

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

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Economics

2017

The Real Effects of Financial Frictions: Amplification, Misallocation, and Instability

Huang, Zongbo

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp01cv43p042d>

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

THE REAL EFFECTS OF FINANCIAL FRICTIONS:
AMPLIFICATION, MISALLOCATION, AND INSTABILITY

Zongbo Huang

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
ECONOMICS
ADVISER: PROFESSOR MARKUS BRUNNERMEIER

JUNE 2017

© Copyright by Zongbo Huang, 2017.

All rights reserved.

Abstract

This dissertation consists of three essays at the intersection of finance and macroeconomics. A common thread is to study the amplification of financial shocks and misallocation due to financial frictions. The first essay studies the role of banks' discretion in managing panics in a dynamic model of credit line run. In downturns, banks tighten liquidity by revoking credit lines. Anticipating this, borrowers run to draw down credit lines in the first place, which imposes further pressure on banks. Thus liquidity rationing and credit line runs form a feedback loop that amplifies bank distress. I fit the model to the U.S. commercial bank data and find that the feedback effects contribute to more than a half of the liquidity contraction during the Great Recession. The second essay is a joint work with Sylvain Catherine, Thomas Chaney, David Sraer, and David Thesmar. We structurally estimate a dynamic model with heterogeneous firms and collateral constraints, based on the causal effect of collateral shocks on firm investment. We then quantify the aggregate impact of financial friction by embedding the model in a general equilibrium framework. The estimates imply that lifting financial frictions would increase welfare by 9.4% and aggregate output by 11%. Half of the output gain is due to an increase in the aggregate stock of capital, one-quarter is due to a larger aggregate labor supply, while the remaining quarter is due to a higher aggregate productivity from a better allocation of inputs across heterogeneous firms. The final essay develops a dynamic general equilibrium model with heterogeneous beliefs and collateral constraints and investigates the cyclicalities of haircuts and default risks jointly. The endogenously determined haircuts are countercyclical and thus lead to a downward margin spiral that exacerbates financial instability. Meanwhile, default risks accumulate in the background, until they materialize during crises.

Acknowledgements

This dissertation would not have been possible without the help of many people. I am grateful to my professors at Fudan University and Toulouse School of Economics. In particular, I thank Yongqin Wang, Zonglai Kou, Emmanuelle Auriol, Jacque Crémer, and Patrick Ray for their advice and help.

At Princeton, I am indebted to my advisor, Markus Brunnermeier, for his encouragement and guidance. I thank him for his guidance, encouragement, and support during the development of this work. His comments and suggestions were always helpful, and he is a real role model. My sincere thanks also go to my other committee members: Wei Xiong and Stephen Morris, for their encouragement, help, and insightful comments, which are precious to me.

I would also like to thank Maryam Farboodi, Valentine Haddad, Nobuhiro Kiyotaki, Atif Mian, Yuliy Sannikov, Hyun Song Shin, and Motohiro Yogo for numerous helpful discussions. Special thanks to my coauthors, Sylvain Catherine, Thomas Chaney, Jason Ravit, Michael Sockin, David Sraer, and David Thesmar, who were sources of great ideas and great conversations. I have learned a lot working with them.

My fellow graduate students and friends at Princeton were constant sources of laughter, joy, and support. Special thanks go to Cristian Alonso, Ioannis Branikas, Qingqing Cao, Tom Carter, Kyle Chauvin, Cheng Chen, Mingyu Chen, Wei Cui, Liang Dai, Xiaoyang Dong, Liyu Dou, Xiaochen Feng, Zhenyu Gao, Alejandro Van der Gote, Daniel Goetz, Mattieu Gomez, Sungki Hong, Ji Huang, Alexander Jakobsen, John Klopfer, Shaowei Ke, So Kubota, Chang Sun, Yulong Wang, Justin Weidner, Jung Eun Yoon, Yu Zhang, and Chen Zhao.

I would like to thank Laura Hedden and Karen Neukirchen for their administrative support and help, and the Graduate School and Department of Economics for generous financial support.

Finally, I want to thank my wife, my parents, and grandparents. This dissertation would have never been written without their everlasting support. I will be forever grateful for their support and love.

To my parents and grandparents, who have always supported me.

Contents

Abstract	iii
Acknowledgements	iv
List of Tables	x
List of Figures	xi
1 Managing Bank Run Risk: The Perils of Discretion	1
1.1 Introduction	1
1.2 Empirical Evidence	7
1.3 The Model	10
1.3.1 Setup	10
1.3.2 Discussions	15
1.4 Equilibrium Analysis	17
1.4.1 Bank Decision Making	17
1.4.2 Borrower Runs	19
1.4.3 Definition of Equilibrium	22
1.4.4 A Positive Feedback Loop	23
1.4.5 A Model with Commitment	26
1.5 Numerical Results	28
1.5.1 Calibration	28
1.5.2 Equilibrium Decision Rules	32
1.6 Counterfactuals	33

1.6.1	Commitment	33
1.6.2	Leverage Ratio Requirements	34
1.6.3	A Commitment Tax on Revocation	35
1.7	Concluding Remarks	36
Appendix 1.A	Proofs	52
Appendix 1.B	Algorithm	55
Appendix 1.C	Additional Figures	57
2	Aggregate Effects of Collateral Constraints	60
2.1	Reduced-form evidence	67
2.2	The model	68
2.2.1	Production technology and demand	68
2.2.2	Input dynamics	69
2.2.3	Financing frictions and capital structure	70
2.2.4	The optimization problem	72
2.3	Structural Estimation	73
2.3.1	Estimation procedure	73
2.3.2	Predefined and Estimated Parameters	74
2.3.3	Data Moments	75
2.3.4	Parameter Identification	77
2.3.5	Estimation results	79
2.3.6	Determinants of financing constraints	81
2.4	General Equilibrium Analysis	83
2.4.1	General equilibrium model	83
2.4.2	The aggregate effect of financing constraints	85
2.4.3	Productivity persistence and misallocation	88
2.5	Discussion	89
2.5.1	Model Identification	89

2.5.2	Robustness: Residual leverage and costly equity issuance	91
2.5.3	Policy Experiments	93
2.6	Conclusion	94
Appendix 2.A	Solving the model and Estimation	112
2.A.1	Solving the model numerically	112
2.A.2	Estimation	114
Appendix 2.B	General Equilibrium Computation	118
Appendix 2.C	Additional figures	120
3	Haircuts and Credit Risk Over the Cycle	124
3.1	Introduction	124
3.2	The Model	129
3.2.1	Set-up	129
3.2.2	Agents' Problem	132
3.3	Equilibrium Analysis	135
3.4	Numerical Examples	140
3.5	Discussions	145
3.6	Conclusion	147
Appendix 3.A	A Discrete-time Setup	155
Appendix 3.B	Proofs	157
3.B.1	Proof of Lemma 5	157
3.B.2	Proof of Proposition 5	157
3.B.3	Proof of Proposition 6	158
3.B.4	Proof of Proposition 7	161

List of Tables

1.1	Borrower's flow payoffs and continuation values	49
1.2	Model calibration	50
1.3	Calibration targets	51
1.4	Counterfactuals	51
2.1	Summary statistics: COMPUSTAT Extract	104
2.2	Elasticity of Moments with respect to Parameters	105
2.3	Parameter estimates (SMM)	106
2.4	Aggregate effects of collateral constraints	107
2.5	Estimating a constrained model on unconstrained data	108
2.6	Robustness: unsecured debt and costly equity issuance	109
2.7	Macro effect of an investment subsidy	110

List of Figures

1.1	Reasons to draw cash from credit lines (%)	41
1.2	Amendments from Dealscan dataset	42
1.3	Total credit growth and loan to total credit ratio	43
1.4	Growth of loans and total credit during the Great Recession	44
1.5	Timeline	44
1.6	Growth of loans and total credit	45
1.7	Law of motion of equity ratio	46
1.8	Policy functions	46
1.9	Policy functions with and without full commitment	47
1.10	Leverage ratio requirements	47
1.11	Tax on cutting credit lines	48
1.12	Sensitivity of moments to the drawdown cost	57
1.13	Sensitivity of moments to the adjustment cost	58
1.14	Sensitivity of moments to the non-interest expenses	59
2.1	Financing constraints as a function of firm characteristics	100
2.2	General equilibrium effect of pledgeability s	101
2.3	Productivity persistence ρ and the dispersion of capital productivity	102
2.4	Non targeted versus targeted investment subsidy	103
2.5	Sensitivity of moments to pledgeability s	120
2.6	Sensitivity of moments to adjustment costs c	121

2.7	Sensitivity of moments to productivity volatility σ	122
2.8	Sensitivity of moments to productivity persistence ρ	123
3.1	Haircuts and Probability of default	150
3.2	Law of Motion of the State Variable	151
3.3	Asset Returns	152
3.4	Risk premium	152
3.5	Loss sprial	153
3.6	Underlying risk	153
3.7	Financial development	154

Chapter 1

Managing Bank Run Risk: The Perils of Discretion¹

1.1 Introduction

Runs on financial institutions played a central role in the 2007-2009 financial crisis. Most existing theories of bank runs focus on the strategic complementarities among depositors but overlook the important role of banks as strategic players, that banks not only respond to bank runs actively but also take run risk control as a key part of liquidity management. This article studies banks' risk management and uncovers a novel type of strategic complementarities between banks and their clients which adds significant fragility to the banking sector. I focus on bank lending through credit lines to address the following questions: How do banks optimally control liquidity risk given the possibility of bank runs on credit lines? How do banks' strategies in turn affect the incentives

¹Princeton University, Email: zongboh@princeton.edu. I am especially indebted to my advisor Markus Brunnermeier for his help and support. I am also grateful to Mark Aguiar, Olivier Darmouni, Maryam Farboodi, Ji Huang, Kai Li, Yunan Li, Xuewen Liu, Adrien Matray, Stephen Morris, Jason Ravit, Michael Sockin, Chang Sun, David Thesmar, Hao Wang, Wei Xiong, Motohiro Yogo, and Yao Zeng for many helpful comments and discussions. All remaining errors are my sole responsibility.

to run on the banks by drawing down the credit lines? Can properly designed policies mitigate and eliminate the run risk?

A credit line is a flexible loan from a bank to a borrower that permits the borrower to borrow up to a certain limit and repay on an unscheduled basis until the contract ends. Credit lines are pervasively used in practice by all kinds of borrowers, corporates, households, and other financial institutions, to obtain immediate liquidity and fund daily operations. Lending through credit lines has been growing fast during the last decades and now becomes a major form of lending by the commercial banks in the U.S. According to the Federal Reserve Survey of Terms of Business Lending, about 80% of all commercial and industrial loans (C&I loans) in the United States were made under credit lines. The increase use of credit lines provides better liquidity insurance to borrowers, however, it also exacerbate the financial instability when banks themselves experience liquidity shortages.

The key mechanism that leads to financial instability is the amplifying feedback effects between liquidity rationing and credit line runs. In this article, I develop a dynamic banking model of credit lines to analyze the mechanism. The model is based on two key features of credit lines. First, credit lines are not fully committed. Banks have legal rights to limit borrowers' access to credit lines, especially when the associated covenants are violated (see, e.g., [Sufi \(2009\)](#) and [Roberts and Sufi \(2009\)](#)); moreover, banks also use this discretion to withhold funds when they themselves experience liquidity shortages (see, e.g., [Acharya, Almeida, Ippolito, and Perez \(2014b\)](#) and [Chaderina and Tengulov \(2015\)](#)). Second, the use of credit lines is flexible. Thus, firms are able to draw down credit lines preemptively, which is effectively a run on the asset side of banks (see, e.g., [Ivashina and Scharfstein \(2010\)](#), [Campello, Graham, and Harvey \(2010\)](#), and [Ippolito, Peydro, Polo, and Sette \(2015\)](#)).

In the model, borrowers have a demand for liquidity. Each borrower invests in a long-term project and suffers a liquidity shock, which would force the borrower to terminate

her project unless she can obtain additional liquidity. For example, firms receive trade receivables instead of cash, but they need cash to cover immediate operating expenses. As a result, borrowers arrange credit lines from a bank to protect themselves from such risks. The bank raises deposits to cover the drawdowns of credit lines and earns a premium by pooling the borrowers' liquidity risk and intermediating funds.

The bank is also exposed to its own liquidity shock, which imposes an additional cost of raising deposits and thus makes lending more costly. Therefore, once hit by the shock the bank has an incentive to ration liquidity by cutting credit lines. This liquidity rationing induces borrowers to draw down preemptively in the first place in case their access to credit lines is limited. I refer the preemptive drawdowns as credit line runs, which are essentially runs of bank liquidity on the asset side of bank balance sheet. The preemptive drawdowns, in turn, impose pressure on banks and lead banks to tighten liquidity further. The process repeats and becomes a downward spiral that amplifies the impact of the initial shock. Furthermore, a dynamic inconsistency problem emerges from this environment. The bank may want to commit to a rule-based liquidity policy. If this commitment were credible, it would dissuade borrowers from running. However, the bank may not be able to commit credibly. If the borrowers do in fact run, the bank's future self would respond to it actively by rationing liquidity, instead of acting according to the plan.

I then fit the model to the U.S. commercial bank data and explore the feedback effects and the dangers of bank discretion quantitatively. I find that having a full commitment by the bank mitigates runs significantly, as the bank internalizes the impact of liquidity rationing on the borrowers. The results have new implications for financial stability policies. I first study the effects of leverage ratio requirements. A tighter leverage ratio requirement alleviates the severity of banking distress at the cost of slower credit growth. The bank liquidity risk is mitigated as a result of the bigger equity buffer. In addition, there is a new insight that the restrictions on leverage ratio also dampen the amplification

effects of liquidity rationing and credit line runs. I then consider a commitment tax on bank cutting credit lines. As an effective commitment device, a tax of 0.4% is enough to eliminate credit line runs. The tax is also a countercyclical policy by design. Especially, unlike quantity requirements, it does not curb credit growth in economic upturns.

The model takes the contractual properties of credit lines as given and sidesteps the deeper reasons. The following two fundamental frictions may be at work, as indicated by the empirical evidence. First, only borrowers themselves can observe whether they are hit by a liquidity shock. Therefore, contracts cannot be contingent on this information, and borrowers are able to tap credit lines at will. In particular, they can do so preemptively when they are not hit, which leads to a run on credit lines. The second friction comes from borrowers' moral hazard. Although the bank can observe whether the borrowers are behaving, it is difficult to verify this information to the outsiders, such as courts. To deter misbehaving, the bank reserves the right to repudiate the contracts, instead of committing fully to them.

Though my modeling of runs is in the spirit of [Diamond and Dybvig \(1983\)](#), it is different from classic bank run models in two aspects. First, instead of focusing on the coordination failure among borrowers, my model underlines the strategic complementarities between a bank and a group of borrowers and traces out the associated amplification mechanism. Second, in my model both the bank and the borrowers optimize over a long rather than a three-period horizon. In particular, a borrower's run decision depends on the expected value of having a credit line from the bank, which in turn depends on the borrower's belief on the bank's future decisions in this fully dynamic model. Third, my model features partial runs that only a fraction of agents would run. Moreover, the severity of runs, i.e. how much fraction of agents run, is affected by the endogenous bank liquidity supply at equilibrium.

This article is related to the literature on credit lines. [Sufi \(2009\)](#) find that credit lines are contingent but not committed sources of liquidity insurance, and shows that

firms use both cash and credit lines in managing liquidity because of the risk of credit line revocation. [Roberts and Sufi \(2009\)](#) explore the consequences of financial covenant violations. They show that creditors obtain the control rights after violations to tight lending terms. On the bank side, [Kashyap, Rajan, and Stein \(2002\)](#) study the synergy between credit lines and deposits regarding liquidity provision by banks. [Acharya and Mora \(2015\)](#) show that this synergy broke down in the first year of the 07-09 crisis and that banks are exposed to double pressures on assets and liabilities. [Acharya, Almeida, Ippolito, and Perez \(2016\)](#) examine the role of bank health and economic conditions in determining the accessibility of credit lines. Recent studies also reveal the existence of credit line runs during the Great Recession. [Ivashina and Scharfstein \(2010\)](#) provide evidence that firms increased the use of credit lines after the failure of Lehman Brothers. [Ippolito, Peydro, Polo, and Sette \(2015\)](#) use Italian Credit Register data to document that firms with multiple credit lines draw down especially from banks with higher exposure to the wholesale funding market. Based on these empirical findings, I identify a new amplification effect in lending through credit lines.

There are also extensive theoretical studies on bank credit lines. [Holmstrom and Tirole \(1998\)](#) demonstrate that banks can use credit lines to provide liquidity insurance to firms as an implementation of the optimal dynamic contract. They also show that when firms' shocks are correlated bank credit lines may not be sufficient. Similarly, in my model, the bank's own risk is not diversifiable and hence disrupts liquidity insurance. The discretion in credit line availability is also highlighted as a fundamental difference from term loan. In [Boot, Greenbaum, and Thakor \(1993\)](#), banks use contracts with discretion to manage jointly financial and reputational capital and to overcome asymmetric information problems by signaling. More recently, [Acharya, Almeida, Ippolito, and Perez \(2014a\)](#) propose a model in which credit line revocation arises endogenously as a result of monitoring. Yet, although the run incentives of credit line borrowers are supported by the empirical studies, it is largely overlooked in the theoretical literature. To

fill this gap, my model combines credit line runs and bank discretion into one framework and quantify the isolated strategic effect.

This article is also related to the vast bank run literature, including the seminal work of [Bryant \(1980\)](#) and [Diamond and Dybvig \(1983\)](#), and more recently [Cooper and Ross \(1998\)](#), [Allen and Gale \(1998\)](#), [Peck and Shell \(2003\)](#), [Rochet and Vives \(2004\)](#), [Bond and Rai \(2009\)](#), [He and Xiong \(2012\)](#), [Vives \(2014\)](#), [Miao and Wang \(2015\)](#), [Benhabib, Miao, and Wang \(2014\)](#), and [Liu \(2016\)](#). This literature largely focuses on the coordination failure among depositors, whereas treats banks passively once runs start. A few exceptions include [Ennis and Keister \(2009, 2010\)](#), [Cheng and Milbradt \(2012\)](#), [Gertler and Kiyotaki \(2015\)](#), and [Zeng \(2016\)](#). [Ennis and Keister \(2009, 2010\)](#) show that when policymakers have limited commitment power suspension of convertibility cannot prevent runs. The benevolent Policymakers would postpone interventions to serve the impatient depositors who haven't withdrawn. In contrast, reducing liquidity supply is ex-post optimal for the self-interested bank in my model, but not ex-ante optimal as it induces runs by the borrowers. [Engineer \(1989\)](#) also show that suspension of convertibility cannot eliminate runs. Furthermore, [Cipriani, Martin, McCabe, and Parigi \(2014\)](#) highlight the danger of suspension that it may create runs. [Gertler and Kiyotaki \(2015\)](#) develop an infinite-horizon model that features financial accelerator effects and roll-over bank runs. The dependency of run probability on fire-sale prices amplifies the aggregate disturbances, even beyond the amplification from the conventional financial accelerator. Considering fire sales resulting from shifts in the composition of assets instead of deleveraging, [Zeng \(2016\)](#) show that cash re-building policies of mutual funds generate a first-mover advantage that leads to shareholder runs. My paper differs from these studies in two aspects. First, my analysis speaks to the impact of bank liquidity on the corporate sector through credit lines, including the reduction of total credit and misallocation of funds. This leads to different welfare analysis and policy implications. Second, credit lines are long-term contracts, hence borrowers' incentive to run is affected by their expectations about bank

liquidity policies in the future. This is important as the bulk of financial relationships are long term.

The quantitative analysis follows the precedent of recent papers that estimate structure models of the banking sector, including [Corbae and D’Erasmus \(2014\)](#) and [Mankart, Michaelides, and Pagratis \(2014\)](#) among others. [Schroth, Suarez, and Taylor \(2014\)](#) estimate a dynamic debt run model based on [He and Xiong \(2012\)](#) and show that runs are sensitive to bank’s balance sheet composition. My paper is also related to [Egan, Hortaçsu, and Matvos \(2015\)](#), who estimate a bank run model with a differentiated deposit market and explore multiple equilibria.

Layout. The remainder of the paper is organized as follows. Section [1.2](#) presents relevant empirical evidence. Section [1.3](#) lays out a benchmark model to study bank liquidity management with run-prone borrowers. Section [1.4](#) discusses the Markov perfect equilibrium of the benchmark model and analyzes an alternative model in which the bank can commit. The benchmark model is then calibrated in section [1.5](#). Section [1.6](#) conducts counterfactual experiments and Section [1.7](#) concludes.

1.2 Empirical Evidence

I begin by providing empirical evidence about credit line usage and availability, and synthesizing the empirical literature. I emphasize three key aspects: (i) banks tighten liquidity by reducing limits of credit lines during the 07-09 crisis, (ii) credit line borrowers have the incentive to draw down early in case banks restrict credit line access in the future, and (iii) on bank balance sheets total loans contracted less than total credit, the sum of loans and unused credit lines, after the Lehman failure.

Credit Line Availability. Figure [1.2](#) shows that banks reduce credit line accessibility during the Great Recession. It plots the two most common outcomes of loan amend-

ments in Dealscan database: credit limit reduction and interest rate increase. The average reduction in credit limits of all amendments in 2009 is 53 million dollars. At the same time, the new margin over LIBOR is about 340 basis points on average in 2009, which makes borrowing more costly and thus limits the use of credit lines indirectly. Consistent with the idea of rationing, the credit limit reductions in 2009 are more prominent than the increases in interest rates.

The presence of covenants in credit line contracts gives creditors the right to limit access conditional on covenant violations. Importantly, banks have the discretion of how to use the right and hence determining the consequences of violations. [Acharya, Almeida, Ippolito, and Perez \(2014b\)](#) find evidence on that accessibility to credit lines, following violations, depends on bank health. In particular, banks are more likely to withdraw credit lines instead of waiving covenant violations during crises. Moreover, most credit lines have the material adverse change covenant, which provides lenders the discretion to determine whether a borrower's credit quality deteriorates significantly enough to trigger a violation.

Credit Line Usage. [Ivashina and Scharfstein \(2010\)](#) is the first to point out that the increase in total lending in 2008Q4 is driven by an increase in drawdowns by existing credit line borrowers. They also provide some evidence on preemptive drawdowns by examining the SEC filings. To identify the effect of expected decline in liquidity supply on credit line usage, [Ippolito, Peydro, Polo, and Sette \(2015\)](#) compare drawdowns by the same firm from different banks. They find that higher exposure to the interbank market leads to more drawdowns. Figure 1.1 presents a more direct evidence on preemptive drawdowns using the CFO survey data provided in Table 8 of [Campello, Graham, and Harvey \(2010\)](#). The 2008Q4 survey explicitly asks about firms' reasons to draw credit lines. There are 17% of constrained firms and 8% of unconstrained firms reporting that

they draw down credit lines in case the bank restricts line access in the future.² Although unfortunately this question is only asked in the 2008Q4 survey hence we cannot tell if firms draw down preemptively beyond the crisis, this fact, nonetheless, shows that firms indeed have the incentives to run.

Bank Balance Sheets. Last, I document facts on bank balance sheets using quarterly data from the Consolidated Report of Condition and Income (known as Call Reports). These bank-level facts are consistent with the observations discussed above.

Given the interest in the effects of liquidity rationing, I first explore the evolution of *bank total credit*, defined as the sum of loans and unused credit lines.³ The upper panel of Figure 1.3 presents the average total credit growth of banking holding companies (BHCs) with asset more than 10 billion dollars over time. There is a clear cyclical pattern; moreover, the quarterly growth rate drops to about -2% in 2009, lower than in previous recessions.

Once borrowers tap credit lines, the drawn part appears on balance sheets as loans. To shed light on the usage of credit lines, I next look at the dynamics of bank loans. I normalize loans by total credit to remove the cyclical component and plot the time series of this ratio in the lower panel of Figure 1.3. The ratio is pretty much stable until it shoots up in 2008. This increase in the loan-to-credit ratio is consistent with the evidence of preemptive drawdowns.

While Figure 1.3 presents the cross-sectional averages across time, Figure 1.4 plots the growth of loans against the growth of total credit of individual BHCs. It shows that loans contract much less than total credit in the year following Lehman failure (above

²The survey also asks whether firms' operations are "not affected", "somewhat affected", or "very affected" by difficulties in accessing the credit markets. Firms that are "very affected" are considered as constrained. The question on reasons to draw credit lines allows multiple choices.

³Loans consist of credit line drawdowns and term loans. For the purpose of measuring liquidity rationing through credit line cuts, it is without loss of generality to treat term loans as a form of credit lines with fixed borrowing and repayment schedules.

the dashed 45% line). It is also statistically significant that the gap between the two growth rates widens as total credit shrinks, which is consistent with my model.

1.3 The Model

In this section I develop a framework with strategic interactions between a bank and a group of borrowers. The model is motivated by the key institutional settings of credit lines: flexibility and bank discretion. Although I consider the specific context of credit lines, the amplification mechanism is fairly general and can be readily extended to other settings with a single large player and a continuum of small players.

1.3.1 Setup

Time is discrete and infinite. There is a continuum of borrowers having credit lines from a single bank. The mass of borrowers varies over time as borrowers enter and exit.

Timeline. Figure 1.5 shows the timeline. In period t , the set of borrowers with access to credit lines is denoted as $\mathcal{I}_t$. Each borrower i has access to her credit line with limit $\phi_{i,t}$, and Φ_t denotes the total available credit. At the beginning of each period, a publicly observed shock hits the bank. This bank liquidity shock $z_t \in \{z_b, z_g\}$ ("bad" and "good" states) evolves as a Markov process $F(z', z) = \text{prob}(z_{t+1} = z' | z_t = z)$. As a consequence, the bank fails at the end of the period with probability $p(z_t)$ such that $p(z_g) \leq p(z_b)$.

In each period, each borrower decides how much to draw down simultaneously with other borrowers. Drawn credit lines show up as loans on the asset side of bank balance sheet, whereas the unused portion remains off balance sheet. I denote the amount tapped by borrower i as $l_{i,t}$ and the total drawdowns as L_t . The bank starts with equity E_t and raises deposits to finance the credit line drawdowns. After that, it may alter the liquidity provision for the next period. On the one hand, the bank may issue new credit lines

to new borrowers. On the other hand, the bank may reduce the credit limits of existing borrowers to $\{\phi_{i,t+1}\}_{i \in \mathcal{I}_t}$ (and total credit to Φ_{t+1}) but not below the outstanding balances because the bank cannot force borrowers to repay. Borrowers exit if they keep a zero balance and their credit limits are reduced to zero.

At the end of the period, three events occur in sequence. First, the bank may fail with probability $p(z_t)$, which is given exogenously as a function of the bank's liquidity shock z_t . This assumption can be micro-founded by a deposit run model, in which $p(z_t)$ is determined by a specific equilibrium selection mechanism. Thus, my model can be viewed as a model of double bank runs – a deposit run and a credit line run. To focus on the feedback effects associated with credit line runs I model bank failure exogenously in this paper and treat it as a shock to bank health. I am also working on an extension with endogenous bank failure. In that setting, the two types of bank runs may reinforce each other.

When the bank fails shareholders receive nothing, and borrowers can no longer tap funds from the credit lines. Therefore, borrowers who keep a zero balance after the bank fails will exit because effectively their credit limits are reduced to zero. Whereas those with positive outstanding balances may keep the drawn funds to finance future needs instead of repaying them immediately.

Second, if the bank does not fail, with an exogenous probability π the bank pays out the whole equity to existing shareholders and raises the same amount of equity from new shareholders. This is a simple way to model dividend payouts. Although the shareholders are replaced, the bank continues to function and the borrowers are not affected.

Finally, each credit line matures independently with probability δ . When their credit lines mature borrowers need to repay their balances and then exit. To sum up, new borrowers enter passively when the bank issues new credit lines. Whereas existing borrowers may exit for two reasons: their credit lines mature, and their credit limits

are reduced to zero. If a borrower exits, she obtains a constant fraction, $1 - \eta$, of what she would otherwise receive.

Borrowers. Each borrower operates a long-term project of constant size that generates return R in each period. At the beginning of each period there is an idiosyncratic and privately-observed liquidity shock that hits each borrower with probability Λ_t . The shock requires borrowers to inject additional funds into the project to avoid their projects being liquidated. Funds are required for one period only and become fully liquid afterward. I normalize the size of the needed funds to be 1\$, and let $\lambda_{i,t} \in \{0,1\}$ denote whether borrower i is affected at period t .

To insure against this liquidity shock, each borrower obtains a credit line from the bank which specifies a credit limit. Borrowers pay a fixed maintenance fee r^ϕ on the credit line limits, regardless of whether the lines are drawn or not. Besides, they also pay a usage fee r_t^l on the amount that they borrow from the bank. The usage fee is set at a fixed margin r_x over a reference rate r_t , such as LIBOR rate.

Borrowers can draw funds up to their credit limits at will, even when they are not hit by the shock. When not hit by the shock, borrowers face a trade-off in making drawdown decisions. The benefit of drawdown is that the drawn funds will be available to borrowers in the next period because the bank cannot force them to payback. In contrast, undrawn credit lines may become unavailable if the access to credit lines is restricted by the bank. The cost of drawdown comes from the usage fee. To generate realistic credit line runs, I also assume that if the funds are drawn preemptively they generate an idiosyncratic and privately-observed return $\kappa_{i,t} \in [\underline{\kappa}, \bar{\kappa}]$. Thus, the effective cost is the difference between the usage fee and the stochastic return. The return is drawn at the beginning of the period from distribution $\Omega(\cdot)$, and it can be negative as a liquidity storage cost.

The assumption of 0-or-1 liquidity shock is essential for tractability. Because of it, borrowers either borrow 1\$ from the bank or do not borrow at all and hold a zero balance. Furthermore, borrowers would obtain credit lines with the limit of 1\$ exactly. A higher limit imposes unnecessary maintenance fees, while credit lines with a limit less than 1\$ are entirely useless. Therefore, this assumption allows us to focus on borrowers' drawdown decisions and abstract away the heterogeneity in the credit line limits.

The bank. The bank lends to borrowers through credit lines. Each period the bank starts with equity E_t and total credit Φ_t , the sum of the limits of all credit lines. The drawn portion of credit lines appears on the balance sheet as loans L_t . The bank then raises deposits D_t to finance the loans, hence the feasibility constraint is as follows:

$$L_t = D_t + E_t. \quad (1.1)$$

Because every credit line has the same limit and the drawdown decision is binary, it is sufficient for the bank to take into account the total credit Φ_t and the total amount drawn L_t , instead of the distribution of credit line limits and usage.

Let Π_t denote the bank's profit from intermediating funds, which is a function of loans L_t , deposits D_t , and total credit Φ_t given by

$$\Pi_t = r^\phi \Phi_t + r_t^l L_t - (r_t + p(z_t)) D_t - c(D_t, \Phi_t). \quad (1.2)$$

The first two terms of Equation 1.2 are the maintenance fee and the usage fee received by the bank. As mentioned above, borrowers need to pay this maintenance fee regardless of whether they borrow from the bank or not. The third term represents interest expenses on deposits. The bank pays a base rate r_t and a premium $p(z_t)$ to compensate depositors

for the risk of bank failure.⁴ The fourth term captures non-interest expenses, such as employee compensation and maintenance of facilities. I assume that $c(D_t, \Phi_t)$ is convex in deposit D_t to introduce curvature into the model.

The bank's profit margin is given by the difference between the return from lending $r^\phi + r_t^l$ and the cost of raising deposits $r_t + p(z_t)$. The margin, together with the convex cost $c(D_t, \Phi_t)$, determines the optimal leverage of the bank. In bad times, because the additional cost that compensates for the bank failure risk, the profit margin is thinner, and in turn the optimal leverage is lower than that in good times. Therefore, in general the bank would prefer to deleverage when hit by the negative bank liquidity shock and leverage up during recoveries.

An increase in drawdowns may lead to an additional pressure on deleveraging, which is a key component of the amplification mechanism. When the marginal non-interest expense of lending exceeds the profit margin, an increase in drawdowns would decrease bank profit and, in turn, reduce next-period equity. Thus, the bank faces further pressure on deleveraging. This may happen when bank leverage is high and, in particular, when the bank is hit by a negative shock after staying in the good state for a long time. In the latter case, lending may become costly because the cost of raising deposit increases when a negative shock arrives.

Different from a typical banking model, the bank cannot deleverage directly by adjusting total loans, which is determined by borrower drawdown decisions. Instead, the bank may reduce next-period credit limits to control leverage indirectly. In doing so, there are three cases to consider. First, a borrower draw down up to the limit of 1\$ and the bank cannot reduce her limit. Second, a borrower does not draw her credit line, and the bank chooses not to reduce her credit limit. Third, a borrower does not draw her credit line, and the bank decides to cut her limit to 0 since any positive amount be-

⁴The usage fee is based on the same interest rate r_t . In practice, the most common reference rate used in credit line contracts is the LIBOR rate, which is the average of interest rates estimated by each of the leading banks that it would be charged were it to borrow from other banks.

tween 0 and 1 is meaningless. To sum up the three cases, because the only heterogeneity among borrowers is in the binary drawdown decisions, it is sufficient for the bank to choose how many undrawn credit lines to cut. For the same reason, it is also sufficient to decide how many new credit lines to originate if the bank wants to leverage up. In addition, the bank would not cut credit lines and issue new credit lines at the same time.

Let Δ denote the change of total credit, which is negative if the bank cuts credit lines and positive if it issues new credit lines. The law of motion of total credit is thus given by

$$\Phi' = (1 - \delta)\Phi + \Delta, \quad (1.3)$$

where δ is the fraction of credit lines maturing at the end of the period. Adjusting total credit incurs a quadratic cost, denoted as $f(\Phi_{t+1}, \Phi_t)$.

I assume that the bank can only accumulate equity via retained profits as in [Gertler and Kiyotaki \(2015\)](#). While this assumption is a reasonable approximation of reality, I do not explicitly model the underlying frictions. Moreover, the bank is risk neutral and only pays out all of its equity as dividends with probability π . The bank maximizes the expected utility of shareholders at the end of period t , which is given by

$$V_t = \mathbb{E}_t \left[\underbrace{\sum_{i=1}^{\infty} \beta^i \pi (1 - \pi)^{i-1} \prod_{j=1}^i (1 - p(z_j))}_{\text{prob. of exiting at period } t+i} E_{t+i} \right]. \quad (1.4)$$

1.3.2 Discussions

The differences from classic bank runs. Classic bank run models emphasize on the strategic complementarities among depositors. Although the similar complementarities also exist in my model among borrowers, the focus of my model is instead on the strategic complementarities between the bank, as a large player, and the borrowers, as a group of small players. More specifically, in classic bank run models liquidity supply is fixed or pre-determined before depositors making withdrawal decisions. In contrast, in my

model the bank controls liquidity supply and reacts to borrower runs optimally. This interaction between the liquidity supply and credit line runs leads to a new amplification mechanism that I explore in this paper.

Model assumptions. First, the bank can reduce credit line limits but cannot force borrowers to pay off outstanding balances. This rules out the case of payment acceleration, which is rarely observed in the data (see, e.g., [Roberts and Sufi \(2009\)](#)). This assumption also captures the idea that borrowers may gain bargaining power in renegotiation by tapping credit lines, and thus rationalizes borrowers' run incentives. Moreover, I also abstract away other dimensions of credit line renegotiation for tractability, such as fees, maturity, and collateral. As long as preemptive drawdowns grant borrowers more bargaining power, the exact channel through which the bank tightens liquidity is nonessential for the existence of the feedback effects between the bank and the borrowers. Incorporating these realistic features of credit lines would greatly complicate the model so that I need to keep track of the distribution of borrowers, but would not undermine the key mechanism.

I also assume that the bank lends through credit lines only, which is a reasonable approximation of reality. In practice, about 80% of all C&I loans are made under credit lines. Moreover, in most times term loans are issued together with credit lines to the same borrowers. Adding term loans into the model does not change my qualitative results because term loan borrowers are passive and do not respond to the bank's decisions. Incorporating term loans would depress the sensitivity of preemptive drawdowns to liquidity rationing, but doing so would not affect my quantitative results in any significant way because I directly calibrate that sensitivity to the data.

Last, the stochastic returns on the credit lines drawn preemptively are introduced to generate realistic equilibrium bank runs. A run is necessarily partial in the data, with

only some borrowers participating. This assumption helps to pin down the fraction of borrowers who run.

1.4 Equilibrium Analysis

In this section, I analyze the game backward in time, first consider the bank's problem and then discuss borrowers' optimal decisions. Since I will use recursive methods to solve a Markov perfect equilibrium, let any variable x_t be denoted by x and x_{t+1} be denoted by x' . All key proofs are left to the appendix.

1.4.1 Bank Decision Making

Before diving into the bank's problem, I first define the borrowers' strategy. Borrowers make drawdown decisions based on their information at the beginning of the period; therefore borrowers' strategy, denoted by σ , is a function of bank equity E , total credit Φ , bank liquidity shock z , and borrower idiosyncratic shocks λ_i and κ_i . Formally, borrower i 's decision is given by $l_i = \sigma(E, \Phi, z; \lambda_i, \kappa_i)$.

Since the bank cannot observe borrowers' idiosyncratic shocks nor infer them from drawdown decisions, the only heterogeneity that the bank takes into account is whether a borrower draws down her credit line or not. Moreover, given that the drawdown decision is binary, the total amount of drawdowns is a sufficient statistics for the bank. After integrating l_i over the two idiosyncratic shocks the total amount of drawdowns L is a function of E , Φ , and z only.

At the beginning of the period, the bank starts with equity E and provides total credit Φ , the sum of its credit line limits. After borrowers make drawdown decisions, the bank chooses the next-period total credit Φ' to maximize expected discounted dividends, as

defined in Equation 1.4. The bank's problem can be represented recursively,

$$\begin{aligned}
V(E, \Phi, z) &= \max_{\Phi' \geq (1-\delta)L} \beta\pi(1-p)E' + \beta(1-\pi)(1-p)\mathbb{E}_{z'|z}V(E', \Phi', z') \quad (1.5) \\
\text{s.t.} \quad D &= L(E, \Phi, z) - E \\
E' &= E + \Pi - f(\Phi', \Phi),
\end{aligned}$$

where $\beta\pi E'$ represents the expected value from dividend payouts, and $(1-\pi)(1-p)$ is the probability of continuing without dividend payouts. The next-period equity E' is given by equity E plus profit Π and minus the adjustment cost $f(\Phi', \Phi)$. In addition, since the bank cannot force repayment of drawn credit lines, Φ' has to exceed the unmatured outstanding balances $(1-\delta)L$.

Let σ^B denote the bank's strategy, then the bank's decision can be expressed as $\Phi' = \sigma^B(E, \Phi, z)$. The optimal strategy satisfies the following first-order condition,

$$\underbrace{\left[\pi + (1-\pi) \frac{\partial \mathbb{E}V(E', \Phi', z')}{\partial E'} \right] \frac{\partial E'}{\partial \Phi'}}_{\text{adjustment costs (in utils)}} + \underbrace{\left[(1-\pi) \frac{\partial \mathbb{E}V(E', \Phi', z')}{\partial \Phi'} \right]}_{\text{benefit from (de)leveraging}} = 0. \quad (1.6)$$

First, note that since dividends are paid out only after the bank exits there is no inter-temporal trade-off on dividends as in [Gertler and Kiyotaki \(2015\)](#). Second, the choice of the next-period total credit is determined by two opposing forces. On the one hand, the adjustment cost reduces next-period equity, which in turn decreases both the continuation value and the dividend payouts. On the other hand, choosing Φ' allows the bank to expand or shrink. In particular, decreasing Φ' in bad times allows the bank to tighten liquidity supply.

1.4.2 Borrower Runs

In this subsection I investigate the borrowers' problem, showing that borrowers have incentives to run by drawing down preemptively. Moreover, credit line runs are more severe when the credit lines are less secure.

The flow payoffs simply depend on borrowers' liquidity shocks and the drawdown decisions, as shown in Table 1.1. Conditional on hit by the shock, a borrower obtains return R from the project and pays maintenance fee r^ϕ and usage fee r^l if she draws down. Otherwise she receives zero as the project being liquidated. If there is no liquidity shock, a borrower gets $R - r^\phi - (r^l - \kappa_i)$ if she taps the credit line, and $R - r^\phi$ if not. The cost of preemptive drawdown lies in the difference in flow payoffs $r^l - \kappa_i$, whereas the benefit comes from her gain in the continuation value.

Borrowers' continuation values depend on whether the bank fails and whether borrowers exit. First, if the bank continues and borrower i 's does not exit, the value at the beginning of the next period W' is determined by bank equity E' , total credit Φ' , bank shock z' , and her own idiosyncratic shocks λ'_i and κ'_i . Looking forward at the end of the current period, borrower i is uncertain about the realizations of the shocks, hence her value is the expectation of W' taken with respect to z' , λ'_i , and κ'_i . Formally, the end-of-period value is given by

$$\hat{W}'(E', \Phi') = \mathbb{E}_{z'|z, \lambda'_i, \kappa'_i}[W'(E', \Phi', z'; \lambda'_i, \kappa'_i)]. \quad (1.7)$$

If the bank continues and borrower i exits, she can only obtain a fraction of this value, $(1 - \eta)\hat{W}'(E', \Phi')$. Let q denote the endogenous probability of exit as a result of bank cutting undrawn credit lines. The total probability of exit is thus δ if borrower i draws down and $\delta + q$ otherwise.

The probability endogenous probability of exit is determined at the equilibrium, which equals 0 if the bank does not cut credit line at all ($\Delta > 0$) and

$$q = \min \left\{ \frac{-\Delta}{\Phi - L}, 1 \right\} \quad \text{if } \Delta < 0, \quad (1.8)$$

where $-\Delta$ is the total credit lines being cut when $\Delta < 0$, and $\Phi - L$ is the total amount of undrawn credit lines. Because each borrower is atomless, borrower i 's decision has no direct impact on q , thus she takes this probability as given when making decisions.

Second, if the bank fails and borrower i does not exit, which happens if she holds a positive balance, borrower i may still hold the borrowed funds to finance future needs until she pays them off upon exit.⁵ In doing so, her end-of-period value can be formulated recursively as

$$W_0 = \underbrace{\mathbb{E}_{\lambda'_i, \kappa'_i} [R - r^\Phi - r^l + (1 - \lambda'_i)\kappa'_i]}_{\text{expected next-period flow payoff}} + \beta(1 - \eta\delta)W_0, \quad (1.9)$$

where the first term represents the expected next-period flow payoff. If borrower i exits after the bank fails, her value drops to $(1 - \eta)W_0$.

