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

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

---

Doctoral Dissertations, 2011-2026 Economics

2015

Essays in Financial Economics

Baron, Matthew David

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp014f16c514n

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

ESSAYS IN FINANCIAL ECONOMICS

Matthew D. Baron

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

Advisor: Wei Xiong

JUNE 2015

© Copyright by Matthew D. Baron, 2015. All rights reserved.

Abstract

This dissertation studies the linkages between financial intermediaries and capital markets. In particular, it analyzes potential sources of instability originating from the intermediary sector.

Chapter 1, "Countercyclical Bank Equity Issuance," studies capital structure dynamics of intermediaries over the credit cycle. I show that during credit booms, banks raise less equity (i.e. countercyclical equity issuance), even though more equity might help banks better absorb shocks. I argue that government guarantees distort the incentives of banks to raise equity and thus play a crucial role in driving banks' countercyclical equity issuance. My findings help explain why banks may resist raising equity during credit expansions, making financial distress more likely.

Chapter 2, "Credit Expansion and Neglected Crash Risk," co-authored with Wei Xiong, studies the role of bank shareholders in credit booms and financial crises. This chapter tests different theories of the causes and consequences of credit expansions through the lens of asset prices and highlights the role of over-optimism and neglect of crash risk by bank shareholders.

Chapter 3, "Risk and Return in High Frequency Trading," co-authored with Jonathan Brogaard and Andrei Kirilenko, studies a very different type of financial intermediary, high-frequency traders. Using a new proprietary, firm-level dataset, we show that high-frequency traders have strong incentives to take liquidity in financial markets and compete over small improvements in speed.

Acknowledgments

Most importantly, I would like to thank my advisor, Wei Xiong, who carefully taught me the craft of research and made my time in graduate school both enjoyable and intellectually rewarding. I am immensely grateful for his wisdom, generosity with his time, and dedication to his students.

I would also like to give my deepest thanks to my committee members, Markus Brunnermeier and Atif Mian, and especially to Tobias Adrian, Jonathan Brogaard, and Andrei Kirilenko, who have all played very important roles in guiding my research and career.

I am grateful to the following people for their support and/or helpful feedback on my work: Yacine Aït-Sahalia, Nina Boyarchenko, Nicola Cetorelli, Valentin Haddad, David Jordan, Mitchell Hoffman, Harrison Hong, Oleg Itskhoki, Jakub Jurek, Asani Sarkar, Hyun Shin, David Sraer, and my fellow PhD students at Princeton. I would also like to thank the Research group at the New York Fed, the Capital Markets and Risk Analysis groups at the Fed Board of Governors, and the Office of the Chief Economist at the Commodity Futures Trading Commission for hosting me over various summers, where much of the research contained in this dissertation was done.

Lastly, I would like to thank all of my friends and, of course, my family for supporting me throughout the PhD; in particular, my parents (Jeff and Patty) and my brother and sister (Dan and Rachel). I am very grateful for everything.

Prior presentations at conferences and meetings

Chapter 2 has been publicly presented at the following conferences and meetings: FSU SunTrust conference (4/2015), INET Conference on Macroeconomic Externalities (4/2015), AEA (1/2015), Cambridge-Princeton finance conference (9/2014), BIS conference on Banking and Asset Management (9/2014), UBC summer finance conference (8/2014), NBER Summer Institute: Monetary Economics (7/2014), NBER Asset Pricing meeting (4/2014).

Chapter 3 has been publicly presented at the following conferences and meetings: CUHK conference (6/2014), Market Microstructure and High Frequency Data conference (04/2014), Advances of Financial Mathematics conference of the Bachelier Finance Society (01/2014), Securities and Exchange Board of India conference (01/2014), World Federation of Exchanges workshop (12/2013), Banff International Research Station for Mathematical Innovation and Discovery workshop (09/2013), WFA (6/2013), Paris / TSE HFT conference (4/2013), Arne Ryde Workshop in Financial Economics (04/2013), The Future of Computer Trading in Financial Markets conference (01/2013), NBER Market Microstructure (12/2012), Institut Louis Bachelier Market Microstructure Seminar (12/2012), Canadian Workshop on the Microstructure of Financial Markets (10/2012).

On February 19, 2014, the U.S. Commodity Futures Trading Commission (CFTC) authorized the paper "Risk and Return in High Frequency Trading" for public disseminating, which appears in this dissertation as Chapter 3. This paper was previously authorized for public disseminating by the CFTC's Office of General Counsel (OGC) on September 27, 2012 and has previously circulated under the titled "The Trading Profits of High Frequency Traders." The CFTC requests the following disclaimer:

Disclaimer

The research presented in Chapter 3 was co-authored by Matthew Baron, a former CFTC contractor who performed work under contracts CFCE-11-CO-0126 and CFOCE-12-CO-0154, and Jonathan Brogaard, a former CFTC contractor who performed work under contracts CFCE 11-CO-0236 and CFOCE-12-CO-0210. Andrei Kirilenko, former CFTC Chief Economist, was also a co-author who wrote this paper in his official capacity with the CFTC. The Office of the Chief Economist and CFTC economists produce original research on a broad range of topics relevant to the CFTC's mandate to regulate commodity future markets, commodity options markets, and the expanded mandate to regulate the swaps markets pursuant to the Dodd-Frank Wall Street Reform and Consumer Protection Act. These papers are often presented at conferences and many of these papers are later published by peer-review and other scholarly outlets. The analyses and conclusions expressed in this paper are those of the authors and do not reflect the views of other members of the Office of Chief Economist, other Commission staff, or the Commission itself.

