employees (Post 1992)

As explained in the Data subsection, I constructed a proxy for full-time equivalent employees as the sum of full-time employees plus part-time employees divided by two. This would in principle be closer to our ideal measure, but coverage starts in 1992. Using this estimation to construct the wage rate, Johansen’s cointegration test shows cointegration only when I specify the model with one lag. For two or higher lags, it is not possible to reject the null of no cointegration. The estimated VECM for the 1992Q2-2012Q1 period is

\[
\Delta u_t = 0.37 \Delta u_{t-1} + 0.01 \Delta p_{t-1} - 0.009 \left( u_{t-1} + 5.01 p_{t-1} - 31.18 \right) \\
\Delta p_t = 0.07 \Delta p_{t-1} - 0.09 \Delta u_{t-1} - 0.017 \left( u_{t-1} + 5.01 p_{t-1} - 31.18 \right)
\]

(3.8) (3.9)

Although some estimates change, the main results hold: the transformed unemployment rate is autoregressive in first differences, and is cointegrated with the stock market in wage units. The latter is a random walk.
Detrending by GDP

Another way of detrending variables is to divide by nominal GDP, to obtain variables expressed in "GDP units". If I detrend the stock market index in this way, the main results do not change. Figure 3.7 compares the FTSEAS index in wage units and in GDP units.

Series move very closely together. Although with mixed results, cointegrations test still show signs of cointegration between the unemployment rate and FTSEAS in GDP units.\footnote{Specifically, the test with 1 and 2 lags rejects cointegration by a thin margin; the test with 3 and 4 lags shows cointegration.} The main results of the VECM still hold: unemployment reverting to the stock market index and not the other way around, autoregressive process for unemployment in first differences, and stock market in GDP units is a random-walk.

\[
\Delta u_t = 0.47\Delta u_{t-1} + 0.02\Delta p_{t-1} - 0.03 (u_{t-1} + 1.22p_{t-1} - 4.60) \tag{3.10}
\]
\[
\Delta p_t = 0.06\Delta p_{t-1} + 0.26\Delta u_{t-1} + 0.005 (u_{t-1} + 1.22p_{t-1} - 4.60) \tag{3.11}
\]

Wealth Index

I construct a wealth index that combines the stock market index with a housing wealth index. I build a housing wealth index multiplying the housing index by the population. Population estimates for the UK are available yearly. I interpolate using a geometric average to estimate quarterly series. Figure 3.8 shows the evolution of FTSEAS and Housing wealth since 1971, both series detrended by wage. In the last decade both series move similarly, but that has not always been the case.

In the UK, roughly 60% of wealth is held in the form of housing, while the remaining 40% is held in the form of industrial buildings, plants and machinery. The wealth index is a weighted average of the housing index and the FTSEAS.
The wealth index thus constructed and detrended by wages is not cointegrated with the transformed unemployment rate, except when I specify the test with one lag. This result holds whether I construct wages with jobs or with persons in employment. The same is true if I restrict the sample to the post-1983 period, where the housing price index is strictly homogeneous and comparable.

3.3.4 A stylized model

We can rationalize the relationship between stock market prices and unemployment with the following model.\(^3\) Assume that aggregate output is produced by many firms with access to the following technology

\[ y_t = A_t L_t^{1-\alpha} K_t^\alpha \]

(3.12)

where, as usual, \(y_t\) is output, \(L_t\) is labor and \(K_t\) is capital. Parameter \(A\) differs from the usual neoclassical productivity parameter in that it reflects externalities originating in the search market for labor. Those externalities can be expressed in reduced form as

\[ A_t = (1 - \overline{L}_t)^{1-\alpha} \]

(3.13)

where \((1 - \overline{L})\) is the aggregate unemployment rate. Firms take \(A_t\) as given, even though their decisions influence \(\overline{L}_t\), hence the externality.