The total probability of exit also depends on drawdown decisions. If borrower i draws down in the period, she exits only if the credit line matures with probability δ . Whereas if she doesn't draw down, the probability of exit is 1 because her balance is zero and she cannot borrow from the failed bank any more.

To sum up, as shown in Table 1.1 borrowers' continuation values are functions of the two end-of-period values $\hat{W}'$ and W_0 , and the probabilities of exit. I also denote the continuation value after liquidation as W_L . Therefore, the value of borrower i at the

⁵The borrower may also repay the balance voluntarily and exit at the end of the period, but she will not choose to do so because of the loss in continuation value. Allowing borrowers to make voluntary exit decisions after observing their shocks would not lead to any significant change.

beginning of the period after observing the shocks is given by

$$W(E, \Phi, z; \lambda_i, \kappa_i) = \max_{l_i} l_i [R - r^\phi - r^l + (1 - \lambda_i) \kappa_i + \beta(1 - p)(1 - \eta\delta) \hat{W}' + \beta p(1 - \eta\delta) W_0] \\ + (1 - l_i) \{ (1 - \lambda_i) [R - r^\phi + \beta(1 - p)(1 - \eta\delta - \eta q) \hat{W}' + \beta p(1 - \eta) W_0] + \lambda_i \beta \omega_L \}, \quad (1.10)$$

where $\hat{W}'$ and W_0 are defined as in Equations 1.7 and 1.9, $l_i \in \{0, 1\}$ denotes the draw-down decision, and $\lambda_i \in \{0, 1\}$ represents whether borrower i is hit by the liquidity shock.

To solve for the borrowers' optimal choices, I first assume that W_L is low enough that borrowers would always draw credit lines if hit by the shock to avoid liquidation. When not hit by the shock borrowers face a key trade-off between the borrowing cost and the loss in continuation value as presented in Table 1.1 and Equation 1.10. Borrower i taps her line if and only if the loss in continuation value exceeds the borrowing cost.

Proposition 1 *Given the value functions and strategies of the bank and other borrowers, the drawdown decision of borrower i who has no liquidity need is given by*

$$l_i^* = \begin{cases} 1 & \text{if } r^l - \kappa_i < \beta[(1 - p)\eta q \hat{W}' + p\eta(1 - \delta)W_0] \\ 0 & \text{otherwise} \end{cases}. \quad (1.11)$$

Moreover, l_i^* is increasing in q .

There exists a threshold value κ^* such that borrower i taps her line if and only if $\kappa_i \geq \kappa^*$. Thus, the stochastic return helps to determine the fraction of borrowers who draw down credit lines.

Finally, since λ_i and κ_i are both i.i.d, the total amount of drawdowns L can be expressed as

$$\begin{aligned} L(E, \Phi, z) &= \mathbb{E}_{\lambda_i, \kappa_i}[\sigma(E, \Phi, z; \lambda_i, \kappa_i)] \Phi \\ &= [\Lambda + (1 - \Lambda)(1 - \Omega(\kappa^*))] \Phi, \end{aligned} \tag{1.12}$$

where Λ is the probability that a borrower being hit by the shock, and $\Omega(\cdot)$ is the cumulative density function of the return of drawdowns. Therefore, $\Lambda + (1 - \Lambda)(1 - \Omega(\kappa^*))$ is the ex-ante probability of drawdown before the idiosyncratic shocks realize, which also represents the fraction of borrowers who draw credit lines.

Following Equations 1.8, 1.11, and 1.12, given the bank's strategy borrowers may coordinate to multiple levels of total drawdowns in some cases. Hence we need a selection mechanism. The exact way in which the stable equilibrium is selected is not crucial to the main point about the amplification mechanism, but in order to show the feedback effects formally I assume the resulting total drawdowns $L(E, \Phi, z)$ being a smooth function. Moreover, in the numerical analysis, to be consistent with the data I assume that the borrowers will coordinate to the best equilibrium whenever multiple equilibria occur.

1.4.3 Definition of Equilibrium

I solve for a Markov perfect equilibrium. A Markov perfect equilibrium is a subgame perfect equilibrium in which the strategies depend only on the payoff-relevant history. The payoff-relevant history can be summarized by the states $\{E_t, \Phi_t, z_t\}$, and both the borrowers and the bank act according to the states. Denote the states as $\{E_t, \Phi_t, z_t\} \in \mathcal{S}$ and borrowers' idiosyncratic shocks as $\{\lambda_{i,t}, \kappa_{i,t}\} \in \mathcal{H}$.

Definition 1 *A Markov perfect equilibrium of the model is a pair of value functions ($V : \mathcal{S} \rightarrow \mathbb{R}^+$, $W : \mathcal{S} \times \mathcal{H} \rightarrow \mathbb{R}^+$) and strategies ($\sigma^B : \mathcal{S} \rightarrow \mathbb{R}$, $\sigma : \mathcal{S} \times \mathcal{H} \rightarrow \{0, 1\}$) such that*

1. *given the policy functions, (V, W) solve the Bellman equations 1.5 and 1.10;*
2. *for any borrower, $\hat{\sigma} = \sigma$ is optimal given W, σ^B , and that all other borrowers follow σ ;*
3. *given σ and V, σ^B is optimal for the bank.*

The Markov perfect equilibrium defined above is subgame perfect, and thus it is dynamically consistent. There is, however, a dynamic inconsistency problem on the bank side. Dynamic inconsistency is a situation where a player's best plan for future periods will not be optimal when the future periods arrive. In my model, the bank wants to commit to not cut credit lines even if in bad times and under the pressure of massive drawdowns. If this commitment were credible, it would stop the borrowers from drawing down in the first place. However, the bank might not be able to commit its future self to the plan because if the borrowers do in fact tapping their lines aggressively, the bank's future self would respond to it actively, instead of sticking to the plan.

Because of this dynamic inconsistency problem, the Markov perfect equilibrium defined above is not constrained efficient. In the rest of this section, I first analyze the feedback effects that lead to this inefficiency and then consider a model in which the bank has commitment power in section 1.4.5.

1.4.4 A Positive Feedback Loop

In this subsection, I demonstrate that a positive feedback loop emerges and exacerbates financial fragility, as a result of strategic complementarities between the bank and the borrowers. In particular, I analyze the impulse responses to a negative bank liquidity shock.

A negative shock to z_t increases the probability of bank failure and, in turn, the cost of raising deposits. Because the shock is persistent, it is also more likely that the cost will also be high in the future. Overall, this negative shock leads the bank to deleverage and ration liquidity by controlling the next-period total credit Φ' . In the following, I discuss

how this negative shock and the resulting liquidity contraction is amplified through the feedback loop.

As demonstrated in Proposition 1, A reduction in total credit Φ' affects drawdown decisions through two channels. First, there is a short-run effect working through the probability of credit line being cut q . A reduction in total credit increases the probability and leads to more preemptive drawdowns. There is also a subtle long-run effect through borrowers' continuation value $\hat{W}$. A reduction in total credit decreases bank leverage. Looking forward the bank becomes healthier and the credit lines are more secure. Thus, the continuation value of having a credit line increases, which in turn induces borrowers to run.

The incentive to run of each borrower results in an increase in total drawdowns. In particular, Proposition 2 considers how more credit line cuts impact the selected stable equilibrium of the coordination stage-game among borrowers. The condition that $\hat{W}'$ increases in credit line cuts guarantees that the long-run effect is in the right direction. In most cases, the short-run effect would dominate; hence, that condition is not essential.

Proposition 2 *If $\hat{W}'$ increases in credit line cuts, borrowers' drawdown L increases in the amount of credit lines being cut at the equilibrium.*

Borrower runs, in turn, may impose further deleveraging pressure on the bank. The bank's trade-off lies between the adjustment cost and the benefit of deleveraging, but how would the bank react to an increase in drawdowns? To facilitate discussion, I first present the normalized bank problem.

Assumption 1 *Both cost functions $c(D, \Phi)$ and $g(\Phi', \Phi)$ are homogeneous of degree one. Moreover, both functions are convex in their first argument.*

Lemma 1 *Under Assumption 1, bank value function is homogeneous of degree one in equity E and total credit Φ . It can be normalized by Φ as*

$$V(E, \Phi, z) = \Phi v(e, z),$$

where e denotes the ratio of equity to total credit. In addition, borrowers' value function can be simplified to

$$W(E, C, z; \lambda_i, \kappa_i) = w(e, z; \lambda_i, \kappa_i).$$

By normalization, essentially the bank is divided into multiple identical banks, each with a total credit of 1. From the borrowers' perspective, since bank strategies are invariant to normalization, having a credit line from one of these identical small banks is the same as having a credit line from the original bank. Therefore, borrowers' value function is homogeneous of degree zero in E and Φ .

When the total drawdowns shoot up, both the cost and the benefit of deleveraging increase. Proposition 3 provides sufficient conditions under which the increase in benefit outweighs the increase in cost. It follows that the bank reduces next-period total credit when drawdowns increase.

Proposition 3 *If bank profit $\Pi(E, \Phi, z)$ is concave in E and $\partial \Pi(E, \Phi, z) / \partial E > -1$ given the borrower strategy $L(E, \Phi, z)$, then (i) the normalized value function v is increasing and strictly concave in equity ratio e , and (ii) the next-period total credit Φ' decreases in drawdowns whenever the marginal non-interest expenses exceed the profit margin.*

Whether the conditions are satisfied depends on the state variables. For example, the marginal non-interest expenses exceeds the profit margin when the bank's leverage is high and, in particular, when the bank is hit by a negative shock after staying in the good state for a long time.

To summarize, after a negative shock to z_t the bank reduces total credit. This reduction induces the first wave of borrower runs. In turn, the preemptive drawdowns push the bank to reduce total credit further, which leads to the second wave of runs. This process repeats and becomes a downward spiral that amplifies the impact of the initial shock. The spiral may lead to multiple equilibria if the strategic complementarity is strong enough; however, it turns out that this is not the case given the calibrated parameters.

1.4.5 A Model with Commitment

In this subsection I consider an alternative model in which the bank can commit to its plan. The main difference of this setting to the benchmark model is that now the bank can influence borrower decisions through the promised bank choices.

In the benchmark model, borrowers make decisions based on their beliefs on the bank's choices, which are determined at the equilibrium as a function of the state variables. In contrast, if the bank can commit to its policy, borrowers would take the promised plan directly into account. Therefore, both borrowers' value function and strategy depend on the promised next-period total credit directly. In particular, the total amount of drawdowns in the model with commitment is given by $L(E, \Phi, z; \Phi')$ instead of $L(E, \Phi, z)$. Taking this into consideration, the bank solves the following problem,

$$\begin{aligned} V(E, \Phi, z) &= \max_{\Phi' \geq (1-\delta)L} \beta\pi(1-p)E' + \beta(1-\pi)(1-p)\mathbb{E}_{z'|z} V(E', \Phi', z'), & (1.13) \\ \text{s.t.} \quad D &= L(E, \Phi, z; \Phi') - E \\ E' &= E + \Pi - f(\Phi', \Phi). \end{aligned}$$

This bank problem is the same as the bank problem in 1.5, except that the total amount of drawdowns is determined differently and, in particular, depends on the

bank's choice of next-period total credit Φ' . Formally, a Markov perfect equilibrium in this model with commitment is defined as follows.

Definition 2 *A Markov perfect equilibrium of the model in which the bank can commit is a pair of value functions $(V : \mathcal{S} \rightarrow \mathbb{R}^+, W : \mathcal{S} \times \mathcal{H} \times \mathbb{R} \rightarrow \mathbb{R}^+)$ and strategies $(\sigma^B : \mathcal{S} \rightarrow \mathbb{R}, \sigma : \mathcal{S} \times \mathcal{H} \times \mathbb{R} \rightarrow \{0, 1\})$ such that*

1. *given the policy functions, (V, W) solve the Bellman equations 1.10 and 1.13;*
2. *for any borrower, $\hat{\sigma} = \sigma$ is optimal given W, σ^B , and that all other borrowers follow σ ;*
3. *given σ and V, σ^B is optimal for the bank.*

Since the bank can promise to its plan and not react to drawdowns, the feedback loop between the bank and the borrowers collapses. Thus, comparing the benchmark model and the model with commitment allows us to tease out the mechanism and better understand the danger of lack of commitment.

Proposition 4 *Assume that drawdowns $L(E, \Phi, z; \Phi')$ is a smooth function. If the bank can commit, (i) it cuts fewer credit lines whenever the marginal non-interest expenses exceeds the profit margin, (ii) there are less preemptive drawdowns, and (iii) both the bank and the borrowers are better off.*

The bank cuts fewer credit lines after internalizing the indirect cost coming from credit line runs, as indicated by the first-order condition given by

$$\left[\pi + (1 - \pi) \frac{\partial \mathbb{E}V(E', \Phi', z')}{\partial E'} \right] \left[\underbrace{\frac{\partial E'}{\partial \Phi'} + \frac{\partial E'}{\partial L} \frac{\partial L}{\partial \Phi'}}_{\text{ind. cost}} \right] + \underbrace{\left[(1 - \pi) \frac{\partial \mathbb{E}V(E', \Phi', z')}{\partial \Phi'} \right]}_{\text{benefit from (de)leveraging}} = 0. \quad (1.14)$$

The new term $\frac{\partial E'}{\partial L} \frac{\partial L}{\partial \Phi'}$ represents the indirect cost from borrower drawdown decisions L as a response to the bank's choice of Φ' , which is absent in the first-order condition 1.6. The

bank is better off because promising the strategy at the equilibrium of the benchmark model is still feasible. Borrowers are better off because their credit lines are more secure.

1.5 Numerical Results

1.5.1 Calibration

I fit the model to U.S. commercial bank data during the 1993-2009 period. The main data source is the bank-level data from the Consolidated Report of Condition and Income (Call Reports). It contains balance sheet information of all U.S. commercial banks at the quarterly frequency. Following the literature, I aggregate the bank-level data to bank holding company level because these ownership ties could foster liquidity sharing across subsidiaries. I drop banks with asset growth greater than 10% during a quarter to mitigate the effect of large mergers and winsorize all variables at 5th and 95th percentile to control the outliers. Also, I focus on large banks with total assets more than 10 billion dollars.

I also use loan-level information from Dealscan dataset, which covers the syndicated corporate loan market in the United States and contains detailed information on credit line issuance.

Parametrization. The prices are exogenously determined and depend on the bank liquidity shock z_t . The usage fee is given by

$$r_t^l = r_t + r_x = r(z_t) + r_x, \quad (1.15)$$

where r denotes the base rate and r_x is the constant premium of usage fee.

I parametrize the non-interest expenses as follows,

$$c(D_t, \Phi_t) = \gamma D_t^2 / \Phi_t. \quad (1.16)$$

This quadratic cost introduces curvature into the model and also leads to well defined bank leverage. The adjustment cost is also quadratic given by

$$g(\Phi_{t+1}, \Phi_t) = \mu \Delta_t^2 / \Phi_t, \quad (1.17)$$

where $\Delta_t = \Phi_{t+1} - (1 - \delta)\Phi_t$ from Equation 1.3. It generates gradual adjustments of total credit and hence gives the model enough flexibility to match the growth of total credit in the sample.

The stochastic process for the return of preemptive drawdown $\Omega(\cdot)$ is simply taken to be the uniform distribution on $[\underline{\kappa}, \bar{\kappa}]$. $\bar{\kappa}$ is normalized to equal to $r^l(z_b)$, such that the highest return cancels with the usage fee in bad times. Given the assumption of uniform distribution, there are two implications. First, the sensitivity of borrowers' response to bank choices depends on the density function, which in turn is determined by the lower bound $\underline{\kappa}$. However, $\underline{\kappa}$ is not directly relevant because borrowers with low returns choose not to run. Second, the stage game among borrowers admits at most two stable equilibria given the bank's strategy. Moreover, if there are two equilibria, one of them corresponds to runs by more than a half of borrowers without liquidity shock, and the other corresponds to runs by less than a half of those borrowers. I select the second equilibrium because in the data I only observe runs by a small fraction of borrowers.

Calibration. A model period is set to be one quarter. I reduce the processes for z_t to a two-state Markov process that $z_t \in \{z_g, z_b\}$. To calibrate the stochastic process $F(z', z)$, I use the NBER recession dates and match the average duration of recessions of 6 quarters and the average time periods between two recessions of 10 years.

I then divide the sample into two subsamples before and after the third quarter of 2008. I map the periods without bank liquidity shock (the good state) in the simulated data to the subsample before 2008Q3, and the bad state to the subsample after 2008Q3.

I calibrate r_t using the average LIBOR rate. The maintenance fee r^ϕ and the premium of usage fee r^l over base rate r_t are set to the corresponding average levels in the Dealscan dataset. Borrower's return R is calibrated to the average profit to liability ratio of U.S. non-financial corporate sector. The loss of continuation values upon exits is taken to be 20% as a normalization. An alternative loss fraction leads to different calibrated parameters, but would not change the quantitative results in any significant way.

The probability of borrower liquidity shock Λ_t in the good state is estimated as the average loan to total credit ratio before 2008Q3, whereas calibrating Λ_t in the bad state requires extra care. During crises, loans may increase because of an increase in the probability of borrower liquidity shock, as well as preemptive drawdowns. To disentangle these two effects, ideally one would compare two banks with different probabilities of failure, but lending to identical borrowers, which requires a sample of firms with multiple simultaneous credit lines held separately at different banks. In the United States, however, most firms borrow from multiple banks through the syndicated loan market. Drawdowns of syndicated credit lines are distributed among the banks according to their shares in the syndication so that firms cannot draw down from a particular bank only. Because of this data limitation, I instead look at the loan-to-credit ratio of "healthy" banks based on the following criteria: (i) banks with liquidity ratios above the median, (ii) banks during the 2001 recession, (iii) banks that co-syndicated credit lines with Lehman but with exposures below the first quartile. These measures are chosen conservatively, yet even based these measures, less than a half of the additional drawdowns after 2008Q3 is due to an increase in Λ_t .

On the bank side, I set the probability of exit $\pi = 5\%$ following [Gertler and Kiyotaki \(2015\)](#). The probability of maturing δ is set to match the net new credit issuance between

2008Q3 and 2009Q4 using the loan origination data from Dealscan. More precisely, it is the effective probability that a credit line matures without being refinanced.

We are left with three parameters $\{\kappa, \mu, \gamma\}$. I calibrate these parameters by matching the following moments. The average growth rate of total credit after 2008Q3 (-1.75%), the average equity to total credit ratio (11.2%), the average loan to total credit ratio after 2008Q3 (76.2%), the average total credit growth rate (2.06%), and the average bank equity ratio after 2008Q3 (14.2%). Table 1.2 shows the calibrated parameters, and Table 1.3 provides the moments generated by the model relative to the data.

Model Fit. The first and second columns of Table 1.3 show that the model has good in-sample fit. It is important to stress that although the parameters are jointly calibrated to match all the targets there is a strong one-to-one link between the targets and the parameters. For example, the average bank equity ratio is mainly determined by the quadratic non-interest cost. Furthermore, $\bar{\kappa}$ controls how sensitive borrowers respond to bank policies, thus impacts the loan-to-credit ratio directly. Figure 1.12-1.14 provide visual evidence of the strong links that are used to identify the parameters.

The model is calibrated to match the average statistics across time, but it also does a reasonable job to capture the cross-sectional patterns. In particular, the model can generate similar patterns presented in Figure 1.4. I simulate a panel of banks from the model starting from a good state and with different leverage ratios. A negative bank liquidity shock hits all banks at the beginning of the second period and lasts for six periods. The solid line in Figure 1.6 reports the relation between the growth of loans and the growth of total credit in the first four quarters calculated from the simulated data. The relation between the two growth rates is nonlinear. Banks with enough equity have little incentive to deleverage; therefore, borrowers have no reason to run, and the growth rates are identical in the model. Whereas banks with low equity ratio choose to

ration liquidity, which induces borrower runs. As a result of preemptive drawdowns, loans decline less relative to the total credit.

Although the model is able to capture the non-linearity, it cannot, nor is designed to match the cross-section distribution of bank statistics, given the only heterogeneity in equity ratio. For instance, the minimum total credit growth is -10% in the simulated data, much higher than that in the sample. Additional heterogeneities are required to match the cross-section distributions, which is out of the scope of this paper.

1.5.2 Equilibrium Decision Rules

For the parameter values in Table 1.2, I find an equilibrium where borrowers run along the equilibrium path. To understand the equilibrium, I first describe the law of motion of the state variables, and Lemma 1 allows us to focus on the equity ratio e_t as the only endogenous state variable. Figure 1.7 shows the law of motion of bank equity ratio. For disposition purpose, I plot the change of bank equity ratio $e_{t+1} - e_t$ against e_t in both the good and the bad state.

The figure shows that equity ratio converges to the optimal level 0.112 if the bank stays in the good state for a long time. Equity ratio increase when below this level, and decreases when above. Above the optimal level, a negative liquidity shock leads to an increase in equity ratio. Thus, bank leverage is pro-cyclical; that is, the equity ratio is higher in state z_b than in z_g . When the equity ratio is too low, it further decreases in the bad state because the loss in equity dominates bank deleveraging. This last case, however, is only relevant if the bank starts with an initial equity ratio below the optimal level 0.112.

Next, I turn to characterizing the strategies. The left panel in Figure 1.8 shows the bank's optimal choice of next-period total credit as a function of the equity ratio. In good times the bank always expands, whereas in bad times the bank reduces liquidity provision when it does not have enough equity buffer. Borrower drawdown decisions

are illustrated in the right panel of Figure 1.8. Borrowers run in the bad state when the equity ratio is low, as a response to the anticipated liquidity rationing. A time-series implication of the strategies is that after a negative liquidity shock bank leverage shoots up immediately, and then decreases gradually if that shock persists.

1.6 Counterfactuals

In this section, I conduct three counterfactual experiments. First, I solve the alternative model in which the bank can commit with the calibrated parameters. Comparing the equilibrium of the alternative model with that of the benchmark model allows us to quantify the amplification mechanism. Then I use the calibrated benchmark model to study two policy interventions: (i) the effects of bank leverage ratio requirements on borrower runs and liquidity rationing, and (ii) the impact of a commitment tax on credit line cuts.

1.6.1 Commitment

If the bank can credibly commit to the plan, it would not respond to drawdown decisions ex-post. Therefore the feedback loop between the bank and the borrowers collapses. Figure 1.9 presents a comparison of the strategies for the bank and the borrowers in the benchmark model (blue lines) and the model with bank commitment (black lines).

In the model with commitment, the bank internalizes the indirect cost through the feedback loop and chooses not to cut credit lines at all. Also, because the bank can control the borrower run risk more effectively with credible commitment in the bad state, it also grows faster by issuing more new credit lines in the good state. On the borrower side, credit line drawdowns are entirely determined by borrowers' liquidity shocks, and there is no run on credit lines.

Table 1.4 compares total credit growth and welfares. Total credit growth is defined in the same way as in Section 1.5.1. For welfare comparison, I compute measures based on the value functions in the benchmark model and the counterfactuals. I use the stationary distribution of equity ratio to calculate the expected value of the borrowers. However, it is not feasible to calculate the bank value in the same way. Although the normalized bank problem is stationary, the original bank problem is not. Therefore, instead, I calculate the value of a bank with 1 unit of equity and at the steady state level of equity ratio.⁶ In the model with commitment, the average total credit growth is 2.04% in the good state, higher than its 2.03% counterpart in the benchmark model. The bank also reduces less total credit in the bad state (−0.6% versus −1.75%). Regarding welfare, the bank’s value increases by 0.1% if the bank is able to commit. At the same time, the value of borrowers also raises by 0.4%. From a social point of view, there is also an additional welfare gain of commitment as more new credit line borrowers entering on the average when the bank grows faster.

1.6.2 Leverage Ratio Requirements

Basel III framework introduced a simple, transparent, non-risk based leverage ratio as a credible measure that complements the risk-based capital requirements. In particular, off-balance sheet items, such as unused credit lines, are included in this measure at 10 to 100% conversion factors. In this subsection I ask the question, how much does a leverage ratio requirement affect credit line runs?

Specifically, I investigate the effects of a leverage ratio requirement with a 100% conversion factor for unused credit lines. Since in my model this leverage measure and the equity ratio are reciprocals of each other, it is the same to work with the equity ratio. In particular, I impose a lower bound of 15% on the equity ratio.

⁶The steady state level is defined as the level of equity ratio after a long history of $z = z_g$.

In the benchmark model, the average equity ratio in the good state is 11.2%, and thus the 15% requirement constraint binds and induces the bank to build more liquidity buffer. Figure 1.10 presents the results of this counterfactual. In the good state, the bank issues less new credit lines with the requirement, especially when its equity ratio is close to the constraint. At the same time, the bank cuts fewer credit lines in the state z_b . This comparison captures the stability-growth tradeoff of quantity requirements.

The traditional rationale for quantity requirements is that individual banks do not take into account the impact of their own leverage decisions on the vulnerability of the system as a whole. Although my model is ready to be extended to incorporate the general equilibrium effects, this rationale does not present in my model. Instead, the benefit of imposing leverage ratio requirements comes from two aspects. It forces the bank to keep more dry powder against its liquidity shock. There is an additional consideration due to the link between liquidity rationing and borrower runs. In particular, the leverage ratio requirement dampens the amplification mechanism and mitigates runs, as shown in the right panel of Figure 1.10.

1.6.3 A Commitment Tax on Revocation

My model also suggests that a commitment tax on cutting credit lines would be efficient to reduce vulnerability. The bank must pay a tax of τ on the amount of reduction in total credit due to credit line cuts.

From the bank's perspective, the tax works in the same way as the adjustment cost in discouraging credit line cuts. Now the marginal cost consists of two component,

$$\frac{\partial E'}{\partial \Phi'} = \tau \mathbb{1}[\Delta < 0] + 2\mu|\Delta|. \quad (1.18)$$

When the tax rate is high enough, the linear cost from taxation alone can deter the bank from cutting credit lines. In Figure 1.11, I plot the policy functions with different levels

of taxes. The commitment tax is effective in dampening the amplification and controlling runs. In particular, a tax of 0.4% is sufficient to eliminate credit line runs.

Commitment taxes are designed to discourage undesirable activities, such as pollution externalities. Recently, [Cochrane \(2014\)](#) suggest a tax on debt to control excessive run-prone liabilities. In a similar spirit, my model proposes a tax on cutting credit lines to mitigate credit line runs. A new insight from the current framework is that the tax provides an efficient commitment device to the bank. Borrowers understand ahead of time that the bank is going to cut fewer credit lines because of the tax. Hence they refrain from preemptive drawdowns. It is also important to note that the tax is a countercyclical policy by design. It reduces run risk in bad times; but, unlike the quantity requirements, it does not curb credit growth in good times.

In practice, setting the tax rate is challenging. First, the tax rate should respond to general economic conditions and bank liquidity shocks promptly. Second, the tax rate should also depend on bank balance sheet information, which requires effective supervision and information disclosure. Therefore, policy makers have an advantage in designing the tax, and it is much more difficult to achieve the same goal by private contracting.

1.7 Concluding Remarks

I have developed a dynamic model that integrates credit line runs with liquidity rationing. I illustrated how introducing bank liquidity management into bank run models is important for characterizing banking instability. In particular, there is a new strategic complementarity between the liquidity-rationing bank and run-prone credit line borrowers, which leads to a feedback loop that amplifies underlying distress. I also demonstrated the lack of commitment problem behind the feedback effects and proposed a commitment tax on cutting credit line to control run risk.

As discussed in the empirical studies, such as [Jiménez, Lopez, and Saurina \(2009\)](#) and [Ippolito, Peydro, Polo, and Sette \(2015\)](#), the usage and availability of credit lines are affected by both bank characteristics and firm characteristics, and hence bank-firm level data is required to disentangle supply and demand effects. Unfortunately, the call report data does not allow me to include other borrower heterogeneities besides their drawdown decisions. Further investigations with a richer modeling of credit line borrowers would be important. It is also worth to explore general equilibrium effects in the framework. I have focused on one bank, but the model is ready to be extended to study the banking system. Finally, my framework can be used to address other dynamic issues with strategic complementarities, such as fund redemption restrictions and partial defaults on sovereign debt.

REFERENCES

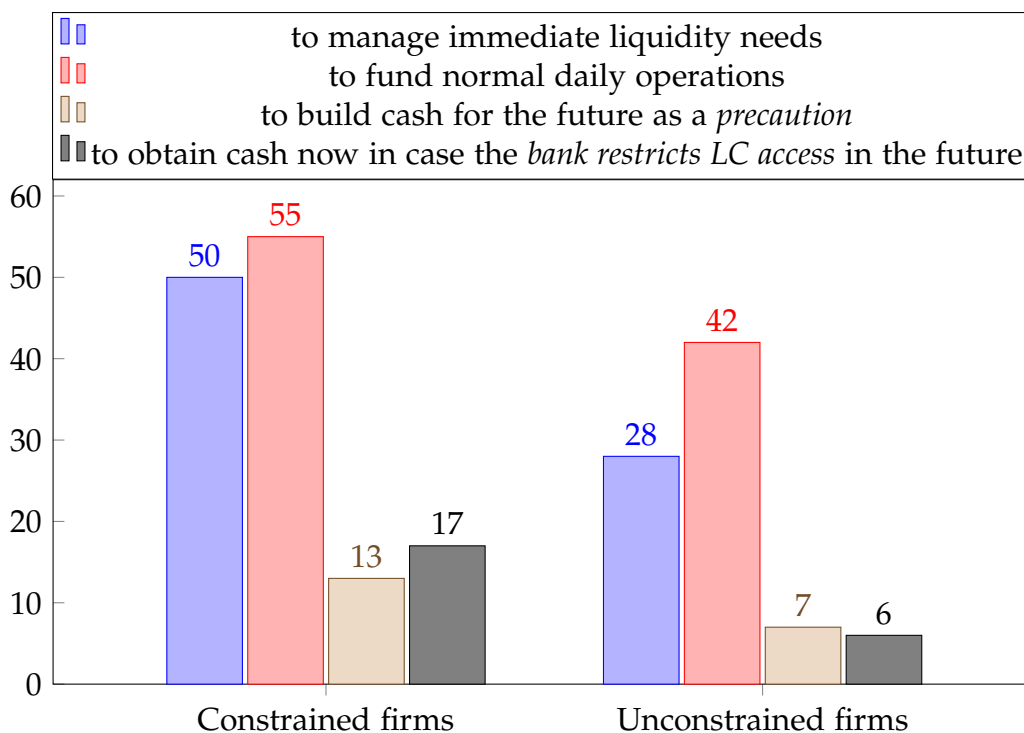
- ACHARYA, V., H. ALMEIDA, F. IPPOLITO, AND A. PEREZ (2014a): "Credit lines as monitored liquidity insurance: Theory and evidence," *Journal of Financial Economics*, 112(3), 287–319.
- (2016): "A new perspective on bank-dependency: The liquidity insurance channel," *Working Paper*.
- ACHARYA, V. V., H. ALMEIDA, F. IPPOLITO, AND A. PEREZ (2014b): "Bank lines of credit as contingent liquidity: A study of covenant violations and their implications," *ECB Working Paper*.
- ACHARYA, V. V., AND N. MORA (2015): "A crisis of banks as liquidity providers," *The Journal of Finance*, 70(1), 1–43.
- ALLEN, F., AND D. GALE (1998): "Optimal financial crises," *The Journal of Finance*, 53(4), 1245–1284.
- BENHABIB, J., J. MIAO, AND P. WANG (2014): "Chaotic banking crises and banking regulations," .
- BOND, P., AND A. S. RAI (2009): "Borrower runs," *Journal of development Economics*, 88(2), 185–191.
- BOOT, A. W., S. I. GREENBAUM, AND A. V. THAKOR (1993): "Reputation and discretion in financial contracting," *The American Economic Review*, pp. 1165–1183.
- BRYANT, J. (1980): "A model of reserves, bank runs, and deposit insurance," *Journal of banking & finance*, 4(4), 335–344.
- CAMPELLO, M., J. R. GRAHAM, AND C. R. HARVEY (2010): "The real effects of financial constraints: Evidence from a financial crisis," *Journal of Financial Economics*, 97(3), 470–487.
- CHADERINA, M., AND A. TENGULOV (2015): "Discretion and Systemic Risk in Credit Line Contracts: Theory and Evidence," *Available at SSRN* 2617223.
- CHENG, H., AND K. MILBRADT (2012): "The hazards of debt: Rollover freezes, incentives, and bailouts," *Review of Financial Studies*, 25(4), 1070–1110.
- CIPRIANI, M., A. MARTIN, P. E. MCCABE, AND B. M. PARIGI (2014): "Gates, fees, and preemptive runs," .
- COCHRANE, J. H. (2014): "Toward a Run-free Financial System," *Across the Great Divide: New Perspectives on the Financial Crisis*, p. 197.
- COOPER, R., AND T. W. ROSS (1998): "Bank runs: Liquidity costs and investment distortions," *Journal of Monetary Economics*, 41(1), 27–38.

- CORBAE, D., AND P. D'ERASMO (2014): "Capital requirements in a quantitative model of banking industry dynamics," .
- DIAMOND, D. W., AND P. H. DYBVIK (1983): "Bank runs, deposit insurance, and liquidity," *Journal of Political Economy*, pp. 401–419.
- EGAN, M., A. HORTAÇSU, AND G. MATVOS (2015): "Deposit competition and financial fragility: Evidence from the us banking sector," *Fama-Miller Working Paper*, pp. 15–33.
- ENGINEER, M. (1989): "Bank runs and the suspension of deposit convertibility," *Journal of monetary Economics*, 24(3), 443–454.
- ENNIS, H. M., AND T. KEISTER (2009): "Bank runs and institutions: The perils of intervention," *The American Economic Review*, 99(4), 1588–1607.
- (2010): "Banking panics and policy responses," *Journal of Monetary Economics*, 57(4), 404–419.
- GERTLER, M., AND N. KIYOTAKI (2015): "Banking, liquidity, and bank runs in an infinite horizon economy," *The American Economic Review*, 105(7), 2011–2043.
- HE, Z., AND W. XIONG (2012): "Dynamic debt runs," *Review of Financial Studies*, 25(6), 1799–1843.
- HOLMSTROM, B. R., AND J. TIROLE (1998): "Private and Public Supply of Liquidity," *Journal of Political Economy*, 106(1), 1–40.
- IPPOLITO, F., J.-L. PEYDRO, A. POLO, AND E. SETTE (2015): "Double Bank Runs and Liquidity Risk Management," *Journal of Financial Economics (JFE)*, *Forthcoming*.
- IVASHINA, V., AND D. SCHARFSTEIN (2010): "Bank lending during the financial crisis of 2008," *Journal of Financial economics*, 97(3), 319–338.
- JIMÉNEZ, G., J. A. LOPEZ, AND J. SAURINA (2009): "Empirical analysis of corporate credit lines," *Review of Financial Studies*, 22(12), 5069–5098.
- KASHYAP, A. K., R. RAJAN, AND J. C. STEIN (2002): "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking," *The Journal of Finance*, pp. 33–73.
- LIU, X. (2016): "Interbank Market Freezes and Creditor Runs," *Review of Financial Studies*, p. hhw017.
- MANKART, J., A. MICHAELIDES, AND S. PAGRATIS (2014): "A Dynamic Model of Banking with Uninsurable Risks and Regulatory Constraints," .
- MIAO, J., AND P. WANG (2015): "Banking bubbles and financial crises," *Journal of Economic Theory*, 157, 763–792.

- PECK, J., AND K. SHELL (2003): "Equilibrium bank runs," *Journal of political Economy*, 111(1), 103–123.
- ROBERTS, M. R., AND A. SUFI (2009): "Control rights and capital structure: An empirical investigation," *The Journal of Finance*, 64(4), 1657–1695.
- ROCHET, J.-C., AND X. VIVES (2004): "Coordination failures and the lender of last resort: was Bagehot right after all?," *Journal of the European Economic Association*, 2(6), 1116–1147.
- SCHROTH, E., G. A. SUAREZ, AND L. A. TAYLOR (2014): "Dynamic debt runs and financial fragility: Evidence from the 2007 ABCP crisis," *Journal of Financial Economics*, 112(2), 164–189.
- SUFI, A. (2009): "Bank lines of credit in corporate finance: An empirical analysis," *Review of Financial Studies*, 22(3), 1057–1088.
- VIVES, X. (2014): "Strategic complementarity, fragility, and regulation," *Review of Financial Studies*, 27(12), 3547–3592.
- ZENG, Y. (2016): "A Dynamic Theory of Mutual Fund Runs and Liquidity Management," .

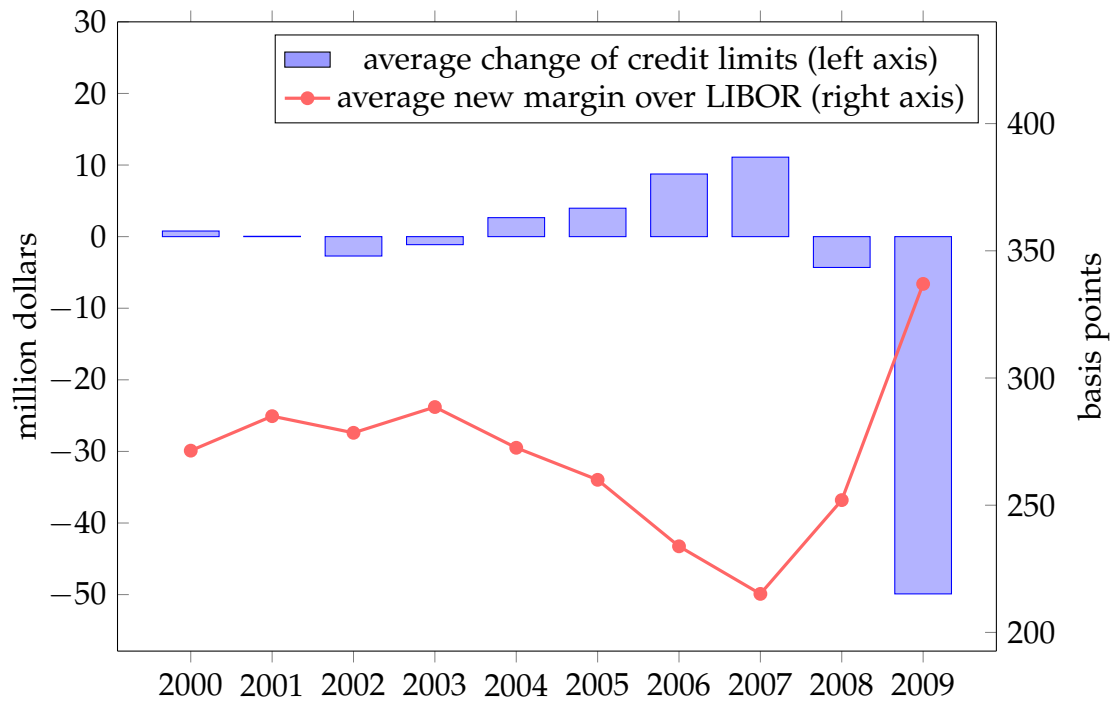
Figures

Figure 1.1: Reasons to draw cash from credit lines (%)



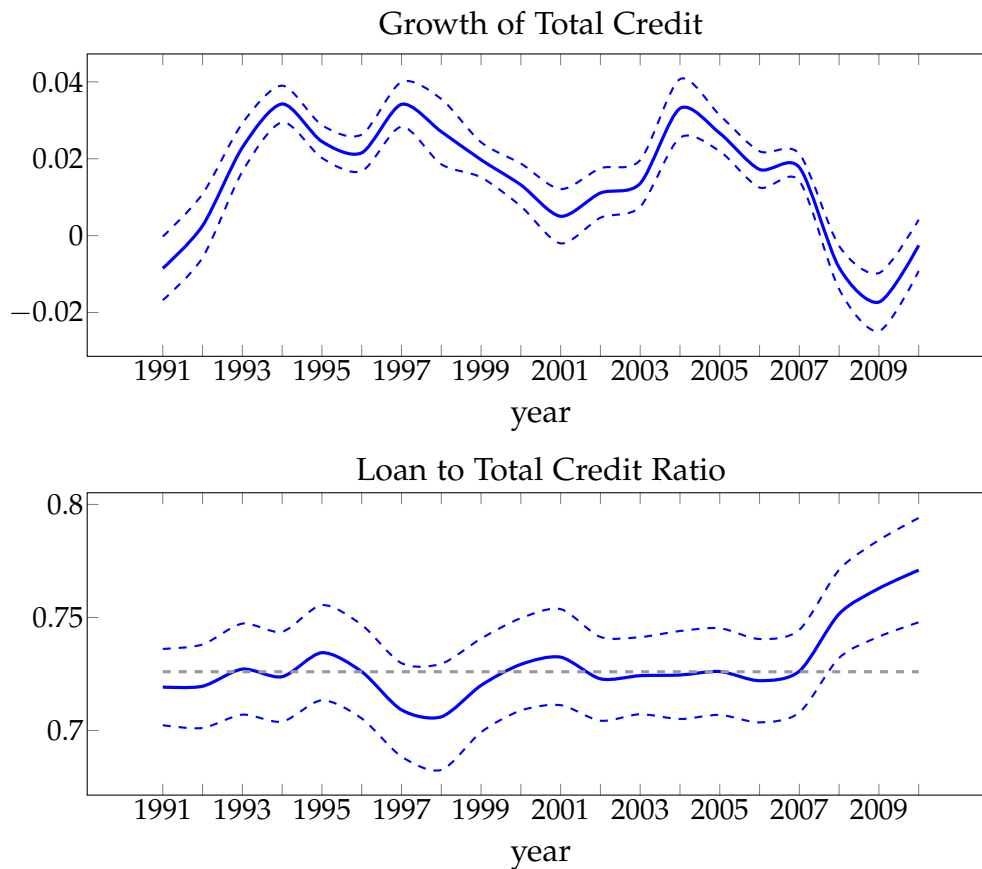
Note: The statistics are taken from Table 8 of [Campello, Graham, and Harvey \(2010\)](#), which is based on a survey conducted by the CFO magazine in 2008Q4. A particular question in the survey asks the reasons to draw cash from existing credit lines. Each column represents the fraction of surveyed firms that report the corresponding reason. The respondents can select all reasons that apply. The figure provides statistics separately for financially constrained firms and unconstrained firms. Whether a firm is financially constrained or not is also based on a survey question asking the respondents to choose from: not constrained, somewhat constrained, or heavily constrained. About a quarter of the 569 U.S. firms report that they are heavily constrained.