Contents

Abstract iii

Acknowledgments iv

Prior presentations at conferences and meetings v

Disclaimer vi

Introduction: Financial intermediaries & capital markets 1

Chapter I: Countercyclical Bank Equity Issuance 5

1 Introduction 5

2 Related Literature 11

3 A model of moral hazard and equity issuance 13

4 Hypothesis and Empirical Strategy 18

5 Data and Summary Statistics 21
5.1 Firm classification . . . 22
5.2 Key variables . . . 22
5.3 New historical data . . . 24
5.4 Summary statistics . . . 25

6 Empirical results 26
6.1 Countercyclical equity issuance for large U.S. banks . . . 27
6.2 Equity issuance across types of firms, historical time periods, and countries 29
6.2.1 Across types of firms . . . 29
6.2.2 Across historical time periods . . . 30
6.2.3 TBTF banks subsequent to bailouts in the 1980s . . . 31
6.2.4 Across countries . . . 33
6.3 Change in bank equity issuance upon adoption of the Euro . . . 35
6.4 Cost of bank equity capital and market timing . . . 36

7 Conclusion 38

Appendix 39
A. Data construction . . . 39
B. Proof of Proposition 1 . . . 41

Tables and Figures 43

Chapter II: Credit Expansion and Neglected Crash Risk 64

1 Introduction 64

2 Related Literature 69

3 Empirical Hypotheses and Methodology 71
3.1 Crash risk . . . 72
3.2 Equity premium . . . 73
3.3 Magnitude of equity premium . . . 76
3.4 Quantile regressions . . . 78
3.5 Standard errors . . . 80

4 Data and Summary Statistics 81
4.1 Key variables . . . 81
4.2 Summary statistics . . . 84

5 Empirical Results 87
5.1 Predicting crash risk . . . 87
5.2 Predicting the equity premium . . . 89
5.3 Excess returns subsequent to credit expansions and contractions . . . 92
5.4 Quantile regression-based analysis . . . 95
5.5 Further analysis . . . 97
5.5.1 Robustness in subsamples . . . 97
5.5.2 Test for small-sample bias . . . 98
5.5.3 Robustness of negative returns subsequent to large credit expansions . . . 100

6 Conclusion 102

Appendix 102
A. Data construction . . . 102
B. Methodology and results for small-sample bias test . . . 106

Tables and Figures 107

Chapter III: Risk and Return in High Frequency Trading 127

1 Introduction 127

2 Related Literature 131

3 Data and Methods 133
3.1 The E-mini S&P 500 futures market . . . 133
3.2 Categorizing Traders . . . 135
3.3 Returns and Profitability . . . 138

4 Risk and Return Performance of HFTs **140**
4.1 Distribution of Returns . . . 141
4.2 HFT Risk and Return in the Cross-Section . . . 142
4.3 Persistence of Returns . . . 144
4.4 Risk and Inventory Management . . . 147
4.5 Spectral Analysis . . . 148
4.6 HFT Profits by Counterparty . . . 151

5 Competition **153**
5.1 HFT Market Concentration over Time . . . 155
5.2 HFT Profitability over Time . . . 156
5.3 Entry and Exit . . . 159

6 Speed **162**

7 Conclusion **164**

Tables and Figures **165**

References **181**

Introduction

Financial intermediaries and capital markets

This dissertation studies the linkages between financial intermediaries and capital markets. In particular, it analyzes potential sources of instability originating from the intermediary sector. Chapters 1 ("Countercyclical Bank Equity Issuance") and 2 ("Credit Expansion and Neglected Crash Risk") study banks as intermediaries. Chapter 3, "Risk and Return in High Frequency Trading," studies a very different type of intermediary, high frequency traders.

Overview of chapters

Chapter 1, "Countercyclical Bank Equity Issuance," studies the capital structure of intermediaries over the credit cycle. It is well established that equity issuance for most non-financial firms is procyclical. This chapter shows that, in contrast, equity issuance for banks is countercyclical across credit cycles after 1980, as is retained income. Thus, during credit booms, banks raise less equity, even though more equity might help banks better absorb shocks.

Chapter 1 also argues that government guarantees play a crucial role in driving banks' countercyclical equity issuance. Countercyclical equity issuance arises in the U.S. during the bailouts of the 1980s, mostly in banks that equity markets perceive as "too-big-to-fail" (TBTF). Across types of firms, historical time periods, and countries, equity issuance is countercyclical when government guarantees are strong and procyclical when government guarantees are weak. Furthermore, bank equity issuance becomes

more countercyclical in Eurozone countries that gain increased implicit guarantees upon adopting the Euro. These findings help explain why banks may resist raising equity during credit expansions, making financial distress more likely.

Chapter 2, "Credit Expansion and Neglected Crash Risk," co-authored with Wei Xiong, studies the causes and consequences of credit expansions through the lens of stock prices. We show that bank shareholders predictably lose during credit expansion and argue that this phenomenon stems from over-optimism and neglect of crash risk by bank shareholders.

In particular, we show that, in a set of 20 developed countries over the years 1920-2012, bank credit expansion predicts increased crash risk in the bank equity index and equity market index. However, despite the elevated crash risk, bank credit expansion predicts lower rather than higher mean returns of these indices in the subsequent one to eight quarters. Furthermore, conditional on bank credit expansion of a country exceeding a 95th percentile threshold, the predicted excess return for the bank equity index in the subsequent eight quarters is -25.8%. We argue that this joint presence of increased crash risk and negative mean returns presents a challenge to the views that financial instability associated with credit expansions are simply caused by either banks acting against the will of shareholders or by elevated risk appetite of shareholders, and instead suggests a need to account for the role of over-optimism and neglect of crash risk by shareholders.

Chapter 3, "Risk and Return in High Frequency Trading," co-authored with Jonathan Brogaard and Andrei Kirilenko, studies the industrial organization of high-frequency trading (HFT). In particular, this paper examines high frequency trading (HFT) in the E-mini S&P 500 futures contract over a two-year period and finds that revenue is concentrated among a small number of HFT firms who achieve greater investment performance through liquidity-taking activity and higher speed. While the median HFT firm realizes an annualized Sharpe ratio of 4.3 and a four-factor annualized alpha

of 22.02%, revenues persistently and disproportionally accumulate to top performing HFTs, consistent with winner-takes-all industry structure. New entrants are less profitable and more likely to exit. Our results imply that HFT firms have strong incentives to take liquidity and compete over small increases in speed.

Common threads

What links these chapters is a focus on the interactions between intermediaries and capital markets. For example, all three chapters highlight how balance sheets of intermediaries (as diverse as commercial banks, broker-dealers, and high-frequency traders) affect outcomes in capital markets, including asset prices, tail risk, volatility, liquidity, and corporate issuance decisions. This dissertation also explores how the reverse may be true: how capital markets may in turn affect intermediaries. For example, the first two chapters highlight issues regarding how credit and equity markets may or may not discipline banks, and how market discipline might break down during credit expansions. Chapter 1 is in some sense about the breakdown of creditor market discipline during credit booms because of government guarantees to creditors, and Chapter 2 is about the breakdown of shareholder market discipline because of over-optimism of shareholders during credit booms.