Farmer drops the usual Nash-bargaining equation present in search models in favor of the assumption that employment is demand determined. In such a model the usual asset pricing equation of neoclassical models hold, i.e.:

\[ y_t = \theta \frac{p_{k,t}}{w_t} \]

(3.14)

where \(p_{k,t}\) is the price of asset \(k\) at time \(t\), and \(\theta\) is a parameter that depends on

\(^3\)See Farmer (2012)
technology and the household’s discount rate.

The model is closed with a belief function that can take the form

$$E_t[p_{k,t+1}^{w_{t+1}}] = X_t$$

where $$X_t$$ are beliefs about the future real value of the stock market. $$X_t$$ could evolve, for example, according to the following law of motion:

$$\log X_t = \log X_{t-1} + f(\Delta U_{t-1}) + \varepsilon_t$$

In a standard neoclassical model the causality in asset pricing function (3.14) goes from output to stock market prices. That is, the fundamentals of the economy determine output which determine asset prices. In Farmer’s stylized model, causation is reversed. Beliefs about the future level of the stock market influence the level of the stock market today (equation (3.15)). Further, those beliefs are influenced by current unemployment rate (equation (3.16)). The real level of the stock market today determines output (equation (3.14)), which in turn determines employment. This model, which is consistent with rational expectations, can explain the observed causal relation between unemployment rate and stock market prices in the UK.

### 3.4 Stock Prices and Consumption in the United Kingdom

The analysis above reveals a robust relationship between stock prices in wage units and the unemployment rate. This relationship, however, is weaker in the UK than in the US. The correlation between both series in the UK is -0.17, while in the US is -0.62 for the same period. Moreover, the point estimate of the error correction coefficient for the unemployment equation in the UK is 0.03, three times smaller than for the US.⁴

One explanation of the weaker relationship in the UK could relate to the relative shares of plants and machinery over total wealth in both countries. As mentioned above, they represent roughly 60% of wealth in the US and 40% of wealth in the UK. Since

---

⁴See Farmer (2012)
the value of plants and machines are ultimately reflected in the stock market, one would expect stock market fluctuations to have more wealth effects in the US than in the UK.

Another explanation relies on the relative inflexibility of UK’s labor markets with respect to US’ labor market, documented in several studies (see for example Nickell (1997), and Huges and McCormick (1991). To understand the role that inflexible labor markets could play, it’s important to remember the mechanisms at work in the model above. In such a model, beliefs about future asset prices determine present wealth, which affects consumption. Since output is demand determined, this ultimately influences employment. That is, the relationship between wealth and employment has two links: the effect of wealth on consumption, and the subsequent effect of consumption on unemployment.

Relative inflexible labor markets, should weaken the effect of consumption on unemployment, hence the relationship between stock market and unemployment. But they shouldn’t get in the way of a strong link between fluctuations in stock market prices and consumption. I investigate this relationship below.

3.4.1 A Framework for the Cointegrating Relationship Between Stock Prices and Consumption

This subsection provides a framework to think about the linkages between wealth and consumption in wage units. I draw from previous work from Campbell and Mankiw (1989) and Lettau and Ludvingson (2001).

We can define the law of motion for aggregate wealth as

$$W_{t+1} = R_t (W_t - C_t)$$  \hspace{1cm} (3.17)

where $W_{t+1}$ is wealth in period $t+1$, $R_t$ is the gross rate of return on wealth, and $C_t$ is consumption. Dividing the above by $W_t$ and taking logs we obtain

$$w_{t+1} - w_t = r_t + \log \left(1 - \frac{C_t}{W_t}\right)$$  \hspace{1cm} (3.18)

$$w_{t+1} - w_t = r_t + \log \left(1 - \exp(\gamma_t - w_t)\right)$$  \hspace{1cm} (3.19)
where lowercase variables are expressed in logarithm. Campbell and Mankiw (1989) show that a linear approximation of the second term in the right hand side of (3.19) can be expressed as