Figure 1.2: Amendments from Dealscan dataset



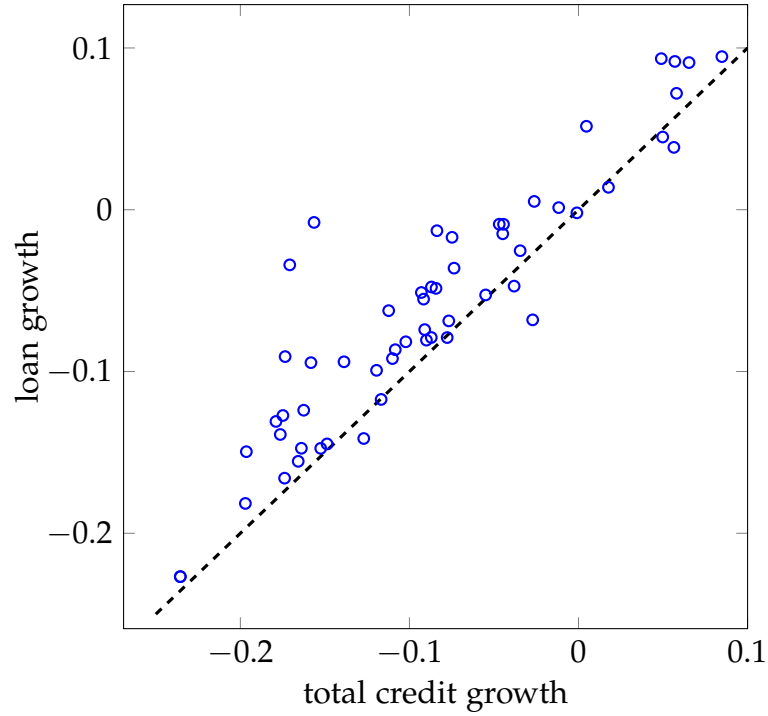
Note: The figure shows the magnitude of the two most common modifications of loan amendments in the Dealscan database from Thomson Reuters: credit limit reductions and margin increases. Average change of credit limits is calculated as the average of all loan amendments. Average new margin over LIBOR is the average of all amendments with a new margin that is based on LIBOR.

Figure 1.3: Total credit growth and loan to total credit ratio



Note: The figure plots the average quarterly growth rate of total credit (upper panel) and the average loan to total credit ratio (lower panel) using data from the Consolidated Report of Condition and Income. Total credit is defined as the sum of loans and unused credit lines. Dashed blue lines correspond to the 95% confidence interval for the averages. The gray dashed line in the lower panel represents the average loan to total credit ratio between 1991 and 2008.

Figure 1.4: Growth of loans and total credit during the Great Recession



Note: This figure plots a scatter plot of total credit growth against loan growth from 2008Q3 to 2009Q3 for banks with assets more than ten billions using data from the Consolidated Report of Condition and Income. Dashed line is the 45° line.

Figure 1.5: Timeline

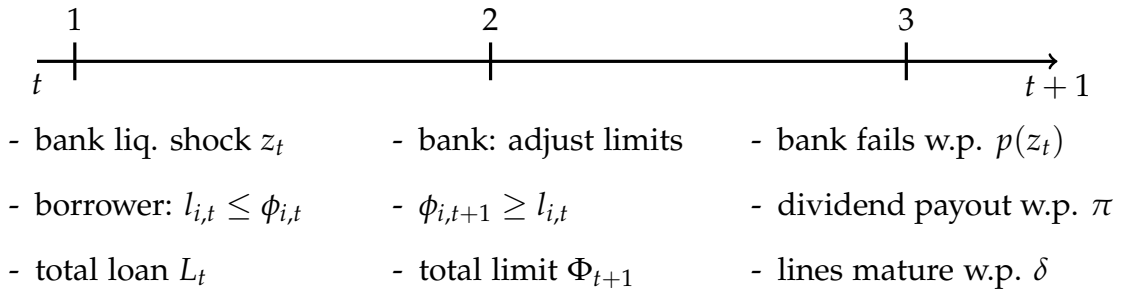
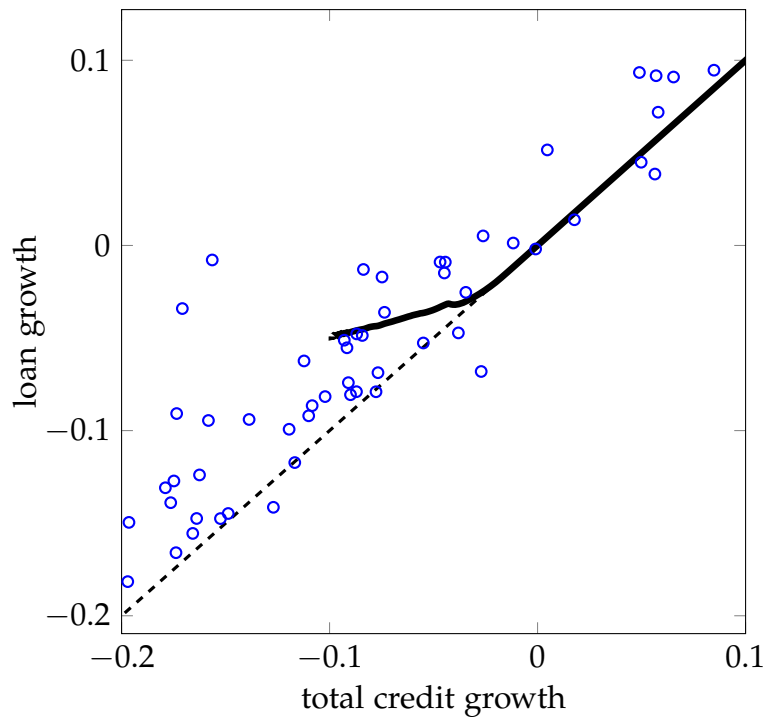
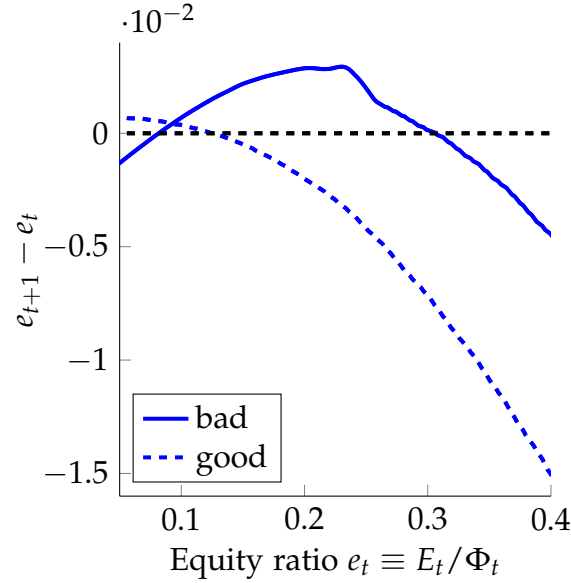


Figure 1.6: Growth of loans and total credit



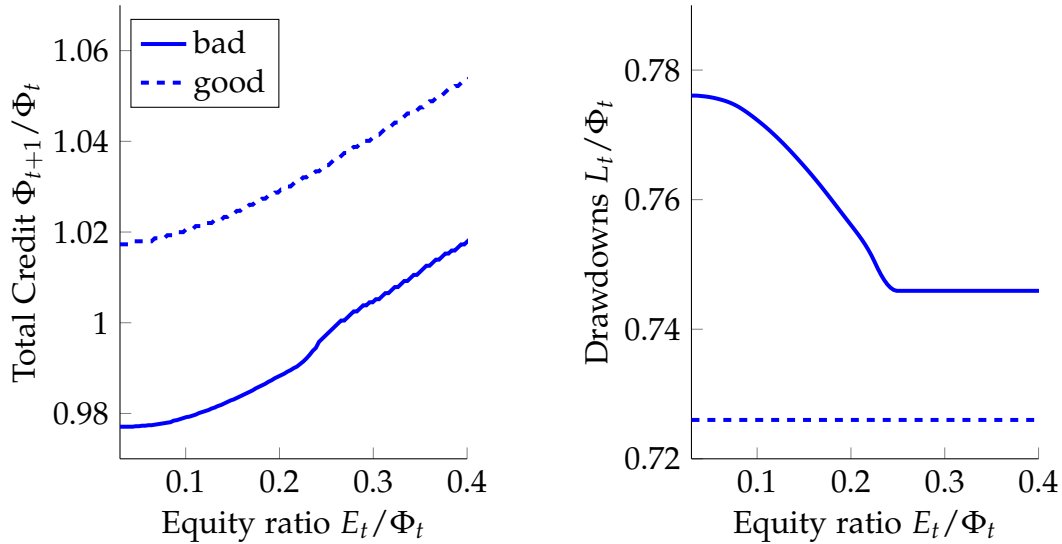
Note: This figure compares the relationship between total credit growth and loan growth in the model versus that in the data. This blue circles plots total credit growth against loan growth from 2008Q3 to 2009Q3 for banks with assets more than ten billions using data from the Consolidated Report of Condition and Income, which are the same as in Figure 1.4. Dashed line is the 45°. For comparison, I simulate a panel of banks with different leverage ratios and compute the growth rate in the first four quarters. The banks are assumed to be staying in the good state for a long time and hit by a negative liquidity shock for more than four consecutive quarters. The black solid line plots total credit growth against loan growth using simulated data.

Figure 1.7: Law of motion of equity ratio



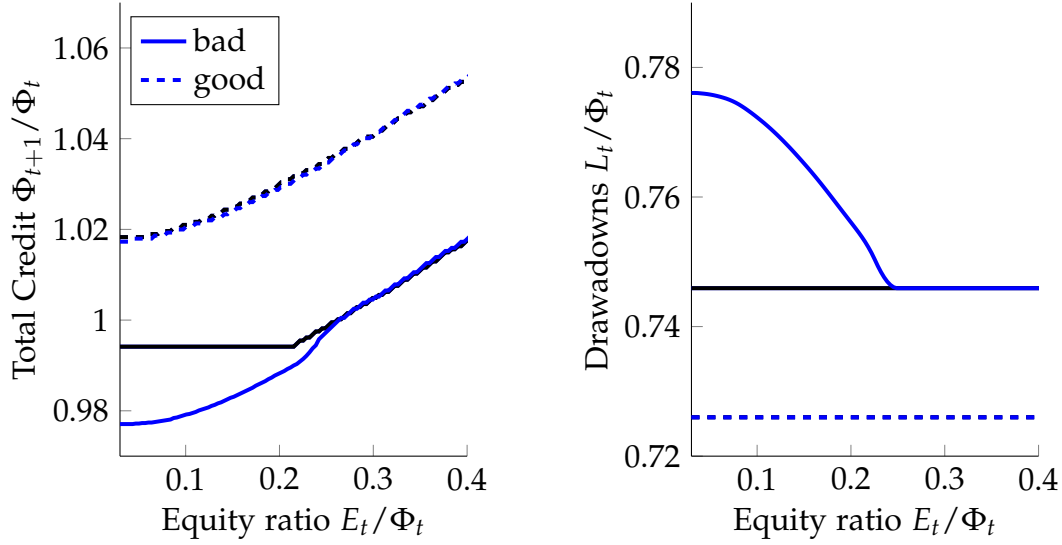
Note: This figure reports the law of motion of the state variable equity ratio (E/Φ). The solid line represents the change of equity ratio as a function of current equity ratio when the bank is hit by the adverse shock; the dashed line represents the change of equity ratio in the good state.

Figure 1.8: Policy functions



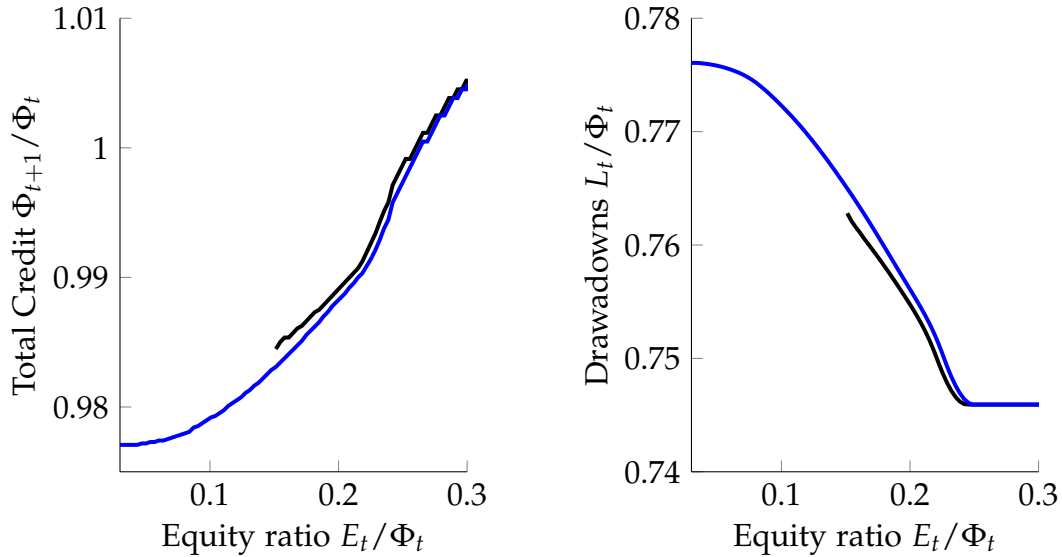
Note: The left panel reports the credit line revocation policy as a function of equity ratio. The right panel plots the total amount of drawdowns as a function of equity ratio. In both panels, the solid lines correspond to the bad state, and the dashed lines correspond to the good state.

Figure 1.9: Policy functions with and without full commitment



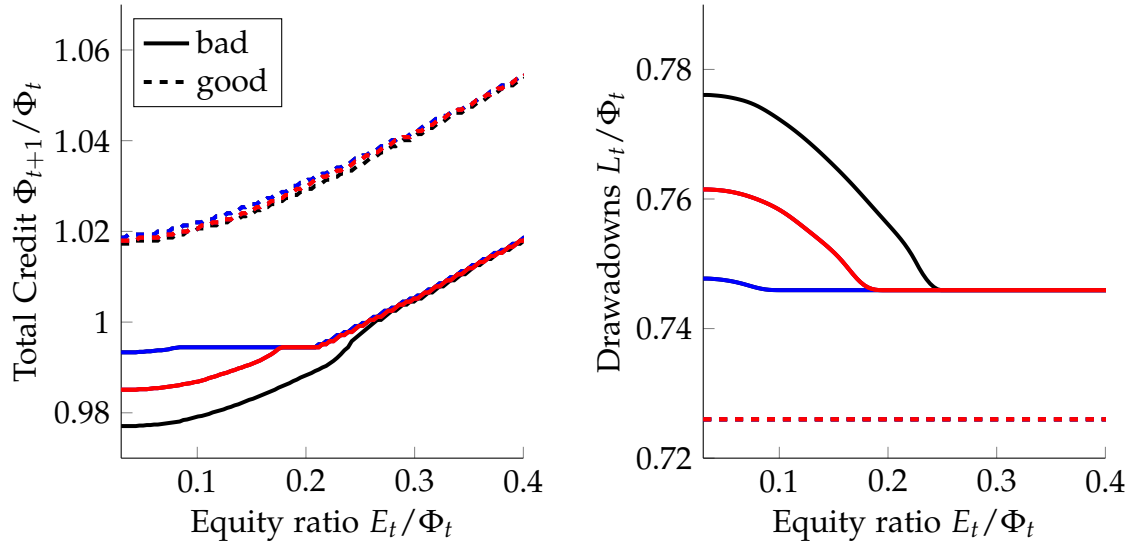
Note: This figure compares the policy functions of the benchmark model (blue lines) and those of the model with commitment (black lines). The left panel reports the credit line revocation policy as a function of equity ratio. The right panel plots the total amount of drawdowns as a function of equity ratio. In both panels, the solid lines correspond to the bad state, and the dashed lines correspond to the good state.

Figure 1.10: Leverage ratio requirements



Note: This figure reports the counterfactual with leverage ratio requirements. Only the policy functions in the bad state is plotted to make the difference visible. The left panel reports the credit line revocation policy as a function of equity ratio. The right panel plots the total amount of drawdowns as a function of equity ratio. The black lines plot the policy functions if the leverage ratio requirement is imposed. The blue lines plot the policy functions of the benchmark model.

Figure 1.11: Tax on cutting credit lines



Note: $\tau = 0, 0.2\%, 0.4\%$ This figure reports the counterfactual with commitment taxes. The left panel reports the credit line revocation policy as a function of equity ratio. The right panel plots the total amount of drawdowns as a function of equity ratio. In both panels, the solid lines correspond to the bad state, and the dashed lines correspond to the good state. The black lines represent the policies under zero tax. The red lines represent the policies under 0.2% tax. The blue lines represent the policies under 0.4% tax.

Tables

Table 1.1: **Borrower's flow payoffs and continuation values**

Liq. Shock	Drawdown	Flow Payoff	Continuation Value
Y	Y	$R - r^\phi - r^l$	$(1 - p)(1 - \eta\delta)\hat{W}' + p(1 - \eta\delta)W_0$
Y	N	0	W_L
N	Y	$R - r^\phi - (r^l - \kappa_i)$	$(1 - p)(1 - \eta\delta)\hat{W}' + p(1 - \eta\delta)W_0$
N	N	$R - r^\phi$	$(1 - p)(1 - \eta\delta - \eta q)\hat{W}' + p(1 - \eta)W_0$

Note: The first column displays whether the borrower needs liquidity (Y) or not (N). The second column displays whether the borrower decides to drawdown (Y) or not (N). Flow payoffs and continuation values contingent on the liquidity shock and drawdown decisions are provided in the last two columns. $\hat{W}'$ represents the expected value at the end of the period if borrower i has access to credit line in the next period. W_0 denotes the expected value of a remaining borrower i after bank failure. $r^l - \kappa_i$ is the effective cost of preemptive drawdowns, and q stands for the endogenous probability of an undrawn credit line being cut.

Table 1.2: **Model calibration**

Parameter		Value	Targets
Transition probability	π_{gg}	0.96	NBER data
Transition probability	π_{bb}	0.81	NBER data
Funding cost (annual)	$r(z_g)$	0.03	Avg. LIBOR pre-08Q3
Funding cost (annual)	$r(z_b)$	0.01	Avg. LIBOR post-08Q3
Usage fee premium (annual)	r_x	224	Avg. usage fee premium (bps)
Maintenance fee (annual)	r^ϕ	37	Avg. maintenance fee (bps)
Bank failure prob. (annual)	$p(z_b)$	0.02	Avg. CDS spread post-08Q3
Liquidity shock	$\Lambda(z_g)$	0.725	Loan-to-credit ratio pre-08Q3
Liquidity shock	$\Lambda(z_b)$	0.74	Loan-to-credit ratio
Discount rate	β	0.98	Conventional value
Prob. of Maturing	δ	0.005	Dealscan
Bank exit rate	π	0.05	Gertler and Kiyotaki (2015)
Min. return on drawdowns	$\underline{\kappa}$	-0.055	Loan to credit ratio post-08Q3
Adj. cost	μ	0.156	Total credit growth
Non-interest Expenses	γ	0.024	Leverage pre-08Q3

Note: This table reports the calibrated parameters. The first set of parameters is calibrated to the data directly. Transition probabilities is taken from the NBER recession data. The annual funding costs are matched to the average LIBOR rates. The fees and maturing rate are calibrated to the average fees in the Dealscan dataset. Bank failure probability is set to the default probability implied by the average CDS spread. The probabilities of borrowers' liquidity shock is calibrated to the average loan-to-credit ratio as in Figure 1.3. $\Lambda(z_b)$ is computed as the average of loan-to-credit ratio after 2008Q3 of banks with liquidity ratios above the median. Discount rate is set to be a conventional value. Bank exit rate is taken from [Gertler and Kiyotaki \(2015\)](#) directly. The last three parameters are set jointly to match a set of empirical moments detailed in Table 1.3. The procedure is described in the text and in Appendix 1.B.

Table 1.3: **Calibration targets**

Empirical Targets		
Statistic (%)	Data	Model
Avg. total credit growth post-2008Q3	−1.75	−1.79
Avg. bank equity ratio pre-2008Q3	11.2	11.4
Avg. Loan-to-credit ratio post-2008Q3	76.3	76.3
Avg. total credit growth pre-2008Q3	2.06	2.03
Avg. bank equity ratio post-2008Q3	14.2	14.0

Note: This table reports the empirical moments used in calibration in the second column and the model-generated moments under the calibrated parameters in the last column. Total credit is defined as the sum of loans and unused credit lines. Equity ratio is defined as the ratio of equity to the total credit.

Table 1.4: **Counterfactuals**

	Benchmark	Commitment
Avg. Total Credit Growth in z_b (%)	−1.79	−0.60
Avg. Total Credit Growth in z_g (%)	2.03	2.04
Bank value	1.18	+0.1%
Borrower value	0.26	+0.4%

Note: This table compares the average credit growth rates and welfare in the benchmark model (the second column) and those in the counterfactual in which the bank has full commitment (the third column). The borrower value is computed as the expected value under the stationary distribution of the state variables. The bank value is calculated as the value of a bank with \$1 equity and at the good steady state.

APPENDIX

1.A Proofs

Proof of Proposition 2. Given Equations 1.8, 1.11, and 1.12, the total drawdowns L at the equilibrium given the bank's strategy can be solved from the following equations.

$$\kappa^* = r^l - \beta \left[(1-p)\eta \hat{W}' \frac{\Delta^-}{\Phi - L} + p\eta(1-\delta)W_0 \right] \quad (1.19)$$

$$L = \hat{L} \equiv [\Lambda + (1-\Lambda)(1-\Omega(\kappa^*))] \Phi, \quad (1.20)$$

where $\Delta^- \equiv -\min\{\Delta, 0\}$. Since the equilibrium is stable, we have $\frac{\partial \hat{L}}{\partial \kappa^*} \frac{\partial \kappa^*}{\partial L} < 1$. Thus, the equilibrium drawdown L increases in Δ^- .

Proof of Lemma 1. Assumption 1 assumes that both $c(D, \Phi)$ and $f(\Phi', \Phi)$ are homogeneous of degree one. Therefore, by an abuse of notation we have $c(D, \Phi) = \Phi c(D/\Phi, 1) = \Phi c(D/\Phi)$ and $f(\Phi', \Phi) = \Phi f(\Phi'/\Phi, 1) = \Phi f(\Phi'/\Phi)$. Also, drawdowns L is linear in Φ as in Equation 1.12.

Dividing both sides of the original bellman equation 1.5 by Φ , we have

$$v(e, z) = \max_{\phi' \geq (1-\delta)l} \beta \pi (1-p) \hat{e} + \beta (1-\pi) (1-p) \phi' \mathbb{E}[v(\hat{e}/\phi', z')] \quad (1.21)$$

$$\text{s.t. } \hat{e} = e + r^\phi + r^l l - (r+p)(l-e) - c(l-e) - f(\phi')$$

where $e \equiv \frac{E}{\Phi}$, $\hat{e} \equiv \frac{E'}{\Phi}$, $\phi' \equiv \frac{\Phi'}{\Phi}$, and $l \equiv \frac{L}{\Phi}$.

Proof of Proposition 3. I first show that the normalized value function $v(e, z)$ is unique, increasing, and concave in e given the borrower strategy. Then I consider how bank optimal policy depends on drawdowns.

Lemma 2 Assume that there is an upper bound $\bar{\phi}$ of total credit growth ϕ' . If $\bar{\phi}\beta(1-\pi) \leq 1$, the bank value function $v(e, z)$ is unique given borrower strategy σ .

Proof. Apply Contraction Mapping Theorem to the bank problem 1.21, and check the Blackwell's sufficient conditions.

First, assume that there is a function $u(e, z) \geq v(e, z)$ for all $v(e, z)$. Let ϕ'_v denote the optimal policy with V .

$$\begin{aligned} T[v(e, z)] &= \max_{\phi'} \beta\pi(1-p)\hat{e} + \beta(1-\pi)(1-p)\phi' \mathbb{E}[v(\hat{e}/\phi', z')] \\ &\leq \beta\pi(1-p)\hat{e}(\phi'_v) + \beta(1-\pi)(1-p)\phi'_v \mathbb{E}[u(\hat{e}(\phi'_v)/\phi'_v, z')] \\ &\leq \max_{\phi'} \beta\pi(1-p)\hat{e} + \beta(1-\pi)(1-p)\phi' \mathbb{E}[u(\hat{e}/\phi', z')] = T[u(e, z)]. \end{aligned}$$

Then we check the discounting condition given $\bar{\phi}\beta(1-\pi) \leq 1$.

$$\begin{aligned} T[v(e, z) + a] &= \max_{\phi'} \beta\pi(1-p)\hat{e} + \beta(1-\pi)(1-p)\phi' \mathbb{E}[v(\hat{e}/\phi', z') + a] \\ &\leq \max_{\phi'} \beta\pi(1-p)\hat{e} + \beta(1-\pi)(1-p)\phi' \mathbb{E}[v(\hat{e}/\phi', z')] + \phi'_{v+a}\beta(1-\pi)(1-p)a \\ &= T[v(e, \phi, z)] + \phi'_{v+a}\beta(1-\pi)(1-p)a. \end{aligned}$$

The constraint that $\bar{\phi}\beta(1-\pi) \leq 1$ can be relaxed with a variant of the Blackwell's sufficient conditions for unbounded functions. ■

Lemma 3 If $\partial\Pi(E, \Phi, z)/\partial E > -1$, bank value function $v(e, z)$ is increasing in e .

Proof. Suppose $e_1 < e_2$ and let ϕ'_1 denote the optimal choice associated with e_1 , need to show that $T[v(e_1, z)] < T[v(e_2, z)]$ if $v(e, z)$ is increasing in e . If $\partial\Pi(E, \Phi, z)/\partial E > -1$, we have that $\hat{e}$ increases in e for the same ϕ' . Thus,

$$\begin{aligned} T[v(e_2, z)] &= \max_{\phi'} \beta\pi(1-p)\hat{e} + \beta(1-\pi)(1-p)\phi' \mathbb{E}[v(\hat{e}/\phi', z')] \\ &\geq \beta\pi(1-p)\hat{e}(\phi' = \phi'_1) + \beta(1-\pi)(1-p)\phi'_1 \mathbb{E}[v(\hat{e}(\phi' = \phi'_1)/\phi'_1, z')] \\ &\geq T[v(e_1, \phi, z)]. \end{aligned}$$

■

Lemma 4 If $f(\Phi', \Phi)$ is concave in Φ' and $\Pi(E, \Phi, z)$ is concave in E , bank value function $v(e, z)$ is concave in e .

Proof. Need to show that the bellman equation maps concave functions into concave functions. Assume that there are two equity ratios $e_1 < e_2$ and the corresponding optimal policies are ϕ'_1 and ϕ'_2 . Denote $e_0 = \eta e_1 + (1-\eta)e_2$ and $\phi'_0 = \eta\phi'_1 + (1-\eta)\phi'_2$ for $\eta \in (0, 1)$, we want to show that if v is concave and increasing in e ,

$$T[v(e_0, z)] \geq \eta T[v(e_1, z)] + (1-\eta)T[v(e_2, z)]. \quad (1.22)$$

The concavity of $f(\Phi', \Phi)$ guarantees that $\hat{e}$ is concave in ϕ' . At the same time since $\Pi(E, \Phi, z)$ is concave in E , $\hat{e}$ is also concave in e . Therefore, as e and ϕ' are separated,

$\hat{e}(e_0, \phi'_0) \geq \eta \hat{e}(e_1, \phi'_1) + (1 - \eta) \hat{e}(e_2, \phi'_2)$. Next, since v is increasing in e ,

$$\begin{aligned} \frac{\hat{e}(e_0, \phi'_0)}{\phi'_0} &\geq \frac{\eta \phi'_1}{\phi'_0} \frac{\hat{e}(e_1, \phi'_1)}{\phi'_1} + \frac{(1 - \eta) \phi'_2}{\phi'_0} \frac{\hat{e}(e_2, \phi'_2)}{\phi'_2} \\ \Rightarrow v\left(\frac{\hat{e}(e_0, \phi'_0)}{\phi'_0}, z'\right) &\geq \frac{\eta \phi'_1}{\phi'_0} v\left(\frac{\hat{e}(e_1, \phi'_1)}{\phi'_1}, z'\right) + \frac{(1 - \eta) \phi'_2}{\phi'_0} v\left(\frac{\hat{e}(e_2, \phi'_2)}{\phi'_2}, z'\right). \end{aligned}$$

The condition that $\Pi(E, \Phi, z)$ is concave in E is satisfied when $L(E, \Phi, z)$ is not too convex in E . ■

The above Lemmas also guarantee that the value function is continuous and differentiable. Now we are ready to prove the second part of Proposition 3. I only consider the interior case when there is a credit line run and the bank withdraws credit lines. I also assume that l is smooth so that the value function is differentiable. The first-order condition is thus given by,

$$-\left[\pi + (1 - \pi) \frac{\partial \mathbb{E}[v(\hat{e}/\phi', z')]}{\partial \hat{e}}\right] f'(\phi') + (1 - \pi) \left[\mathbb{E}[v(\hat{e}/\phi', z')] + \phi' \frac{\partial \mathbb{E}[v(\hat{e}/\phi', z')]}{\partial \phi'}\right] = 0.$$

Let v' denote the first-order derivative of $v(\hat{e}/\phi', z')$ with respect to its first argument. The condition can be rewritten as

$$-(\pi \phi' + (1 - \pi) \mathbb{E}[v']) f'(\phi') + (1 - \pi) (\phi' \mathbb{E}[v] - \hat{e} \mathbb{E}[v']) = 0.$$

Hence we have $\hat{e} + f'(\phi') > 0$.

To derive the effect of an increase in drawdowns, I total differentiate the first-order condition at the optimal choice, which gives us,

$$\frac{d\phi'}{dl} = - \frac{(1 - \pi) \left[-\frac{\partial^2 \mathbb{E}[v']}{\partial (\hat{e})^2} f'(\phi') + \frac{\partial \mathbb{E}[v']}{\partial \hat{e}} + \phi' \frac{\partial^2 \mathbb{E}[v']}{\partial \phi' \partial \hat{e}} \right] [r^l - r - p - c'(l - e)]}{-\left[\pi + (1 - \pi) \frac{\partial \mathbb{E}[v']}{\partial \hat{e}}\right] f''(\phi') + (1 - \pi) \left[\frac{\partial^2 \mathbb{E}[v']}{\partial (\hat{e})^2} (f'(\phi'))^2 - \frac{\partial^2 \mathbb{E}[v']}{\partial \phi' \partial \hat{e}} f'(\phi') + \phi' \frac{\partial^2 \mathbb{E}[v']}{\partial (\phi')^2} + \frac{\partial \mathbb{E}[v']}{\partial \phi'} \right]}.$$

First, note that the denominator is negative, since the bank solves a maximization problem. Second, the numerator can be simplified to

$$-(1 - \pi) \frac{\partial^2 \mathbb{E}[v']}{\partial (\hat{e})^2} \frac{f'(\phi') + \hat{e}}{\phi'} [r^l - r - p - c'(l - e)].$$

Since v is concave, when the profit margin $r^l - r - p$ is smaller than the marginal non-interest cost, i.e. $r^l - r - p - c'(l - e) < 0$, the numerator is negative as well. Hence bank choice ϕ' decreases in drawdowns l .

Proof of Proposition 4. Below I show that the bank cuts fewer credit lines when it can commit. The other results follow. I focus on the case when the bank chooses to cut credit lines in the model with commitment in the bad state. Otherwise, the Proposition

is trivial. The first-order condition in the model with commitment is as follows:

$$\left[\pi + (1 - \pi) \frac{\partial \mathbb{E}[v_C(e', z')]}{\phi' \partial e'} \right] \left[-f'(\phi') + \frac{\partial \hat{e}}{\partial l} \frac{\partial l}{\partial \phi'} \right] + (1 - \pi) \left[\mathbb{E}[v_C(e', z')] - e' \frac{\partial \mathbb{E}[v_C(e', z')]}{\partial e'} \right] = 0.$$

Since the value functions in the two models are different, it is hard to compare the two policy functions directly. Instead, I consider a continuous change in bank's commitment power. In particular, I assume that α fraction of borrowers believe that the bank is going to stick to the plan. The comparative statics with respect to α is consistent for all $\alpha \in (0, 1)$.

$$\left[\pi + (1 - \pi) \frac{\partial \mathbb{E}[v_\alpha(e', z')]}{\phi' \partial e'} \right] \left[-f'(\phi') + \alpha \frac{\partial \hat{e}}{\partial l} \frac{\partial l}{\partial \phi'} \right] + (1 - \pi) \left[\mathbb{E}[v_\alpha(e', z')] - e' \frac{\partial \mathbb{E}[v_\alpha(e', z')]}{\partial e'} \right] = 0.$$

When α increases, the optimal bank choice ϕ' increases. This can be shown by total differentiating the first-order condition. The numerator of $d\phi'/d\alpha$ is positive since $\frac{\partial \hat{e}}{\partial l} < 0$ and $\frac{\partial l}{\partial \phi'} < 0$. Given that the value functions are continuous, we can take α to the limits, and $\phi'_{\alpha=1} \geq \phi'_{\alpha=0}$ follows.

1.B Algorithm

This appendix describes the algorithms used to compute the equilibrium and calibrate the model. To calibrate the model, one needs to find the set of parameters such that model-generated moments fit a pre-determined set of data moments. I solve the model numerically and compute model-generated moments with simulated data.

Equilibrium. In order to solve the equilibrium under a given set of parameters, we need to discretize the state space (e, z) . The grid of equity ratio e contains 200 grid points with equal-space between 0 and 0.7. The exogenous state z has two grid points, high or low. Revocation choice is discretized over a range from -0.1 (which means credit growth of 10%) to 0.05. Larger grid size or wider domain would not affect the numerical results in any significant way. The algorithm is as follows:

1. Guess a pair of value functions of borrowers and the bank (V, W) .
2. Iterate on the policy functions
 - (a) Guess a policy function of credit line drawdowns σ . Solve the optimal revocation policy σ^B given the bank's value function and the borrowers' drawdown decisions σ .
 - (b) Given the revocation policy and borrowers' value function, solve for the optimal drawdown decisions $\hat{\sigma}$. Update the policy function by setting the new policy function as $\lambda \hat{\sigma} + (1 - \lambda) \sigma$ for some small positive constant λ .
 - (c) Go to bulletin point (a) until the policy functions converge.

3. Update the pair of value functions and iterate until the value functions converge.

Calibration. We pin down the parameters κ, μ, γ by simulated method of moments, which minimizes the distance between data moments and simulated moments. Let us call $\mathbf{m}$ the vector of moments computed from the actual data, and let us call $\mathbf{m}(v)$ the moments generated by the model with parameters v . The SMM procedure searches the set of parameters that minimizes the weighted deviations between the actual and simulated moments,

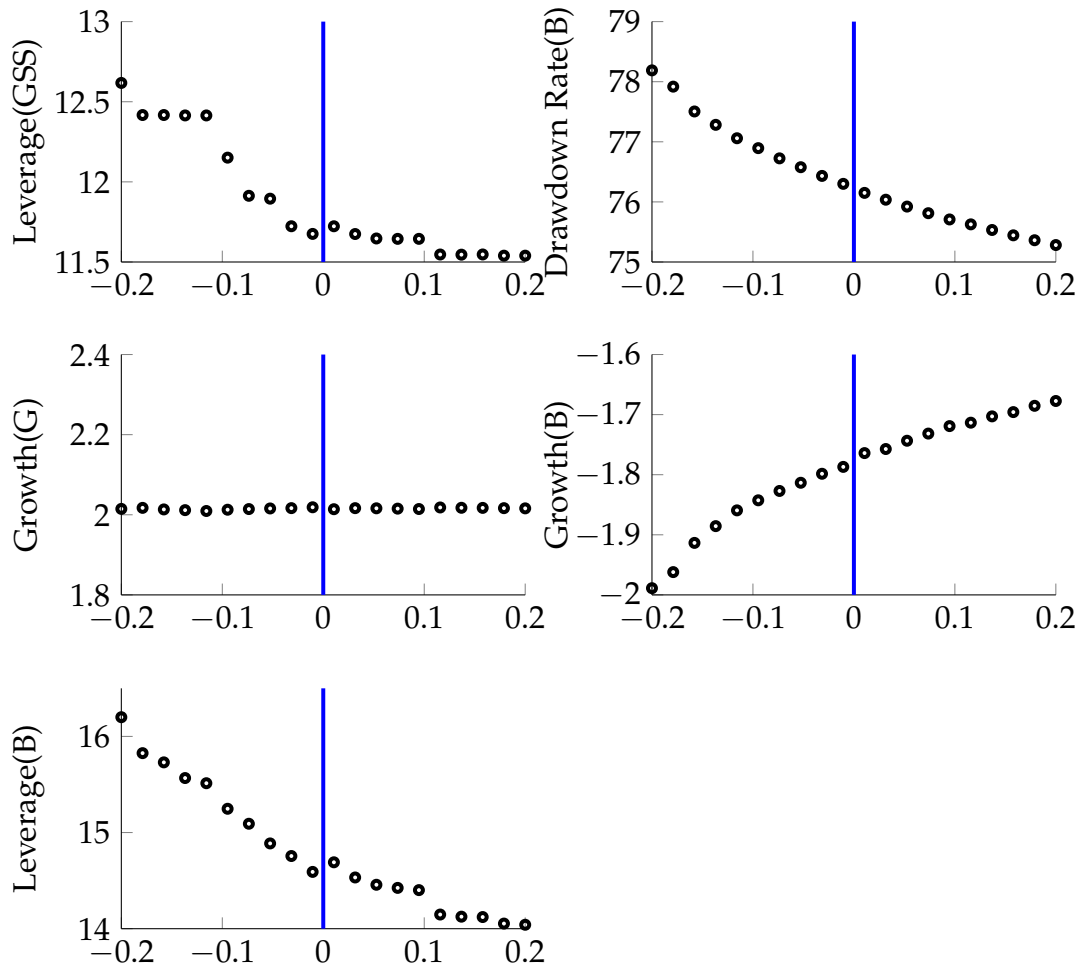
$$(\mathbf{m} - \hat{\mathbf{m}}(v))' \mathbf{W} (\mathbf{m} - \hat{\mathbf{m}}(v)) \quad (1.23)$$

The weight matrix $\mathbf{W}$ adjusts for the fact that some moments are more precisely estimated than others. It is computed as the inverse of the variance-covariance matrix of actual moments estimated by bootstrap with replacement on the actual data.

We use the CMA-ES algorithm to minimize the distance. The CMA-ES algorithm is a derivative-free evolutionary algorithm for non-linear optimization problems. In each round we draw n off-springs in the parameter space which are normally distributed around the mean inherited from last round. Then we compute the distance between simulated data with each off-spring and actual data as its fitness. With the fitness at hand, we rank the off-springs by their fitness, and update the mean by the weighted average of the upper half off-springs. We also update the variance-covariance matrix of multi-dimensional normal distribution that will be used in the next round to draw off-springs. Note that we use the fitness-weighted average of solved value function as the initial guess in the next round for efficiency. We iterate this procedure until the off-springs are distributed close enough to the mean.