Another common thread tying these chapters together is that they analyze potential sources of instability from different perspectives and related to different types of market participants: for example, financial crises when it comes to banks and events like the flash crash when it comes to high frequency trading. Chapters 1 and 2 are also about credit expansions in the broader macroeconomy, which can also be a key source of instability.

A final link between these chapters is the use to new and non-traditional data

sources. Chapter 3 uses a new, proprietary data set from the Commodity Futures Trading Commission to study the trading strategies of individual HFT firms. And for Chapters 1 and 2, I have collected new historical data on asset prices and balance sheets of securities firms and banks going back to the early 20th century in a variety of countries. Using long historical time series is important in this research for understanding the role of crises and rare events and their impact on financial markets. Besides being useful for understanding crises, historical datasets may give us a richer understanding of the connections between economic growth and banking in good times. This work can also help us understand the linkages between financial markets, intermediaries, and monetary policy, both across history and in emerging economies today, in addition to giving us early warning indicators of future crises.

Chapter I

Countercyclical Bank Equity Issuance

1 Introduction

A central issue in banking finance concerns the optimal capital requirements of banks and whether bank equity is expensive from both a social and private point of view. It is often argued that higher capital requirements act to reduce banks' risk-taking incentives and help banks withstand adverse shocks (see, for example, Admati et al., 2013a). As a result, higher capital requirements have been a key focus of recent financial reforms around the globe. To analyze optimal capital requirements of banks, it is important to understand why banks issue equity at some times and pay out equity at other times.

This paper investigates the drivers of bank equity over the credit cycle. A robust fact for equity issuance by non-financial firms is that they issue more equity in good times, i.e. equity issuance is procyclical.[1] This tendency makes sense because both investment opportunities for firms and their stock prices are high in good times, making equity issuance attractive for managers of non-financial companies. However, this paper demonstrates that the logic of this argument does not hold for banks: bank equity issuance is countercyclical.[2] During credit booms, banks raise less equity, even though more equity might help banks better absorb shocks. Banks instead wait until after the downturn, when it is potentially more costly to raise equity and too late to mitigate a crisis.

[1] See, for example, Marsh (1982), Pagano, Panetta, and Zingales (1998), Baker and Wurgler (2002), Hong, Wang, and Yu (2008), and Covas and Den Han (2011).

[2] This paper analyzes cyclicality in terms of the credit cycle (i.e. aggregate lending by banks to households and non-financial firms), which is most relevant for banks. Similar patterns of equity issuance hold when defining cyclicality in terms of GDP or other variables.

Why does the logic of procyclical equity issuance not hold for banks? The contribution of this paper is to demonstrate how bailout distortions and government guarantees may lead banks to choose countercyclical equity issuance. This paper shows that across types of firms, historical time periods, and countries, equity issuance is countercyclical when government guarantees are strong and procyclical when government guarantees are weak or absent. Importantly, countercyclical equity issuance arises in the U.S. during the bailouts of the 1980s and is driven by banks that equity markets perceive as having large implicit guarantees. In a difference-in-differences test, bank equity issuance becomes more strongly countercyclical in Eurozone countries that gain increased implicit guarantees upon adopting the Euro, relative to Eurozone countries without increased implicit guarantees. These results suggest that implicit guarantees distort incentives of banks to raise new equity and affect the dynamics of bank capital structure over the credit cycle.

The paper proceeds as follows. First, I develop an illustrative model of moral hazard to understand why bailout distortions and government guarantees may lead banks to choose countercyclical equity issuance. The model features a conflict between bankers (working in the interests of shareholders) and creditors. Bankers would like to take advantage of the option-like payoff to shareholders by paying out rather than raising equity. Though payouts enrich shareholders, they increase the probability of firm default. Foreseeing the possibility of payouts, creditors threaten to increase the interest rate if shareholders do not raise sufficient equity, the threat of which incentivizes bankers to raise equity. This threat represents market discipline. The key friction in the model is the presence of government guarantees to creditors: by partially shielding creditors from losses in the event of default, government guarantees dull creditor market discipline. While the model is static, we can think about the cyclicality of equity issuance using comparative statics, letting downside risk vary cyclically. As a result, when government guarantees are strong and therefore creditor market discipline weak, equity

issuance is countercyclical, since countercyclical issuance is privately optimal from the point of view of bankers and shareholders. In contrast, in the absence of government guarantees, creditor market discipline produces procyclical issuance.

Turning to the data, I look systematically at bank equity issuance, share repurchases, dividend payouts, and retained income across credit cycles. Bank equity issuance, measured in a variety of ways, is countercyclical over the period 1980-2012, while equity issuance by non-financial firms is procyclical. The data further indicate that countercyclical bank equity issuance is not driven by market timing or simply by time-varying cash-flows[3] or provisioning across the credit cycle.

I then present three key pieces of evidence showing that equity issuance is highly countercyclical when government guarantees are strongest and procyclical when government guarantees are weak or absent. First, over the time period 1980-2012, large banks have the strongest countercyclical equity issuance, followed by investment banks and then small banks, while non-financial firms have procyclical issuance, consistent with their different levels of government guarantees as reported in Gandhi and Lustig (2013).

Second, equity issuance is strongly procyclical in the U.S. before the creation of the FDIC (pre-1935), moderately procyclical from that time until the start of the Savings and Loan failures (1935-1979), and countercyclical thereafter (1980-2012). This pattern is consistent with the fact that there was little, if any, perception of government guarantees (with the exception of Fed discount window assistance) before the creation of the FDIC in 1933, along with the fact that TBTF government guarantees increased in the U.S. in the 1980s (for example, after the bailouts of Continental Illinois in 1984 and of a variety of large banks in the late 1980s and 1990s). Moreover, this shift in the 1980s is driven by banks that equity markets perceive as having large implicit guarantees. [4]

[3]Meaning that during credit expansions, banks may earn high profits; if they lack adequate investment opportunities to employ these profits, they may simply pay out excess profits (and vice versa for negative cash flow shocks).