\[
\log(1 - \exp(c_t - w_t)) \approx k + \left(1 - \frac{1}{\rho}\right)(c_t - w_t)
\]  
\(3.20\)

where \(\rho = 1 - \exp(c - w)\), and \(k = \log(\rho) - \left(1 - \frac{1}{\rho}\right)\log(1 - \rho)\). Parameter \(\rho\) can be interpreted as the average ratio of invested wealth \((W - C)\) to total wealth \(W\). Then (3.19) becomes

\[
w_{t+1} - w_t = k + r_t + \left(1 - \frac{1}{\rho}\right)(c_t - w_t)
\]  
\(3.21\)

solving this equation forward, getting rid of uninteresting approximation constants, and imposing a transversality condition we obtain

\[
c_t - w_t = \sum_{i=1}^{\infty} \rho^i (r_{t+i} - \Delta c_{t+i})
\]  
\(3.22\)

Wealth, in turn, can be divided into human wealth \((H)\) and asset wealth \((A)\)

\[
W_t = H_t + A_t
\]  
\(3.23\)

Since human wealth is not observable, the usual approach in the literature (see for example Lettau and Ludvingson (2001)) is to assume that it is proportional to current labor income, so that \(h_t = \kappa + y_t + z_t\) where \(y_t\) is wage, \(\kappa\) is a constant, and \(z_t\) is a stationary shock. Then (3.23) can be expressed in logs as

\[
w_t \approx \omega h_t + (1 - \omega)a_t
\]  
\(3.24\)

where \(\omega\) is the share of assets in total wealth. Plugging this in (3.22) (and again getting rid of linearization constants)

\[
c_t - \omega y_t - (1 - \omega)a_t = \sum_{i=1}^{\infty} \rho^i (r_{t+i} - \Delta c_{t+i}) + \omega z_t
\]  
\(3.25\)
Now adding and subtracting $y_t$ from the left hand side we obtain

$$\tilde{c}_t - (1 - \omega)\tilde{a}_t = \sum_{i=1}^{\infty} \rho^i (r_{t+i} - \Delta c_{t+i}) + \omega z_t \quad (3.26)$$

where tilda variables are expressed in wage units (i.e. $\tilde{x}_t = x_t - y_t$)

Since the right hand side of equation (3.26) is presumed stationary, there exists a cointegrating relationship between consumption in wage units and asset wealth in wage units. Deviations from that cointegrating relationship (the cointegrating error) will be eliminated by “error corrections” from either consumption in wage units, assets in wage units, or both. I analyze this in the next subsection.

### 3.4.2 The empirics

Following the usual approach in the literature (see for example Blinder et al. (1985)), my measure of consumption is expenditure on nondurables and services. This is a better variable than total consumption since $C_t$ in equation (3.17) refers to period’s $t$ consumption flow. One would ideally like to include estimates of the flow of services delivered by durable goods in my definition of consumption. Due to the difficulties of such an estimation, I do not pursue that path. Nonetheless, the main results in the following analysis are robust to the use of total consumption instead.

Following Section 3, my benchmark specification will focus on the stock market portion of wealth. Figure 3.9 plots the evolution of the logarithm of consumption (of nondurables and services) and the logarithm of stock market prices for the UK (both in wage units). The co-movement is evident, with a correlation coefficient of 0.71. This suggests that there is a stronger relationship between stock market prices and consumption than between stock market prices and unemployment. The null hypothesis of unit root in the consumption series cannot be rejected by either augmented Dickey-Fuller tests or Phillips-Perron unit-root tests.