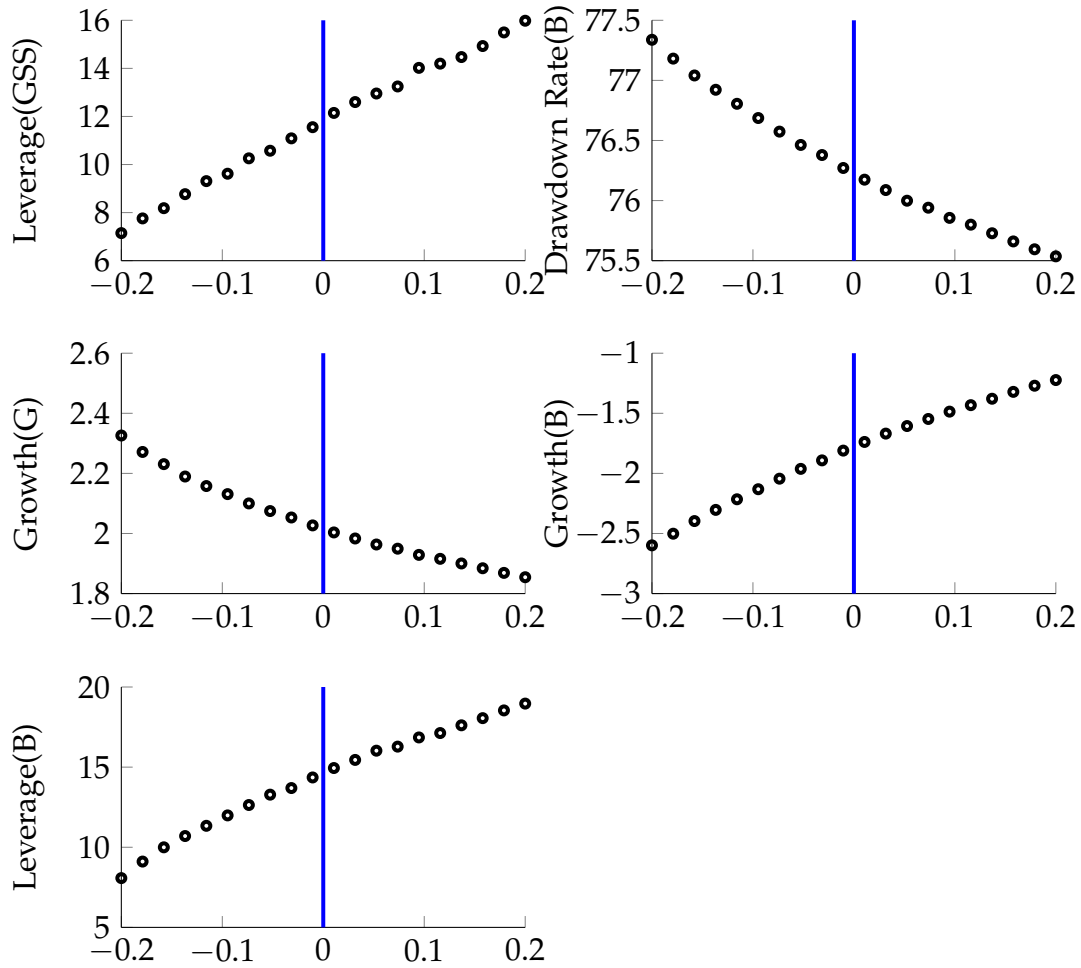
1.C Additional Figures

Figure 1.12: Sensitivity of moments to the drawdown cost



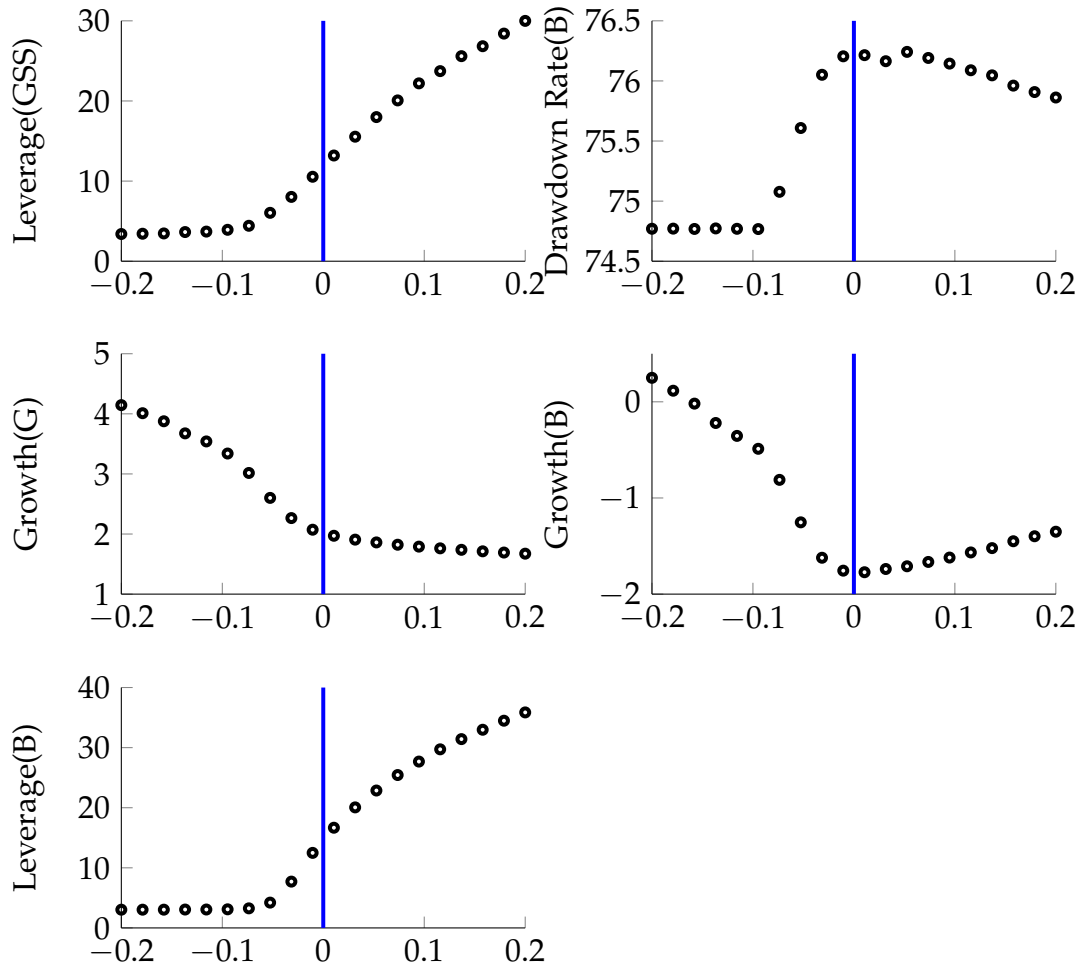
Note: This figure shows the sensitivity of moments to the drawdown cost κ . The adjustment cost varies from 80% to 120% of the calibrated value, which is normalized to zero on the x-axis.

Figure 1.13: Sensitivity of moments to the adjustment cost



Note: This figure shows the sensitivity of moments to the adjustment cost μ . The adjustment cost varies from 80% to 120% of the calibrated value, which is normalized to zero on the x-axis.

Figure 1.14: Sensitivity of moments to the non-interest expenses



Note: This figure shows the sensitivity of moments to the non-interest expenses γ . The adjustment cost varies from 80% to 120% of the calibrated value, which is normalized to zero on the x-axis.

Chapter 2

Aggregate Effects of Collateral Constraints¹

This chapter is co-authored with Sylvain Catherine (HEC Paris), Thomas Chaney (Sciences Po and CEPR), David Sraer (UC Berkeley, CEPR and NBER), and David Thesmar (MIT and CEPR)

There is an accumulating body of evidence showing the causal effect of financing frictions on firms' investment decisions *at the micro-level*.² While this literature safely rejects the null hypothesis that firms are unconstrained financially, it does not measure whether these constraints matter quantitatively. In this paper, we use a quantitative model that matches these findings to investigate the *aggregate effects* of financing frictions. We focus on a pervasive source of financing friction – collateral constraints. Our approach expands on the existing literature by (i) estimating our structural model using

¹This is a substantially revised version of our earlier paper with the same title. We are grateful to conference and seminar participants in Berkeley, Capri, Duke, HBS, Kellogg, NYU-Stern, Stanford, the LSE, the Chicago Fed, Zurich, WFA, the FED Board for their comments. We warmly thank Toni Whited for sharing her fortran code with us and for her insightful discussion at WFA. Sraer is grateful for financial support from the Fisher Center for Real Estate & Urban Economics. Thesmar is grateful to the Fondation Banque de France for its financial support. All errors are our own.

²See, among many others, [Lamont \(1997\)](#), [Rauh \(2006\)](#), [Chaney, Sraer, and Thesmar \(2012\)](#), [Blanchard, Lopez-de Silanes, and Shleifer \(1994\)](#) for the effect of financial frictions on investment and [Benmelech, Bergman, and Seru \(2010\)](#) or [Chodorow-Reich \(2013\)](#) for the effect of financial frictions on employment

well-identified firm-level evidence that collateral constraints causally affect investment and (ii) nesting this model in a general equilibrium framework with heterogeneous firms to study the aggregate effect of collateral constraints. Our estimated model shows that even in a developed country like the U.S., collateral constraints can have a large effect on welfare. Compared to a counterfactual economy without financing constraints, welfare in our constrained economy is lower by 9.4%, and output by 11%. Of this output loss, only a quarter can be attributed to lower aggregate TFP due to input misallocation.³ The remaining output loss is due to lower aggregate inputs, mostly capital. Thus, collateral constraints induce significant misallocation, but their impact on the aggregate capital stock is larger.

We estimate our structural model by targeting the sensitivity of investment to exogenous shocks to firms' real estate value. Starting with [Gan \(2007\)](#) and [Chaney, Sraer, and Thesmar \(2012\)](#), a large literature documents how corporate investment responds to real estate shocks and argues that such sensitivity is evidence of financing constraints, insofar that real estate shocks are shocks to debt capacity that are uncorrelated with investment opportunity. Relying on this insight, we use this sensitivity to identify the parameter governing financing constraints in our model. The existing literature that estimates similar models (e.g., [Hennessy and Whited \(2007\)](#)) typically targets capital structure decisions such as the average debt to capital ratio. However, this moment is driven by many forces (e.g., trade credit, inventory, unsecured debt capacity) that may not be all captured by the model. As a result, estimates of the parameters driving financial constraints will be influenced by these additional forces. In contrast to leverage, causal estimates coming from the reduced-form literature are in principal purely attributable to financing constraints. Targeting these reduced-form moments should lead to more reliable estimates of financing constraints parameters. We show that, in our data, targeting firms' leverage leads to *underestimating* the effect of financing constraints. The intuition is that the

³The costs of input misallocation is the focus of [Hsieh and Klenow \(2009\)](#), [Moll \(2014\)](#), [Midrigan and Xu \(2014\)](#).

sensitivity of investment to real estate value is relatively low in the data, indicating a relatively low pledgeability of capital. Leverage is, on the other hand, relatively large empirically, so that an estimation procedure that seeks to match leverage will assume that capital is easily collateralized. This makes financing constraints less binding. At the aggregate level, when targeting leverage, the estimated aggregate output loss is only half as large as when targeting the sensitivity of investment to real estate shocks.

We start by documenting how, on a panel of U.S. firms, corporate investment and leverage respond to shocks to real estate value. Repeating earlier analysis ([Chaney, Sraer, and Thesmar, 2012](#)) with slightly different specifications, we find that a \$1 increase in real estate value leads to a \$0.04 increase in investment and a \$0.04 increase in financial debt. While these estimates allow to comfortably reject the null that firms are not financially constrained, they do not tell us whether these constraints matter *quantitatively* and in the *aggregate*.

To assess whether these micro-level elasticities have significant aggregate implications, we proceed in two steps. First, we set-up a structural model of firms dynamics. The model builds on the standard neo-classical model of investment with adjustment costs ([Jorgenson, 1963](#); [Lucas, 1967](#); [Hayashi, 1982](#)). To this standard model, we add one simple amendment. We assume that firms face a collateral constraint: the amount they can borrow every period is limited by how much tangible assets –including real estate– they own. Each period, the value of real estate assets fluctuates randomly, creating variations in the collateral constraint, thus mimicking our reduced-form empirical design.⁴ We estimate this model through a Simulated Method of Moments. In addition to the standard moments used in the structural corporate finance literature, our estimation procedure explicitly targets the sensitivity of investment to variations in local real estate prices. We show that the model manages to fit the targeted moments and some non-targeted ones precisely. It also has well-behaved comparative statics properties, which

⁴While we do not explicitly micro-found the collateral constraint, it emanates naturally from limited enforcement models ([Hart and Moore, 1994](#)).

ensures a precise parameter estimation. We also show that a simple ratio of sales to capital is a good measure of financing constraints, as argued in the development literature (Hsieh and Klenow, 2009).

In a second step, the estimated model is nested in a simple general equilibrium where firms compete for customers, workers and for capital goods. We simulate two economies: one in which firms face the estimated collateral constraints, and a counterfactual economy where firms are unconstrained. We compute output and welfare losses from financing constraints by comparing the two economies. We find aggregate welfare loss from financing constraints of 9.4% and output loss of 11%. Such losses arise in part from the *misallocation* of inputs across heterogeneous producers (Hsieh and Klenow, 2009; Moll, 2014; Midrigan and Xu, 2014) and in part from a sub-optimal aggregate capital stock. While both channels matter, aggregate capital matters twice as much as misallocation. It is important to note that, in line with the macroeconomic literature, we formally quantify the *cost* of financing frictions, but not their potential *benefit*. We model collateral constraints in a reduced-form way and do not take a stance on whether the rationale behind these collateral constraints is efficient or not.

Related Literature. Our focus on collateral constraints is rooted in a large array of empirical evidence on the importance of collateral constraints. It is well documented that collateral plays a key role in financial contracting. More redeployable assets receive larger loans and loans with lower interest rates (Benmelech, Garmaise, and Moskowitz, 2005). The value of collateral affects the relative ex post bargaining power of borrowers and lenders (Benmelech and Bergman, 2008). Beyond these effects on financial contracting, collateral values also affect real outcomes at the micro-economic level: Firms with more valuable collateral invest more (Gan, 2007; Chaney, Sraer, and Thesmar, 2012); individuals with more valuable collateral are more likely to start up new businesses (Schmalz, Sraer, and Thesmar, Forthcoming; Adelino, Schoar, and Severino, 2015). In

addition, many empirical evidence point to the prevalence of real estate collateral in loan contracts ([Davydenko and Franks, 2008](#); [Calomiris, Larrain, Liberti, and Sturgess, 2015](#)). Our paper adds to the literature by bridging the gap between microeconomic evidence on the role of collateral constraints and the macroeconomic effect of financial frictions.

Our paper also contributes to the long-standing literature in corporate finance investigating the real effects of financing frictions. This literature has traditionally explored the effect of financing frictions on corporate investment. A key challenge is to find exogenous variations in financing capacity that are not correlated with investment opportunities. For instance, [Lamont \(1997\)](#) overcomes this challenge by showing that non-oil divisions of oil conglomerates increase their investment when oil prices increase. [Rauh \(2006\)](#) shows that firms with underfunded defined benefit plans need to make financial contributions to their pension fund, depriving them of available cash-flows and leading to reduced investment.⁵⁶

Several important papers have developed a structural quantitative approach to estimate the effect of financing frictions. This literature is reviewed in [Strebulaev and Whited \(2012\)](#). In a seminal contribution, [Hennessy and Whited \(2007\)](#) use SMM to estimate a dynamic model of investment and infer the magnitude of financing costs. They find that for small firms, the estimated marginal equity flotation costs is about 10.7% of capital and bankruptcy costs 15.1%. [Hennessy and Whited \(2005\)](#) develop a dynamic trade-off model, which they structurally estimate to explain several empirical findings inconsistent with the static trade-off theory. [Lin, Ma, and Xuan \(2011\)](#) examines the impact of the divergence between corporate insiders' control rights and cash-flow rights on firms' external finance constraints from a generalized method of moments estimation of

⁵See [Bakke and Whited \(2012\)](#) for a discussion of this identification strategy.

⁶The literature on this topic is extensive. For some important contributions, see [Fazzari, Hubbard, Petersen, Blinder, and Poterba \(1988\)](#), [Erickson and Whited \(2000\)](#), [Kaplan and Zingales \(1997\)](#), [Almeida and Campello \(2007\)](#), [Blanchard, Lopez-de Silanes, and Shleifer \(1994\)](#), [Campello, Graham, and Harvey \(2010\)](#), [Chaney, Sraer, and Thesmar \(2012\)](#), [Kaplan and Zingales \(2000\)](#), [Peek and Rosengren \(2000\)](#), [Campello, Gimabona, Graham, and Harvey \(2011\)](#).

an investment Euler equation and show that the agency problems associated with the control-ownership divergence can have a real impact on corporate financial and investment outcomes. [Nikolov and Whited \(2014\)](#) estimate a dynamic model of finance and investment with different sources of agency conflicts between managers and shareholders to analyze the role of agency conflicts in corporate policies and investment. Our contribution to this literature is twofold. First, we include coefficient estimates from a reduced-form regression identifying the effect of collateral constraints on investment and debt as targeted moments. We show that these moments are crucial in identifying the strength of financial frictions in our data. Second, we nest our investment model into a general equilibrium model, which allows us to account for general equilibrium effects in our counterfactuals. In contrast, the literature typically only considers partial equilibrium counterfactuals. In that sense, our model is close to [Gourio and Miao \(2010\)](#) who focus on taxation. Compared to their paper, we focus on model estimation and the effect of financing constraints.

Finally, our paper contributes to the important macroeconomic literature on the aggregate effects of financial frictions. [Restuccia and Rogerson \(2008\)](#), [Hsieh and Klenow \(2009\)](#) and [Bartelsman, Haltiwanger, and Scarpetta \(2013\)](#) emphasize the effect of misallocation of resources across heterogeneous firms on aggregate TFP and welfare. [Midrigan and Xu \(2014\)](#) focus on financing frictions as a source of misallocation. They calibrate a model of establishment dynamics with financing constraints and find that financing frictions cannot explain large aggregate TFP losses from misallocation. In contrast, [Moll \(2014\)](#) shows that for a TFP persistence parameter in the empirically relevant range, financial frictions can matter in both the short and the long run. [Buera, Kaboski, and Shin \(2011\)](#) develop a quantitative framework to explain the relationship between aggregate/sector-level TFP and financial development across countries and show that financial frictions account for a substantial part of the observed cross-country differences in output per worker, aggregate TFP, sector-level relative productivity, and capital-to-

output ratios. Beyond misallocation, a large literature has investigated the effects of financing friction on aggregate TFP growth and welfare. [Jeong and Townsend \(2007\)](#) develop a method of growth accounting based on an integrated use of transitional growth models and micro data and find that in Thailand, between 1976 and 1996, 73 percent of TFP growth is explained by occupational shifts and financial deepening. [Amaral and Quintin \(2010\)](#) present calibrated simulations of a model of economic development with limited enforcement and find that the average scale of production rise with the quality of enforcement. [Riddick and Whited \(2009\)](#) study the costly reallocation of capital across heterogeneous firms. They infer the cost of reallocation from a calibrated model and show that reallocation costs need to be strongly countercyclical to be consistent with the observed dispersion of productivity. Our contribution to this literature is that we base our quantification exercise on an estimation procedure that targets moments from a reduced-form analysis exploiting exogenous shocks to financing capacity. Second, our paper combines adjustment costs with financing frictions. [Asker, Collard-Wexler, and Loecker \(2014\)](#) consider the effect of adjustment costs on static misallocation measures, but their economy does not feature a financing friction. In contrast, our approach delivers interesting implications on the interaction between adjustment costs and credit frictions.

We present reduced-form evidence of the effect of collateral values on both investment and employment in Section 2.1. We present our formal model of firm dynamics with collateral constraints in Section 2.2. We structurally estimate the model using US firm level data in Section 2.3. Section 2.4 describes and implements the general equilibrium analysis. Section 2.5 discusses robustness and implements a policy experiment.

2.1 Reduced-form evidence

We estimate the investment and borrowing sensitivity to real estate value as in [Chaney, Sraer, and Thesmar \(2012\)](#). The construction of the data is detailed in that paper. The dataset is a panel of publicly listed firms from 1993 to 2006 extracted from COMPUSTAT. We require that these firms supply information about the accounting value and cumulative depreciation of land and buildings (items ppenb, ppenli, dpacb, dpacli) in 1993. We then combine this information with office prices in the city where headquarters are located, in order to obtain a measure of the market value of firms' real estate holdings, which we normalize by the previous year property, plant and equipment. We call this measure $REValue_{it}$ for firm i at date t . We require that this variable is available for all firms, so that we end up with a panel of 20,074 observations corresponding to 2,218 firms which are followed from 1993 until 2006 unless they drop out of the panel before (only 676 firms are still present in 2006).

We then run the following regression:

$$\frac{Y_{it}}{k_{it-1}} = \frac{\alpha + \beta \cdot REValue_{it}}{k_{it-1}} + Offprice_{it} + a_i + \epsilon_{it},$$

where k_{it-1} is the lagged stock of productive capital (item ppent). $Offprice_{it}$ is an index for office prices in the city where firm i 's headquarters are located. This index is available from Global Real Analytics for 64 MSAs. We further add a firm fixed effect (a_i) and cluster error terms ϵ_{it} at the firm level. We are interested in β , the sensitivity of Y_{it} to real estate value. We report descriptive statistics for these variables in [Table 2.1](#).

We look at two different left hand-side variables Y_{it} : capital expenditures (item capx) and net debt increase (sum of changes in long term debt – item dltt – and short term debt – item dlc). The estimated sensitivity of investment to real estate value, $\hat{\beta}$, is equal to 0.04 with a t-stat of 6.1. This can be interpreted as a \$0.04 investment response per \$1 increase in real estate value. The sensitivity of net borrowing to real estate value is also

estimated at 0.04, with a t-stat of 4.5. These numbers are close to the main estimate of [Chaney, Sraer, and Thesmar \(2012\)](#), the difference coming from the set of controls used. We opt here for a simpler specification with *fewer* controls, in order to restrict ourselves to variables available in the simulations of the model we present in the next section. This model will be estimated using the first coefficient (the investment sensitivity) as a targeted moment, while the second coefficient (the borrowing sensitivity) will serve as a non-targeted moment.

2.2 The model

In this section, we lay out our model of investment dynamics under collateral constraints. The economy is populated with heterogeneous, financially constrained firms, which combine capital and labor to produce differentiated goods. Those differentiated goods are then combined into a final good, consumed by a representative consumer and used as capital good.

2.2.1 Production technology and demand

The firm-level model is close to [Hennessy and Whited \(2007\)](#) in the sense that it includes a tax shield for debt and a large cost of equity issuance (in our case, infinite⁷) and [Midrigan and Xu \(2014\)](#) in the sense that firms face a collateral constraint. The firm's shareholder is assumed risk-neutral and has a time discount rate of r . Firm i produces output q_{it} combining capital k_{it} and efficiency units of labor l_{it} into a Cobb-Douglas production function with capital share α

$$q_{it} = F(e^{z_{it}}, k_{it}, l_{it}) = e^{z_{it}} \left(k_{it}^{\alpha} l_{it}^{1-\alpha} \right), \quad (2.1)$$

⁷This infinite equity issuance cost simplifies the model and clarifies its exposition. We show in section 2.5 how the quantitative features of the model are changed when we assume a finite issuance cost within the range of the literature's estimates.

with z_{it} the firm's log total factor productivity which is assumed to follow an AR(1) process:

$$z_{it} = \rho z_{it-1} + \epsilon_{it},$$

where we denote σ^2 the variance of the innovation ϵ_{it} . The firm faces a downward sloping demand curve with constant elasticity $\phi > 1$,

$$q_{it} = Q p_{it}^{-\phi}, \quad (2.2)$$

where Q is aggregate spending and will be determined in equilibrium (see Section 2.4).

Labor is fully flexible, and w is the wage – also determined in equilibrium. As labor is a static input, the total revenue of the firm net of labor input is

$$r(z_{it}; k_{it}) = \max_{l_{it}} p_{it} q_{it} - w l_{it} = b Q^{1-\theta} w^{-\frac{(1-\alpha)}{\alpha} \theta} e^{\frac{\theta}{\alpha} z_{it}} k_{it}^{\theta}, \quad (2.3)$$

with b a scaling constant and $\theta \equiv \frac{\alpha(\phi-1)}{1+\alpha(\phi-1)} < 1$.

2.2.2 Input dynamics

Capital accumulation is subject to depreciation, time to build, and adjustment costs. At date t , gross investment i_{it} is given by

$$k_{it+1} = k_{it} + i_{it} - \delta k_{it}, \quad (2.4)$$

where δ is the depreciation rate. In period t , investing i_{it} entails a convex cost of $\frac{c}{2} \frac{i_{it}^2}{k_{it}}$. Additionally, the firm pays in period t for capital that will only be used in production in period $t + 1$: this one period time to build for capital is conventional in the macro literature (Hall, 2004; Bloom, 2009) and acts as an additional adjustment cost. Introducing adjustment costs to capital is important in our estimation exercise, since they generate patterns similar to financing constraints and could thus be a natural confounding factor

in our estimation procedure. For instance, adjustment costs make capital vary less than firm output, which generates a natural dispersion in capital productivities, exactly like financing constraints do (Asker, Collard-Wexler, and Loecker, 2014). As we will show below, using the reduced-form moments presented in Section 2.1 allow us to identify both frictions separately.

We do not, however, include fixed adjustment costs to our model, a choice also made by Gourio and Kashyap (2007): our estimation targets firm-level data at an annual frequency, for which investment is not very lumpy. In our sample (described in Section 2.1), only 4% of the observations have an investment rate smaller than 2% of capital.⁸

2.2.3 Financing frictions and capital structure

The firm finances investment out of retained earnings and debt issuance to outside investors. d_{it} is net debt, so that $d_{it} < 0$ means that the firm holds cash. As is standard in the structural corporate finance literature (Hennessy and Whited, 2005), we only consider short-term debt contracts with a one period maturity. We set up the model so that debt is risk-free and pays an interest rate r^9 – determined in equilibrium in Section 2.4. For an amount d_{it} of debt issued at date t , the firm commits to repay $(1 + r)d_{it+1}$ at date $t + 1$. Finally, we also assume that the interest rate the firm receives on cash is lower than the interest rate it has to pay on its debt: if the firm has negative net debt, it receives a positive cash inflow of $-(1 + (1 - m)r)d_{it+1}$ with $0 < m < 1$.

Consistently with the corporate finance literature, we also assume that firm's profits net of interest payments and of capital depreciation, δk_{it} , are taxed at rate τ . As a result, debt is tax free, which creates an incentive for firms to increase their leverage. Other papers make alternative assumptions to make debt attractive to firms, either by assuming that debt holders are intrinsically more patient than shareholders, or that the sharehold-

⁸To compute the investment rate, we divide item capx by lagged item ppent

⁹While this risk-free interest rate could be time-varying, i.e. r_t , it will always be constant in our model and we thus omit the t subscript for simplicity.

ers seek to smooth consumption, for instance through log utility as in [Midrigan and Xu \(2014\)](#). Finally, note that all tax proceeds are rebated to the representative consumer – see Section 2.4.

The financing frictions come from the combination of two constraints. First, firms cannot issue equity, an assumption we relax in Section 2.5 where we instead consider a finite cost of equity issuance in line with parameter estimates from the literature. Second, firms face a collateral constraint, which emanates from limited enforcement ([Hart and Moore, 1994](#)). We follow [Liu, Wang, and Zha \(2013\)](#) and adopt the following specification for the collateral constraint:

$$(1 + r)d_{it+1} \leq s((1 - \delta)k_{it+1} + \mathbb{E}[p_{t+1}|p_t] \times h), \quad (2.5)$$

The total collateral available to the creditor at the end of period $t + 1$ consists of depreciated productive capital $(1 - \delta)k_{it+1}$ and real estate assets with value $p_{t+1}h$. We assume $\log p_t$ to be a discretized AR(1) process. s , the fraction of the collateral value realized by creditors, captures the quality of debt enforcement, but also the extent to which collateral can be redeployed and sold.¹⁰

In assuming that the quantity of real estate h is the same across firms and time, we abstract from issues related to real estate ownership heterogeneity, which is an important limitation of this paper. In reality, we recognize that firms decision to buy or lease real estate assets can potentially depend on expected productivity, investment opportunities and financing constraints. However, we leave the analysis of how the endogeneity of real estate ownership affects current investment decisions for future research and focus this paper on measuring and aggregating financial frictions given the observed levels of real estate ownership in the data.

¹⁰The formulation of the collateral using the expected future value of collateral is standard in macroeconomics. It can be justified as an optimal contract in a set-up where (1) the firm has the entire bargaining power in its relationship with creditors (2) it cannot commit not to renegotiate the debt contract at the end of period t and (3) collateral can only be seized at the end of period $t + 1$.

2.2.4 The optimization problem

The firm is subject to a death shock with probability d , but infinitely lived otherwise. Every period, physical capital and debt are chosen optimally to maximize a discounted sum of per period cash flows, subject to the financing constraint. The firm takes as given its productivity, local real estate prices, and forms correct expectations for future productivities and real estate prices.

Define as $V(S_{it}; X_{it})$ the value of the discounted sum of cash flows given the exogenous state variables $X_{it} = \{z_{it}, p_t\}$ and the past endogenous state variables $S_{it} = \{k_{it}, d_{it}\}$. Shareholders are assumed to be perfectly diversified so their discount rate is the same as risk-free debt r .

This value function V is the solution to the following Bellman equation,

$$\left\{ \begin{array}{l} V(S_{it}; X_{it}) = \max_{S_{it+1}} \left\{ e(S_{it}, S_{it+1}; X_{it}) + \frac{1-d}{1+r} \mathbb{E}[V(S_{it+1}; X_{it+1}) | X_{it}] + \frac{d}{1+r} (k_{it+1} - (1 + \tilde{r}_{it})d_{it+1}) \right\} \\ \text{s.t.} \quad (1+r)d_{it+1} \leq s((1-\delta)k_{it+1} + \mathbb{E}[p_{t+1}|p_t] \times h) \\ \quad \quad \quad e(S_{it}, S_{it+1}; X_{it}) \geq 0 \\ \text{with:} \quad e(S_{it}, S_{it+1}; X_{it}) = (1-\tau) \left(r(z_{it}; k_{it}) - i_{it} - \frac{c}{2} \frac{i_{it}^2}{k_{it}} + d_{it+1} - (1 + \tilde{r}_{it})d_{it} \right) \\ \quad \quad \quad + \tau (\mathbf{1}_{d_{it}>0} \times r d_{it} + \delta k_{it}) \\ \quad \quad \quad i_{it} = k_{it+1} - (1-\delta)k_{it} \\ \quad \quad \quad \tilde{r}_{it} = r \text{ if } d_{it} > 0 \text{ and } (1-m)r \text{ if } d_{it} \leq 0 \end{array} \right. \quad (2.6)$$

where the second term in the maximand ($\frac{d}{1+r} (k_{it+1} - (1 + \tilde{r}_{it})d_{it+1})$) corresponds to the shareholder's payoff in case of firm death. This term avoids a bias towards borrowing. If we assume instead that bankers can recover capital when a firm exit, shareholders then have an incentive to borrow more in order to transfer value from the states of nature where they cannot consume to states where the firm survives. By assuming that shareholders receive the remaining capital when the firm exit, we ensure that this risk-shifting behavior does not drive the capital structure decisions of firms in our model.

Aggregate demand Q and the real wage w are equilibrium variables that the firms takes as given when optimizing inputs. Given the absence of aggregate uncertainty and the steady state assumption, they are fixed over time. Due to downward sloping demand, firms have an optimal scale of production. A firm initially below this level accumulates capital, but only gradually because of convex adjustment costs and time to build. Once the target scale is reached, firms replace depleted capital. Finally, spending on adjusting capital is bound by the collateral constraint. When the value of a firm's real estate assets increases, the collateral constraint is relaxed, and the firm finances more of the cost of adjusting towards its desired scale. This will generate the sensitivity of investment to real estate value that we have documented in Section 2.1.

2.3 Structural Estimation

2.3.1 Estimation procedure

We estimate the key parameters of the model via a Simulated Method of Moments. The entire procedure is described in detail in Appendix 2.A. We look for the set of parameters $\hat{\Omega}$ such that model-generated moments $\mathbf{m}(\hat{\Omega})$ on simulated data fit a pre-determined set of data moments $\mathbf{m}$. If we could solve the model analytically, we could just invert the system of equations given by model-based moments. Because our model does not have an analytic solution, we need to use indirect inference to perform the estimation. Such inference is done in two steps:

1. For a given set of parameters, we solve the Bellman problem (2.6) numerically and obtain the policy function $S_{it+1} = (d_{it+1}, k_{it+1})$ as a function of $S_{it} = (d_{it}, k_{it})$ and exogenous variables $X_{it} = (z_{it}, p_t)$. We discretize the state space (S, X) into a grid that is as fine as possible to minimize numerical errors in the presence of hard financing constraints. This is critical: a 1-2% numerically generated error would be too large to quantify aggregate effects of this order of magnitude. Solving the

model repeatedly to estimate our structural parameters would not be feasible on a conventional CPU (several hours per iteration), so we use a GPU instead (a few minutes per iteration), as described in Appendix 2.A.1.

2. Our parameter estimates $\hat{\Omega}$ minimize the distance from simulated to data moments $\mathbf{m}$,

$$\hat{\Omega} = \arg \min_{\Omega} (\mathbf{m} - \hat{\mathbf{m}}(\Omega))' \mathbf{W} (\mathbf{m} - \hat{\mathbf{m}}(\Omega)),$$

where the weighting matrix $\mathbf{W}$ is the inverse of the variance-covariance matrix of data moments. Standard errors are calculated by bootstrapping. Appendix 2.A.2 describes how we escape the many local minima present from estimating a large number of parameters.

2.3.2 Predefined and Estimated Parameters

The model has 14 parameters. We calibrate 9 of them using estimates from the literature or the data, and estimate the 5 remaining ones.

Predefined parameters.— Our 9 calibrated parameters are as follows. We set the capital share $\alpha = 1/3$ from Bartelsman, Haltiwanger, and Scarpetta (2013) and the demand elasticity $\sigma = 5$ from Broda and Weinstein (2006) (which would lead to mark-ups of 25% in the absence of adjustment costs). Real estate prices $\log p_t$ follow a discretized AR(1) process. We estimate this AR(1) process on de-trended logged real estate prices and find a persistence 0.62 and innovation volatility 0.06. Both AR(1) processes for $\log z_t$ and $\log p_t$ are discretized using Tauchen’s method. The rate of obsolescence of capital is set at $\delta = 6\%$ as in Midrigan and Xu (2014). The risk-free borrowing rate r is fixed at 3%, while the lending rate is set to $(1 - m)r = 2\%$. We fix the death rate d to 8% which corresponds to the turnover rate of firms in our data. We set the corporate tax rate τ at 33%. Finally, we set $w = 0.03$ (\$30,000) and $Q = 1$ for the estimation. They

will, however, be endogenously determined in general equilibrium in our counterfactual analyses —see Section 2.4.

Estimated parameters.— We estimate 5 deep parameters but focus the discussion on 4 of them: The persistence ρ and innovation volatility σ of log productivity, the collateral parameter s and the adjustment cost c . The fifth parameter, the amount of real estate collateral available h , allows us to match the average ratio of real estate to capital h/k_t , and is essentially a normalization.

2.3.3 Data Moments

We compute the moments on the COMPUSTAT sample described in Section 2.1. We describe them here with a short heuristic discussion about their “identifying” power. In the next section, we discuss identification more systematically and show how simulated moments vary with parameters.

First, in the spirit of [Midrigan and Xu \(2014\)](#), we use the short- and long-term volatility of output to estimate the persistence and volatility of the productivity process. In our sample, the volatility of change in log sale ($\log \text{sales}_{it} - \log \text{sales}_{it-1}$, COMPUSTAT item: sale) equals 0.327. The volatility of 5-year change in log sales ($\log \text{sales}_{it} - \log \text{sales}_{it-5}$) equals 0.911. The fact that 5-year growth is less than 5 times more volatile than 1-year growth indicates mean-reversion and contributes to the identification of the persistence parameter. Targeting these two moments instead of directly matching the persistence coefficient of log sales makes our estimation less sensitive to model misspecification, e.g. for a true process with a longer memory than an AR(1).

Second, we use the autocorrelation of investment to identify adjustment costs ([Bloom \(2009\)](#)). For each firm in our panel we compute the ratio $\frac{i_{it}}{k_{it-1}}$ of capital expenditures (COMPUSTAT item: capx) to *lagged* capital stock (COMPUSTAT item: ppent). The correlation between $\frac{i_{it}}{k_{it-1}}$ and $\frac{i_{it-1}}{k_{it-2}}$ in our data is 0.43. Adjustment costs are needed to match this large correlation: they compel the firm to smooth its investment policy in response

to a productivity shock (Asker, Collard-Wexler, and Loecker, 2014). Financing frictions add to this smoothing motive.

Third, we use two alternative moments to estimate the collateral constraint parameter s . The first moment is net book leverage, a moment typically used in the literature (Hennessy and Whited, 2007; Midrigan and Xu, 2014). Book leverage is computed as financial debt (COMPUSTAT items: `dlc` + `dltt`) minus cash holdings (COMPUSTAT item: `che`), normalized by total assets (COMPUSTAT item: `at`). This definition reflects the notion that cash is equivalent to negative debt, as it is the case in our model. We obtain an average of 0.313 in our data. In our model, leverage directly identifies the collateral parameter s as higher collateral values unambiguously lead to more borrowing. Yet, as we discuss more extensively below, this moment (leverage) is not ideal to identify financing constraints for two reasons. First, from an identification standpoint, leverage may be an ambiguous moment. For instance, a firm may not be financially constrained yet choose to lever up for tax purposes. This behavior would lead to mis-attribute corporate leverage to collateral constraints (see Section 2.5.1 for a formal analysis of this identification problem). Second, financial leverage may be a noisy measure of a firm's indebtedness. For instance, financial debt typically includes unsecured debt, which is not part of our model (see Section 2.5.2 for such an extension), and which would lead to overstate the extent to which collateral can be pledged. For all these limitations of the leverage moment, we use a more direct measure of financing constraints instead, the sensitivity of investment to real estate value, computed in Section 2.1. Because it is also an informative and natural moment, we also look at the sensitivity of debt issuance to real estate value. We never target this second moment in our estimation, but it turns out our main model matches it very well (more on this below).

Finally, we compute the quantity of real estate held by the average firm, by taking the ratio of real estate holdings (COMPUSTAT item `land` + `buildings`) in 1993 normalized by total assets (COMPUSTAT item: `at`), and obtain 0.14. By adjusting h , our estimation

procedure matches this moment perfectly; we view this part of the estimation as a normalization more than anything else. As a result, we omit discussion of this parameter from this point on.

2.3.4 Parameter Identification

This section discusses identification of the parameters of the model. In Appendix Figures 2.5-2.8, we reproduce how moments vary as a function of model parameters. We also show, in Table 2.2, the elasticities of each moments with respect to estimated parameters – a simple transformation of the Jacobian matrix. All this analysis is about *local* identification, in the sense that we operate around our main SMM estimate for (s, c, ρ, σ) – which we discuss in detail in the next section.

We first discuss the graphical evidence. In Figures 2.5-2.8, we offer visual evidence of how the different moments we use in our estimation help identify the model’s parameters. To construct these figures, we first set all parameters (s, c, ρ, σ) at their estimated value, and then vary one of these parameters in partial equilibrium, i.e. holding fixed w and Q . All figures are reported using the same scale for each moment. Importantly, the comparative statics we report on these figures are *direct simulation output*: The relative smoothness of these plots gives us confidence in the robustness of our numerical procedure, which we attribute to the dense grid for capital (about 300 points), debt (29 points) and productivity (51 points) we use, as well as to the large number of simulated observations (1,000,000 firms over 10 years). See Appendix 2.A for details.

Figure 2.5 shows that the collateral parameter s influences mostly the leverage moment as well as the investment and debt sensitivities to real estate prices. This result is intuitive. Obviously, a higher s unambiguously leads to higher leverage: In our setting, the firm takes on more debt if it is allowed to. The sensitivity moments are non-monotonic with s . Intuitively, for low values of s , firms investment decisions are constrained by collateral availability: In this range of values for s , an increase in s allows firms to extract

more debt and investment capacity out of a \$1 increase in collateral values. For higher values of s , however, firms become less financially constrained, so that their investment policies become less driven by collateral values. At the limit, when s grows close to 1, the firm becomes unconstrained and investment is no longer sensitive to fluctuations in house prices. We also see in Figure 2.5 that around the SMM estimate (represented by a vertical line), both sensitivity moments are smooth and increasing functions of s . The second panel of Figure 2.5 also shows that an increase in s leads to an increase in the long-term volatility of production: when the firm is less constrained, its capital stock responds more to productivity shocks, which increases the volatility of output.

Figure 2.6 shows that the adjustment cost parameter c is mostly identified by the autocorrelation of investment: Large adjustment costs lead the firm to smooth investment across time, which lead to a large autocorrelation of investment. Larger adjustment costs to capital also lead to lower short-term output volatility: Similar to financing constraints, adjustment costs prevent firms from adjusting their capital stock to productivity shocks, making output less volatile. Figures 2.7 and 2.8 shows that (1) the volatility of log-productivity σ has a nearly linear impact on the short-term volatility of output (2) the persistence ρ of productivity shocks strongly influences the long-term volatility of output, but has no first-order effect on short-term volatility. Combined together, these two observations are consistent with the idea that the ratio of the 1-year to 5-year output volatility allows to identify the persistence parameter ρ . Note also that the persistence of productivity shocks has a sizable positive effect on the autocorrelation of investment: Firms can afford to delay their response to productivity shocks, since these shocks are more persistent.

In Table 2.2, we quantify how the various simulated moments vary as a function of the estimated parameters. More precisely, we compute for each moment m_n , and each

parameter ω_k , the following elasticity (Hennessy and Whited (2007)):

$$\epsilon_{n,k} = \frac{m_n^+ - m_n^-}{\omega_k^+ - \omega_k^-} \times \frac{\hat{\omega}_k}{\hat{m}_n} \approx \frac{\partial \log(\hat{m}_n)}{\partial \log(\hat{\omega}_k)},$$

where $\hat{\omega}_k$ is the parameter value at the SMM estimate and $\hat{m}_n$ the corresponding value for moment n . $\hat{\omega}_k^+$ (respectively $\hat{\omega}_k^-$) is the parameter value located right above (resp. below) on the grid used to plot Figures 2.5-2.8. m_n^+ (resp. m_n^-) is the corresponding moment obtained using parameter $\hat{\omega}_k^+$ (resp. $\hat{\omega}_k^-$), keeping the other parameters $\hat{\omega}_{k'}$ at their SMM estimate.

Table 2.2 confirms formally the results we discussed from Figure 2.5-2.8.

2.3.5 Estimation results

We report the results of the SMM estimation in Table 2.3. One key contribution of the paper is to use the sensitivity of investment to real estate value as a targeted moment in this estimation. To highlight the contribution of this moment, we thus report two sets of results: One estimation where the SMM targets the mean leverage to identify financing constraints – as the existing literature does – and one set of results where the SMM instead targets the sensitivity moment. Each column corresponds to a model specification (with adjustment costs, Columns (3) and (4), and without adjustment cost, Columns (1) and (2)) and a set of targeted moments including leverage (Columns (1) and (3)) or the sensitivity of investment to house prices (Columns (2) and (4)). Column (5) corresponds to the data.

We first study the version of the model without adjustment cost ($c = 0$). There are 3 parameters to estimate: The persistence (ρ) and volatility (σ) of log-productivity, as well as the pledgeability parameter s . In Column (1) of Table 2.3, the SMM targets “traditional moments”, i.e. the short- and long-term volatilities of log sales, and mean leverage. At the estimated parameters, the model matches all the targeted moments up to the second

decimal, but does poorly on non targeted moments. The sensitivity of investment and debt to real estate value is high (three times their empirical value: 0.12 instead of 0.04 in both cases). The autocorrelation of investment is negative, instead of positive in the data, due to the absence of adjustment costs.

In Column (2), the estimation targets the sensitivity of investment to real estate prices instead of leverage. As a result, the estimated pledgeability parameter, s , is smaller than in the estimation of Column (1) (0.133 instead of 0.495). As was explicit on Figure 2.5, the sensitivity of investment to real estate prices is an increasing function of s in this range of parameters: As a result, to reduce the sensitivity of investment to real estate prices relative to the one delivered by the estimation of Column (1), a smaller value for s is estimated. A lower estimated s implies a lower debt capacity, so that mean leverage in this model is much smaller, and in particular, much smaller than its empirical value (0.013 vs. 0.313 in the data). Since this model does not include adjustment costs to capital, the average autocorrelation of investment in the simulated model of Column (2) remains distant from its empirical counterpart (0.064 vs. 0.436 in the data).