[4]The equity market measure of implicit guarantees that I construct is based on equity returns

Third, among developed countries, countries with large relative banking sectors (bank assets greater than 200% of GDP, e.g., Switzerland, Hong Kong, the Netherlands) generally have procyclical bank equity issuance, while countries with small relative banking sectors (bank assets less than 200% of GDP, e.g., the U.S., Canada, Japan) have countercyclical bank equity issuance. This evidence is consistent with the idea that implicit government guarantees might be perceived to be weaker in countries with bank assets greater than 200% of GDP due to the government's limited fiscal capacity to fully protect creditors in the event of a large bank failure.[5] If investors anticipate in advance that some banks in these countries might be "too-big-to-save," creditor market discipline might be stronger and equity issuance might be less countercyclical or even procyclical.

While the above evidence suggests the potential importance of government guarantees and bailout distortions, these results alone are not sufficient to establish a direct link. For example, there might be other differences between large and small banks[6], over time[7], and between large and small countries.[8] Therefore, the heart of my evidence is

subsequent to announcements that give banks increased implicit guarantees (drawing on work by O'Hara and Shaw, 1990). The measure captures increased perceptions of implicit guarantees conferred to large banks; small banks do not see any effect. Details on constructing this measure are presented in Section 4.

[5]This perception is borne out by reality: Iceland and Cyprus had to impose deep haircuts on bondholders and foreign depositors. Ireland, in contrast, provided their large banking sector with a blanket guarantee for all creditors, which eventually forced the Irish government nearly into insolvency alongside the banks. Bailouts in many countries have grown so large that they have strained the public finances in many countries (Brown and Dinç 2011; Demirgüç-Kunt and Huizinga 2010; Acharya, Drechsler, and Schnabl, 2013). In particular, Spain, Portugal, and Ireland all had to ask the IMF and EU for emergency loans with the joint purpose of recapitalizing their largest banks and forestalling sovereign debt crises.

[6]See, for example, Hanson, Kashyap, and Stein (2011) on capital structure and competition for funding.

[7]Some events in the 1980s that might affect bank equity issuance include: deregulation (e.g., the Garn-St Germain Act of 1982, relaxation of branching and interstate bank restrictions) that gave banks more risk-taking ability and increased competition, increased regulation (Prompt Corrective Action, Basel I), which mandates recapitalization for distressed banks, and the rise of securitization and universal banking.

[8]For example, shocks to smaller countries might be larger (e.g., trade shocks, interest rate shocks, capital outflows), leading to greater market discipline. Alternatively, regulators, foreseeing both larger possible shocks and lack of fiscal capacity to backstop the largest banks, might be more vigilant in regulating the banking sector.

a difference-in-differences test demonstrating that bank equity issuance becomes more strongly countercyclical in Eurozone countries that gain increased implicit guarantees upon adopting the Euro, relative to Eurozone countries that do not gain increased implicit guarantees upon adopting the Euro.

This difference-in-differences test exploits the introduction of the Euro as an exogenous shock with differential effects on the banking sectors of different countries. The idea is that bank creditors in some Eurozone countries (e.g., Italy and Ireland) gained implicit protections in adopting the Euro, perhaps from the European Central Bank (ECB) or from other Eurozone countries.[9] I exploit the variation in implicit guarantees gained by different Eurozone countries upon adopting the Euro, which may make bank equity issuance more countercyclical in some countries than in others. This identification strategy assumes that the introduction of the Euro was a political project unrelated to developments in the banking sector and also that the only differential effect the Euro had on the banking sectors of its member countries was through its impact on implicit guarantees. I discuss my identification strategy and potential caveats in Section 3.

I use interest rate convergence as a proxy for implicit guarantees — specifically, the change between 1996 and 1999 in the real short-term interest rate spread relative to Germany. The fall in interest rates in countries such as Greece, in part, reflected lower perceived default risk, as creditors believed that other Eurozone countries and the ECB would protect Greek sovereign debt and the Greek banking system in the case of financial distress, because of the possibility of spillover in a monetary union (which is exactly what did happen).

To analyze whether changes in implicit guarantees lead to changes in the cyclicality of bank equity issuance, I divide the initial group of Euro-adopting countries into two

[9]Even if the implicit guarantees in joining the Euro are conferred mostly upon these countries' sovereign debt rather than on bank debt, implicit guarantees for sovereign debt can indirectly support banks because they provide the fiscal backing for individual governments to support their banks in the event of a crisis. The fact that the EU lent money to Ireland, Spain, and Portugal in 2011-2012 to help these governments jointly recapitalize their banks and forestall a sovereign debt crisis illustrates how implicit guarantees for sovereign debt can indirectly confer protections to a country's banking sector.

groups based on interest rate convergence. I classify countries as having low interest rate convergence if the change in the spread is less than one percentage point (e.g., Germany, the Netherlands, and Finland) or high if the change in the spread is greater than 1 percentage point (e.g., Ireland, Italy, and Portgual). Using the above classification, I estimate a difference-in-differences regression to test whether countries with high interest rate convergence have increased countercyclical issuance when joining the Euro, compared to countries with low interest rate convergence. The difference-in-differences coefficient is a statistically significant -0.026 (t = -2.561, dually clustered), meaning that, conditional on a one-standard deviation increase in credit expansion, net equity issuance decreases by 2.6 percentage points (annualized, as a fraction of book equity) in countries that gained implicit guarantees upon adopting the Euro, relative to countries that didn't gain implicit guarantees upon adopting the Euro. This result suggests that changes in implicit guarantees lead to economically-important changes in the cyclicality of bank equity issuance.

Finally, I investigate the role of market timing in bank equity issuance over the credit cycle, in light of the fact that credit variables can predict future returns, especially for bank stocks (Baron and Xiong, 2014).[10] Using credit expansion to predict future equity returns, I find that the cost of equity capital for both banks and non-financial firms is countercyclical over the credit cycle, with low cost-of-capital (i.e. high stock prices) during expansions and high cost-of-capital (i.e. low stock prices) during contractions. According to market timing theory, equity issuance should therefore increase during credit expansions alongside the lower cost of capital, and issuance should be procyclical. While procyclical issuance for non-financial firms is consistent with market timing, countercyclical issuance by banks indicates that banks do not appear to be simply timing the market.