The Johansen test shows evidence of cointegration between both series. Based on the Akaike and Hannan-Quinn information criteria, my benchmark VECM specification includes two first-difference lags. The results of my estimation are $^5$

\footnote{Note that we have dropped the tildas above the variables. Just to be clear, in this section $p_t$ is the}
The long term cointegrating relationship between consumption and stock prices is given by $c = 0.22p + 0.69$. Since the cointegrating coefficient is significantly different from zero in equations (3.27) and (3.28), both variables adjust to deviations in their common trend. Note that in both equations, the error correction coefficients are higher than what I found in my estimates for stock prices and unemployment. This is another sign that the connection between stock prices and consumption is stronger than that between stock prices and unemployment, in the sense that any deviation from the long-term trend is corrected faster in the former.

The fact that the stock market reverts to a cointegrating equation for consumption is in line with previous findings for US data. Lettau and Ludvigson (2004) for example, find that asset wealth reverts to its cointegrating relationship with consumption and income wealth, although they don’t find evidence of consumption reverting to the cointegrating relationship.

The results of limiting the estimation window to the first half of the sample are shown in equations (3.29) and (3.30). It can be seen that the existence of a cointegrating equation is robust, with larger coefficients. So is the error correction term being significantly different from zero in equation (3.27), although the point estimate and its standard error increase significantly. The estimates on the stock market prices equations are much more noisy, hence nothing is significantly different from zero.

$\Delta c_t = -0.06 \Delta c_{t-1} + 0.27 \Delta c_{t-2} - 0.01 \Delta p_{t-1} + 0.02 \Delta p_{t-2}$ \hspace{1cm} (3.27)

$\Delta p_t = 0.15 \Delta p_{t-1} + 0.04 \Delta p_{t-2} - 1.15 \Delta c_{t-1} + 0.28 \Delta c_{t-2}$ \hspace{1cm} (3.28)

$\Delta p_t = 0.15 \Delta p_{t-1} + 0.04 \Delta p_{t-2} - 1.15 \Delta c_{t-1} + 0.28 \Delta c_{t-2}$ \hspace{1cm} (3.28)

$\Delta c_t = -0.06 \Delta c_{t-1} + 0.27 \Delta c_{t-2} - 0.01 \Delta p_{t-1} + 0.02 \Delta p_{t-2}$ \hspace{1cm} (3.27)

$\Delta p_t = 0.15 \Delta p_{t-1} + 0.04 \Delta p_{t-2} - 1.15 \Delta c_{t-1} + 0.28 \Delta c_{t-2}$ \hspace{1cm} (3.28)

$\Delta p_t = 0.15 \Delta p_{t-1} + 0.04 \Delta p_{t-2} - 1.15 \Delta c_{t-1} + 0.28 \Delta c_{t-2}$ \hspace{1cm} (3.28)
Causality

What can we say about causality? In my story, beliefs about the level of the stock market are self-fulfilling, and they affect consumption. Hence we expect causation to go from stock market prices to consumption, but not the other way around. To test this hypothesis, I perform a Granger noncausality test.

Sims et al. (1990), show that the usual Wald statistic for Granger noncausality has a nonstandard asymptotic distribution. So the standard test in statistical packages is inaccurate for my purposes. Further, it is incorrect to perform a standard Granger Noncausality Test following a VECM estimation. This is true since a Granger Noncausality test performed after a VECM estimation is a test conditional on a cointegration test being positive (see Toda and Yamamoto (1995)) and the statistics’ conditional distributions are non-trivial.

I follow the procedure suggested by Toda and Yamamoto (1995). Their approach is based on the estimation of a VAR in levels with $p$ lags, where $p$ equals the optimal number of lags according to the usual information criteria, plus the highest integrating order of the series. Once the model is estimated, I perform a Wald test on the joint significance of the first $(p-1)$ coefficients of the exogenous variables for every equation.\footnote{See Appendix 2 on Granger noncausality test for more details on this procedure.}

In our case $p = 4$. This follows from the fact that according to the Akaike and Hannan-Quinn information criteria the optimal VAR lag is 3; and both variables are integrated of order one, hence the higher order of integration is one.