We introduce these adjustment costs to capital in Columns (3) and (4). With these costs, the estimated model matches the autocorrelation of investment exactly, whether we target mean leverage (Column (3)) or the investment sensitivity coefficient (Column (4)). However, when the estimation targets the sensitivity of investment to real estate prices instead of mean leverage, we estimate a much smaller pledgeability parameter s (0.189 vs 0.422), for the same reason as mentioned in the discussion of the estimated models of Column (1) and (2). The introduction of adjustment costs to the model leads to a higher estimated pledgeability parameter (0.189 in Column (4) vs. 0.133 in Column (2)): In the presence of collateral constraints, adjustment costs to capital make investment less responsive to collateral values; as a result, to match the sensitivity of investment to real estate prices, the estimated s has to increase. With adjustment costs to capital and this sensitivity as a targeted moment (Column (4)), we are able to match perfectly not

only the sensitivity of investment to real estate prices, but also the sensitivity of debt, not targeted in the estimation. The leverage ratio in the estimated model of Column (4) is larger than in the model with no adjustment costs (0.095 in Column (4) vs. 0.013 in Column (2)) – since the firm now has to pay for these adjustment costs – but it remains, however, below its empirical value (0.095 in Column (4) vs. 0.313 in the data). We do not view this discrepancy as a major source of concern. The corporate finance literature has put forth a number of determinants of leverage that are not included in our model (working capital management, moral hazard etc), but that would not necessarily interact with the *real outcomes* from the model. We thus take Column (4) as our preferred specification. We propose an extension to our model in Section 2.5.2, which allows us to simultaneously match the sensitivity of investment to real estate prices and mean leverage.

The calculation of standard errors is done by bootstrapping and is detailed in Appendix 2.A. We draw 100 data samples and compute the set of targeted moments for each of these sample. We then run our SMM procedure for each one of these samples, and compute standard errors as the empirical s.d. of these parameters. To save on computing time, we estimate these 100 SMMs in parallel. Each time we solve the model with a new set of parameters, we check whether these parameters improve the matching of each one of the 100 moments. All parameters are estimated with a t-stat between 15 and 100. Such precision is not rare in SMM estimation. The collateral coefficient s is however, less precisely estimated (with a t-stat slightly above 3).

2.3.6 Determinants of financing constraints

In this section, we briefly discuss how firm characteristics covary with financing constraints. We use our preferred specification of Column (4), Table 2.3. We define a firm to be financially constrained when its capital stock is lower than 80% of its frictionless capital stock. To compute the frictionless capital stock, we solve the model using the same

parameters but remove the no equity issuance constraint. We then consider various firm characteristics x , sort the simulated firms into 20 equal-sized bins of x and compute the fraction of constrained firms in each bin.¹¹ This methodology allows to see how, in the cross-section of firms, financing constraint covary with firm characteristics.

We report the results of this investigation in Figure 2.1. Panel A shows that more productive firms are more constrained: they are typically firms that experienced a positive productivity shock, but inherited a small capital stock, preventing them from growing as much as they would in the absence of collateral constraints. Panels B-E investigate the relationship between constraints and characteristics that are typically observable in firm-level data. Panel B shows a weak link between firm size and financing constraints: Larger firms are typically more productive (and therefore more constrained), but they also have more collateral (and are thus less constrained). Panel C shows that growing firms are typically more constrained, which is not surprising since they are likely to have experienced recent positive productivity shocks. Panel D shows that firms with high leverage are more likely to be constrained: Since there is no heterogeneity in s in our model, a firm with a high leverage ratio is typically a firm that experiences a large positive productivity shock and exhausts its debt capacity without being able to reach its first-best level of investment. Panel E shows a sharply increasing relation between the ratio of sales to capital and the fraction of constrained firms in the simulated data: This ratio captures the marginal revenue product of capital and captures the effective capital wedge firms face when optimizing investment (Hsieh and Klenow (2009)). Panel F illustrates the non-monotonic relation between the market-to-book ratio and the fraction of firms constrained: A low market-to-book ratio implies that firms have few investment opportunities and are thus less constrained; firms with a large stock of capital are close to unconstrained and as a result, have a large market-to-book ratio.

¹¹As we do in our estimation procedure, we simulate firms over 100 years, but only use the last 10 years to compute the fraction of constrained firms, so as to make sure each firm has reached its steady-state.

2.4 General Equilibrium Analysis

We now have a fully estimated model of firm behavior under financial constraints. To estimate the quantitative effect of this model on aggregate production and TFP, we embed it into a simple macro-model that accounts for general equilibrium feedbacks.

2.4.1 General equilibrium model

By clearing the goods and labor markets, the model endogenizes aggregate demand Q and the real wage w introduced in the firm-level model of Section 2.2. The model consists of the following simple assumptions.

Firms. A large number N of firms indexed by i produce intermediate inputs, in quantity q_{it} , at price p_{it} . All intermediate inputs are combined into a CES-composite final good

$$Q_t = \left(\sum_{i=1}^N q_{it}^{\frac{\phi-1}{\phi}} \right)^{\frac{\phi}{\phi-1}}. \quad (2.7)$$

The final good is produced competitively. The demand for input i is thus given by $q_{it} = Q_t \left(\frac{p_{it}}{P_t} \right)^{-\phi}$, with $P_t = \left(\sum_i p_{it}^{1-\phi} \right)^{\frac{1}{1-\phi}}$. We normalize P_t to 1 and derive the demand function in equation (2.2).

Consumption and consumer behavior. The final good is used for (i) consumption, (ii) investment and (iii) to pay for adjustment costs. The final good market equilibrium thus writes:

$$Q_t = C_t + \text{Adj. Cost}_t + I_t \quad (2.8)$$

with C_t being aggregate consumption, $\text{Adj. Cost}_t = \sum_i \frac{\epsilon}{2} i_{it}^2 / k_{it}$ is the sum of all adjustment costs, and $I_t = \sum_i i_{it}$ is aggregate investment.

Consumption goes to a representative consumer that maximizes inter-temporal utility over consumption and labor supply:

$$U_s = \sum_{t \geq s} \beta^{t-s} u_t \text{ with } u_t = C_t - \bar{L}^{-\frac{1}{\epsilon}} \frac{L_t^{1+\frac{1}{\epsilon}}}{1 + \frac{1}{\epsilon}} \quad (2.9)$$

where L_t are aggregate hours worked, $\bar{L}$ is a simple scaling constant, and ϵ is the Frisch elasticity of labor supply. With quasi-linear preferences, the Hicksian, Marshallian and Frisch elasticities of labor supply are all equal to ϵ . Labor supply is a static decision given by

$$L_t^s = \bar{L} w_t^\epsilon. \quad (2.10)$$

The consumption Euler equation ties the equilibrium interest rate r_t to the discount rate β , and so we take the interest rate $r_t = 1/\beta - 1$ as fixed throughout all counterfactuals.

Steady state assumption and equilibrium definition. We assume that the economy is in steady state. Intermediate good producers produce according to the technology described and estimated in the previous section. The log productivity shocks z_{it} that they face have no aggregate component. Given our assumption that the number of firms is large, aggregate output Q and wage w are constant over time. We are thus exactly in the case studied in Section 2.2.

The equilibrium (Q, w) of this economy is defined by two equations: the labor market equilibrium and the final good aggregator:

$$\bar{L} w^\epsilon = \sum_{i=1}^N l^d((Q, w); z_{it}, k_{it}(Q, w)) \quad (2.11)$$

$$Q = \sum_{i=1}^N p_{it} q((Q, w); z_{it}, k_{it}(Q, w)) \quad (2.12)$$

where $l^d(\cdot)$ is the numerically obtained labor demand function which is a function of each firm state variable and aggregate equilibrium (Q, w) . Similarly $pq(\cdot)$ is the supply function, which, for each firm, associates state variables and macroeconomic conditions to its dollar sales. The equilibrium (Q, w) is the solution of these two conditions. We solve this problem by iteration, using a variant of the Newton-Raphson algorithm. We describe our methodology in detail in Appendix 2.B.

In our quantitative exercise, we focus on the following aggregate quantities. Aggregate output Q and real wage w are direct outcomes of the algorithm. Aggregate employment is given by the supply curve: $L = \bar{L}w^\epsilon$. Aggregate log TFP is classically given by $\log Q - \alpha \log K - (1 - \alpha) \log L$, where K , the aggregate capital stock in the steady state, is computed as the sum of capital stocks over all firms. Finally, welfare is a function of (Q, w) , the aggregate capital stock K and aggregate adjustment cost

$$U = \frac{1}{1 - \beta} \left((Q - \delta K - \text{Adj. Cost}) - \frac{\bar{L}w^{1+\epsilon}}{1 + \frac{1}{\epsilon}} \right).$$

2.4.2 The aggregate effect of financing constraints

We are now in a position to evaluate the aggregate effect of financing constraints. Compared to the firm-level model, the macroeconomic model has a few additional free parameters. Following Chetty (2012), we set the labor elasticity $\epsilon = 0.50$. We adjust $\bar{L}$ and the number of firms N so that the equilibrium parameter chosen for the estimation process ($Q = 1$ and $w = 0.03$) are actual equilibrium parameters when firm parameters are at the SMM estimate.

To measure the aggregate impact of financing constraints, we present all aggregates (output Q , wage w , TFP and welfare) in log deviations from the “unconstrained” benchmark. The appropriate way to define the unconstrained benchmark in our model is to lift the no equity issuance constraint, rather than the collateral constraint. With no equity issuance constraint, investment is unconstrained since equity is freely available

to all firms and fairly priced at r . With no collateral constraint (unlimited debt), firms would raise infinite debt because it gives them a tax advantage. So strictly speaking, our unconstrained benchmark corresponds to a model with no equity issuance constraint, a collateral constraint, and all structural parameters otherwise unchanged.¹²

We first ask how the estimation method affects the aggregate effect of financing constraints. We implement this exercise in Table 2.4. First, we see that estimations targeting the sensitivity of investment to real estate prices (Columns (2) and (4)) generate a TFP loss twice as large as estimations targeting the leverage ratio (Columns (1) and (3)). In our preferred specification (Column (4)), we find a TFP loss of 2.7%, compared to 1.5% in Column (3). This discrepancy is at the core of our analysis: When the estimation targets the mean leverage ratio, it maps all the leverage in the data to collateralized debt in the model. This estimation thus implies a large level of pledgeability s so that the simulated model can match the high level of mean leverage in the data (0.42). This estimate is larger than actual net leverage (0.31 in the data), since in the model firms maintain some debt capacity and therefore issue less debt than they actually can. By contrast, when matching the rather low sensitivity of investment to collateral value, a moment that characterizes how real outcomes are affected by the collateral constraint, the estimated pledgeability parameter is smaller ($s = 0.189$ in Column (4)). In this estimation of Column (4), the collateral constraint is thus tighter than in the estimated model of Column (2), and as a result, losses from financing constraints are larger. In our context, the estimated TFP loss from financing constraints depends strongly on the choice of moment selected to reflect the importance of these constraints. Our paper argues that targeting the average investment response to shocks to collateral values provides more identifying power than targeting the mean leverage ratio (see the formal discussion in Section 2.5.1), and that as a result, we obtain larger TFP losses from financing constraints.

¹²We show below that lifting the collateral constraint (increasing s to a large yet *finite* level) gives results similar to removing the no equity constraint.

Second, Column (4) shows that output loss from financing constraints are as large as 11%. More than half of this output loss is accounted for by a smaller aggregate stock of capital in the constrained economy ($0.192 \times 0.3 = 6.5\%$). About a quarter of this output loss comes from misallocation, since, as we discussed above, TFP in the constrained economy is lower by 3% relative to the unconstrained benchmark. These two effects combined reduce the productivity of labor, which in turn depresses labor supply. The labor supply response accounts for the remaining quarter of the overall output loss. Hence, even though misallocation is non-negligible, the total output loss from financing constraints mostly arises from aggregate under-investment: Firms are constrained, so that the representative consumer under-saves and supplies too little labor relative to the unconstrained economy. Overall, removing financing constraints has a large effect on welfare, 9.4% higher in the unconstrained relative to the actual economy. Consistent with the discussion on TFP losses, we find that the welfare loss from financing constraints is halved (5.1% in Column (3)) when using the estimated s obtained by targeting the average leverage ratio. We also see in Table 2.4 that adjustment costs tend to attenuate the welfare losses from financing constraints. In the presence of adjustment costs, firms smooth out investment by responding partially to productivity shocks. As a result, financing constraints bind less often. Note, however, that this effect of adjustment cost on the estimated welfare loss from financing constraints is quantitatively small.

In Figure 2.2, we show how these general equilibrium quantities are affected when we vary the pledgeability parameter s from 0 to 1. We start from the estimated model of Column (4), Table 2.3, which includes adjustment costs and target the sensitivity of investment to real estate prices. We then change s relative to its estimated value, determine the new general equilibrium of the model and compute the general equilibrium quantities reported in Table 2.4. As in Table 2.4, we report these quantities as deviations from the corresponding unconstrained benchmark. Finally, Figure 2.2 also reports the estimated pledgeability parameter s (vertical dark line), as well as the 95% confidence

band for this parameter (light blue bar). The precision of our estimate – a standard error of 0.008 for a point estimate of 0.189 – implies that for values of s in the 95% confidence interval, aggregate effects remain close to their value reported in Table 2.4. The TFP loss from financing constraints vary by 0.5 percentage point, the output loss by about 2 percentage points and the capital loss by about 5 percentage points.

Overall, Figure 2.2 shows clearly how aggregate outcomes are affected by the pledgeability parameter s . In an economy with no pledgeability ($s = 0$) – and therefore where financing is done entirely through cash-flows – and relative to the unconstrained economy, output is smaller by about 15%, welfare by about 15% as well, employment by about 5%, capital by about 25% and aggregate TFP by about 4%. The effect of pledgeability on these aggregate quantities in general equilibrium is approximately linear. The limited response of aggregate employment to variations in s stems from the relatively small elasticity of labor supply we use. Finally, in the last panel of Figure 2.2, we report the cross-sectional dispersion of log MRPK ($\log p_i q_i / k_i$), the measure of distortions used in Hsieh and Klenow (2009). Note that in the presence of adjustment costs and time-to-build in investment, this dispersion is not 0 (Asker, Collard-Wexler, and Loecker, 2014). However, Figure 2.2 reports the dispersion of log MRPK relative to the unconstrained economy, which features the same adjustment costs and thus account for the effect of adjustment costs in the dispersion of log MRPK. When collateral cannot be pledged ($s = 0$), the dispersion in log MRPK is about 14% higher than in the unconstrained economy.

2.4.3 Productivity persistence and misallocation

Recent papers emphasize that the persistence of productivity shocks should reduce the aggregate effect of financing frictions (Moll, 2014; Buera, Kaboski, and Shin, 2011). Intuitively, if productivity shocks are persistent, firms “grow out” of their financing constraints: productive firms are likely to remain productive and can accumulate cash holdings necessary to fund future investment. To measure this effect in our quantita-

tive model, we start from the estimated model of Column (4), Table 2.3, pick alternative values for the parameter ρ , and compute the equilibrium dispersion of log MRPK ($\log p_i q_i / k_i$) (Hsieh and Klenow, 2009; Midrigan and Xu, 2014). When varying ρ , we keep $\text{Var}(z) = \sigma^2 / (1 - \rho^2)$ constant, varying σ^2 accordingly, as in Moll (2014).

Figure 2.3 shows that the amount of misallocation in equilibrium is significantly reduced when productivity shocks become very persistent. When ρ is set to 0.35 – about one third of its estimated value – the dispersion of log MRPK is more than 50% larger (0.43 vs. 0.66). At the estimated persistence (0.895 in Table 2.4, Column (4)), misallocation as measured through this dispersion is quite sensitive to variations in the persistence parameter.

2.5 Discussion

2.5.1 Model Identification

An important contribution of this paper is to base the estimation of a model of dynamic investment with collateral constraints on a well-identified, reduced-form moment that evaluates how real outcomes respond to shocks to collateral value – the sensitivity of investment to real estate prices. In contrast, most of the literature relies on moments related to financial leverage. Table 2.5 shows why our approach provides a better identification. To obtain this table, we simply simulate data from a model where firms are fully unconstrained. We then show that an estimation targeting the empirical mean leverage would fail to reject that firms are constrained; in contrast, an estimation targeting the sensitivity of investment to house prices would correctly reject this hypothesis.

More precisely, we start from the estimated model of Column (4), Table 2.3. We then simulate a sample of firms from these estimated parameters, but remove the no equity issuance constraint. These simulated firms are unconstrained, by definition. We then compute the following moments on this synthetic dataset: the long- and short-

run volatility of log sales, the autocorrelation of investment, mean leverage, and the sensitivity of investment to real estate prices. Table 2.5, Column (3) show these moments. Unsurprisingly, the sensitivity of investment to real estate prices is -0.001: Firms are unconstrained, investment is efficient and unaffected by real estate shocks, which, by construction, are uncorrelated with productivity shocks.

Using this simulated sample – where the data generating process is such that firms are unconstrained – we estimate our model from Section 2.2 using either mean leverage (Column (1) of Table 2.5) or the sensitivity of investment to real estate prices (Column (2) of Table 2.5) as a targeted moment. When the estimation targets leverage (Column (1)), the pledgeability parameter is in part determined to match leverage, 0.168 in the data. As a result, the estimated pledgeability parameter is low, $s = 0.436$, and in particular lower than one. This estimated s leads to wrongfully conclude that the economy suffers from substantial losses due to financing constraints: the estimated model implies, relative to the unconstrained economy, a 3.1% TFP loss, a 13.0% output loss and a 10.9% welfare loss) when the true model feature no such losses.

When we instead target the sensitivity of investment to real estate prices (Column (2)), the pledgeability parameter is estimated close to 1 ($s = 0.953$): The data used to compute the moments is such that firms are unconstrained so that their investment does not covary with real estate prices; to match this moment, the estimation has to find that the pledgeability of collateral is very high, so that firms' investment is close to its first best. As a result, the estimation based on this moment rightly concludes that there are no aggregate losses from financing constraints.

In other words, in this exercise, both models are misspecified, as they wrongly assume no equity issuance, while the firms in our synthetic dataset are free to issue equity. However, the estimated model targeting the leverage moment completely misses the fact that firms are unconstrained, while the model targeting the sensitivity of investment to real estate prices correctly infers negligible financing constraints.

Of course, our approach could also be invalidated using a similar exercise. One simply needs to find an alternative model where land-holding firms invest more following increases in house prices *relative* to firms not holding land for reasons other than collateral constraints. Finding such a model is equivalent to rejecting the identifying assumption in [Chaney, Sraer, and Thesmar \(2012\)](#). Under their identifying assumption, however, the reduced-form moment purely arises from the existence of collateral constraints, and therefore cannot be falsified. In this sense, the point we make in this paper is generic, and goes beyond this particular reduced-form moment: A valid reduced-form moment identifying the effect of financing constraints on investment – valid in the sense that it estimates the causal effect of financing frictions under a reasonable identifying assumption – will provide a better source of identification in the structural estimation than generic financial moments such as leverage.

2.5.2 Robustness: Residual leverage and costly equity issuance

In this section and [Table 2.6](#), we discuss the robustness of our findings to either a setting where firms have spare debt capacity in addition to the collateralized debt that is the focus of our study, or to assuming firms are allowed to issue equity at a finite cost.

Residual leverage. A potential concern with our baseline specification is that it fails to match the mean leverage ratio (see [Table 2.3](#), Panel A, column 4). The reason for this mis-match is the inherent tension in our baseline model between leverage and the sensitivity of investment to real estate prices. If one targets the leverage ratio (0.313 in our data), s has to be large (0.422), which leads to a counterfactually large investment to real estate sensitivity. If one targets the investment moment (0.04 in our data), s has to be small (0.189) and leverage is counterfactually small. We defend the choice of the investment moment in [section 2.5.1](#) as a better way to detect the presence of financing constraints. In addition, leverage may be determined by a host of firm characteristics

(unsecured debt capacity, trade credit, inventories etc) that we omit in our model. It is possible that once these other sources of external funding are accounted for, firms have enough debt capacity to escape financing constraints. We show here this is not the case.

We modify the baseline model and add a debt capacity $\bar{d}$ to the borrowing constraint,

$$(1 + r)d_{it+1} \leq \bar{d} + s((1 - \delta)k_{it+1} + \mathbb{E}[p_{t+1}|p_t]h). \quad (2.13)$$

This coefficient $\bar{d}$ captures un-collateralized debt capacity left out of the model.

We estimate this new model and report the results in Table 2.6, Column (2). The estimation targets both leverage and the sensitivity of investment to real estate prices, as well as the short- and long-term volatilities and autocorrelation of investment. To match the high level of leverage in the data, $\bar{d}$ is estimated to be high (0.45), while s remains close to our previous estimate (0.254 instead of 0.189). The productivity shock process remains similar. Interestingly, however, the aggregate impact of financing constraints (-10% welfare) is not smaller than in the model that does not fit leverage, i.e. where $\bar{d} = 0$ (-9.4% welfare). Firms do have a higher debt capacity, but this extra debt capacity is similar to free cash, i.e. cash that is not penalized in terms of returns. As a result, firms lever up more in order to minimize taxes, which is why the estimation now matches the leverage ratio almost perfectly. However, the overall borrowing constraint does not bind less because the extra debt capacity is used for tax optimization and not investment in physical capital. Overall, this simple addition to the model – an unsecured debt capacity $\bar{d}$ – allows us to match firms leverage, without changing our inference on the aggregate effects of financing constraints.

Costly equity issuance. We conclude this section by allowing for costly equity issuance. We assume a variable equity issuance cost of 15%, within the range estimated by [Hennessy and Whited \(2007\)](#). The results are presented in Table 2.6, column 3. Neither the parameter estimates nor the fit between simulated and actual moments vary much

compared to our baseline specification. As firms now have the ability to issue equity, the aggregate effects of financing frictions are naturally reduced (5.8% aggregate output loss compared to 11% in our baseline specification).

2.5.3 Policy Experiments

In this section, we use our model to investigate the effect of an investment tax credit (ITC). We consider two types of policies. The first is a non targeted investment subsidy, where each firm in our sample receives a subsidy equal to $x \times i_{it}$, where i_{it} is the firm's investment and x is a fraction equal to 5, 10 and 15%. The second is a targeted investment subsidy, aimed only at capital poor firms, i.e. firm with a high MRPK ($\log(p_i q_i / k_i) > 0.4$). This second policy is motivated by the evidence in Figure 2.1 that most firms with a sales to capital ratio below 0.4 are unconstrained, while most firms above this threshold have a sub-optimally low level of capital.

In both cases, the subsidy is a linear function of investment, i.e. it becomes a tax when investment is negative. This feature avoids the emergence of short capital cycles where firms buy capital to enjoy the subsidy, and sell it the following year. Finally, this subsidy is financed via a lump-sum tax raised on household income. We make this assumption in order to focus on the effect of the ITC.

We report the results of these policy experiments in Table 2.7. With a non-targeted tax credit of 5% the capital stock increases by 11% and aggregate employment by about 1.4%. As a result, output rises by 4.3% and welfare by 2.9%. This large effect of the ITC occurs in our model because corporate profits are taxed at a high rate (33%), which depresses investment significantly: The ITC partially undoes the depressing effect of the corporate income tax.

Interestingly and perhaps surprisingly, the non-targeted investment subsidy increases welfare by about as much as a subsidy specifically targeted to capital poor firms, but at *a much lower cost* in terms of the aggregate amount spent to finance this

subsidy. With a 15% subsidy, welfare increases by 9% for both the targeted and the non-targeted program. This increase in welfare corresponds to almost all the welfare loss from financing constraints estimated in our model (+9.4% in Table 2.4 column 4). However, the non-targeted subsidy requires a tax from household of about 2.4% of total output, while for the targeted subsidy, the cost of the program ends up being larger (about 4.4% of total output). The reason for this differential cost of the two subsidies is that a targeted program induces an opportunistic investment strategy: To benefit from the subsidy, firms invest little (disinvest) as long as their sales to capital ratio is below the policy threshold, and their investment rate jumps discontinuously as soon as they cross the policy threshold. Figure 2.4 makes clear this unintended effect of the targeted subsidy, by showing how the investment rate varies as a function of the sales to capital ratio, in both experiments. With the un-targeted subsidy, investment increases smoothly with the variable used to assign the subsidy (the sales to capital ratio); with the targeted one, investment increases sharply right at the policy threshold.

2.6 Conclusion

This paper provides a quantification of the aggregate effects of a specific source of financing frictions, collateral constraints. We build a simple dynamic general equilibrium model with heterogeneous firms and collateral constraints. To estimate this model structurally, we match not only key features of firm-level dynamics, but also a well identified reduced-form evidence that an increase in the value of a firm's collateral leads to an increase in investment. The estimated model is then used to simulate a counterfactual economy where financing frictions are lifted. Welfare increases by 9.4% and aggregate output by 11%. Quantitatively, only one quarter of these gains can be attributed to a more efficient allocation of inputs across heterogeneous firms – more productive firms are able to obtain more financing and expand – while half of these gains are due to a

higher aggregate stock of capital, and the remaining quarter to a higher aggregate labor supply.

One limitation of this analysis is that the shocks to collateral value that we use to identify the effect of collateral constraints at the firm-level are exogenous in the model. Yet in equilibrium, increased investment and hiring at the local level will clearly feed back into local real estate prices. In addition, since households are not fully mobile across regions, variations in real estate prices will induce variations in wages faced by firms, which will affect their local input choices. Endogenizing the housing market and its feedback effect on local labor markets, and incorporating it into our quantitative analysis is an important step that we plan to tackle in future research.

REFERENCES

- ADELINO, M., A. SCHOAR, AND F. SEVERINO (2015): "House prices, collateral, and self-employment," *Journal of Financial Economics*, 117(2), 288–306.
- ALMEIDA, H., AND M. CAMPELLO (2007): "Financial Constraints, Asset Tangibility, and Corporate Investment," *Review of Financial Studies*, 20(5), 1429–1460.
- AMARAL, P. S., AND E. QUINTIN (2010): "Limited Enforcement, Financial Intermediation, And Economic Development: A Quantitative Assessment," *International Economic Review*, 51(3), 785–811.
- ASKER, J., A. COLLARD-WEXLER, AND J. D. LOECKER (2014): "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 122(5), 1013 – 1063.
- BAKKE, T.-E., AND T. M. WHITED (2012): "Threshold Events and Identification: A Study of Cash Shortfalls," *The Journal of Finance*, 67(3), 1083–1111.
- BARTELSMAN, E., J. HALTIWANGER, AND S. SCARPETTA (2013): "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 103(1), 305–334.
- BENMELECH, E., N. BERGMAN, AND A. SERU (2010): "Financing Labor," Discussion paper, Northwestern University.
- BENMELECH, E., AND N. K. BERGMAN (2008): "Liquidation Values and the Credibility of Financial Contract Renegotiation: Evidence from U.S. Airlines," *The Quarterly Journal of Economics*, 123(4), 1635–1677.
- BENMELECH, E., M. J. GARMAISE, AND T. J. MOSKOWITZ (2005): "Do Liquidation Values Affect Financial Contracts? Evidence from Commercial Loan Contracts and Zoning Regulation," *The Quarterly Journal of Economics*, 120(3), 1121–1154.
- BLANCHARD, O. J., F. LOPEZ-DE SILANES, AND A. SHLEIFER (1994): "What do firms do with cash windfalls?," *Journal of Financial Economics*, 36(3), 337–360.
- BLOOM, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.
- BRODA, C., AND D. WEINSTEIN (2006): "Globalization and the Gains from Variety," *Quarterly Journal of Economics*, 121, 541–585.
- BUERA, F. J., J. P. KABOSKI, AND Y. SHIN (2011): "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 101(5), 1964–2002.
- CALOMIRIS, C., M. LARRAIN, J.-M. LIBERTI, AND J. STURGESS (2015): "How Collateral Laws Shape Lending and Sectoral Activity," .
- CAMPELLO, M., E. GIMABONA, J. GRAHAM, AND C. HARVEY (2011): "Liquidity Management and Corporate Investment During a Financial Crisis," *Review of Financial Studies*, 24(6).

- CAMPELLO, M., J. R. GRAHAM, AND C. R. HARVEY (2010): "The real effects of financial constraints: Evidence from a financial crisis," *Journal of Financial Economics*, 97(3), 470–487.
- CHANEY, T., D. SRAER, AND D. THESMAR (2012): "The Collateral Channel: How Real Estate Shocks affect Corporate Investment," *American Economic Review*, 102(6), 2381–2409.
- CHETTY, R. (2012): "Bounds on Elasticities with Optimization Frictions: A Synthesis of Micro and Macro Evidence on Labor Supply," *Econometrica*, 80(3), 969–1018.
- CHODOROW-REICH, G. (2013): "The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008-09 Financial Crisis," *Quarterly Journal of Economics*, Forthcoming.
- DAVYDENKO, S. A., AND J. R. FRANKS (2008): "Do Bankruptcy Codes Matter? A Study of Defaults in France, Germany, and the U.K.," *Journal of Finance*, 63(2), 565–608.
- ERICKSON, T., AND T. M. WHITED (2000): "Measurement Error and the Relationship between Investment and q ," *Journal of Political Economy*, 108(5), pp. 1027–1057.
- FAZZARI, S. M., R. G. HUBBARD, B. C. PETERSEN, A. S. BLINDER, AND J. M. POTERBA (1988): "Financing Constraints and Corporate Investment," *Brookings Papers on Economic Activity*, 1988(1), pp. 141–206.
- GAN, J. (2007): "Collateral, debt capacity, and corporate investment: Evidence from a natural experiment," *Journal of Financial Economics*, 85(3), 709–734.
- GOURIO, C., AND A. KASHYAP (2007): "Investment spikes: New facts and a general equilibrium exploration," *Journal of Monetary Economics*, 54, 1–22.
- GOURIO, C., AND J. MIAO (2010): "Firm Heterogeneity and the Long-Run Effects of Dividend Tax Reform," *American Economic Journal: Macroeconomics*, 2, 131–168.
- GUVENEN, F., S. OZKAN, AND J. SONG (2014): "The Nature of Countercyclical Income Risk," *Journal of Political Economy*, 122(3), 612.
- HALL, R. (2004): "On the Nature of Capital Adjustment Costs," *Quarterly Journal of Economics*, 119(3), 899–927.
- HART, O., AND J. MOORE (1994): "A Theory of Debt Based on the Inalienability of Human Capital," *The Quarterly Journal of Economics*, 109(4), pp. 841–879.
- HAYASHI, F. (1982): "Tobin's Marginal q and Average q : A Neoclassical Interpretation," *Econometrica*, 50(1), pp. 213–224.
- HENNESSY, C., AND T. WHITED (2007): "How Costly is External Financing? Evidence From a Structural Estimation," *Journal of Finance*.
- HENNESSY, C. A., AND T. M. WHITED (2005): "Debt Dynamics," *Journal of Finance*, 60(3), 1129–1165.

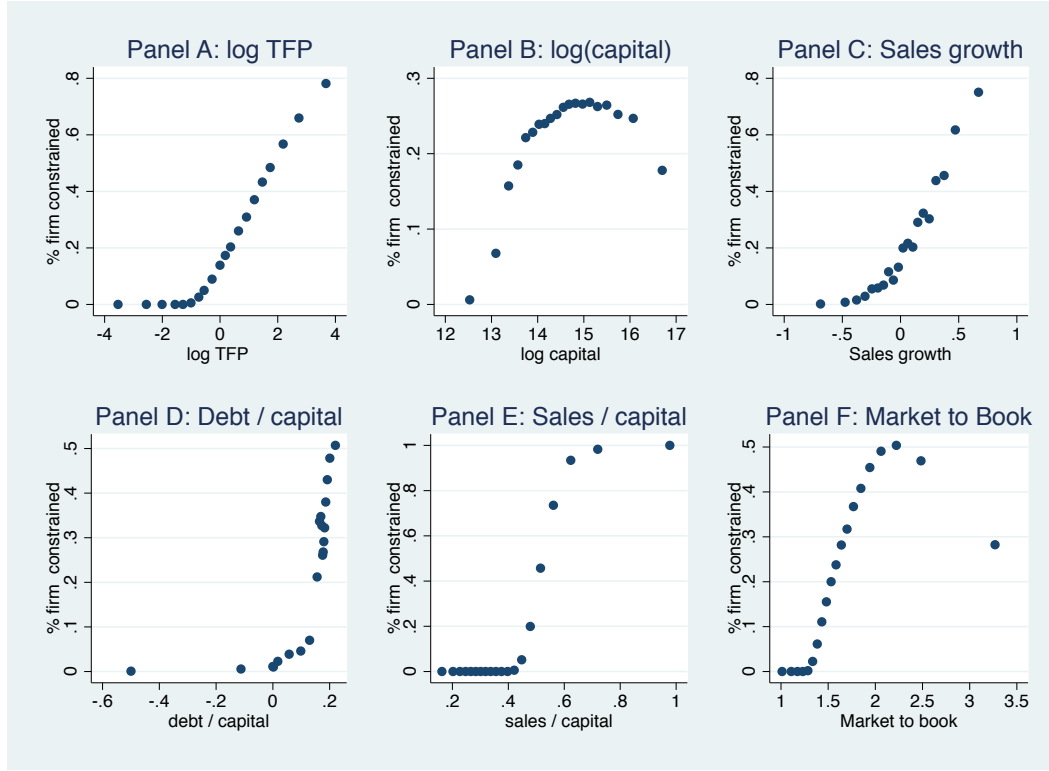
- HSIEH, C.-T., AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124(4), 1403–1448.
- JEONG, H., AND R. TOWNSEND (2007): "Sources of TFP growth: occupational choice and financial deepening," *Economic Theory*, 32(1), 179–221.
- JORGENSEN, D. W. (1963): "Capital Theory and Investment Behavior," *The American Economic Review*, 53(2), pp. 247–259.
- KAPLAN, S. N., AND L. ZINGALES (1997): "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints," *The Quarterly Journal of Economics*, 112(1), 169–215.
- KAPLAN, S. N., AND L. ZINGALES (2000): "Investment-Cash Flow Sensitivities Are Not Valid Measures Of Financing Constraints," *The Quarterly Journal of Economics*, 115(2), 707–712.
- LAMONT, O. (1997): "Cash Flow and Investment: Evidence from Internal Capital Markets," *The Journal of Finance*, 52(1), pp. 83–109.
- LIN, C., Y. MA, AND Y. XUAN (2011): "Ownership structure and financial constraints: Evidence from a structural estimation," *Journal of Financial Economics*, 102(2), 416–431.
- LIU, Z., P. WANG, AND T. ZHA (2013): "Land-Price Dynamics and Macroeconomic Fluctuations," *Econometrica*, 81(3), 1147–1184.
- LUCAS, R. E. (1967): "Adjustment Costs and the Theory of Supply," *The Journal of Political Economy*, 75(4), 321–334.
- MIDRIGAN, V., AND D. Y. XU (2014): "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, 104(2), 422–58.
- MOLL, B. (2014): "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, 104(10), 3186–3221.
- NIKOLOV, B., AND T. M. WHITED (2014): "Agency Conflicts and Cash: Estimates from a Dynamic Model," *Journal of Finance*, 69(5), 1883–1921.
- PEEK, J., AND E. S. ROSENGREN (2000): "Collateral Damage: Effects of the Japanese Bank Crisis on Real Activity in the United States," *American Economic Review*, 90(1), 30–45.
- RAUH, J. D. (2006): "Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans," *The Journal of Finance*, 61(1), pp. 33–71.
- RESTUCCIA, D., AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11(4), 707–720.
- RIDDICK, L. A., AND T. M. WHITED (2009): "The Corporate Propensity to Save," *Journal of Finance*, 64(4), 1729–1766.

SCHMALZ, M., D. SRAER, AND D. THESMAR (Forthcoming): "Housing Collateral and Entrepreneurship," *Journal of Finance*.

STREBULAEV, I. A., AND T. M. WHITED (2012): "Dynamic Models and Structural Estimation in Corporate Finance," *Foundations and Trends(R) in Finance*, 6(1–2), 1–163.

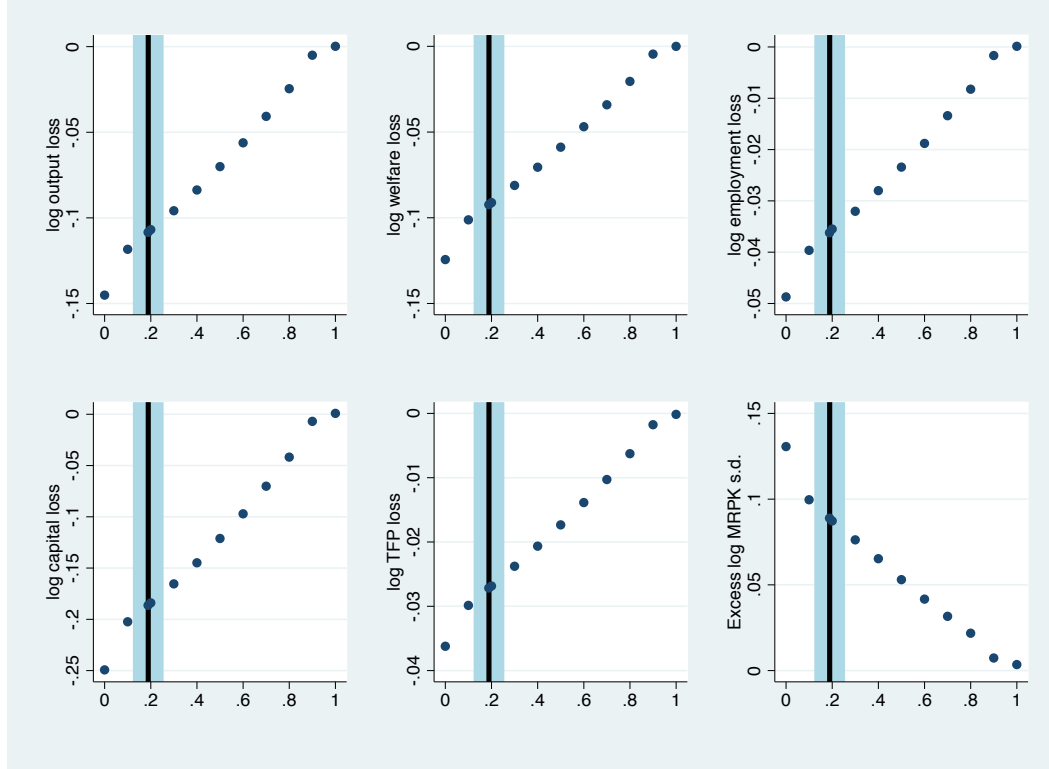
Figures

Figure 2.1: Financing constraints as a function of firm characteristics



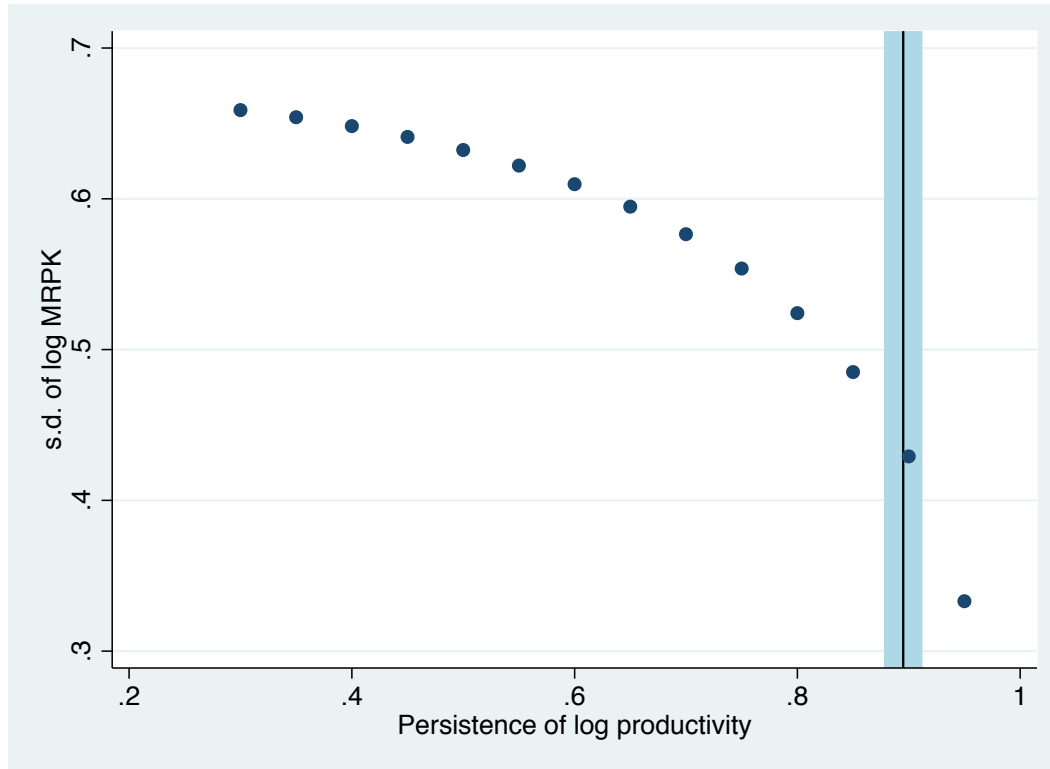
Note: This Figure shows how the extent of financing constraints covaries with firm characteristics, in the cross-section of simulated firms. We simulate a dataset of 1,000,000 firms over 10 years using parameters from our preferred specification (Table 2.3, Panel A, column 4). We remove the first 90 years to make sure firms are in steady state. For each characteristic x , we then sort firms into 20 quantiles of x , and for each quantile compute the average fraction of constrained firms in our simulated data. We label a firm constrained if its capital stock is less than 80% of its unconstrained capital stock. Unconstrained capital stock is computed after solving the same model, with the same parameters but without the no equity issuance constraint. The conditioning variable x is given by z (Panel A), $\log k$ (Panel B), $\log pq_t - \log pq_{t-1}$ (Panel C), $\frac{d}{k}$ (Panel D), $\frac{pq}{k}$ (Panel E), and $\frac{V}{k}$ (Panel F).