This final result also ties into the ongoing debate on the cost of bank capital (see

[10]Similarly, but from a different theoretical perspective, Adrian, Etula, and Muir (2013) and Muir (2014) show how intermediary book leverage can price a variety of assets.

Admati, DeMarzo, Hellwig, and Pfleiderer, 2013b). While previous work (e.g., Baker and Wurgler, 2013; Kisin and Manela, 2014) estimates a time-averaged cost of bank capital, my result suggests that, due to the time-varying cost of capital, the cost of increased regulatory capital may be substantially lower than previously estimated, if regulators are careful to implement increased capital requirements during expansions.

The paper proceeds as follows. Section 1 discusses related literature. Section 2 presents an illustrative model. Section 3 presents my hypothesis and empirical strategy. Section 4 discusses the data and summary statistics. Section 5 presents the empirical results. Section 6 concludes.

2 Related Literature

In related papers, Acharya, Le, and Shin (2013) and Acharya, Gujral, Kulkarni, and Shin (2012) study dividend payments of selected large banks and broker-dealers in 2007-2009 and show how increased risk-shifting can take place via capital structure and dividend policies.[11] Similar to my paper, these papers highlight the conflict of interest between shareholders and bondholders of banks, leading bankers to prefer immediate payouts, even though these payouts substantially reduced banks' equity cushions in 2007-2009. I extend their work by systematically studying new equity issuance, share repurchases, and dividend payouts of all banks and large broker-dealers over credit cycles from 1925-2012. I find that banks reduce new issuance and retain less income during the peak of the credit cycle, not just when banks are on the verge of failing.

Another closely related paper concerned with agency frictions between shareholders and creditors is Admati, DeMarzo, Hellwig, Pfleiderer (2013a), which argues that shareholder-creditor conflicts can lead to leverage "ratchet effects"; shareholders may

[11]Similarly, Scharfstein and Stein, in a New York Times Op-Ed dated Oct. 20, 2008, point out that TARP recipients planned to pay out a substantial fraction of injected capital from TARP as dividends. Scharfstein and Stein argue that these policies benefit shareholders at the expense of the debt holders, while depleting bank capital at a time when banks crucially needed more capital.

avoid reducing leverage when leverage is already high, even if leverage reduction might increase the total value of the bank. Complementing their theory, my paper provides empirical evidence demonstrating that banks raise less equity and pay out more capital during credit expansions, when their leverage is already high.

My paper is also related to the extensive literature on capital structure and capital requirements of banks.[12] Related in focus to my paper, Hanson, Kashyap, and Stein (2011) argue that the high leverage of financial firms is driven by competition on the cost of funding. They show a strong inverse relationship between bank size and capital ratios. In addition, using data on state banking deregulation, they show that when banks are faced with more competition, banks tend towards both higher and more uniform levels of leverage. In contrast, my paper argues for the role of shareholder-creditor conflicts and government guarantees in driving capital structure decisions. However, it is important to note that both competition between banks and shareholder-creditor conflicts in the presence of government guarantees may be jointly present and mutually reinforcing. Competition, by reducing the franchise value of a bank, may further encourage banks to take advantage of creditors and government guarantees by issuing less equity. Consistent with this argument, Keeley (1990) finds that banks with more market power (and thus presumably more franchise value) tend to hold more capital and have lower default risk.

Theories of bank capital structure have generally focused on the role of capital structure in disciplining bank managers (e.g., Calomiris and Kahn, 1991; Diamond and Rajan, 2000, 2001), the benefits of liquidity creation (e.g., Kashyap, Rajan, and Stein, 2002; DeAngelo and Stulz, 2013), asset substitution and monitoring of borrowers (e.g., Holmstrom and Tirole, 1997; Mehran and Thakor, 2010), and competition on the lending side (e.g., Allen, Carletti and Marquez, 2011). Macroeconomic models incorporating the role of financial sector leverage and net worth have generally avoided the role of new equity issuance (e.g., Adrian and Boyarchenko, 2012; Adrian and Shin,

[12]Kashyap, Rajan, Stein (2008) provide an overview of various perspectives motivating bank capital requirements.

2013; He and Krishnamurthy, 2013; Brunnermeier and Sannikov, 2014). My study adds to this literature on bank capital structure by demonstrating the dynamic nature of equity issuance and payouts over the credit cycle and the key importance of government guarantees.

Finally, this paper is also related to the expansive corporate finance literature on equity issuance, share repurchases, and dividend decisions of firms, though this literature mostly focuses on non-financial firms. The literature highlights the roles of adverse selection (e.g., Ross, 1977; Myers and Majluf, 1984), debt overhang (e.g., Myers, 1977), tax benefits of debt (e.g., Miller, 1977; Graham, 2000; Fama and French, 2002), market timing (Baker and Wurgler, 2002), optimism (e.g., Malmendier and Tate, 2005), and disagreement between managers and shareholders (e.g., Dittmar and Thakor, 2007; Huang and Thakor, 2013). In contrast to these papers, my paper focuses on the determinants of equity issuance and payouts over time specifically for banks. The contrast between procyclical non-financial issuance and countercyclical bank issuance suggests that theories for non-financial firms may be less applicable for banks.

3 A model of moral hazard and equity issuance

Consider the following illustrative model, which borrows from the framework of Acharya, Le, and Shin (2013) and highlights the link between moral hazard, government guarantees, and equity issuance.

There are two periods, today and tomorrow. The bank's balance sheet consists of A_0 risky assets today, which turn into $\tilde{A}$ assets tomorrow, where $\tilde{A}$ is a random variable. The bank is financed by L liabilities, with the rest of the assets financed by shareholder common equity. The bank defaults if the realization of assets plus any new cash raised (or minus any cash paid out) is less than liabilities L. Furthermore, assume that $\tilde{A}$ is uniformly distributed over the interval $[\underline{A}, \overline{A}]$, such that $\underline{A} < L < \overline{A}$. Let $\sigma = \frac{\overline{A}-\underline{A}}{A_0}$

represent total asset risk, and let $\sigma_D = \frac{A_0 - \underline{A}}{A_0}$ represent downside or tail risk. The discount rate is zero.