When I test the null hypothesis that the first 3 (that is, $p-1$) lags of (first-differenced) stock market prices are jointly equal to zero in the equation for consumption, the p-
value is 0.000. When I test the null hypothesis that the first 3 lags of (first differenced) consumption are jointly equal to zero in the equation for stock prices, the p-value is 0.20. These results indicate that stock market prices Granger-cause consumption at a 1% confidence level, but consumption does not Granger-cause stock prices. Thus, in line with our causality story, there is evidence in UK data that stock market prices Granger-cause consumption, but not the opposite.

Since there is evidence that stock market fluctuations drive shifts in consumption, how much of an impact does it have? To answer this question, Figures 3.10 and 3.11 show the impulse-response functions to a 10% increase in stock prices (in wage units). Figure 3.10 shows the response of consumption, Figure 3.11 the response of Stock Prices. The vertical axis show the percent variation with respect to time zero, the time of the shock. In Figure 3.11 I see that the initial increase in stock prices gradually declines, with prices stabilizing at a level around 4% higher than in period 0 after 20 quarters. The response of consumption is gradual, but monotonically increasing. In the end, it stabilizes at a level about 1% higher than in period zero, making half of that adjustment in the first 4 quarters.

**Robustness Check: Consumption and Wealth**

Do the results above hold when I use my Wealth Index (see Section 3.3.4) Figure 3.12 shows both series. I am using again consumption of nondurables and services. The correlation is 0.81, even higher than the correlation between stock prices and consumption.

Evidence for cointegration is harder to find, unless we are willing to accept a 15% significance level. If we are, the VECM estimates above hold with striking similarity, as the following equations show (where $w_i$ is the Wealth Index)
\[
\Delta c_t = -0.04 \Delta c_{t-1} + 0.30 \Delta c_{t-2} - 0.01 \Delta w_{t-1} + 0.04 \Delta w_{t-2}
\]
\[
\Delta w_t = -0.05 \left( c_{t-1} - 0.30 w_{t-1} - 0.53 \right)
\]

Granger Noncausality tests as explained in the previous section reach the same conclusion: the Wealth Index Granger-causes consumption, but not the opposite.

### 3.5 Conclusion

There is evidence of a cointegrating relationship between the stock market index in wage units and the unemployment rate in the US and UK. In the UK, this is explained by the causal effect of stock market fluctuations on consumption. This is evidence against the statement that the unemployment rate is mean-reverting. These are not standard results in the literature, and can be rationalized in a model where output is demand-determined, and beliefs about the future evolution of asset prices are a fundamental of the economy. This approach seems more capable of explaining events like the Great Recession than both neoclassical and new-keynesian models. This is a small contribution to this agenda, but much more research is needed in order to uncover the full implications of this approach.
3.6 Tables and Figures

Figure 3.1: Real GDP (deviations from HP trend)

![Graph showing Real GDP deviations from HP trend and Unemployment Rate over time.](image)

Source: Office for National Statistics.

Figure 3.2: Real GDP in Wage Units

![Graph showing Real GDP in Wage Units and Unemployment Rate over time.](image)

Source: Office for National Statistics.
Figure 3.3: FTSE100 vs. FTSEAS

Figure 3.4: Different Measures of Employment
Figure 3.5: Unemployment Rate and the Stock Market

Source: Office for National Statistics and Bloomberg

Figure 3.6: Actual and Forecasted Unemployment Rate

Source: Office for National Statistics and own computations.
Figure 3.7: FTSEAS in Wage and GDP Units

Source: Office for National Statistics and Bloomberg.