Figure 2.2: General equilibrium effect of pledgeability s



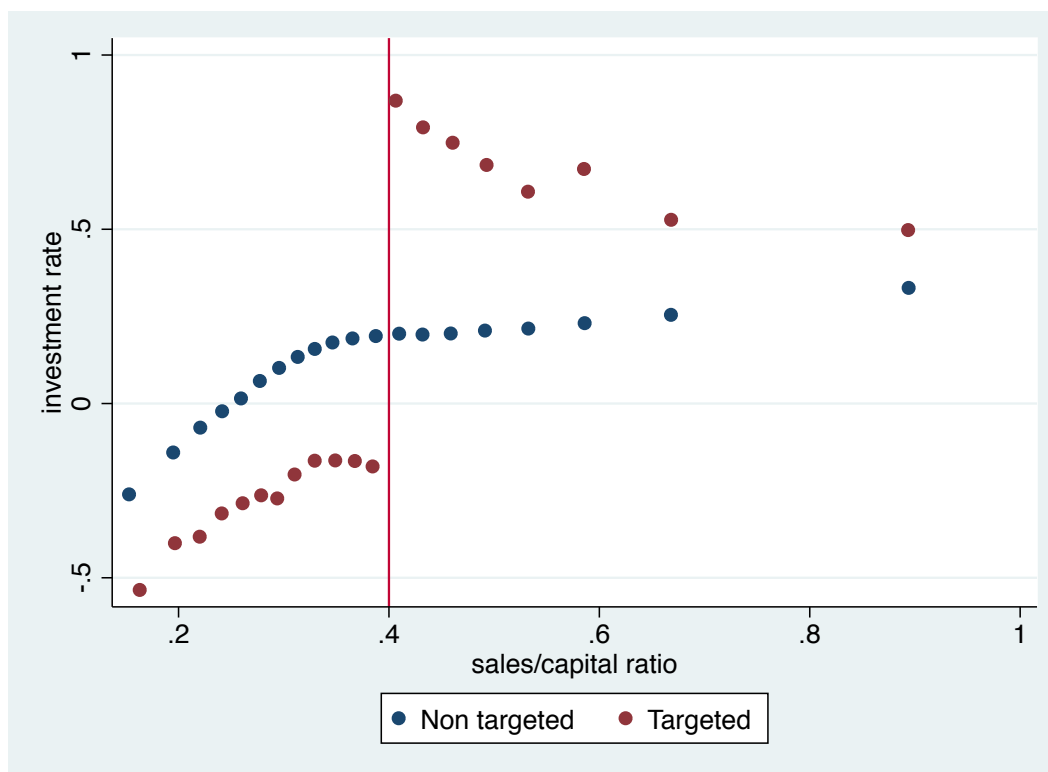
Note: This figure reports the general equilibrium effects of changing the collateral parameter s from 0 (full financial constraints) to 1 (100% of the capital stock can be pledged to lenders). We use the model with adjustment costs and estimated targeting the investment sensitivity moment (thus using the parameters reported in Table 2.3, Panel A, column 4). All aggregates are represented in deviation with respect to the unconstrained benchmark: For each value of s , we compute the general equilibrium of the economy populated with constrained firms, and also the GE of the economy populated by firms with the same parameters, but without the no equity issuance constraint. We then compute the log difference of output, welfare, employment, capital stock, TFP and the difference in the s.d. of log sales to capital ratio (MRPK). We then try all values of s from 0 to 1, spaced by .1. The vertical red line correspond to the SMM estimate of s (.189). Reading: When s increases from .1 to .6, the loss of log capital stock w.r.t. the unconstrained benchmark goes from -.2 to -.1.

Figure 2.3: Productivity persistence ρ and the dispersion of capital productivity



Note: This figure reports the effect on capital misallocation of changing the log productivity persistence ρ from 0.35 (low persistence) to .95 (high persistence). We use the model with adjustment costs and estimated targeting the investment sensitivity moment (Table 2.3, Panel A, column 4). Following [Hsieh and Klenow \(2009\)](#), we measure misallocation as the s.d. of log sales to capital ratio (MRPK).

Figure 2.4: Non targeted versus targeted investment subsidy



Note: This figure shows the relation between the sales to capital ratio and investment for two types of subsidies on investment – a non targeted 10% subsidy for all firms, and a targeted 10% subsidy aimed only at capital poor firms, i.e. firms with a high MRPK ($p_i q_i / k_i > 0.4$). The data is simulated using our SMM parameter estimates from Table 2.3 panel A column 4.

Tables

Table 2.1: **Summary statistics: COMPUSTAT Extract**

	Mean	s.d.	Obs
Investment _{it} / k_{it-1}	.37	.42	20,074
Net borrowing / k_{it-1}	.05	.48	19,998
Real estate value _{it}	.77	1.27	20,074
$\frac{1}{k_{it}}$	.42	.65	20,074
Office price	.67	.21	20,074

Source: COMPUSTAT for accounting items and Global RealAnalytics for office prices. The construction of this data is described in detail in [Chaney, Sraer, and Thesmar \(2012\)](#). The dataset is an extract of COMPUSTAT. It contains all firms present in 1993 who report accounting value and cumulative depreciation of land and buildings. These firms are then followed until they exit the sample or until 2006. We also require that office price data are available in the city where these firms have their headquarter in 1993. The variables shown are used in the two regressions presented in Section [2.1](#).

Table 2.2: Elasticity of Moments with respect to Parameters

	s.d. $\Delta \log q$	s.d. $\Delta_5 \log q$	$\frac{d_t}{k_t}$	$\beta(Inv, RE)$	$corr(\frac{i_t}{k_{t-1}}, \frac{i_{t+1}}{k_t})$	$\beta(Debt, RE)$
Pledgeability s	.077	.16	1.3	-1.3	-.044	-.99
Adjustment cost c	-.041	-.063	.34	.014	.42	.011
Volatility σ	.97	.92	-1.4	-.48	-.15	-.76
Persistence ρ	.081	1	-2.2	-.99	2	-2

Note: This table reports the elasticity of various moments with respect to the structural parameters that we estimate. First, we start with the SMM estimate $\hat{\Omega}$ of the parameters Ω . For each $k = 1, \dots, 4$, we set $\omega_l = \hat{\omega}_l$ for all $l \neq k$, and vary the parameter ω_k around the estimated $\hat{\omega}_k$ in order to compute the elasticity of moments to parameters in the vicinity of the SMM estimate. For each moment m_n , we compute

$$\epsilon_{n,k} = \frac{m_n^+ - m_n^-}{\omega_k^+ - \omega_k^-} \times \frac{\hat{\omega}_k}{\hat{m}_n} \approx \frac{\partial \log(\hat{m}_n)}{\partial \log(\hat{\omega}_k)},$$

where $\hat{m}_n$ is the n^{th} data moment. m_n^+ is the moment based on data simulated with parameter $\hat{\omega}_k^+$. Likewise, m_n^- is the average of moments based on data simulated with parameters ω_k^- . ω_k^+ and ω_k^- are parameter values right above and right below the SMM estimate $\hat{\omega}_k$, when the interval of definition of ω is graded on a scale going from 0 to 10 as in Figures 2.5-2.8. *Reading:* Around the SMM estimate, a 1% increase in s is associated with a 1.3% decrease in the sensitivity of investment to real estate and a 1.3% increase in leverage.

Table 2.3: **Parameter estimates (SMM)**

	(1)	(2)	(3)	(4)	(5)
	No adj cost, Lev. target	No adj cost, Inv. target	Adj cost, Lev. target	Adj cost, Inv. target	Data
<i>Panel A: Estimated Parameters</i>					
ρ	0.917 (0.011)	0.919 (0.008)	0.865 (0.008)	0.895 (0.008)	
σ	0.623 (0.010)	0.725 (0.017)	0.818 (0.013)	0.820 (0.012)	
s	0.495 (0.024)	0.133 (0.030)	0.422 (0.020)	0.189 (0.014)	
c			0.050 (0.003)	0.045 (0.003)	
<i>Panel B: Moments (targeted in bold)</i>					
Std of 1-year sales growth	0.327	0.327	0.327	0.327	0.327
Std of 5-year sales growth	0.909	0.910	0.910	0.911	0.911
Real-Estate to assets	0.140	0.140	0.140	0.140	0.140
Net debt to assets	0.300	0.013	0.315	0.095	0.313
$\beta(Inv, RE)$	0.126	0.038	0.082	0.040	0.040
Autocorrelation of investment	-0.057	0.064	0.436	0.436	0.436
$\beta(Debt, RE)$	0.124	0.037	0.084	0.038	0.039

Note: This table reports the results of our SMM estimations. The estimation procedure is described in the text and in Appendix 2.A. Columns (1)-(4) correspond to SMMs using different sets of parameters and targeting different sets of moments. Columns (1) and (2) assume not adjustment cost ($c = 0$), while Columns (3) and (4) introduce adjustment costs to the model. Estimations reported in Columns (1) and (3) target the short- and long-term volatilities of log sales, mean leverage, and the autocorrelation of investment. Columns (2) and (4) target the sensitivity of investment to real estate prices instead of mean leverage. For each of these estimations, Panel A shows the estimated parameters, along with standard errors (obtained via bootstrapping) in parenthesis. Panel B shows the value of a set of moments, measured on simulated data (with 1,000,000 observations). Moments in bold are the ones that are targeted by the estimation.

Table 2.4: **Aggregate effects of collateral constraints**

	(1)	(2)	(3)	(4)
	No adj cost, Lev. target	No adj cost, Inv. target	Adj cost, Lev. target	Adj cost, Inv. target
<i>Panel A: Targeted moments</i>				
Std of 1-year sales growth	Y	Y	Y	Y
Std of 5-year sales growth	Y	Y	Y	Y
Real-Estate to assets	Y	Y	Y	Y
Net debt to assets	Y	N	Y	N
$\beta(Inv, RE)$	N	Y	N	Y
Autocorrelation of investment	N	N	Y	Y
<i>Panel B: Loss from financial constraint in general equilibrium</i>				
log(TFP)	0.015	0.034	0.015	0.027
log(output)	0.081	0.160	0.061	0.110
log(wage)	0.054	0.106	0.040	0.073
log(L)	0.027	0.053	0.020	0.036
log(K)	0.157	0.296	0.107	0.192
log(welfare)	0.063	0.131	0.051	0.094

Note: This table reports the results of the general equilibrium counterfactual analysis for different SMM parameter estimates. The general equilibrium analysis is described in Section 2.4 and the procedure detailed in Appendix 2.B. Columns (1)-(4) correspond to parameters from SMMs assuming different parameter restrictions and targeting different sets of moments. Columns (1) and (2) assume not adjustment cost ($c = 0$), while Columns (3) and (4) allow for them. Parameters in Columns (1) and (3) correspond to SMMs which target “classic” moments, including mean leverage, while Columns (2) and (4) target the sensitivity of investment to real estate value instead of mean leverage. For each one of these estimations, panel A simply recalls the targeted moments. Panel B reports the result of the GE counterfactual analysis. All results are shown as log deviations with respect to the unconstrained benchmark. The unconstrained benchmark correspond to an equilibrium where firms face the same set of parameters as in the SMM estimate – as reported in the same column, Table 2.3, panel A – but no constraint on equity issuance. In this unconstrained benchmark, investment reaches first best. *Reading:* In column 1 (targeted leverage, no adjustment cost), the aggregate TFP loss compared to a benchmark without financing constraints is $e^{0.015} \approx 1.5\%$.

Table 2.5: Estimating a constrained model on unconstrained data

	(1)	(2)	(3)
	Leverage	Investment	Unconstrained
	target	target	simulated model
<i>Panel A: Parameters</i>			
ρ	0.943	0.900	0.895
σ	0.886	0.811	0.820
s	0.436	0.953	0.189
c	0.042	0.042	0.045
Cost of equity	$+\infty$	$+\infty$	0
<i>Panel B: Moments (matched in bold fonts)</i>			
Std of 1-year sales growth	0.377	0.366	0.367
Std of 5-year sales growth	1.164	1.178	1.171
Real-Estate to assets	0.133	0.156	0.152
Net debt to assets	0.167	0.885	0.168
$\beta(Inv, RE)$	0.037	0.003	-0.001
Autocorrelation of investment	0.420	0.431	0.426
$\beta(Debt, RE)$	0.034	0.361	0.069
<i>Panel C: Loss from financial constraint in general equilibrium</i>			
log(TFP)	0.031	0.000	0
log(output)	0.130	0.002	0
log(wage)	0.086	0.001	0
log(welfare)	0.109	0.002	0

Note: This table reports the result of our SMM estimation on a synthetic dataset simulated by a model *without* financing friction. We start with our baseline parameters (Table 2.3 Panel A Column (4)). We remove the no equity constraint, and simulate a synthetic panel dataset of unconstrained firms. We compute various moments and report them in column 3. We then perform two SMM estimations of a model *with* no equity issuance constraint. The estimation procedure and general equilibrium analysis are described in the text and in Appendices 2.A and 2.B. In Column (1), we match short- and long-term log sales volatility, the autocorrelation of investment, and mean leverage. In Column (2), we match short- and long-term log sales volatility, the autocorrelation of investment, and the sensitivity of investment to real estate value. Panel A reports the estimated parameters (to be compared with the true parameters used for the simulation in Column (3)), Panel B the moments, and Panel C computes the implied GE losses from financial constraints.

Table 2.6: **Robustness: unsecured debt and costly equity issuance**

	(1)	(2)	(3)	(4)
	Baseline	Unsecured Debt Capacity	Costly Equity Issuance	Data
<i>Panel A: Estimated Parameters</i>				
ρ	0.895	0.877	0.868	
σ	0.820	0.821	0.806	
s	0.189	0.254	0.235	
$\bar{d}$	-	0.450	-	
c	0.045	0.052	0.048	
<i>Panel B: Targeted moments</i>				
Std of 1-year sales growth	0.327	0.327	0.324	0.327
Std of 5-year sales growth	0.911	0.906	0.920	0.911
Real-Estate to assets	0.140	0.140	0.145	0.140
Net debt to assets	0.095	0.300	0.181	0.313
$\beta(Inv, RE)$	0.040	0.040	0.039	0.040
Autocorrelation of investment	0.436	0.439	0.458	0.436
$\beta(Debt, RE)$	0.038	0.039	0.054	0.039
<i>Panel C: Loss from financial constraint in general equilibrium</i>				
log(TFP)	0.027	0.027	0.013	
log(output)	0.110	0.122	0.058	
log(wage)	0.073	0.081	0.038	
log(welfare)	0.094	0.100	0.048	

Note: This table reports the SMM estimation result and GE counterfactual experiments of two alternative versions of our baseline model. In Column (2), we assume that in addition to collateralized debt, the firm has access to an extra fixed debt capacity ($\bar{d}$) as in equation (2.13). In column 3, we relax the zero equity constraint, and allow for costly equity issuance, with a 15% variable cost. The estimation procedure and general equilibrium analysis are described in the text and in Appendices 2.A and 2.B. We target the following moments: short- and long-term volatilities of log sales, the autocorrelation of investment, investment sensitivity to real estate value, and mean leverage in Column (2) only. Panel A shows the estimated parameters. Panel B shows the value of simulated moments. Moments in bold are the ones that are targeted by the estimation. Panel C reports the result of a GE counterfactual experiment. All results are shown as % losses with respect to the unconstrained benchmark. The unconstrained benchmark correspond to an equilibrium where firms face the same set of parameters as in the SMM estimate – as reported in Panel A – but no constraint on equity issuance.

Table 2.7: Macro effect of an investment subsidy

	(1)	(2)	(3)	(4)	(5)	(6)
	<i>Untargeted</i>			<i>Targeted</i>		
Subsidy (share of investment)	5%	10%	15%	5%	10%	15%
Aggregate subsidy (% of output)	.007	.015	.024	.027	.035	.044
$\Delta \log$ Output	.043	.089	.14	.069	.093	.12
$\Delta \log$ Capital	.11	.23	.36	.13	.18	.24
$\Delta \log$ Labor	.014	.03	.046	.023	.031	.039
$\Delta \log$ TFP	0	0	-.001	.014	.016	.017
$\Delta \log$ Welfare	.029	.059	.089	.058	.074	.091

Note: This Table reports the aggregate equilibrium impact of tax subsidies. We start with the model and parameter estimates of Table 2.3, Column (4). To cash flows of firm i at date t , we add a tax free subsidy equal to xI_{it} where I_{it} is the investment of firm i at date t and x is a fraction equal to 5,10 and 15%. Note that this subsidy becomes a tax when the firm's investment becomes negative. This subsidy is financed by a non distortionary tax on households. In columns 1-3, the tax is not targeted. In Columns (4)-(6), the tax is targeted only towards capital-poor firms, i.e. firms with a high MRPK ($\log(p_i q_i / k_i) > 0.4$). For each one of these six policies, we compute the equilibrium and report the change in log aggregates compared to the case without subsidy. For instance, we find that giving firms a non-targeted subsidy equal to 5% of their investment leads to an increase in aggregate output of 4.3%.

APPENDIX

This Appendix contains: the method used to to solve and estimate the model (Section 2.A), the method we use to compute the general equilibrium of our model (Section 2.B) and the additional comparative static results in partial equilibrium designed to show that the model is well behaved around the estimate (Section 2.C).

2.A Solving the model and Estimation

This Appendix details the algorithms used to solve the model and estimate it. To estimate the model, one needs to find the set of parameters such that model-generated moments fit a pre-determined set of data moments. Because our model does not have an analytic solution, we need to use indirect inference to perform the estimation. Such inference is done in two steps:

- For a given set of parameters, we need to solve the model numerically, which means solving the Bellman problem (2.6) and obtain the policy function $S_{t+1} = (d_{t+1}, k_{t+1})$ as a function of $S_t = (d_t, k_t)$ and exogenous variables $X_t = (z_t, p_t)$.
- We then use this resolution technique to estimate the parameters that match best a set of moments chosen from the data. We explain the methodology (Simulated Method of Moments) and the numerical algorithm that we use to implement it.

2.A.1 Solving the model numerically

In this section, we describe how we numerically solve the firm's problem with given parameters.

Grid definition

In order to solve the model numerically, we need to discretize the state space $(S; X)$. Let us start with the two exogenous variables. The log productivity process z is discretized using the standard Tauchen method on 51 grid points. Log real estate prices are also an AR(1), discretized using the Tauchen method on 11 grid points. For both variables, we set the bounds of the grid at -2.5 and 2.5 standard deviations.

Capital choice is discretized over a range going from k_{min} to k_{max} . k_{min} is the smallest level of capital chosen by a firm without adjustment costs and financing constraint. For this particular case, we can solve the capital decision analytically. In most cases, this number should serve as a lower bound because adjustment costs would prevent firms from adjusting all the way down to this level; and financial constraints would push them to keep more capital as precautionary savings. Since we did not, however, establish this result analytically, we check that k_{min} is always "far enough" from the lowest simulated value of capital. Similarly, k_{max} is the capital stock chosen by unconstrained firms, without adjustment cost, facing the highest productivity level on the grid. Again, we expect this level to be above the upper bound of capital for a constrained firm with adjustment costs. We check that this is the case in our simulations. We then form an equally spaced grid for log capital between $\log k_{min}$ and $\log k_{max}$, with increment of $\log(1 + \delta/2)$. Thus, the capital grid is geometrically spaced using $(1 + \delta/2)$ as the multiplying coefficient, i.e. the n^{th} point is equal to $k_{min} \times (1 + \delta/2)^n$ until k_{max} . Given that k_{min} and k_{max} are functions of productivity, the grid thus depends on the persistence ρ and volatility σ of log productivity. Larger persistence or volatility leads to wider grid. In our preferred specification, capital evolves on a grid containing 270 points. We will take this number

as a reference when we later discuss grid size, bearing in mind that, in fact, the capital grid is a function of parameter values.

Finally, the debt grid d_t is defined as a function of the amount of capital k_t . This adaptive feature of the debt grid comes from the fact that the amount of debt is bounded above by a function of capital: larger firms can borrow more. We restrict future period debt d' to the $[-4\bar{d}; \bar{d}]$ interval, where $\bar{d} = s((1 - \delta)k + p_{max}h)$ and p_{max} is the maximum house price level. The grid interval is thus a function of the model parameters s but also ρ and σ via the grid of k . The upper bound is a natural consequence of the collateral constraint: the model imposes that it cannot be exceeded. The lower bound is somewhat arbitrary as there is in theory no upper bound as to how much cash the firm may decide to hold. We check that there is no accumulation of cash at this bound during the estimation process. Within this interval, the grid is geometrically spaced so that it is more dense when debt becomes closer to the constraint, i.e. right below $\bar{d}$. We implement this by setting the n^{th} grid point at $\bar{d}(1 - 0.001 \times e^{3n})$ until it reaches $-4\bar{d}$. Thus, the grid size for debt does not depend on parameters (in contrast to the capital grid size) and always has 29 points.

Bellman resolution algorithm

We solve the firm's problem using policy iteration. This algorithm is based on the fact that the value function is the solution of a fixed point problem generated by a contraction mapping.

Before starting to iterate, we compute profit flows $e(S, S'; X)$ using the production and cost functions, for all possible values of S and X on the grid. We set e to "missing" when $(S, S'; X)$ are such that $e < 0$ – the no equity issuance constraint is violated, or when the borrowing constraint is violated. Profits are only defined when both financing constraints are satisfied.

To initiate the process, we start with the value function $V_0(S; X) = 1$. We then look for the policy function $(k'_0, d'_0) = P_0(S; X)$ which solves:

$$P_0(S; X) = \operatorname{argmax}_{S'} \{e(S, S'; X) + \frac{1}{1+r}\}$$

for each state of $(S; X)$. Then, we iterate the following loop (where $n \geq 1$ denotes the step in the loop):

1. Start from $(k'_{n-1}, d'_{n-1}) = P_{n-1}(S; X)$, the policy function obtained from the previous round; and $V_{n-1}(S; X)$, the value function obtained from the previous round. For every point $(S; X)$ on the grid, we compute the value function V_n that satisfies:

$$\begin{aligned} V_n(S; X) = & e(S, P_{n-1}(S; X); X) + \frac{1-d}{1+r} \mathbb{E}_{X'} [V_{n-1}(P_{n-1}(S; X'); X') | X] \\ & + \frac{d}{1+r} (k'_{n-1} - (1 + \tilde{r}_t)d'_{n-1}) \end{aligned}$$

2. We then use the new value function V_n and compute the optimal policy given this value function ($P_n(S; X)$):

$$P_n(S; X) = \operatorname{argmax}_{S'} \{e(S, S'; X) + \frac{1-d}{1+r} \mathbb{E}_{X'} [V_n(S'; X') | X'] + \frac{d}{1+r} (k' - (1 + \tilde{r}_t)d')\}$$

3. We stop when $P_n = P_{n-1}$.

Thanks to the contraction mapping theorem, we are guaranteed to find a good approximation of the value function $V(S; X)$ and the policy function $S' = P(S; X)$ defined over the grid. The computationally costly step is the determination of the policy function in step 2 with respect to S' . This consists of $29 \times 270 \times 51 \times 11 = 4,392,630$ optimizations of vectors with $29 \times 270 = 7,830$ points. This is where parallelization achieved through a GPU accelerates the process. For the range of parameters we explore, we typically solve the model in about 2 minutes with a GPU (Nvidia K80), compared to several hours with a CPU. What prevents us from having a finer grid is the RAM of the GPU, since the computer needs to create the maximand in step 2, a $29 \times 270 \times 29 \times 270 \times 51 \times 11 \approx 34$ billion numbers array.

The above algorithm is the standard policy function iteration algorithm. We make two adjustments to adapt it to our setting. First, in order to reduce computing time, we first solve the model with a coarser grid, and then solve it again on the grid describe above. To define this coarser grid, we divide the resolution of the control (k and d) grids by two. This divides computing time by four in the first step but only gives us the value and policy functions on the coarser grid. We then re-run the algorithm on the finer grid with the “coarser” policy and value functions as starting point. Convergence occurs much more quickly.

The other adjustment is related to the treatment of missing values, which in our set-up occur when one of the two financing constraints are violated (i.e. the no-equity constraint or the collateral constraint). Without modification, the policy iteration algorithm does not behave well in the presence of missing values. This is because, for some given value functions V_{n-1} , there may exist some (S, X) for which there is no acceptable policy S' . In this case, the optimal policy function $P_n(S, X)$ is not defined everywhere on the grid (note (S_0, X_0) such states for which the policy is not defined). When this happens, the next iteration value $V_n(S; X)$ is non-defined for all (S, X) , which leads with non-zero probability to states (S_0, X_0) . As we iterate, missing value progressively spread to the entire grid and the algorithm is blocked. To solve this problem, we modify step 1. of the algorithm by requiring that $V_n(S; X)$ replaces $V_{n-1}(S; X)$ if and only if $V_n(S; X)$ is non missing. This prevents missing values from spreading to the entire grid of states $(S; X)$.

2.A.2 Estimation

We now proceed to estimate the parameters (s, c, ρ, σ) for which the model best matches a predefined set of moment (we experiment with different set of moments and models in the main text).

Estimation method: SMM

We estimate the key parameters of the model by simulated method of moments (SMM), which minimizes the distance between moments from real data and simulated data. Let us call $\mathbf{m}$ the vector of moments computed from the actual data, and let us call Ω the moments generated by the model with parameters Ω . The SMM procedure searches the set of parameters that minimizes the weighted deviations between the actual and simulated moments,

$$(\mathbf{m} - \hat{\mathbf{m}}(\Omega))' \mathbf{W} (\mathbf{m} - \hat{\mathbf{m}}(\Omega)) \quad (2.14)$$

We detail the various components of our implementation in the following sections.

Empirical moments $\mathbf{m}$ and Weight matrix $\mathbf{W}$

The empirical moments are computed in a simple way, and the definitions are given in the main text, in Section 2.3.3.

The weight matrix $\mathbf{W}$ adjusts for the fact that some moments are more precisely estimated than others. It is computed as the inverse of the variance-covariance matrix of actual moments estimated by bootstrap with replacement on the actual data. To compute the elements of this matrix, we repeat 100 times the following procedure. Using our dataset, we draw, with replacement, N firms with their entire history where N is the number of firms in the sample (we use the `bsample` command in Stata, clustered at the firm level). We then compute the moments, and store them. Once we have performed this procedure 100 times, we compute the empirical variance-covariance matrix of the moments, and invert it.

Model-generated moments $\mathbf{m}$

Once we have solved the model for a given set of parameter Ω (Appendix 2.A.1), we need to simulate data in order to compute the simulated moments. We simulate a balanced panel of 1,000,000 firms over 100 years, and only keep the last 10 years to ensure each firm has reached steady state. For each firm, we simulate a path of log productivities z_{it} and a path of log real estate prices p_{it} . This makes the variability of real estate prices larger than in the data, where prices only vary at the city (MSA) level. Recall however that our objective in this simulation is not to replicate the variability of the data, but ideally to estimate model-generated moments. If we had closed forms for the model, we could measure these moments without infinite precision. The problem here comes from the fact that we cannot directly compute these moments but have to “estimate” them. Ideally, we would want to generate an infinitely large simulated dataset in order to compute the model-generated moments exactly, but computational constraints make it infeasible. 100,000 firms over 100 years already generate arrays with 10m entries. Allowing real estate prices to vary at the firm-level is a way to make sure the sensitivity to prices model-generated moments are estimated as well as possible.

Optimization algorithm

We now have all the ingredients necessary to compute the objective function (2.14). In this Section, we explain how to minimize it. Since in our most preferred specification we have 5 parameters, we need to make sure that we are indeed reaching a global minimum. We do this by implementing the following two-step procedure, which follows Guvenen, Ozkan, and Song (2014):

- We generate 1,000 quasi-random vectors of parameters Ω taken from a Halton sequence. The Halton sequence is a deterministic sequence of numbers that has the property of covering the parameter space evenly. For each of these parameters, we solve the model to obtain the policy function, simulate a dataset, compute the moments and therefore the distance to data moments (2.14).
- We then use the lowest points (in terms of objective function) as starting points for minimization. We iterate on the following loop. We begin with parameter estimate $\hat{\Omega}_1$ for which the objective function is the lowest. We then use the Nelder-Mead method (command *fminsearch* in Matlab) to perform a local optimization starting from this point. We then compute the objective function O_1 . We then move to the second lowest parameter estimate ($\hat{\Omega}_2$) and compute the objective function O_2 . We iterate on this, and stop as soon as $O_n = 0$. Among the lowest parameters, a large fraction typically leads to the same parameters for which the objective function is equal to 0. This gives an indication that our objective function is well-behaved.

There is no general theoretical results arguing that this technique dominates other popular algorithms adapted for large dimension optimization. In our setting however, we found that the genetic algorithm and simulated annealing were much slower at converging. Also, this approach allows to “control” the smoothness of the objective function. For instance, within the lowest 20 parameters isolated after step 1., it would be worrisome if minimizations starting from each of these parameters gave inconsistent parameters. On the contrary, they tend to be very consistent. The only cases where convergence goes to alternative choice of parameters than the one we present are cases where the objective function is much bigger than zero (i.e. other local optima). Finally, the best argument in favor of our selected estimates is the well-behaved comparative statics we present in Appendix 2.C.

Standard errors

We estimate our standard errors using a block-bootstrap procedure. As for the computation of the variance-covariance matrix, we start by generating $B = 100$ datasets of N firms drawn without replacement from the data, and then compute the vector of targeted moments for each dataset. To preserve the panel structure we make sure to draw firms and not observations (hence the “block” in block-bootstrapping). The result is a set of 100 vectors m_b , for each of whom we seek the vector of model parameters Ω_b that minimizes

$$f_b(\Omega_b) = (m_b - \hat{m}(\Omega_b))' W (m_b - \hat{m}(\Omega_b)). \quad (2.15)$$

To reduce computing time we estimate the 100 parameters $\Omega_{\mathbf{b}}$ in parallel. We use the following algorithm. We define a new objective function as the sum of all 100 objective functions, that is

$$F(\Omega_1, \dots, \Omega_B) = \sum_{b=1}^B f_b(\Omega_{\mathbf{b}}). \quad (2.16)$$

1. $\hat{M}$ is initialized using our SMM estimate $\hat{\Omega}$. As a result, each parameter $\Omega_{\mathbf{b}}$ is equal to $\hat{\Omega}$ (so they are all identical). Let b^* be the sample for which $f_b(\hat{\Omega})$ is that highest. This corresponds to the bootstrapped sample for which the main SMM estimate fits the moments the worst.
2. We use the Nelder-Mead simplex algorithm to improve the estimate $\Omega_{\mathbf{b}^*}$ of the least well matched sample b^* . Specifically, we use Matlab `fminsearch` function with the following options:
 - The initial simplex Δ_{b^*} is computed using the current estimate of $\Omega_{\mathbf{b}^*}$ as an “initial guess”
 - The local optimization is stopped as the soon as b^* is no longer the sample with the worst fit.
 - If `fminsearch` reaches a maximum of 50 iterations, Δ_{b^*} is reinitialized using the best available estimate of $\Omega_{\mathbf{b}^*}$ as an “initial guess”.

We then use the outcome of this procedure to update the parameter estimate of sample b^* in the list $\hat{M}$

3. For each vector m_b , we find in $\hat{M}$ the vector $\Omega_{\mathbf{b}}$ that minimizes $f_b(\Omega_{\mathbf{b}})$. We then find the new sample b^* for which the objective function f_{b^*} has the highest value.
4. If the standard deviations of $\Omega_{\mathbf{b}}$ have moved by less than 1% over the last 500 evaluations, and if the value of F is less than one tenth of its initial value, then the procedure stops. Otherwise, it goes back to step 2.

Standard errors of Ω are estimated using the standard deviation of the $\Omega_{\mathbf{b}}$. The fact the value of F is divided by at least ten indicates that the dispersion of $\Omega_{\mathbf{b}}$ is sufficient to explain 90% of the (weighted) dispersion of $\Omega_{\mathbf{b}}$. To reach that point, our procedure typically takes the equivalent of 2-3 SMMs to converge, and is thus about 30 times faster than running all 100 SMMs sequentially.

2.B General Equilibrium Computation

In this Section, we describe how we compute the general equilibrium of an economy populated by firms whose behavior is described by the model estimated and solved in Appendix 2.A. First, recall that this model is estimated assuming aggregate demand $Q = 1$ and aggregate wage $w = 0.03$.

The economy is described in detail in Section 2.4 in the main text. There is a large number of firms (a continuum in the model), each of them facing an idiosyncratic path of productivity and of real estate prices. The behavior of each of these firms is described by the dynamic model with adjustment costs, time-to-build capital, the collateral constraint and the no-equity constraint. All firms are monopolists that produce intermediate inputs combined in a CES-aggregate with elasticity of substitution ϕ . As a result ϕ measures the intensity of competition between intermediate producers ($\phi = +\infty$ means perfect competition). The final good is then consumed by a representative producer with linear utility, Frisch elasticity of labor supply ϵ and subjective discount rate r . Consumption equals production minus adjustment costs and investment. The price of the final good is normalized to 1 without loss of generality. This economy has no aggregate uncertainty and the equilibrium is uniquely described by aggregate production Q and real wage w , which are fixed over time.

Start from a set of SMM estimates $\hat{\Omega}$. Our goal is to investigate the GE consequences of a change in parameter Ω from its estimated value $\hat{\Omega}$ to another Ω' . This change affects firm's behavior, hence aggregate labor demand and aggregate production. This, in turn, affects the wage and aggregate demand which, in turn, changes firm behavior. The following algorithm finds the fixed point of this problem such that: (1) aggregate production of all firms equal aggregate demand Q in firms' problems and (2) the labor market clears such that aggregate labor demand equals labor supply at prevailing wage. Our approach broadly consists of postulating a given equilibrium (Q_n, w_n) , then check if aggregate labor and product supply given these values is above or below (Q_n, w_n) . We then adjust (Q_{n+1}, w_{n+1}) accordingly. This approach assumes that there is a unique fixed point and that the contraction mapping theorem applies in our setting.

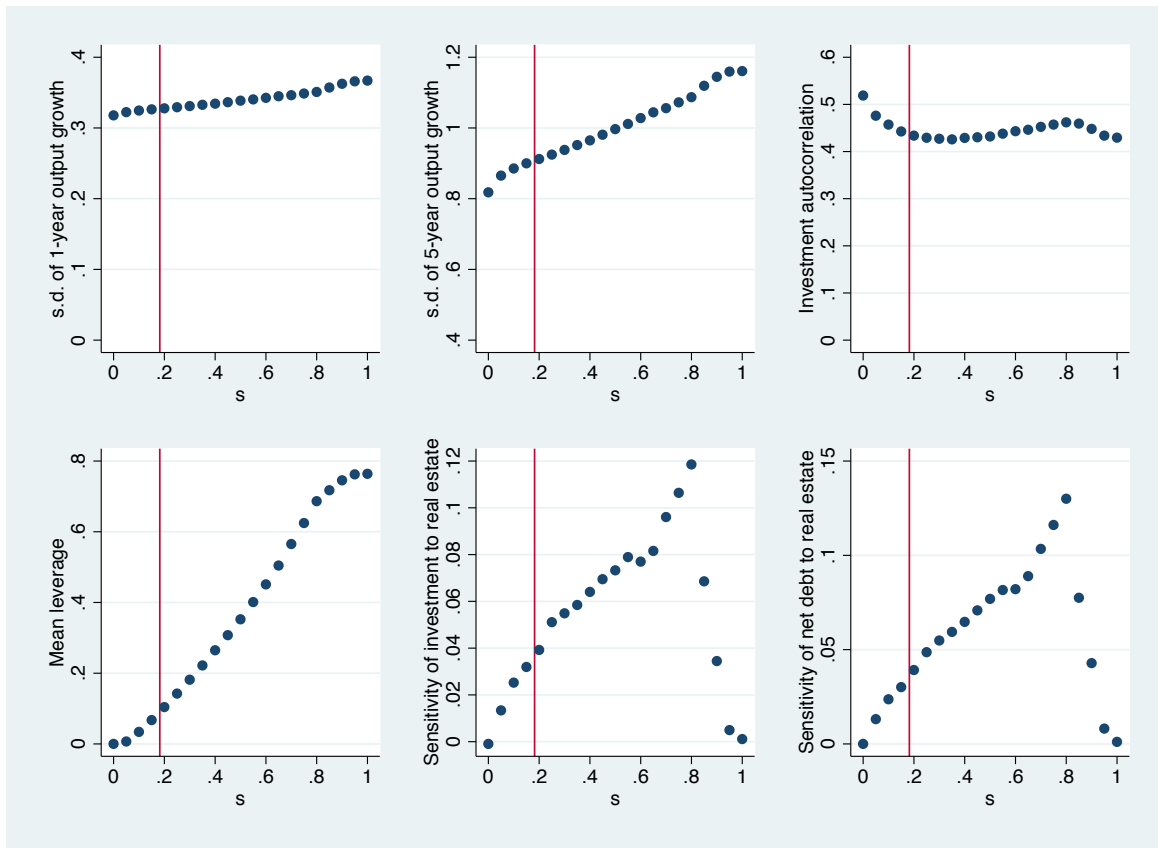
Formally, we proceed in three main steps:

1. Find the number of firms N and the labor supply L_0 at wage $w_0 = .03$, so that the estimated model is at equilibrium with wage $w_0 = 0.03$ and aggregate production $Q = 1$. This will become part of the structure of the economy.
 - (a) Simulate the data with 100,000 firms, $w = .03$, $Q = 1$ and parameters $\hat{\Omega}$.
 - (b) Compute *mean* labor demand l and *mean* revenue pq .
 - (c) Set $N = \frac{1}{pq}$ and $L_0 = \frac{l}{pq}$. With such parameters, the economy with N firms and labor supply parameter L_0 is at equilibrium with $w_0 = 0.03$ and $Q = 1$.
2. Change one of the parameters to its new value Ω' . Given this, we loop to find the new equilibrium w and Q .
 - (a) Set $w_0 = 0.03$ and $Q_0 = 1$.

- (b) Initiate round number $n = 1$. Then,
- i. Solve the model with w_{n-1} and Q_{n-1} and simulate 100,000 firms.
 - ii. Compute average revenue $p q_n$ and average labor demand l_n and multiply both by N to obtain aggregate production Q_n^* and aggregate labor demand L_n .
 - iii. Compute labor market clearing wage $w_n^* = w_0 (L_n / L_0)^{1/\epsilon}$
 - iv. Take $w_n = (w_{n-1})^\lambda (w_n^*)^{1-\lambda}$ and $Q_n = (Q_{n-1})^\lambda (Q_n^*)^{1-\lambda}$
 - v. go back to step (iii), until convergence in Q and w .
- (c) compute aggregates:
- $Q, w, K = \sum_i k_i, L = \sum_i l_i, \text{Adj. Cost}_t = \sum_i i_{it}^2 / k_{it}$.
 - $\log \text{TFP} = \log Y - \alpha \log K - (1 - \alpha) \log L$.
 - $\text{Welfare} = (Q - \delta K - \text{Adj. Cost}) - \frac{\bar{L} w^{1+\epsilon}}{1+1/\epsilon}$

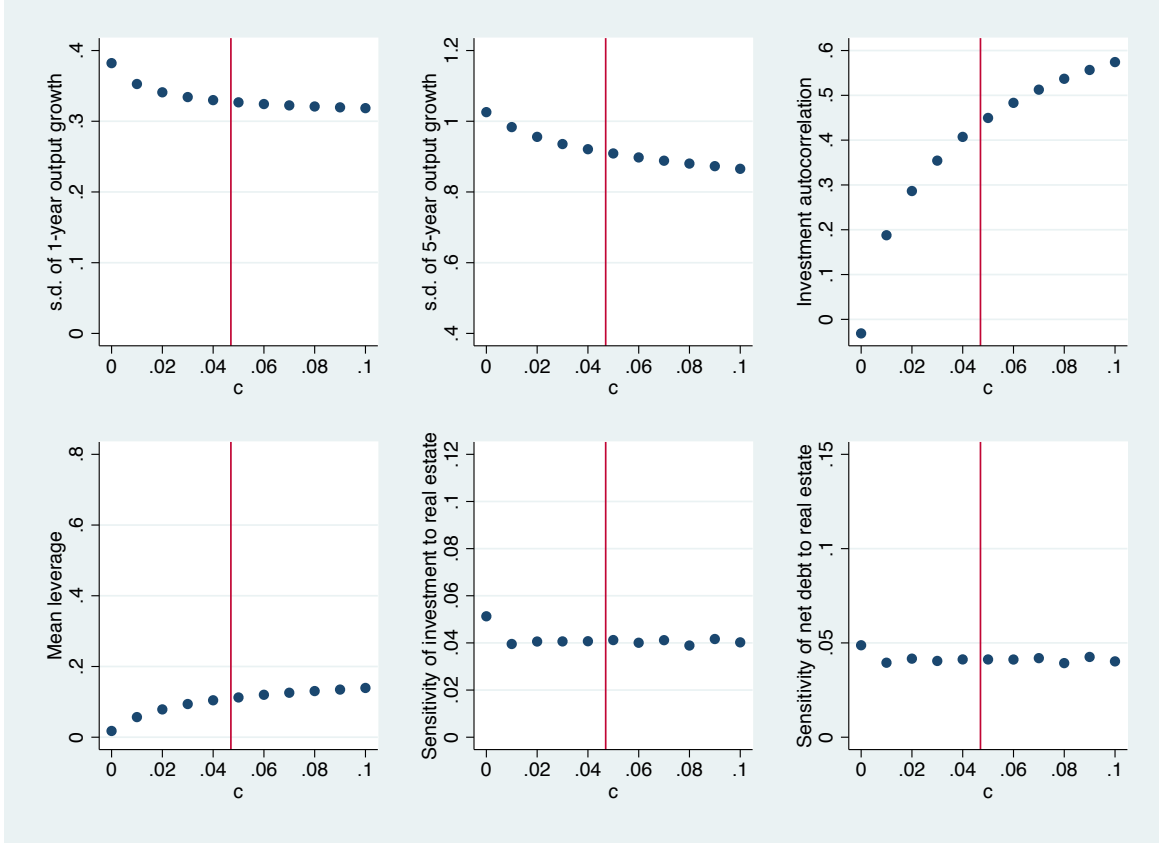
2.C Additional figures

Figure 2.5: Sensitivity of moments to pledgeability s



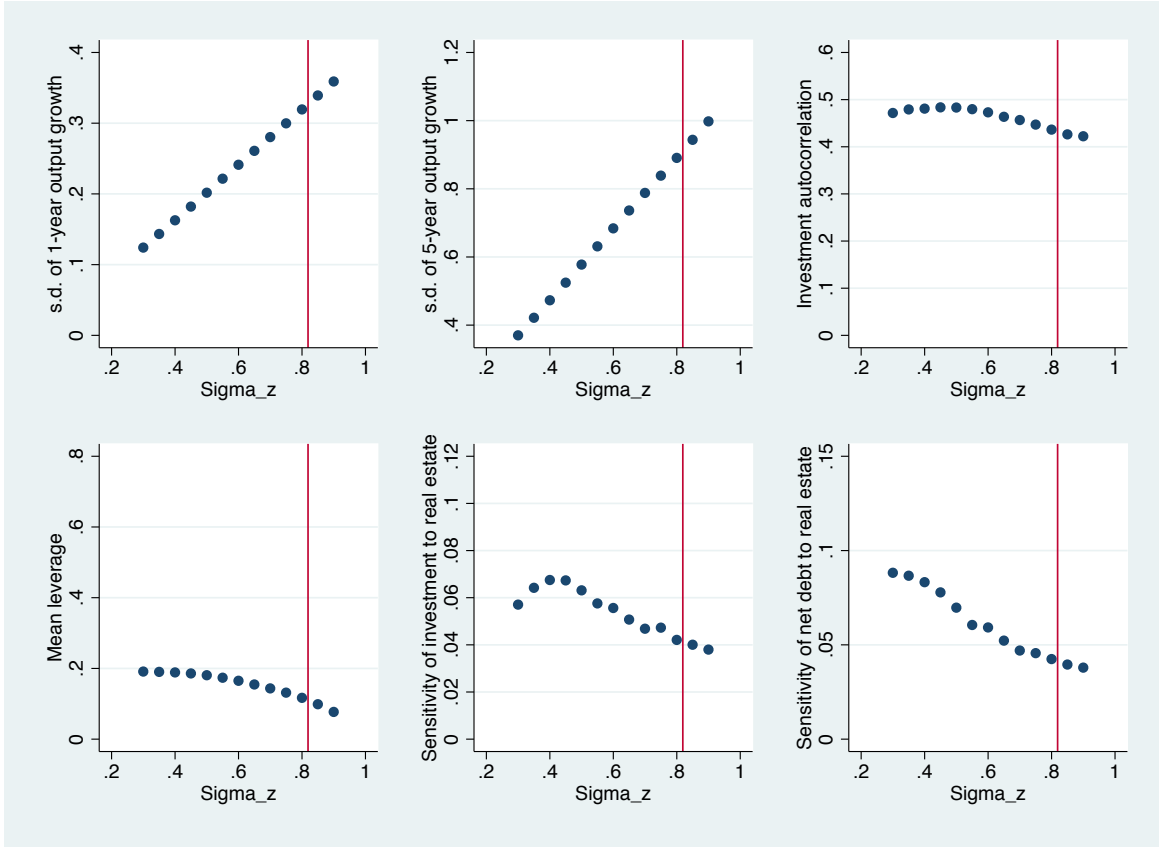
Note: In this figure, we set all estimated parameters (s, c, ρ, σ and H) at their SMM estimate in our preferred specification – as per Column (4), Table 2.3. We fix w and Q at their reference levels: $w = 0.03$ and $Q = 1$. We then vary s from 0 to 1. For each value of s that we choose, we solve the model, simulate the data, and compute four target moments, plus the average leverage ratio and the sensitivity of debt issuance to real estate value. Each panel corresponds to one moment. The red vertical line corresponds to the SMM estimate of s .