The model adopts the point of view of a risk-neutral bank manager working in the interest of shareholders to maximize their expected payoff. The banker chooses n, the value of new capital to raise through equity issuance (with $n < 0$ corresponding to payouts), with N the maximum amount of capital that the bank can raise or pay out. Given the option-like payoff to equity holders due to limited liability, the banker wishes to maximize shareholder value by taking advantage of creditors and government guarantees, while also trying to protect the bank's franchise value V (which can be thought of as the present value of future cash flows), which would be lost in the event of default. The creditors, anticipating that the banker is trying to exploit limited liability, thus threaten to charge a higher interest rate r in the presence of increased moral hazard to compensate themselves for a higher probability of default. This threatened higher interest rate is the key channel here for market discipline — it serves to incentivize increased equity issuance in line with the "first-best."

The key friction is the presence of government guarantees: fraction γ of all creditor losses of principal will be reimbursed by the government in the event of default. By partially shielding creditors from losses in the event of default, the government guarantees result in less market discipline.[13]

Lastly, I also include the possibility of market timing to demonstrate how moral hazard can counteract market timing. If the bank issues shares whose total market value should be n (which is inclusive of the dilution from issuing these new shares), the firm receives $(1 + a) \cdot n$ in cash, where a is exogenous and reflects the premium or

[13]This friction reflects extensive evidence that implicit guarantees diminish the relationship between bank risk-taking and funding costs (e.g., Flannery and Sorescu, 1996; Flannery and Nikolova, 2004; Morgan and Stiroh, 2005; Balasubramanian and Cyree, 2011; Strahan, 2013; and Acharya, Anginer and Warburton, 2014) and that this diminished creditor market discipline translates into increased risk-taking by banks (e.g., Keeley, 1990; Duchin and Sosyura 2014; Demirguc-Kunt and Detragiache 2002, 2006; Panageas, 2010; Fahlenbrach, Prilmeier, and Stulz, 2012).

discount from equity markets.[14]

The model is static, but it is still useful to think about how the model's predictions might vary over the credit cycle using comparative statics. One way to do capture variation over the credit cycle is to think of perceived downside risk, σ_D, as varying countercyclically (which in turn drives countercyclical market discipline in the model). While the model assumes that downside risk σ_D is a fixed parameter known by all agents in the model, we can think of it more generally as reflecting time-varying beliefs about downside risk or as reflecting measured value-at-risk of banks (which is backward looking) or the VIX (which is forward looking) and which are countercyclical over the credit cycle (Adrian and Shin, 2013).[15]

Returning to the model, the expected payoff to shareholders is given below by Equation 1.[16]

$$\begin{aligned} \text{Expected Shareholder Value} \quad = \quad & E[\tilde{A} - (1+r)L + (1+a)\cdot n & (1) \\ & \mid \tilde{A} + (1+a)\cdot n \geq L] & (2) \\ & + Pr[\tilde{A} + (1+a)\cdot n \geq L]\cdot V \; - \; n \end{aligned}$$

The first term represents the value of assets plus new equity raised, minus liabilities, conditional on the bank's survival. The second term is the franchise value of the bank, minus the interest paid to creditors, conditional on the bank's survival. The third term is the cash that shareholders give up when they purchase new equity from the bank.

The creditors' breakeven (or participation) constraint is given in Equation 3. Antic-

[14] In the model, equity payouts take the form of share repurchases rather than dividends, in order to capture market timing.

[15] Another possible driver of countercyclical equity issuance could be time-varying regulation. We can think of regulation being lax during credit expansions — perhaps reflecting institutional frictions like regulatory capture, political lobbying, and lax monitoring of regulators by the public, or behavioral frictions like forgetting about past crises — and more vigilant after the crisis, when regulators are more closely watched by the public, and stress tests, higher capital requirements, and financial reform legislation become more common.

[16] For calculational simplicity, I make the approximation that the condition of the bank's survival is that $\tilde{A} + n \geq L$, even though it is really $\tilde{A} + n \geq (1+r)L$.

ipating that bankers may exploit limited liability and government guarantees to creditors, creditors set the interest rate to compensate themselves for expected losses.

$$\begin{aligned} 0 &= Pr[\tilde{A} + (1+a) \cdot n \geq L] \cdot rL \\ &\quad +(1-\gamma)\, E[\tilde{A} - L + (1+a) \cdot n \,|\, \tilde{A} + (1+a) \cdot n < L] \end{aligned} \tag{3}$$

The first term represents interest payments the creditors receive, conditional on the bank's survival. The second term represents expected losses of the creditors, conditional on the bank's failure.

Assume for now that equity is priced correctly ($a = 0$). The "first-best" solution, the one that takes into account the total value of the firm (shareholders and creditors together) plus the net costs to the government, is that the bank should always raise the maximum allowed equity (n = +N) in order to minimize the probability of default, which is a pure loss from a social point of view. To see this, add expected shareholder value from Equation 1 to expected creditor value from Equation 3:

$$\begin{aligned} \text{Expected Total Value} &= E[\tilde{A} - (1+r)L + n \,|\, \tilde{A} + n \geq L] && (4) \\ &\quad +Pr[\tilde{A} + n \geq L] \cdot V && (5) \\ &\quad -n + Pr[\tilde{A} + n \geq L] \cdot rL + E[\tilde{A} - L + n \,|\, \tilde{A} + n < L] \\ &= E[\tilde{A} - L] + Pr[\tilde{A} + n \geq L] \cdot V \end{aligned}$$

Thus, it is clear that expected total value is increasing in n, since more equity decreases the probability of default. Since default results in forfeiting the bank's franchise value, it is a pure social loss. All other private costs and benefits of equity cancel out, merely transferring value between equity holders and creditors.