Figure 3.8: FTSEAS and Housing Wealth (Wage Units)

Source: Bloomberg and Nationwide.
Figure 3.9: FTSEAS and Consumption

Figure 3.10: Response of Consumption to a 10% Shock in Stock Prices

Figure 3.11: Response of Stock Prices to a 10% Shock in Stock Prices
Figure 3.12: Wealth Index and Consumption

Source: Office for National Statistics, Bloomberg and Nationwide.
3.7 Appendices

3.7.1 Appendix 1: Data Sources

Table A1 describes the data and sources.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Source</th>
<th>ONS Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Economically Active: UK: All: Aged 16-64</td>
<td>ONS</td>
<td>LF2K</td>
</tr>
<tr>
<td>Total in Employment: UK: All: Aged 16-64</td>
<td>ONS</td>
<td>MGRZ</td>
</tr>
<tr>
<td>Employee Jobs: UK: All: Aged 16-64</td>
<td>ONS</td>
<td>BCAJ</td>
</tr>
<tr>
<td>Employees: Full-time: UK: All: Thousands: SA</td>
<td>ONS</td>
<td>YCBK</td>
</tr>
<tr>
<td>Employees: Part-time: UK: All: Thousands: SA</td>
<td>ONS</td>
<td>YCBN</td>
</tr>
<tr>
<td>Total compensation of employees</td>
<td>ONS</td>
<td>DTWM</td>
</tr>
<tr>
<td>Unemployment rate: UK: All: Aged 16 and over: %: SA</td>
<td>ONS</td>
<td>MGSX</td>
</tr>
<tr>
<td>Gross Domestic Product at market prices: Current price: SA</td>
<td>ONS</td>
<td>YBHA</td>
</tr>
<tr>
<td>Gross Domestic Product: Chained volume measures: Current price: SA</td>
<td>ONS</td>
<td>ABMI</td>
</tr>
<tr>
<td>Population. Mid-year Estimates</td>
<td>ONS</td>
<td>DYAY</td>
</tr>
<tr>
<td>Expenditure on Non-Durable Goods</td>
<td>ONS</td>
<td>UTHI</td>
</tr>
<tr>
<td>Expenditure on Services</td>
<td>ONS</td>
<td>UTIN</td>
</tr>
<tr>
<td>FTSE All Share Index</td>
<td>Bloomberg</td>
<td></td>
</tr>
<tr>
<td>House Price Index</td>
<td>Nationwide</td>
<td></td>
</tr>
</tbody>
</table>

To construct the benchmark wage variable I divided ”total compensation of employees” by ”employee jobs”. As a robustness check I defined two alternative measures. One that divides ”total compensation of employees” by ”Total in Employment”, and another that divides ”total compensation of employees by a proxy of full-time equivalent employees. The proxy was built as ”full time employees’ plus ”part-time employees” divided by two.

The transformed unemployment rate is defined as

\[ ur_{\text{trans}} = \ln\left(\frac{100 * ur_t}{100 - ur_t}\right) \]

where \( ur \) is the percentage unemployment rate.

The housing index is constructed as

\[ H_{wealth_t} = \frac{HPI_t * Pop_t}{w_t} \]
where $HPI$ is the house price index, $Pop_t$ is the population, and $w_t$ is the wage.

The wealth index is constructed as

$$WI_t = 0.6 \ln(H\text{wealth}_t) + 0.4 \ln(FTSEASw_t)$$

where $FTSEASw_t$ is the FTSE All Share index in wage units. The index is set equal to 100 for the first quarter of 1983.
3.7.2 Appendix 2: Granger Non-causality Test

To perform the Granger noncausality test when series potentially have unit roots, I follow the procedure suggested by Toda and Yamamoto (1996), that can be summarized in a series of steps.

1) Test each time series to determine their order of integration.

2) Fit a VAR model in levels. Set the number of lags to be equal to \( p \), where \( p \) is equal to the optimal lag determined by the usual information criteria, plus the maximum order of integration found in 1)

3) Test for serial autocorrelation of the residuals. If present, increase the lags.

4) Apply the Granger Non-causality Test only to the coefficients of the first \( p - m \) lags, where \( m \) is the maximum order of integration found in 1).

5) Test the null hypothesis of Granger Non-causality for each equation.
References