Figure 2.6: Sensitivity of moments to adjustment costs c



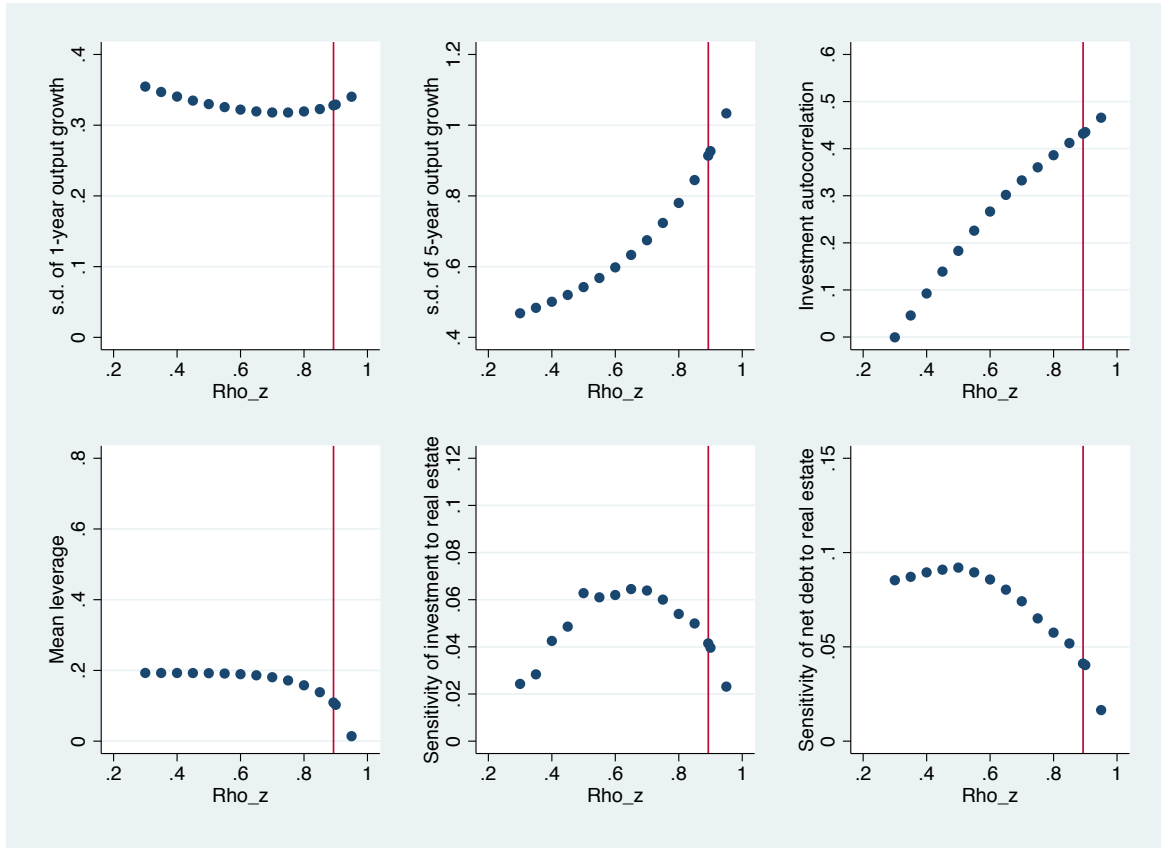
Note: In this figure, we set all estimated parameters (s, c, ρ, σ and H) at their SMM estimate in our preferred specification – as per Column (4), Table 2.3. We fix w and Q at their reference levels: $w = 0.03$ and $Q = 1$. We then vary c from 0 to .1. For each value of c that we choose, we solve the model, simulate the data, and compute four target moments, plus the average leverage ratio and the sensitivity of debt issuance to real estate value. Each panel corresponds to one moment. The red vertical line corresponds to the SMM estimate of c .

Figure 2.7: Sensitivity of moments to productivity volatility σ



Note: In this figure, we set all estimated parameters (s, c, ρ, σ and H) at their SMM estimate in our preferred specification – as per Column (4), Table 2.3. We fix w and Q at their reference levels: $w = 0.03$ and $Q = 1$. We then vary σ from 0 to 1. For each value of σ that we choose, we solve the model, simulate the data, and compute four target moments, plus the average leverage ratio and the sensitivity of debt issuance to real estate value. Each panel corresponds to one moment. The red vertical line corresponds to the SMM estimate of σ .

Figure 2.8: Sensitivity of moments to productivity persistence ρ



Note: In this figure, we set all estimated parameters (s, c, ρ, σ and H) at their SMM estimate in our preferred specification – as per column 4, Table 2.3. We fix w and Q at their reference levels: $w = 0.03$ and $Q = 1$. We then vary ρ from 0 to 1. For each value of ρ that we choose, we solve the model, simulate the data, and compute four target moments, plus the average leverage ratio and the sensitivity of debt issuance to real estate value. Each panel corresponds to one moment. The red vertical line corresponds to the SMM estimate of ρ .

Chapter 3

Haircuts and Credit Risk Over the Cycle¹

3.1 Introduction

Collateralized lending is a multi-trillion dollar market that plays a central role in the modern finance system: mortgages are secured by the mortgaged property, collateralized debt obligations and repurchase agreements and other instruments are secured by pools of securities.² In this paper I study the pledgeability of collateral and its impact on asset pricing and amplification of fundamental risks.

A key measure of pledgeability of collateral is the haircut, the difference between the market value of an asset used as loan collateral and the amount of the loan. For example, if \$90 million is borrowed against collateral worth \$100 million, then there is a haircut of 10 percent. Haircuts are in general countercyclical, high in booms and low in busts. The average haircut in the bilateral repo market rises from zero in early 2007 to nearly 50 percent at the peak of the crisis, after Lehman's bankruptcy in late 2008 ([Gorton and Metrick \(2012\)](#)). A sudden and significant tightening of haircuts exacerbates deleveraging,

¹I am grateful to Markus Brunnermeier for his guidance and support. I also thank Olivier Darmouni, Zhenyu Gao, Valentin Haddad, Ji Huang, Nobuhiro Kiyotaki, Jason Ravit, Yuliy Sannikov, Hyun Song Shin, Wei Xiong, and Yu Zhang for their enlightening feedback and comments. The views and errors of this paper are my own.

²Practitioners have long recognized the importance of collateral in lending. Collateralized lending contracts on the Amsterdam stock market dates back to the 1770s . See [Koudijs and Voth \(2013\)](#).

leads to a contraction, or even shut down, of collateralized lending markets. Furthermore, the countercyclical haircuts amplify fundamental risks and further depresses the real economy, as demonstrated during the last decade.

In line with the leverage cycle literature, investors in the economy may buy or sell collateral lending contracts with different haircuts, but only one of the contracts is traded with positive volume. Thus, the haircut of lending is well-defined. I go beyond the literature by focusing on risky loans instead of riskless loans and analyzing the haircuts and credit risks jointly. I also distinguish two pledgeability concepts. Investors may not be able to fully pledge their assets as collateral due to information frictions. Investors could strategically default and divert a fraction of the assets they hold. Hence, only the fraction that cannot be diverted is pledgeable. On top of this informational pledgeability, there is also the financial pledgeability that investors can only borrow with haircuts to protect lenders against the asset valuation risks. Hence, the amount of the loan must be below the market value of the fraction of assets that cannot be diverted.

First, I find that the informational pledgeability creates a wedge that diminishes the collateral value of assets and dampens the incentive to leverage up. The wedge is wider in downturns when borrowers' net worth is lower and the collateral is scarcer. As a result, borrowers prefer to use a higher haircut instead of a higher leverage to absorb more risks and earn a higher risk premium in downturns relative to normal periods. The information wedge and the change of relative net worth position of borrowers can generate cyclical haircuts, even if the fundamental risk is acyclical.

Second, the countercyclical haircuts lead to financial instability. In good times, haircuts keep decreasing with absence of shocks. The buffer provided by the excessive collateral becomes thinner and thinner, which translates to more credit risk both in terms of a higher probability of default and a larger credit supply. During contraction periods, a decline in the value of collateral erodes net worth of highly leveraged agents. In addition, haircut spikes due to tightening credit conditions further force investors to

deleverage. This two liquidity spirals reinforce each other and exaggerate the financial distress.³

I consider an endowment economy with collateralized lending. In this economy a single tree produces consumption goods. It grows at a constant rate, but is exposed to an adverse aggregate shock. Two groups of agents hold stock shares of the tree, and they disagree on the frequency of aggregate shock. I call them pessimists and optimists. They trade the shares, as well as being able to issue collateralized loans to each other. Any such financial claim specifies a triplet: par value, haircut, and interest rate. The claim gives the holder no claim on other assets of the issuer except the associated collateral in the event of default. Rather than assuming haircuts of any arbitrary form, I endogenize the choice of haircuts in a collateral equilibrium à la [Geanakoplos and Zame \(2014\)](#). In the equilibrium, although all financial contracts with different collateral requirements are tradable at competitive interest rates, market competition determines which contracts are actively traded with positive volume. The selection of contracts further pins down the size of haircuts in equilibrium. Adopting the concept of collateral equilibrium is appealing for two reasons. First, it features endogenously incomplete markets, which is motivated by real world practice of collateralized lending. Moreover, it permits a sharp prediction of haircuts and credit risk over business cycles.

There are two key risk measures of borrowers' portfolios: the leverage and the haircut. The higher the haircut is the more collateral is required and the contract is more risky from the borrowers' perspective. Meanwhile, higher leverage also leads to higher risk due to the loan volume. Although the equilibrium allocations may differ in both dimensions, the optimality conditions force the leverage and the haircut to move in the opposite directions because of *the scarcity of collateral*. Borrowers will fully explore the collateral value of assets and pledge all their assets as collateral. If they post a higher haircut per unit of loan, they have to decrease the total amount of the loans and thus

³See [Brunnermeier and Pedersen \(2009\)](#) for further discussions on liquidity spirals.

use a lower leverage; and vice versa. In other words, the cyclicalities of haircuts has to be opposite to that of the leverage.

In the model, the optimists are natural borrowers who borrow with leverage. They are also hit more heavily by the adverse shock due to their leveraged position, and their wealth drops much faster relative to pessimists. Thus, in downturns, defined as periods preceded by a sequence of negative shocks, optimists hold a relatively small fraction of the total wealth. Hence, they have a strong incentive to load more risk to earn a high premium. In order to do so, the optimists can either leverage up and thus post a lower haircut, or they can post additional collateral and decrease leverage. At the equilibrium, they will follow the latter strategy because of the wedge coming from the information friction. By increasing leverage the optimists can only use a fraction of the leveraged assets as collateral to secure the loans. In contrast, by decreasing leverage and posting a higher haircut they are immune from such information discount on collateral values.

To illustrate the key mechanism I focus on the case where the fundamental risk is acyclical. I consider an extension where countercyclical fundamental risk provides an additional and independent mechanism for countercyclical haircuts. The two mechanisms reinforce each other in a unified framework. I also analyze the case where investors can learn about the true frequency of the adverse shock as a way to endogenously generate countercyclical (perceived) risk. A belief spiral emerges in the environment and magnifies the margin spiral.

The rest of the paper is organized as follows. I set up the model in [Section 3.2](#). In [Section 3.3](#) I characterize the Markov equilibrium with net worth share as state variable and prove that at any equilibrium there is effectively a single financial contract being actively traded in the market at a time. I then use numerical examples to show the main results in [Section 3.4](#). Discussion and conclusion follows.

Related Literature. My paper is closely related to the work of [Geanakoplos \(1997, 2010\)](#), which pioneered the general equilibrium analysis of collateralized lending and asset pricing. I focus on a dynamic extension of [Geanakoplos \(2010\)](#) with continuum of states and only two groups of agents. [Fostel and Geanakoplos \(2008\)](#) develops a formal theory of asset pricing incorporating liquidity and collateral. They show that asset prices have a component determined by the assets' ability to be used as collateral. This additional component is named as liquidity wedge, which guarantees that every market of financial contracts can be cleared.

One distinctive feature of my model is that the loans traded in equilibrium are risky. One can use this framework to study economic issues associated with default of collateralized loans, which were at the forefront during the recent crisis. [Fostel and Geanakoplos \(2015\)](#) proves that in an economy with two future states any equilibrium is equivalent to another equilibrium in which there is no default. There are two ways to break this No-Default Theorem: introduce heterogeneity in agents' valuation of dividend, or have more than two future states. I follow the second approach and also [Simsek \(2013\)](#), which focuses on the role of different types of belief disagreements in a static setting.

This paper is also related to studies of financial frictions in dynamic general equilibrium models with seminal works of [Bernanke and Gertler \(1989\)](#), [Kiyotaki and Moore \(1997\)](#), and [Bernanke, Gertler, and Gilchrist \(1999\)](#). These papers showed how credit market frictions could amplify the impact of exogenous shocks through adverse feedback loops. An important distinction of my paper is that collateral requirement in my economy is determined endogenously by the selection of actively traded financial contracts at equilibrium, rather than assumed exogenously.

The determination of collateral requirement provides a precise prediction on cyclicity of haircut. A handful of papers also yield countercyclical haircuts for a different mechanism, including [Geanakoplos \(2010\)](#), [Brunnermeier and Pedersen \(2009\)](#), [Jurek and Stafford \(2010\)](#). Two exogenous assumptions guarantee the results in most of these

paper. Collateralized lending is assumed to be safe thus haircut is solely determined by the fundamental risk. Moreover, the fundamental risk is higher in bad times than in good time. Loans are endogenously risk-free in [Geanakoplos \(2010\)](#) and the following papers by adopting a richer market structure and the explicit role of collateralized lending. The main departure of my paper is to show that haircuts can be countercyclical even if the fundamental risk is acyclical.

[Kubler and Schmedders \(2003\)](#) first introduces collateral constraints into a dynamic general equilibrium model with aggregate shocks and heterogeneous agents. [Cao \(2013\)](#) extends their framework to models with heterogeneous beliefs. He follows the two-state setting in [Geanakoplos \(2010\)](#) and focuses on risk-free debt. In general, dynamic models with collateral constraints and incomplete markets are very difficult to analyse (see [Kubler and Schmedders \(2003\)](#) for a discussion). As opposed to these works, I use continuous-time methodology to solve for the full dynamics of the model and explicitly link the collateral requirements to the overall risk and state of the economy.⁴ I follow the emerging literature of models with financial frictions in continuous-time settings, starting from [Brunnermeier and Sannikov \(2014\)](#) and [He and Krishnamurthy \(2012, 2013\)](#).

3.2 The Model

3.2.1 Set-up

In this section I develop a model of the economy with collateralized lending in continuous time. This economy is populated by two groups of agents. One group is more optimistic about the economy and they believe that negative shocks are very rare, whereas the other group is less optimistic and worries more about negative shocks. I denote these two groups by optimists and pessimists. The agents may also have different discount rates.

⁴A simple analysis with discrete-time set-up is provided in [Section 3.A](#).

Preferences. Both groups of investors have logarithmic preference

$$\mathbb{E}_0 \left[\int_0^{+\infty} \exp(-\rho t) \log(c_t) dt \right],$$

where c_t is the individual consumption, and ρ is the discount rate.

A Tree. There is one tree in this endowment economy, which produces a certain amount a of apples at every instant. Denote k_t as the size of the tree at time t , and Y_t the total output,

$$Y_t = ak_t.$$

The tree grows at a constant rate g and also suffers from a negative shock. The shock arrives at a Poisson rate λ . Once hit by the shock, the tree shrinks by $1 - u$ fraction, which gives

$$dk_t/k_{t-} = gdt - (1 - u)M(dt).$$

u has a continuous distribution $G(u)$ on its domain $[\underline{u}, 1)$.

Heterogeneous Beliefs. Agents disagree on the Poisson intensity λ . Optimists believe that the true intensity is λ^o , whereas pessimists believe it to be λ^p . I assume that $\lambda^p > \lambda > \lambda^o$. Except this they agree on everything else, including the growth rate of the tree and how much the tree shrinks once hit by a shock, which are publicly observable.

Returns. Agents hold shares of the tree, which is the only asset in this economy and pays dividend in apples. The price of those shares q_t follows

$$dq_t/q_t = \mu_t^q dt - (1 - v_t)M(dt),$$

where both its drift μ_t^q and jump size v_t are determined endogenously and varying over time. Hence the return for holding the shares follows

$$dr_t^k = \frac{a}{q_t}dt + (\mu_t^q + g)dt + (uv_t - 1)M(dt).$$

Define the drift and jump parts of it as

$$R_t^k \equiv \frac{a}{q_t} + \mu_t^q + g, \quad z_t \equiv uv_t(u),$$

also denote $F_t(\cdot)$ as the cumulative distribution function of z_t . To facilitate the exposition, I denote $\inf(\text{supp}(z_t))$ as $\underline{z}_t$ and $\sup(\text{supp}(z_t))$ as $\bar{z}_t$.

Collateralized lending. At every time t , there are a continuum of markets for different contracts. These contracts only last from t to $t + dt$ when the relationship is broken.⁵

A simple debt contract $C \in \mathcal{C}_t$ is an instantaneous bond issued at time t that promises 1 unit of consumption good at $t + dt$, using stock shares that worth C units of consumption good as collateral. The interest rate of contract with collateral requirement C is $R_t(C)$. After default the most a borrower can lose is the associated collateral,⁶ hence the actual delivery of contract C is

$$\min\{1, Cz_t\},⁷$$

where Cz_t is the value of collateral after the tree is hit by a negative shock. Interest rate doesn't show up here because borrowing is instantaneous. However one must pay this interest rate in order to roll over the debt.

⁵Short-term financing is wildly used in real world, such as over-night REPO markets. Also see [He and Xiong \(2012\)](#) for a motivation why longer term collateralized financial contracts are not used in equilibrium.

⁶There is no additional punishment for default besides losing the collateral. There is also a literature on "limited enforcement" where default is punished by exclusion from markets.

⁷For short, I use $1 \wedge Cz_t$ instead in later exposition.

The set of collateral requirements of available contracts is $\mathcal{C}_t = [1, +\infty)$. If the collateral requirement is close to one, the contract is almost an equity. Whereas, the contract with $C = 1/\underline{z}_t$ is risk-free.

Borrowing Constraint. Borrowers can divert a fraction $1 - \alpha$ of total shares they hold and hide it from the lenders. Only the pledgeable asset can be used as collateral. This constraint is similar to the ones in [Brunnermeier and Sannikov \(2014\)](#) and [Brumm, Grill, Kubler, and Schmedders \(2015\)](#). In other word,

$$\text{Value of Collateralizable Asset} = \alpha \cdot \text{Value of Total Asset}$$

Alternatively this constraint can be motivated as working capital requirement. Borrowers lose certain control right over their assets that are used as collateral. To keep running their business requires full control right over a fraction of total capital, and only the rest can be used as collateral.

Moreover there is a second layer between how much borrowers can borrow and the value of collateral.

$$\text{Value of A Loan} \leq 1/C \cdot \text{Value of The Associated Collateral}$$

where C is the collateral requirement and $(C - 1)/C$ is the haircut, which can be different for each loan and will be determined endogenously in this model.

3.2.2 Agents' Problem

Both groups of agents make decisions on consumption and portfolio choice in order to maximize their life-time discounted utility from consumption. They face budget constraint and borrowing constraint when issuing financial contracts.

Net Worth. Following [Kiyotaki and Moore \(1997\)](#) and [Bernanke, Gertler, and Gilchrist \(1999\)](#), I emphasize the key role of net worth as the state variable. The net worth N_t^i of an agent of group i evolves as follows

$$dN_t^i = x_t N_t^i dr_t^k + N_t^i \int_C [R_t(C)dt + ((Cz_t - 1) \wedge 0)M(dt)] (d\mu_+^i - d\mu_-^i) - c_t^i N_t^i dt \quad (3.1)$$

where the agent invests fraction x_t of his net worth in the asset, and consume c_t fraction of his net worth. μ_+^i and μ_-^i are measures represent his positive(lending) and negative(borrowing) positions on financial contracts. They are also normalized as a fraction of net worth.

Holding a positive position on contract C , i.e. lending to others, can earn a flow of interest rate $R_t(C)$, and obtain a repayment of $(Cz_t \wedge 1)$. If there is no default, lender gives out one unit of consumption good and gets back the same amount; whereas if there is default, lender can only take all the associated collateral and get only Cz_t rather than 1 unit of consumption good back. Borrowers take the opposite position.

Agents who borrow with any financial contract must satisfy the following borrowing constraint,

$$\alpha x^i - \int_C C d\mu_-^i \geq 0. \quad (3.2)$$

Agents are also restricted by how much they can borrow. Only α fraction of their total asset holdings can be used as collateral, and further more there is a non-negative haircut on any financial contract. I will show that pessimists, as natural lender, don't hold a negative position on any financial contract. So effectively they are not restricted by this constraint.

Agents of group $i \in \{o, p\}$ solve the following optimization problem, subject to the law of motion of net worth, the borrowing constraint, and the resource constraint.

$$V^i(N_t^i, t) = \sup_{x_t^i, c_t^i, \mu_{+, -}^i} \mathbb{E}_0 \left[\int_0^{+\infty} \exp(-\rho^i t) \log(c_t^i N_t^i) dt \right]$$

s.t. (3.1) (3.2) and

$$x^i + \int_C (d\mu_+^i - d\mu_-^i) = 1. \quad (3.3)$$

Rewrite the problem in recursive form by guessing $V^i(N_t, t) = \exp(-\rho^i t) V^i(N^i)$.

$$\begin{aligned} \rho^i V^i(N^i) = & \max_{c^i, x^i, \mu(\cdot)} U(c^i N^i) + (V^i)'(N^i) N^i \left[x^i R^{k,i} + \int_C R(C) (d\mu_+^i - d\mu_-^i) - c^i \right] \\ & + \int_z \left[V^i(N^i [zx^i + \int_C (Cz \wedge 1) (d\mu_+^i - d\mu_-^i)]) - V^i(N^i) \right] dF^i(z) \end{aligned}$$

s.t. (3.2) (3.3)

Denote the Lagrangian multipliers associated to the budget constraint and the borrowing constraint by γ^i, ξ^i respectively. To simplify the notation, define

$$D^i(z) \equiv x^i + \int_C (C \wedge 1/z) (d\mu_+^i - d\mu_-^i).$$

With logarithm utility function, one can further guess that the value function is also a logarithmic function of net worth, i.e. $V^i(N^i) = \log(N^i)/\rho^i + b$, where b is a function of the state variable but not a function of net worth. With logarithm utility it is also optimal for the agents to consume a constant fraction of net worth at any time t , that is $c_t = \rho^i$.

Then consider the portfolio choices. As shown below, there is a gap between FOCs for taking positive and negative positions of any financial contract. This is a key feature of the market structure proposed by [Geanakoplos \(2010\)](#). In [Fostel and Geanakoplos \(2008\)](#) they focus on the spread between the interest rate optimists would be willing to

pay and the rate pessimists would be willing to take, and call it “liquidity wedge”. It is the same in my model, a spread exists between buyers and sellers of any financial contract.

$$\partial x^i \quad R^{k,i} + \int_z \frac{1}{D^i(z)} dF^i(z) - \gamma^i + \xi^i \alpha = 0 \quad (3.4)$$

$$\partial \mu_-^i \quad R(C) + \int_z \frac{(C \wedge 1/z)}{D^i(z)} dF^i(z) - \gamma^i + \xi^i C \geq 0; = \text{ for } \forall C \in \text{supp}(\mu_-^i) \quad (3.5)$$

$$\partial \mu_+^i \quad R(C) + \int_z \frac{(C \wedge 1/z)}{D^i(z)} dF^i(z) - \gamma^i \leq 0; = \text{ for } \forall C \in \text{supp}(\mu_+^i). \quad (3.6)$$

Equation 3.5 binds only when agents take a negative position of any financial contract, meanwhile **Equation 3.6** binds only when they take a positive position of any financial contract. Otherwise both FOCs are slack. Although both FOCs may be strictly larger/smaller than zero these markets can still be cleared, because both investors prefer not to sell nor to buy. If their position is positive, the corresponding FOC is negative hence they want to reduce their position. On the other hand, if their position is negative, the corresponding FOC is positive hence they want to increase their position. As a result, they don't trade this financial contract.

As in [Geanakoplos \(2010\)](#) and [Simsek \(2013\)](#), the first result that I am going to prove below is that at any equilibrium only a single financial contract is actively traded. Moreover, only optimists buy that contract from pessimists. For any other financial contract, both optimists and pessimists take zero position and that market is inactive.

3.3 Equilibrium Analysis

Equilibrium in this economy is characterized by optimists optimization, pessimists optimization, and market clearing conditions. Formally,

Definition 3 For any initial endowments of assets $\{k_0^o, k_0^p\}$, a collateral equilibrium is described by stochastic processes on the probability space $\mathcal{F}$: the price process of asset $\{q_t\}$, net worths $\{N_t^o, N_t^p\}$, asset holdings $\{k_t^o, k_t^p\}$, consumption choices $\{c_t^o, c_t^p\}$, and the distribution of jump sizes $F_t(z_t)$ such that

1. initial net worth satisfies $N_0^i = k_0^i q_0$ for $i \in \{o, p\}$,
2. both groups of investors solve their optimization problem given prices and the distribution of jump sizes,
3. markets for consumption goods, capital, and all financial contracts clear

$$N_t^o d\mu_+^o(C) + N_t^p d\mu_-^p(C) = 0$$

$$N_t^o d\mu_-^o(C) + N_t^p d\mu_+^p(C) = 0$$

$$(N_t^o a + N_t^p a) / q_t = N_t^o \rho^o + N_t^p \rho^p.$$

I consider only Markov equilibria. Since the model is scale-invariant there is a single state variable : the fraction of wealth that belongs to optimists,

$$\eta_t = \frac{N_t^o}{N_t^o + N_t^p}.$$

Definition 4 A Markov equilibrium is a collateral equilibrium in which the price of real and financial assets and the allocation of consumption and of real and financial asset holdings in each history depend only on the endogenous determined wealth share η_t .

I derive the process of wealth share below, all other equilibrium prices and quantities can be expressed as a function of the state variable.

Lemma 5 *The law of motion of η_t is given by*

$$\begin{aligned} \frac{d\eta_t}{\eta_{t-}} = & \left[(x_t^o - 1)R_t^k + \int_C R(C)(d\mu_+^o - d\mu_-^o) + \frac{a}{q_t} - \rho^o \right] dt \\ & - \left[1 - x_t^o - \int_C (C \wedge 1/z)(d\mu_+^o - d\mu_-^o) \right] M(dt). \end{aligned} \quad (3.7)$$

Proof. See [Appendix 3.B.1](#). ■

As stated in the definition, this paper explores collateral equilibria of this economy. A collateral equilibrium requires that any financial contract must specify explicitly its corresponding collateral. Moreover a collateral equilibrium doesn't restrict haircuts of financial contracts, the markets of contracts with different haircuts must be cleared at the same time.

Collateral equilibrium degenerates to general equilibrium with incomplete markets (GEI) when neither borrowing constraint binds. Intuitively since collateral constraints are slack there is no reason for haircuts to be relevant in GEI. Formally, when collateral equilibrium degenerates to GEI the GEI is indetermined in terms of haircuts. Any of the continuum of combination of leverage and haircut can support the equilibrium prices and allocations. In order to analysis the dynamics of haircuts, this indeterminacy has to be ruled out. The following assumption guarantees that collateral equilibrium doesn't degenerates.

Assumption 2 $\sup(\text{supp}(u)) = 1$

Collateral equilibrium could be undetermined for a second reason, when there is a gap in the distribution of shock size u . In this case, average haircut (weighted by portfolio of contracts) is a sufficient statistics for equilibrium prices and quantities hence the composition of financial contracts are undetermined. A particular example is that u is a constant (has a degenerated distribution)⁸. Under this assumption for any collateral

⁸This example corresponds to the setting of binary distribution of future states in [Geanakoplos \(1997\)](#)

equilibrium there exists another collateral equilibrium, in which only risk-free debt is available, leading to the same equilibrium prices and quantities. This is a standard and intuitive result. With simple Poisson shock any financial contract can be decomposed into two parts: risk-free contract and risky assets. In another word, risk-free debt and risky asset can span the whole contract space.

This second indeterminacy affects no result in this paper as long as Assumption 2 is satisfied, i.e. there is no gap at the right end of the domain. Given the average haircut, composition of the financial contracts is irrelevant to the equilibrium prices and allocations. The following assumption rules out the possibility of gaps, which is not required for the results but helps to better understand the equilibrium structure .

Assumption 3 *u is continuously distributed on $[\underline{u}, 1)$ with a cumulative distribution function $G > 0$.*

Remark. The pledgeability condition is essential to have a non-degenerated collateral equilibrium. Without this pledgeability condition collateral equilibrium always degenerates to GEI. (Assumption 2 is sufficient only after assuming $\alpha < 1$.) This is due to the fact that agent's utility function satisfies the Inada condition $u'(0) = -\infty$. Suppose $\alpha = 1$ and the borrowing constraint of some agent binds, once the borrower defaults due to a large negative shock all his assets are taken away as collaterals and his net worth drops to zero. With Inada condition this is not optimal for the borrower. Hence by contradiction we know that the borrowing constraint is always slack and collateral equilibrium reduces to GEI.

Proposition 5 *Under Assumption 2, in any equilibrium the borrowing constraint of optimists' binds, and that of pessimists' slacks.*

Proof. See Appendix 3.B.2. ■

This proposition has two layers. First, given Assumption 2 the collateral equilibrium doesn't degenerate to GEI, hence at least one of the borrowing constraint binds.

Moreover, as natural borrowers optimists face binding borrowing constraint. For pessimists, as natural lenders, the borrowing constraint slacks. Hence the borrowing constraint of pessimists always slacks. Combined with the first part, the borrowing constraint of optimists has to bind.

Now I am ready to prove the first key property of the correlated equilibrium: there is only a single financial contract being actively traded at the equilibrium. The proof consists of two steps. First I show that optimists don't lend to pessimists through any contract.

Proposition 6 *In any equilibrium optimists take non-positive position in any financial contract. That is $\mu_o^+ = 0$.*

Proof. See [Appendix 3.B.3](#). ■

Since pessimists are the natural lenders, the only possible incentive for them to borrow is to lend more than their total shares. In another word, they may want to borrow through contracts that are riskier (with lower haircut) and then lend through contracts that are safer, in order to dump more risk on the optimists. Without no-short-sell constraint, this scenario doesn't happen at equilibrium. By allowing short selling of the shares, pessimists could instead short sell the shares, rather than borrow, and lend to optimists, without posting any collateral.

Then I show that although potentially trading can be taken at multiple or infinite number of markets at any equilibrium only one of the continuum of markets is active at one time.

Proposition 7 *Effectively there is only a single contract that is actively traded in non-zero quantities at any time in any equilibrium.*

Proof. See [Appendix 3.B.4](#). ■

There are two counter forces which jointly determine which contracts are actively traded. On one hand, since one needs to post collateral in order to borrow there is a collateral value associated to financial contracts. This collateral value is captured by the Lagrangian multiplier associated to the binding borrowing constraint. On the other hand, by borrowing and lending agents transfer risk to other agents (one can think this as insurance). The higher is the haircut associated to a contract the safer the contract is, and further the more risk is transferred from lenders to borrowers. Mathematically we can express the two effects by combining the FOCs of both optimists and pessimists,

$$\int_z \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] \underbrace{(C \wedge 1/z)}_{\text{risk transfer}} dF(z) + \gamma^p - \gamma^o + \underbrace{\xi^o C}_{\text{collateral value}} \geq 0 \quad (3.8)$$

It is obvious that the first effect is linear in collateral requirement, whereas the second effect is strictly convex in collateral requirement. The marginal risk transfer increases in collateral requirement because that the level of collateral is relevant in more future states. The left hand side of [Equation 3.8](#) has a unique minimum where it equals zero, this leads to the uniqueness of actively traded contract.

[Proposition 7](#) allows us to derive the optimal portfolio choices at any time given the state variable and equilibrium prices. In the numerical examples the equilibrium prices only depend on the state variable, hence the dynamic problem degenerates to a comparative statics of the static problem with respect to the state variable. In the general framework, however, prices are endogenously determined.

3.4 Numerical Examples

[Proposition 7](#) allows us to characterize the full stochastic equilibrium dynamics of the economy. For this numerical example I assume that u is uniformly distributed on $[0.8, 1)$.

I fix $a = 0.05$, $\rho^o = \rho^p = 0.05$, $\lambda^o = 0.1$, $\lambda = 0.2$, $\lambda^p = 0.3$, $g = 0.021$ to compensate the Poisson shock, and $\alpha = 0.8$. Numerically I set a discrete grid on $[0.8, 1)$ with equal space 0.001 for u . This approximation is good enough to ensure that collateral equilibrium doesn't degenerate. All the results are qualitatively robust for alternative settings.

In this example price of the shares with respect to consumption goods is

$$a\eta_t + a(1 - \eta_t) = q_t(\rho^o\eta_t + \rho^p(1 - \eta_t)) \quad (3.9)$$

The dynamic of prices is abstracted in this example, but will be added into the discussion later.

Haircuts. Panel A of [Figure 3.1](#) plots the haircuts, i.e. $(C^* - 1)/C^*$. The figure shows clearly the countercyclicality of haircuts. The wealth share of optimists η_t can be taken as a proxy of the status of an economy. High wealth share corresponds to good times of an economy, whereas low wealth share corresponds to bad times. An economy in its good time when it keeps growing without being hit by any negative shock. To the contrary, the economy in its bad time after being hit by a sequence of negative shocks. One can take the trading between optimists and pessimists as insurance, that optimists ensures pessimists against the negative shocks by borrowing with haircuts. In good periods optimists keep accumulating the premium and as a result their wealth share increases. Whereas in bad periods optimists has to compensate the losses hence their wealth share decreases.

The main message of this figures is simply that haircuts are higher when wealth share of optimists is lower (bad times) and lower when wealth share is higher (good times). [Brunnermeier and Pedersen \(2009\)](#) discusses the amplification effect of loss spiral joint with margin spiral. This two liquidity spirals present in my model as well, but in this numerical example one can not explicit see it since prices are constant. I will discuss the spirals and the endogenous risk in next section.

Credit Risk. Panel B of [Figure 3.1](#) plots the probability of default. Given that the asset price is constant, z_t equals u and has the same distribution. Borrowers default when the shock is large enough such that they are willing to give up the associated (excessive) collateral, i.e. when $Cz_t < 1$. Haircuts decrease in the wealth share of optimists, hence default triggers in more states z_t and the probability of default increases.

The dashed red line depicts the probability of default under subjective measure and weighted by net worth of agents. It captures the aggregate market belief. In bad times pessimistic agents dominate in the economy in terms of wealth, so the market belief is more pessimistic relative to that in good times. This change in composition of market belief acts as a counterforce to the increasing probability of default in objective measure. Hence the dashed red line is non-monotonic in the wealth share.

The comparison of the two lines captures the idea that credit risk accumulates during normal and good times of an economy but materializes after a sudden shock.

State Variable. [Figure 3.2](#) shows the stochastic dynamics of the state variable η_t as formalized in [Equation 3.7](#). Panel A plots the drift, and Panel B plots the wealth share before and after a negative shock, which is large enough to trigger default. These two plots provide a full characterization of the dynamics for any starting point. Both the drift and the jump size are non-monotonic in the wealth share, in particular reach zero at the two ends where $\eta = 0$ and $\eta = 1$. The two ends correspond to two absorbing states of the economy. When the economy consists of only one type of agents there is no borrowing and lending going on and the state variable stays at the same place forever. For smaller shocks which don't lead to defaults the jump size of wealth share is smaller. As long as a shock triggers default, however, the jump size is independent of the size of the shock. The limited liability nature of borrowing, meaning no further penalty besides giving up the collateral, imposes a cap on the change of relative net worth. The term $(C \wedge 1/z)$ in [Equation 3.7](#) captures this non-linear effect of shock size on jump size of

wealth share directly. For an extreme case, the wealth share drops from 0.7 to 0.45 by about 36% when the economy is hit by a large negative shock. A sequence of two large negative shocks is enough to lead to a jump from 0.7 to 0.3.

Survival. Given the dynamics of the state variable, one can compute its stationary distribution. The stationary distribution is highly sensitive to the objective Poisson intensity. Friedman's market selection hypothesis suggests that if some agents hold correct belief the financial wealth trends towards zero in the long run for the other agents with incorrect beliefs. The same is true in my model. However, if neither optimists nor pessimists are correct they can all survive in this economy, at least for a fairly long time. The single peak of the stationary distribution is closer to zero when the objective probability is closer to the pessimists' belief.

Risk Premium. Figure 3.3 illustrates the returns of asset and financial contracts. Since asset price is constant in this example, so is the return of asset. The red dotted line depicts the returns of the active financial contract, this contract determines the haircut of trading. The blue dashed line represents the returns of risk-free contract, i.e. the contract requires high enough haircut such that it is risk free. The interest rates specified in contracts is presented in Panel D. Interest rates of the actively traded contract increase in the wealth share from 4% to 7%. These interest rates cannot be used directly to compute the risk premiums, instead we need to compute the expected returns including the compensation of the Poisson shocks. Panel A to C plots the expected returns under the optimists' belief, the pessimists' belief, and the objective intensity of Poisson shocks. Under the optimists' belief the returns of asset is always higher than the risk-free rate, however it is not true in Panel B and C when agents are less optimistic. Empirically negative equity premium is a pervasive phenomenon but it is very hard to be reconciled

in a rational expectation framework⁹. In a model with heterogeneous beliefs like this one it is natural to yield negative equity premium.

Figure 3.4 depicts the equity premiums under the objective intensity. From the perspective of the optimists, the asset offers a premium due to shocks, which they believes rarely occurs. As a result they share risk with the pessimists. In bad times, optimists' net worth is low relative to pessimists hence their capacity to sharing risk is limited, i.e. the supply of risk sharing is low. Combined with a high demand of risk sharing from pessimists, the low supply leads to a high premium. To the contrary, in good time optimists' net worth is high relative to pessimists. Hence the equity premium is low as a result of a high supply of risk sharing (or betting as the other side of the same coin) and a low demand.

Loss Spiral. Now I introduce time-varying asset prices into the picture, by assuming that optimists are more patient than pessimists¹⁰. **Figure 3.5** highlights the role of time-varying asset prices as plotted in Panel A. The amplification through fire sale is illustrated in Panel B. A negative shock forces borrowers to fire sale their assets, causing a price drop, leading to a further loss of the borrowers. As a result, the size of shock on asset value (further on net worth) exceeds that on quantity, the extra loss comes from drops in asset price. The underlying risk is defined as the sum of exogenous risk from Poisson shocks and the endogenous risk from price drops.

Brunnermeier and Pedersen (2009) discusses two liquidity spirals reinforcing each other in terms of amplification: a loss spiral arises from drops in asset price, and a margin spiral emerges due to the countercyclicality of haircuts. The loss spiral amplifies the exogenous shock and leads a lower haircut. On the other hand, the margin spiral

⁹See [Xiong and Baron \(forthcoming\)](#) for further discussion.

¹⁰[Hall \(2014\)](#) assumes that the discount rate rises in recessions to explain the high unemployment rate. Heterogeneous discount rates, combined with the changing composition of agents, can lead to countercyclical discount rate at the aggregate level.

forces agents to de-lever during downturns hence exacerbates fire sales. This two spirals together could lead to a much larger drop in the wealth share.

Compared with [Brunnermeier and Pedersen \(2009\)](#), there is an extra layer of subtlety here since haircuts is endogenously determined in my model. The higher is the endogenous risk from price drops, the higher is the underlying risk and the less risk sharing given haircut unchanged. As a result haircut has to increase in order to compensate the decrease in risk sharing. Panel A of [Figure 3.6](#) shows the difference between the haircuts with endogenous risk and that of the benchmark example. Compared with Panel B of [Figure 3.5](#) it is clear that haircuts increase in the underlying risk.