We can now solve generally for the banker's optimal choice of equity issuance or payout, taking into account the presence of government guarantees. As a result of

the uniform distribution, the shareholder's expected value and the creditor's breakeven constraint are:

$$\begin{aligned} \text{Expected Shareholder Value} \;=\; & \frac{\frac{1}{2}(\overline{A} - L + (1+a)\cdot n)^2}{\overline{A} - \underline{A}} + \\ & V\frac{(\overline{A} - L + (1+a)\cdot n)}{\overline{A} - \underline{A}} \\ & - n \;-\; (rL)\cdot Pr[\tilde{A} + (1+a)\cdot n \geq L] \end{aligned} \tag{6}$$

$$0 = (rL)\cdot Pr[\tilde{A} + (1+a)\cdot n \geq L] + (1-\gamma)\,\frac{\frac{1}{2}(\underline{A} - L + (1+a)\cdot n)^2}{\overline{A} - \underline{A}} \tag{7}$$

Substituting the creditor's breakeven constraint (Equation 7) into the shareholders' expected value (Equation 6) results in an upward-facing quadratic equation with respect to n, and therefore no interior solution exists. Thus, a bank manager who is maximizing shareholder value will choose either n = +N (issuing the maximum amount) or n = -N (paying out the maximum amount), depending on whether equity issuance or payout is more desirable. Mathematically, the optimum n = +N or -N depends on whether the vertex of Equation 6 is greater or less than n = 0.

The result is the following proposition, a proof of which can be found in the Appendix:

Proposition 1: Let $\Psi = \frac{1}{\gamma}\left[\frac{V}{A_0} + \sigma\cdot\frac{a}{1+a}\right] + (1-\sigma_D) - \frac{L}{A_0}$.

Shareholders' optimal equity issuance or payout is a corner solution:

n = +N, if $\Psi > 0$,

n = -N, if $\Psi < 0$,

with indifference between +N and -N, if $\Psi = 0$.

Thus, issuance is more likely when: government guarantees γ are low, franchise value

of the bank $\frac{V}{A_0}$ is high, perceived downside risk σ_D is high, the price of equity is high (i.e. $a > 0$), and book leverage $\frac{L}{A_0}$ is high.[17]

If, as described earlier, we think of perceived downside risk σ_D as countercyclical across the credit cycle, we can compare different types of firms and implicit guarantee regimes. For banks with a high level of government guarantees γ, net equity issuance may be countercyclical, as countercyclical variation in σ_D can drive Ψ above and below the zero threshold, generating countercyclical net issuance. However, for banks with a low level of government guarantees γ, net equity issuance may be substantially less countercyclical, perhaps nearly acyclical, because a low γ keeps Ψ positive almost all the time. After adding in market timing (i.e. letting a be procyclical), net equity issuance can be procyclical with low government guarantees γ.

4 Hypothesis and Empirical Strategy

The model suggests that government guarantees, by partially shielding creditors from losses in the event of default, may result in less creditor market discipline and thus stronger countercyclical equity issuance. This section introduces a hypothesis motivated by the model and my empirical strategy to test it.

This paper exploits the fact that government guarantees vary widely across firms, countries, and historical time periods. For example, Gandhi and Lustig (2013) find a tail-risk subsidy that is greatest for large commercial banks, less for large investment banks, and negative for small commercial banks. Non-financials are generally not perceived to have government guarantees protecting creditors. Little, if any, perception of government guarantees (with the exception of Fed discount window assistance) existed before the creation of the FDIC in 1933. Perceptions of implicit government guarantees are generally regarded to have increased in the U.S. in the 1980s and 1990s — for exam-

[17]The fact that payouts are even more likely when leverage is already high is the "ratchet effect" analyzed in Admati et al. (2013a).

ple, after large-scale interventions, such as the bailouts of Continental Illinois in 1984 and LTCM in 1998. Similarly, perceptions of implicit guarantees may have changed when certain Euro countries (e.g., Ireland, Italy, Portugal) adopted the Euro. Across this wide variation in government guarantees, do we see differences in equity issuance associated with different levels of government guarantees?

Motivated by the model and the considerations discussed above, I put forward the following hypothesis:

Hypothesis: Equity issuance is more strongly countercyclical in the presence of higher levels of government guarantees. In contrast, equity issuance is procyclical when government guarantees are weak or absent.

At the heart of my analysis is a difference-in-differences test that exploits the introduction of the Euro as an exogenous shock that has differential effects on the banking sectors of different countries. As mentioned earlier, the identification strategy assumes that the introduction of the Euro was a political project unrelated to developments in the banking sector and that the only differential effect the Euro systematically had on the banking sectors of Eurozone countries is through implicit guarantees. One caveat of this approach, however, needs to be pointed out: to the extent that adopting the Euro also reduced various forms of risk (e.g., exchange rate and inflation risk), potentially leading to laxer market and regulatory discipline of banks, increased countercyclical issuance may stem from reduced risk rather than increased implicit guarantees. However, the ongoing Eurozone crisis illustrates quite dramatically how lower funding costs for banks in Eurozone periphery countries did not reflect lower risks (risks were quite high, as it turned out) but rather the correct perception that EU fiscal and monetary support would be used to rescue and recapitalize important banks in countries such as Greece,

Spain, Portugal, Italy, and Ireland (sometimes indirectly, through supporting sovereign debt).

My regression methodology is as follows. To analyze the cyclicality of equity issuance, payouts, and retained earnings, I estimate the following regression:

$$\begin{aligned} Net\,Equity\,Issuance\,minus\,Dividends_{i,t} &= \alpha + \delta \cdot CreditExpansion_t \\ &\quad + \gamma' \cdot Controls_t + \epsilon_t \end{aligned} \tag{8}$$

The standard dependent variable is *net equity issuance minus dividends* aggregated over all large banks, though I also look at other dependent variables, such as *retained income*. *Credit expansion*, i.e. $\Delta(bank\,credit/GDP)$, the standardized year-over-year change in aggregate bank credit to GDP, is my standard measure of the credit cycle. The time-series regression is estimated with Newey-West standard errors on quarterly data.

To test differences between two groups (e.g., small vs. large countries), I estimate the following regression, where variable i can denote either individual firms or countries:

$$\begin{aligned} Net\,Equity\,Issuance\,minus\,Dividends_{i,t} &= \alpha_i + CreditExpansion_{i,t} \cdot \\ &\quad (\delta_i + \delta_1 \cdot 1_{TreatmentGroup}) \\ &\quad + \gamma' \cdot Controls_{i,t} + \epsilon_{i,t} \end{aligned} \tag{9}$$

$$\tag{10}$$

The above estimation approach is conceptually similar using Equation 8 to estimate δ firm-by-firm or country-by-country with a time-series regression and then testing the difference in the average δ between the two groups (i.e. small vs. large countries). However, special care must be taken to test differences in countercyclical issuance across firms or countries because innovations to equity net issuance may be correlated across firms or countries (due to common global shocks) and over time (due to persistence of firm- or country-specific shocks). Therefore, I adopt a panel regression

approach for Equation 9 and estimate standard errors following Thompson (2011) that are dually clustered on time and firm/country, to account both for correlations across firms/countries and over time, which is the approach recommended by Petersen (2009).