Although this example goes against the countercyclicality of haircut, it provides a clear intuition on how underlying risk affects haircuts. Underlying risk could be countercyclical if asset price increases in wealth share but at a decreasing rate. Learning through Bayesian updating is another potential mechanism. In bad times although the objective probability of shocks stays the same agents perceive a much higher probability after observing a sequence of negative shocks. They will behave as if the underlying risk is higher in bad times and lower in good time.

3.5 Discussions

Financial Development. [Kalemli-Ozcan, Sorensen, and Yesiltas \(2012\)](#) presented new stylized facts on bank and firm leverage during the period of 2000-2009 using micro-level data from many countries. They found that banks in markets with tighter bank regulation and less developed financial system experienced significantly less deleveraging during the crisis. For example, they observed only a slight tendency for leverage to be procyclical for large European banks, with a much smaller slope than that for large US banks. A comparative statics exercise can be done within the model, taking α as a proxy of the tightness of banking regulation, or the development level of financial mar-

ket. [Figure 3.7](#) plots the haircut for different tightness. Other parameters are the same as those in the first example. It is easy to see that with the tightest constraint, the haircut is first flatter and then becomes steeper. It is still consistent with the fact, since developing markets, by definition, are far away from the right boundary $\eta = 1$. The right panel shows the probability of default on the collateralized lending. The tighter the regulation is the lower probability of default will be. This makes sense, although it is not consistent with the cross-country comparison if one think tighter regulation corresponds to less developed financial market.

This comparative statics also has its policy implication on economic stability. With tighter borrowing constraint in general collateral requirements are higher and leverage is lower. So with tighter borrowing constraint, not only leverage is less procyclical but also the economy is more stable. However every coin has two sides, tighter borrowing constraint leads to lower liquidity and less risk-sharing. Potentially agents with higher productivity cannot lever up as much, which dampens growth. This is the long lasting trade-off between economic growth and stability.

Belief Dynamics. The last point that I want to raise goes back to [Geanakoplos \(2010\)](#). In his model there are continuum of investors with different beliefs. The identity of the marginal investor varies endogenously, so does the belief of the marginal investors. Similarly [He and Xiong \(2012\)](#) considered a model allowing two groups of agents to have time-varying beliefs. There are also empirical works looking at how beliefs of investors evolve before and during the crisis. So an immediate extension of this current model is to add time-varying beliefs, which also captures naturally the fact that investors are learning all the time. My guess is that this extension will make the model much richer and favor procyclical leverage. However a challenge to work it out is that there is (at least) an additional state variable to be taken care of, which describes the differences between beliefs of both groups of agents. A crude exercise is to assume that agents can

change their type. If the economy grows steadily, there is a possibility that pessimists become optimists. Whereas right after a shock arrives, some optimists realize that the world is more risky than they thought and turn themselves into pessimists. This “belief spiral” acts a new amplification mechanism.

3.6 Conclusion

I present a dynamic general equilibrium model of leverage cycle with risky debt. I first show that although a continuum of markets indexed by different collateral requirements are available, effectively only a single market is active. Which market is active is endogenously determined and varying across time, so is the collateral requirement associated to the active market. Haircuts of the defaultable debt are countercyclical as a result of scarce collaterals and financial pledgeability. At the same time, default risks accumulate in the background until materialize when a crisis erupts.

My study highlights the importance of taking haircut and leverage explicitly into macroeconomic models. A fruitful future direction would be to take the model to the data and to quantify the amplification mechanism associated with the endogenously determined haircuts.

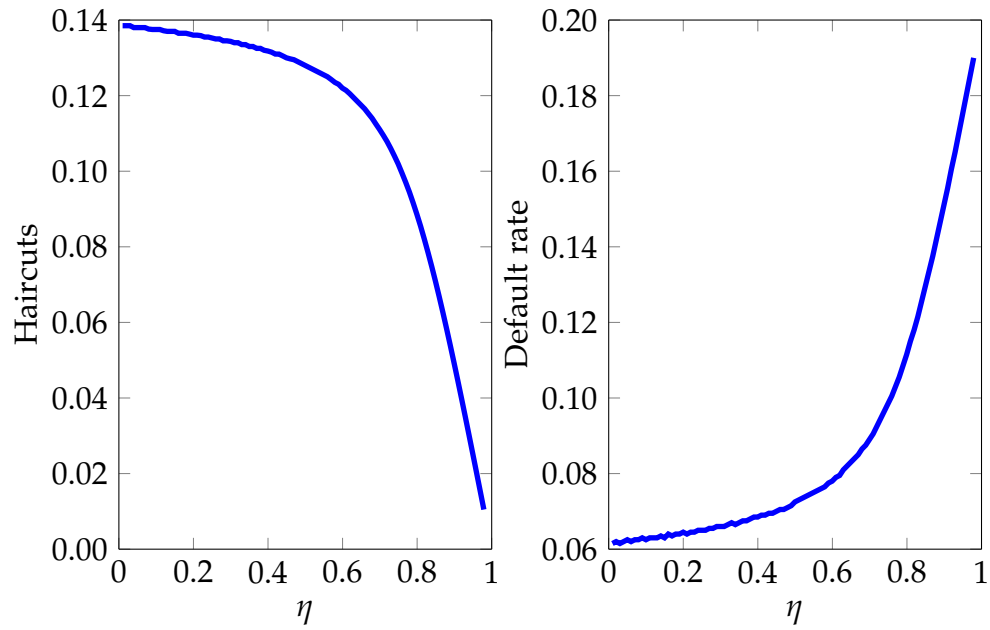
REFERENCES

- BERNANKE, B., AND M. GERTLER (1989): "Agency Costs, Net Worth, and Business Fluctuations," *The American Economic Review*, 79(1), pp. 14–31.
- BERNANKE, B. S., M. GERTLER, AND S. GILCHRIST (1999): "The financial accelerator in a quantitative business cycle framework," *Handbook of macroeconomics*, 1, 1341–1393.
- BRUMM, J., M. GRILL, F. KUBLER, AND K. SCHMEDDERS (2015): "Collateral requirements and asset prices," *International Economic Review*, 56(1), 1–25.
- BRUNNERMEIER, M. K., AND L. H. PEDERSEN (2009): "Market liquidity and funding liquidity," *Review of Financial studies*, 22(6), 2201–2238.
- BRUNNERMEIER, M. K., AND Y. SANNIKOV (2014): "A Macroeconomic Model with a Financial Sector," *The American Economic Review*.
- CAO, D. (2013): "Belief Heterogeneity, Collateral Constraint, and Asset Prices," .
- FOSTEL, A., AND J. GEANAKOPLOS (2008): "Leverage cycles and the anxious economy," *The American Economic Review*, 98(4), 1211–1244.
- (2015): "Leverage and default in binomial economies: a complete characterization," *Econometrica*, 83(6), 2191–2229.
- GEANAKOPLOS, J. (1997): "Promises promises," *Cowles Foundation Discussion Paper*, (1143).
- (2010): "The leverage cycle," in *NBER Macroeconomics Annual 2009, Volume 24*, pp. 1–65. University of Chicago Press.
- GEANAKOPLOS, J., AND W. R. ZAME (2014): "Collateral equilibrium, I: a basic framework," *Economic Theory*, 56(3), 443–492.
- GORTON, G., AND A. METRICK (2012): "Securitized banking and the run on repo," *Journal of Financial Economics*, 104(3), 425–451.
- HALL, R. E. (2014): "High Discounts and High Unemployment," Discussion paper, National Bureau of Economic Research.
- HE, Z., AND A. KRISHNAMURTHY (2012): "A Model of Capital and Crises," *The Review of Economic Studies*.
- (2013): "Intermediary Asset Pricing," *American Economic Review*, 103(2), 732–70.
- HE, Z., AND W. XIONG (2012): "Debt Financing in Asset Markets," *American Economic Review*, 102(3), 88–94.
- JUREK, J. W., AND E. STAFFORD (2010): "Haircut Dynamics," Available at SSRN 1658935.
- KALEMLI-OZCAN, S., B. SORENSEN, AND S. YESILTAS (2012): "Leverage across firms, banks, and countries," *Journal of International Economics*, 88(2), 284–298.

- KIYOTAKI, N., AND J. MOORE (1997): "Credit Cycles," *Journal of Political Economy*, 105(2), pp. 211–248.
- KOUDIJS, P., AND H.-J. VOTH (2013): "Leverage and Beliefs: Risk Taking and Personal Experience in Collateralized Lending Transactions on the Amsterdam Stock Market, 1770-1775," .
- KUBLER, F., AND K. SCHMEDDERS (2003): "Stationary Equilibria in Asset-Pricing Models with Incomplete Markets and Collateral," *Econometrica*, 71(6), pp. 1767–1793.
- SIMSEK, A. (2013): "Belief disagreements and collateral constraints," *Econometrica*, 81(1), 1–53.
- XIONG, W., AND M. BARON (forthcoming): "Credit Expansion and Financial Instability: Evidence from Stock Prices," *Quarterly Journal of Economics*.

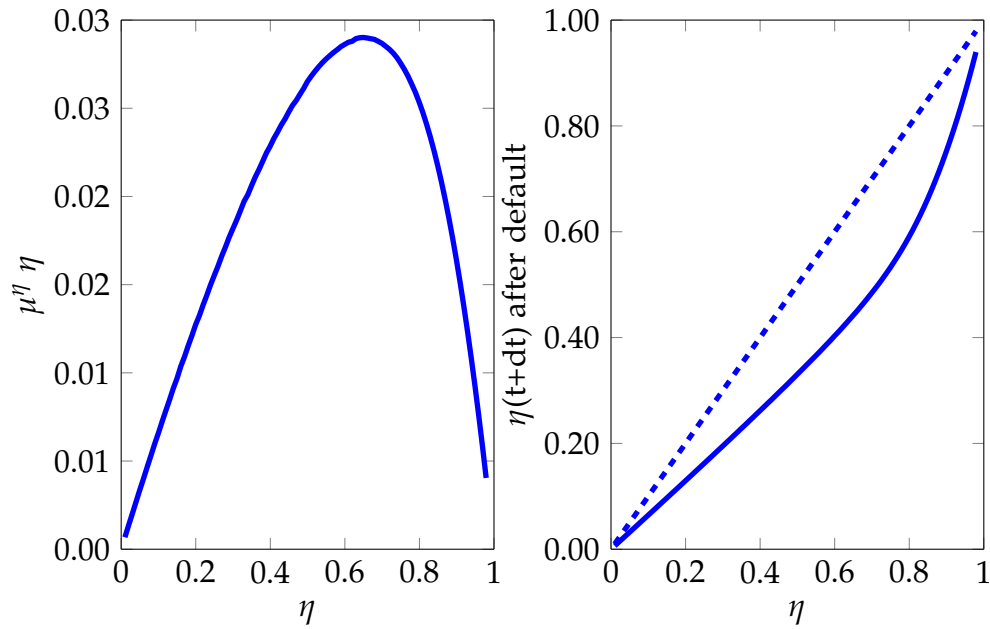
Figures

Figure 3.1: Haircuts and Probability of default



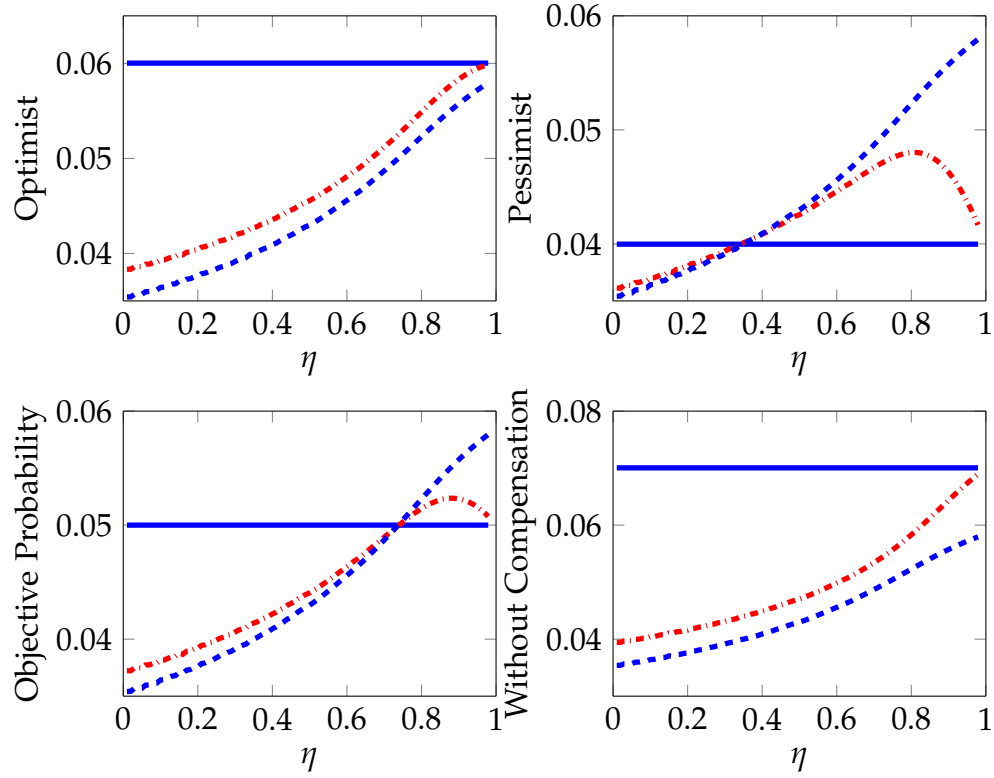
Note: Panel A plots the **haircuts** and Panel B plots the **probability of default** as a function of **wealth share** η .

Figure 3.2: Law of Motion of the State Variable



Note: Panel A plots the **drift of η** , Panel B plots the **jump size of η** as a function of **wealth share η** . Dashed line is the 45° line, and solid line plots wealth share after defaults as a function of the wealth share before defaults.

Figure 3.3: Asset Returns



Note: Returns of asset in solid blue, returns of financial contract which is actively traded in dotted red, and returns of financial contract which is risk-free in dashed blue. Panel A plots the returns under the optimists' belief, Panel B (up-right) under the pessimists', Panel C (down-left) under the objective probability, and Panel D plots the returns before compensating the shock.

Figure 3.4: Risk premium

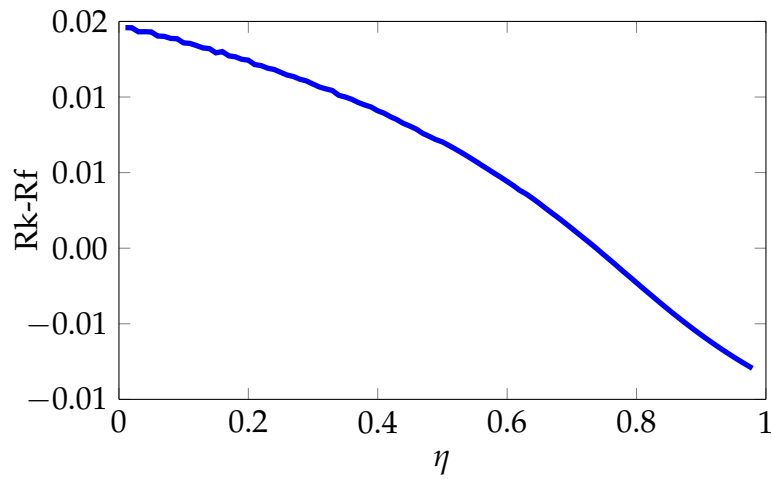
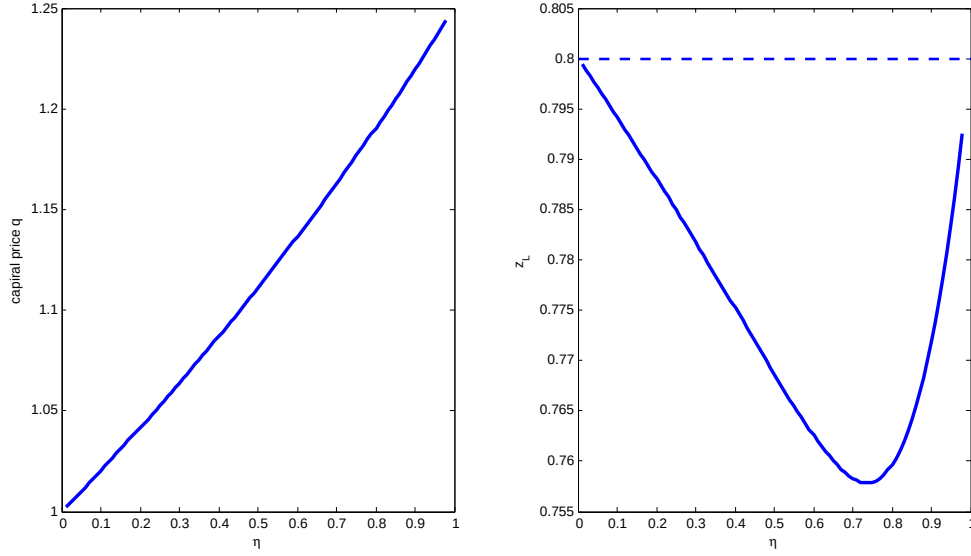
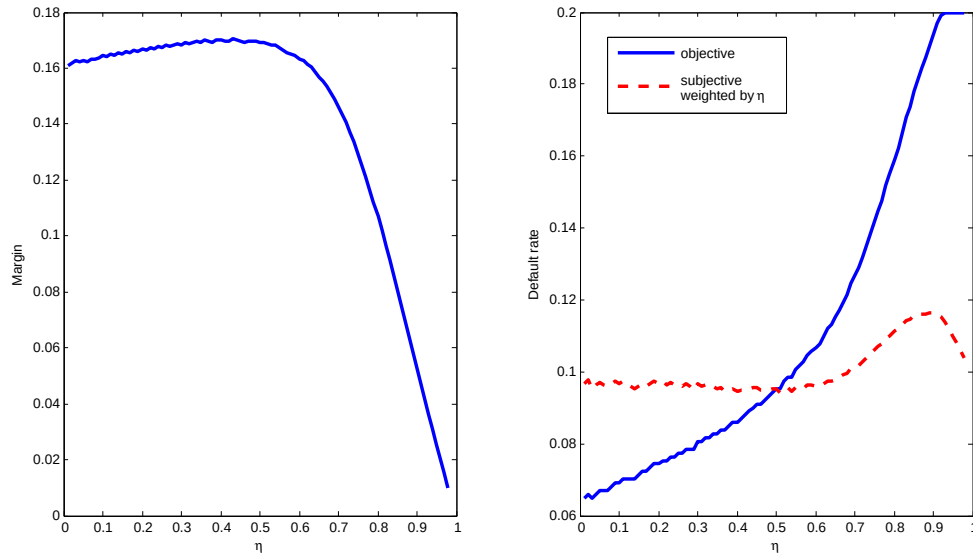


Figure 3.5: Loss sprial



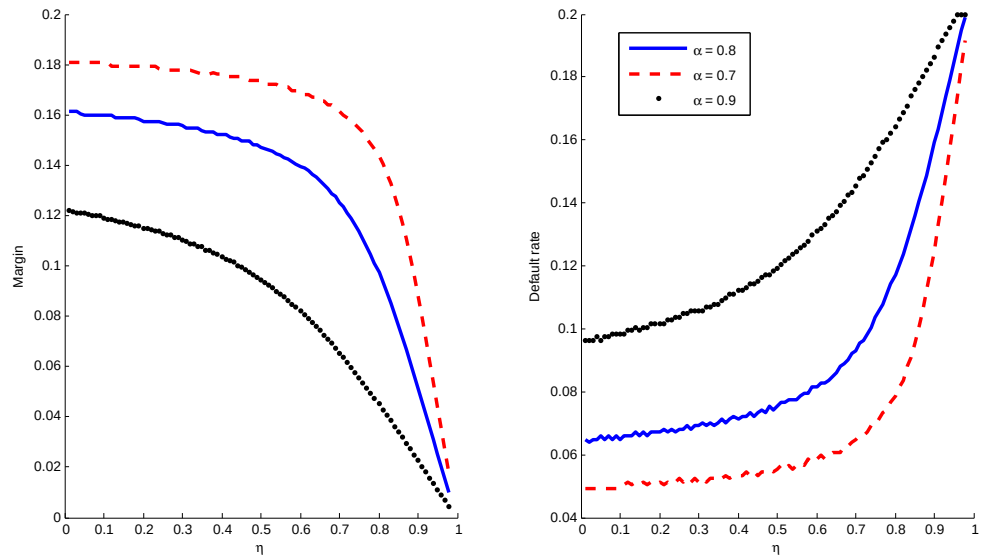
Note: Panel A plots the **asset price**, Panel B plots the **worst shock** $\min(z)$ ($\min(u)$ as the dashed line) as a function of **wealth share** η . For parameter value $\rho^o = 0.04$; $\rho^p = 0.05$.

Figure 3.6: Underlying risk



Panel A plots the **haircuts**, Panel B plots the **probability of default** as a function of **wealth share** η . For parameter value $\rho^o = 0.04$; $\rho^p = 0.05$.

Figure 3.7: Financial development



Note: Panel A plots the **haircut**, Panel B plots the **probability of default** as a function of **wealth share** η .

APPENDIX

3.A A Discrete-time Setup

Here I briefly set up the same model in discrete time. Every choice variable is denoted as a fraction of agents' net worth. The time-line is as follows:

1. Period t starts, agent starts with net worth N_t ;
2. Agent decides his portfolio choice, puts $x_t N_t$ into capital stock, and the rest $\int_C p_t(C) d\mu_t(C) N_t$ into risky debt ($p = 1/R$);
3. Agent uses their capital stock to produce consumption goods, consumes, and accumulates net worth;
4. Agent's capital holding is hit by a shock.
5. Period $t + 1$ arrives, agent repays his debt from last period.
6. Agent now has his updated net worth N_{t+1} .

The key object in this discrete-time framework is the law of motion of net worth. Let $z_t = \frac{q_{t+1}(u_t)}{q} u_t$.

$$N_{t+1}^i = N_t^i \left[x_t^i \left(z_t + \frac{a^i}{q_t} \right) + \int_{C, z_t} \min\{Cz_t, 1\} d\mu_t^i(C) dm_t(z_t) - \rho^i \right]$$

where the agent invests fraction x_t^i of his net worth in capital, and consume ρ^i fraction of his net worth. $\mu^i = \mu_+^i - \mu_-^i$ is the measure that represents positions on financial contracts. They are also normalized as a fraction of net worth.

Value function here is still of the same form. We omit the superscript i .

$$\begin{aligned} V(N, \eta) &= \max_{x, d\mu, c} \log(cN) + \beta \mathbb{E}[V(N')] \\ \text{s.t. } & x + \int_C p(C) d\mu(C) = 1; \quad \alpha x + \int_C C d\mu_-(C) \geq 0 \\ & N' = N[x(z + a/q) + \int_{C, z} \min\{Cz, 1\} d\mu(C) dm(z) - \rho] \end{aligned}$$

Plug in $V(N, \eta) = \frac{1}{1-\beta} [\log(N) + h(\eta)]$, we get

$$\begin{aligned} \frac{1}{1-\beta} [\log(N) + h(\eta)] &= \max_{x, d\mu, c} \log(cN) + \frac{\beta}{1-\beta} \mathbb{E}[\log(N) + \log[x \left(z + \frac{a}{q} \right) \\ &\quad + \int_{C,z} \min\{Cz, 1\} d\mu(C) dm(z) - \rho] + h(\eta')] \end{aligned}$$

$$\begin{aligned} \frac{1}{1-\beta} [h(\eta)] &= \max_{x, d\mu, c} \log(c) + \frac{\beta}{1-\beta} \mathbb{E}[\log[x \left(z + \frac{a}{q} \right) \\ &\quad + \int_{C,z} \min\{Cz, 1\} d\mu(C) dm(z) - \rho^i] + h(\eta')] \end{aligned}$$

The FOC here is different to what we obtained in continuous-time model. There jump component and drift component are separable, whereas in discrete-time there is no such thing called “jump”.

Below I analyse a simple case with fixed size jump $z \in \{z_L, 1\}$. Denote

$$D^o(z) \equiv x^o \left(z + \frac{a^o}{q} \right) + \int_{C,z} \min\{Cz, 1\} d\mu^o(C) dm(z) - \rho^o$$

$$D^p(z) \equiv x^p \left(z + \frac{a^p}{q} \right) + \int_{C,z} \min\{Cz, 1\} d\mu^p(C) dm(z) - \rho^p.$$

At steady state if there is no shock $u = 1$, the capital price stays the same. Hence $q'(1) = q$. Moreover the net worth of either group of agents doesn't move as well, that is $D^i(1) = 0$. We can simplify the FOCs: for any $\tilde{C} \in \mathcal{C}$ and $i \in \{o, p\}$

$$\begin{aligned} \frac{z_L + a^i/q}{D^i(z_L)} \lambda^i + (1 - \lambda^i) \frac{a^i}{q} - \gamma^i + \xi^i \alpha &= 0 \\ \frac{\tilde{C} z_L}{D^i(z_L)} \lambda^i + 1 - \lambda^i - \gamma^i p(\tilde{C}) + \xi^i \tilde{C} &\geq 0 \\ \frac{\tilde{C} z_L}{D^i(z_L)} \lambda^i + 1 - \lambda^i - \gamma^i p(\tilde{C}) &\leq 0. \end{aligned}$$

The liquidity wedge still presents in discrete time. When $x^p > 0$, define

$$\begin{aligned} H(C) &= \frac{\tilde{C} z_L \lambda^o}{D^o(z_L)} - \frac{\tilde{C} z_L \lambda^p}{D^p(z_L)} - (\gamma^o - \gamma^p) p(\tilde{C}) + \xi^o \tilde{C} \geq 0 \\ H'(C) &= \frac{z_L \lambda^o}{D^o(z_L)} - \frac{z_L \lambda^p}{D^p(z_L)} - (\gamma^o - \gamma^p) p'(\tilde{C}) + \xi^o \\ H''(C) &= -(\gamma^o - \gamma^p) p''(\tilde{C}). \end{aligned}$$

From the pessimists' FOC we have

$$\frac{z_L \lambda^p}{D^p(z_L)} - \gamma^p p'(\tilde{C}) = 0$$

and

$$-\gamma^p p''(\tilde{C}) = 0$$

which means that $H(C)$ is linear in C and $p'(C)$ is constant.

3.B Proofs

3.B.1 Proof of Lemma 5

Proof. The law of motion of net worth can be written as

$$\begin{aligned} \frac{dN_t^o}{N_{t-}^o} &= x_t^o R_t^{k,o} dt + \int_C R(C)(d\mu_+^o - d\mu_-^o) dt - \rho^o dt + \\ &\quad \left(x_t^o z_t + \int_C (Cz_t \wedge 1)(d\mu_+^o - d\mu_-^o) - 1 \right) M(dt), \\ \frac{d(N_t^o + N_t^p)}{N_{t-}^o + N_{t-}^p} &= (\mu_t^q + \Phi(\iota_t) - \delta) dt + (z_t - 1) M(dt). \end{aligned}$$

The law of motion of η_t follows Ito's Lemma. ■

We can express the capital price as a function of the state variable. Let $\tilde{\eta}(u)$ be the state after a shock u hit the economy at η , similar for $\tilde{q}(u)$, then we have

$$\begin{aligned} \mu_t^q &= q'(\eta) \mu_\eta \eta = q'(\eta) \eta [(x_t^o - 1) R_t^{k,o} + \int_C R(C)(d\mu_+^o - d\mu_-^o) + \frac{a - \iota_t}{q_t} - \rho^o] \\ \tilde{\eta}(u)/\eta &= x_t^o + \int_C \min\{C, q/(\tilde{q}(u)u)\} (d\mu_+^o - d\mu_-^o), \quad \forall u. \end{aligned}$$

3.B.2 Proof of Proposition 5

We show that at least one of the two borrowing constraint must bind at any time in any equilibrium. Suppose that no borrowing constraint is binding, we have the following FOCs.

$$\partial x^i \quad R^k + \int_z \frac{1}{D^i(z)} dF^i(z) - \gamma^i = 0 \quad (3.10)$$

$$\partial \mu_-^i, \partial \mu_+^i \quad R(C) + \int_z \frac{(C \wedge 1/z)}{D^i(z)} dF^i(z) - \gamma^i = 0 \quad \text{for } \forall C. \quad (3.11)$$

We combine the FOC of two groups of agents on the same financial contract C :

$$\int_z \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] (C \wedge 1/z) dF(z) - \gamma^o + \gamma^p = 0; \quad \forall C$$

The left hand side is continuous hence we can take the first order derivative w.r.t. C

$$\int_{z < 1/C} \left[\frac{\lambda^o}{x^o + \int_{C'} (C' \wedge 1/z) (d\mu_+^o - d\mu_-^o)} - \frac{\lambda^p}{x^p + \int_{C'} (C' \wedge 1/z) (d\mu_+^p - d\mu_-^p)} \right] dF(z) = 0; \quad \forall C$$

Further more the following equation has to be satisfied given that $f(z) > 0$.

$$\left[\frac{\lambda^o}{x^o + \int_{C'} (C' \wedge C) (d\mu_+^o - d\mu_-^o)} - \frac{\lambda^p}{x^p + \int_{C'} (C' \wedge C) (d\mu_+^p - d\mu_-^p)} \right] = 0; \quad \forall C$$

Then let us consider the extreme case that $C = 1$, the equation above requires

$$\frac{\lambda^o}{x^o + 1(1 - x^o)} = \frac{\lambda^p}{x^p + 1(1 - x^p)},$$

which contradicts with the assumption that $\lambda^o < \lambda^p$.

3.B.3 Proof of Proposition 6

Lemma 6 *The optimists has to take negative position of some financial contract. That is $\mu_-^o \neq 0$.*

Proof. We prove this lemma by contradiction. If $\mu_-^o = 0$ a.s., then actually overall the optimists are lending to pessimists. Moreover, pessimists have to hold some capital and optimists are not constraint by the borrowing constraint ($d\mu_+^p = d\mu_-^o = 0$). From Equation 3.5 and 3.6, we have $\forall C$

$$\int_z \frac{(C \wedge 1/z) - 1}{D^o(z)} dF^o(z) - \int_z \frac{(C \wedge 1/z) - 1}{D^p(z)} dF^p(z) = 0.$$

However since $\forall(z), D^o(z) \geq 1 \geq D^p(z)$, the LHS of the equation above is negative. Formally,

$$LHS \leq \int_z \frac{(C \wedge 1/z) - 1}{D^p(z)} dF^o(z) - \int_z \frac{(C \wedge 1/z) - 1}{D^p(z)} dF^p(z) < 0$$

The strict inequality comes from the fact that $\lambda^o < \lambda^p$. ■

Lemma 7 $\text{supp}(\mu_+^i) \cap \text{supp}(\mu_-^i) = \emptyset$, for $i \in \{o, p\}$.

Proof. $R(C)$ is a continuous function of C , otherwise there will be an arbitrage opportunity. Optimists' FOCs are also continuous w.r.t. C .

$$R(C) + \int_z \frac{(C \wedge 1/z)}{D^o} dF^o(z) - \gamma^o + \xi^o C \geq 0 \quad = \text{ for } \quad \forall C \in \text{supp}(\mu_-^o) \quad (3.12)$$

$$R(C) + \int_z \frac{(C \wedge 1/z)}{D^o} dF^o(z) - \gamma^o \leq 0 \quad = \text{ for } \quad \forall C \in \text{supp}(\mu_+^o). \quad (3.13)$$

The two FOCs cannot equal to zero at the same C . For pessimists the argument is exactly the same. ■

Lemma 8 $\text{supp}(\mu_-^o) = \text{supp}(\mu_+^p)$ and $\text{supp}(\mu_+^o) = \text{supp}(\mu_-^p)$.

This lemma is directly from **Lemma 7** and market clearing conditions.

Lemma 9 For any $C_1 \in \text{supp}(\mu_-^i)$ and $C_2 \in \text{supp}(\mu_+^i)$, there exists C_L, C_H ($C_1 \leq C_L < C_H \leq C_2$) such that $\forall C \in (C_L, C_H)$ we have $C \notin \text{supp}(\mu_-^i) \cup \text{supp}(\mu_+^i)$, for $i \in \{o, p\}$.

Proof. WLOG we assume that $C_1 < C_2$ and focus on optimists.

Define $C_H \equiv \inf\{C | C \geq C_1, C \in \text{supp}(\mu_+^o)\}$, and $C_L \equiv \sup\{C | C \leq C_H, C \in \text{supp}(\mu_-^o)\}$. Given Lemma 3, $C_1 \leq C_L < C_H \leq C_2$. Given the definition of C_L, C_H , $\forall C \in (C_L, C_H)$ we have $C \notin \text{supp}(\mu_-^o) \cup \text{supp}(\mu_+^o)$. ■

Lemma 10 Let

$$A(1/z) \equiv \frac{\lambda^o}{x^o + \int_C (C \wedge 1/z)(d\mu_+^o - d\mu_-^o)} - \frac{\lambda^p}{x^p + \int_C (C \wedge 1/z)(d\mu_+^p - d\mu_-^p)},$$

then $A(1/z)$ is monotonic in any $[C_L, C_H]$, defined as in Lemma 4 such that $\forall C \in (C_L, C_H)$ $C \notin \text{supp}(\mu_-^o)$ and $C \notin \text{supp}(\mu_+^o)$.

Proof. $\forall C' \in (C_L, C_H)$

$$D^o \equiv x^o + \int_C (C \wedge C')(d\mu_+^o - d\mu_-^o) = x^o + \int_1^{C_L} C(d\mu_+^o - d\mu_-^o) + \int_{C_H}^{C'} C'(d\mu_+^o - d\mu_-^o)$$

$$\partial D^o / \partial C' = \int_{C_H}^{C'} (d\mu_+^o - d\mu_-^o),$$

which is a constant in (C_L, C_H) . Similar argument applies to D^p . Moreover since at any equilibrium $D^o \eta + D^p(1 - \eta) = 1$, exact one of D^o, D^p is increasing and the other is decreasing, although both D^o and D^p can be constant. ■

Given these lemmas we prove **Proposition 6** by contradiction. **Lemma 6** tells us that there is some $\tilde{C}$ such that $\tilde{C} \in \text{supp}(\mu_-^o)$. Suppose there exists $\bar{C} \neq \tilde{C}$ such that $\bar{C} \in \text{supp}(\mu_+^o)$.

As preparation, we combine the FOCs of both pessimists and optimists and zoom into these inequalities.

$$\begin{aligned} \int_z \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] (C \wedge 1/z) dF(z) - \gamma^o + \gamma^p + \xi^o C &\geq 0, &= \text{ for } \forall C \in \text{supp}(\mu_-^o); \\ \int_z \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] (C \wedge 1/z) dF(z) - \gamma^o + \gamma^p &\leq 0, &= \text{ for } \forall C \in \text{supp}(\mu_+^o). \end{aligned}$$

Let us define

$$H(C) \equiv \int_z \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] (C \wedge 1/z) dF(z) - \gamma^o + \gamma^p.$$

Then

$$\begin{aligned} H'(C_-) &= \int_{z \leq 1/C} \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] dF(z), \quad H'(C_+) = \int_{z < 1/C} \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] dF(z), \\ H'(C_+) - H'(C_-) &= -A(C)dF(1/C). \end{aligned}$$

Now we are ready to discuss the cases as follows:

Case 1: $\tilde{C} = 1$. Combine the FOCs w.r.t. the risky asset,

$$\int_z \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] (1 \wedge 1/z) dF(z) - \gamma^o + \gamma^p + \alpha \xi^o = 0.$$

Comparing it with $H(1)$ we have $H(1) = -\alpha \xi^o \in (-\xi^o, 0)$, hence $\tilde{C} \neq 1$.

Case 2: $\tilde{C} \geq C^f$. $H(C) + \xi^o C$ is linearly increasing in C when $C > C^f$. Suppose $H(C_0) + \xi^o C_0 = 0$ for some $C_0 > C^f$, we know that $H(C_0 - \epsilon) + \xi^o(C_0 - \epsilon) < 0$ for some small positive ϵ , which contradicts with the equilibrium condition. Hence we can focus on $\tilde{C} = C^f$.

A necessary condition for $\tilde{C} = C^f$ is $A(C^f) < 0$, because that $[H(C_-^f) + \xi^o C_-^f]'$

$$[H(C_-^f) + \xi^o C_-^f]' = \int_{z \leq 1/C^f} \left[\frac{\lambda^o}{D^o(z)} - \frac{\lambda^p}{D^p(z)} \right] dF(z) + \xi^o$$

must be non-positive to make sure $H(C_-^f) + \xi^o C_-^f \geq 0$.

Since $\tilde{C} = C^f$, $H(C^f) + \xi^o C^f = 0 > H(C^f)$. Moreover $H(C) = H(C^f)$ for $C \geq C^f$. We can rule out $\tilde{C} \geq C^f$. Next we show the contradiction of $\tilde{C} \in (1, C^f)$.

Define $C_L \equiv \sup\{C | C \in [\tilde{C}, C^f), H(C) = 0\}$, and $C_H \equiv \sup\{C | C \in (C_L, C^f], H(C) + \xi^o C\}$. For any $C \in (C_L, C_H)$, $C \notin \text{supp}(\mu_-^o) \cup \text{supp}(\mu_+^o)$.

For C_L being a local maxima of $H(C)$ we need $A(C_L) \geq 0$. Whereas for C_H being a local minima of $H(C) + \xi^0 C$ we need $A(C_H) \leq 0$.¹¹ Given **Lemma 10** we know that $A(\cdot)$ is monotonic in $[C_L, C_H]$. Hence

$$\int_{C_H}^{C^f} (d\mu_+^0 - d\mu_-^0) \geq 0.$$

$A(C_L) = A(C_H) = 0$ violates the definition of C_L ¹², so

$$\int_{C_H}^{C^f} (d\mu_+^0 - d\mu_-^0) > 0,$$

which again contradicts with the definition of C_L .

Case 3: $\tilde{C} \in (1, C^f)$. From the discussion of Case 1 we know that $\tilde{C} \neq 1$. Moreover since $H'(C) = 0$ for $C \geq C^f$ if $\tilde{C} > C^f$ we know that $H(C^f) = 0$. So we can focus on the region $(1, C^f]$.

Next we show that if there exists $\tilde{C}$ such that $H(\tilde{C}) = 0$ we must have $A(\tilde{C}) \geq 0$. It is trivial hold if $\tilde{C}$ is a local maxima in $(1, C^f)$. For $\tilde{C} = C^f$ we require that $H'(C_-^f)$ to be nonnegative. Given $H'(C_+^f) = 0$ we have $A(C^f) \geq 0$.

Finally we argue in a similar way as Case 2. We consider here only $\tilde{C} \in (1, \tilde{C})$, one can prove the other half by the same token.

Define $C_L \equiv \sup\{C | C \in [\tilde{C}, \tilde{C}], H(C) = 0\}$, and $C_H \equiv \sup\{C | C \in (C_L, \tilde{C}], C = H(C) + \xi^0 C\}$. For any $C \in (C_L, C_H)$, $C \notin \text{supp}(\mu_-^0) \cup \text{supp}(\mu_+^0)$.

Given **Lemma 10** we know that $A(\cdot)$ is monotonic in $[C_L, C_H]$. Hence

$$\int_{C_H}^{C^f} (d\mu_+^0 - d\mu_-^0) \geq 0.$$

$A(C_L) = A(C_H) = 0$ violates the definition of C_L , so

$$\int_{C_H}^{C^f} (d\mu_+^0 - d\mu_-^0) > 0,$$

which again contradicts with the definition of C_L .

3.B.4 Proof of **Proposition 7**

Lemma 11 $A(1/z)$ is decreasing in $(1, 1/\inf(\text{supp}(z))]$.

Proposition 6 allows us to focus on the contracts through which optimists borrow from pessimists. Hence the lemma above is directly from **Lemma 10**. We define $h(C) =$

¹¹If $C_H = C^f$ the local minima is not enough for $A(C) \leq 0$, but we have already proved $A(C^f) < 0$ above.

¹²If this is true, then $H'(C)$ is constant in (C_L, C_H) . However, given $H(C_L) = 0$ and $H(C_H) + \xi^0 C_H = 0$, $H'(C)$ has to be negative in (C_L, C_H) . This contradicts with $A(C_L) = A(C_H) = 0$, since neither $H'((C_L)_-)$ nor $H'((C_H)_+)$ can be negative.

$H(C) + \xi^0 C$, more explicitly

$$\begin{aligned}
h(C) &= \int_z \left[\frac{\lambda^0}{D^0(z)} - \frac{\lambda^p}{D^p(z)} \right] (C \wedge 1/z) dF(z) + \gamma^p - \gamma^0 + \xi^0 C \geq 0, \\
h'(C_-) &= \int_{z \leq 1/C} \left[\frac{\lambda^0}{D^0(z)} - \frac{\lambda^p}{D^p(z)} \right] dF(z) + \xi^0, \\
h'(C_+) &= \int_{z < 1/C} \left[\frac{\lambda^0}{D^0(z)} - \frac{\lambda^p}{D^p(z)} \right] dF(z) + \xi^0, \\
h'(C_+) - h'(C_-) &= -A(C) dF(1/C).
\end{aligned}$$

First observe that $A(1) < 0$. Given **Lemma 11**, $A(C) < 0$ for any $C \in [1, C^f]$. Hence $h(C)$ is convex in $C \in [1, C^f]$. The function $h(C)$ can reach zero only at the two boundaries or at the critical point. We discuss the three cases in turn.

If $h(1) = 0$ then $h'(1) \geq 0$, otherwise $h(1 + \varepsilon) < 0$ for some small $\varepsilon > 0$. Given **Lemma 11**, $h''(C)$ increases hence both $h''(C)$ and $h'(C)$ is always positive in $(1, 1/z_L]$. The only actively traded contract in this case is the one with $C = 1$.

Denote $\inf(\text{supp}(z))$ as z_L . If $h(1/z_L) = 0$ then $h'(1/z_L) \leq 0$, otherwise $h(1/z_L - \varepsilon) < 0$ for some small $\varepsilon > 0$, and $h''(1/z_L) \geq 0$. Given **Lemma 11** $h'(C)$ must increase weakly in $[1, 1/z_L]$. So there doesn't exist a critical point where $h'(C) = 0$. The only actively traded contract in this case is the one with $C = 1/z_L$.

Otherwise since $h(C)$ is a convex function in $[1, C^f]$ there can exist at most one critical point where $h'(C) = 0$ and $h''(C) > 0$, and that market is active.