To test difference-in-differences (for example, countries with increased implicit guarantees versus those without, before and after adopting the Euro), I estimate the following regression, where variable i can denote either individual firm or country:

$$\begin{aligned} Net\,Equity\,Issuance\,minus\,Dividends_{i,t} &= \alpha_i + CreditExpansion_{i,t} \cdot \\ &(\delta_i + \delta_1 \cdot 1_{Treat} + \delta_2 \cdot 1_{year>\overline{y}} \\ &+ \delta_3 \cdot 1_{Treat} \cdot 1_{year>\overline{y}}) \\ &+ \gamma' \cdot Controls_{i,t} + \epsilon_{i,t} \end{aligned} \tag{11}$$

As before, I estimate dually-clustered standard errors following Thompson (2011). The coefficient of interest is the difference-in-differences coefficient, δ_3, which tests the change in the cyclicality of equity issuance in the treatment group (the group that received increased implicit guarantees) compared to the change in the control group (the group that did not).

5 Data and Summary Statistics

I construct several related data sets: 1) a time series data set of aggregate credit expansion, equity issuance, payouts, and other variables for different subsets of firms (large banks, small banks, investment banks, non-financials) over the sample periods 1925-2012; 2) a panel data set of U.S. banks (a de-aggregated version of the data set above); and 3) a cross-country panel data set of aggregate bank lending, equity issuance, payouts, and other variables for 26 developed countries over the sample period 1980-2012.

5.1 Firm classification

I classify firms into large commercial banks, small commercial banks, investment banks, and non-financial firms. Commercial banks (which I also take to also include savings and trust institutions) are selected by SIC codes 6020-6036, then further filtered by NAICS and GIC codes to remove erroneous firms; holding companies that are primarily banks (e.g., Citigroup) with SIC codes 6712 are also included as commercial banks. Investment banks are selected by SIC codes 6712 and 6211, further filtered by NAICS and GIC codes, and then manually filtered to include only firms whose primary line of business (based on Compustat descriptions and web searches) is investment banking. Non-financial firms are Compustat manufacturing firms (SIC codes 2000-3999). To the extent that manufacturing firms are not representative of non-financial firms, the results in this paper should be viewed accordingly.

For each quarter, I categorize large commercial banks as the largest 20 banks by assets, and small commercial banks as banks ranked between 50 and 100 by assets.[18] The data is quarterly data from 1925 to 2012 for large commercial banks and from 1980-2012 for small commercial banks, investment banks, and non-financial firms.

5.2 Key variables

The main variable analyzed in this paper is *net equity issuance* (new equity issuance minus share repurchases) in dollar values, normalized by beginning-of-quarter *book equity*. To construct *net equity issuance*, I look at the quarterly net change in *common shares outstanding* from Compustat (after adjusting for stock splits and effective dilutions) and value the equity based on the end-of-quarter price for the equity. Stephens and Weisbach (1998) and Jagannathan, Stephens, & Weisbach (2000) suggest that this

[18]The term "small" is a misnomer, since these banks are actually quite large. These banks tend to be regional banks and savings and loans. The term "small" is simply used to distinguish these banks from "large" banks.

method is generally accurate in measuring net issuance. They also suggest a method for separating equity issuance and share repurchases, which I follow (see Appendix). I similarly construct a measure for dividend payouts normalized by book equity. As a robustness check, I also use alternative approaches for measuring equity issuances and repurchases.

As the key variable representing the credit cycle, I define the variable *credit expansion*, also denoted Δ(bank credit / GDP), as the year-over-year change in aggregate bank credit to GDP, where bank credit is defined as credit from banks to domestic households and private non-financial corporations, constructed with data from BIS and from Schularick and Taylor (2012). See the Appendix for details.

For data post-1965, I collect financial characteristics of firms from standard sources. For the U.S., I use Compustat for *shares outstanding, assets, book equity*, and two measures of income (*net income before changes to loan-loss reserves* and *net income after changes to loan-loss reserves*, normalized by beginning-of-quarter book equity). For measuring the cost of capital for the banking sector, I use quarterly *stock returns* and *dividend yields* from CRSP, from which I construct quarterly excess returns (total returns minus the three-month T-bill rate) for each of the sectors (market-cap-weighted indices for large banks, small banks, investment banks, and non-financial firms).

For international data on 26 developed countries (1980-2012), I use Thompson-Reuters Datastream for bank *net equity issuance*, *payouts*, *equity returns*, *income*, *book equity*, and *market equity*. For international data on bank assets to GDP, I use OECD data when available or data from each country's central bank otherwise. For Eurozone economies, I also calculate real short-term interest rate spreads relative to Germany, using data from Eurostat, Global Financial Data, IMF, and OECD. See the Appendix for information on control variables for the regressions.

Lastly, to construct a measure of market-perceived changes in implicit TBTF guarantees for U.S. banks, I look at the one-day abnormal returns corresponding to specific

announcements, following O'Hara and Shaw (1990). Those announcements are: 1) The OCC announcement related to Continental Illinois (Sept. 20, 1984); and 2) President George H.W. Bush's proposal for a Savings and Loan bailout (Feb. 7, 1989). See the Appendix and Section 5.2 for historical details of these events and their relationship to TBTF. For each bank, I average the abnormal returns over the two announcements to get the TBTF measure for each bank.

5.3 New historical data

I hand-collect new data for banks over the period 1925-1965. It is important to my study to use data going back to 1925 in order to establish that bank equity issuance was procyclical prior to 1980 and highly procyclical before the establishment of the FDIC in 1933. Hand-collected data is needed because bank stocks do not appear in CRSP before 1965, since they were traded almost exclusively OTC before around 1965. Likewise, Compustat begins in 1965 for all firms.[19] Given the challenges of hand-collecting data, I only analyze the 15 largest banks ranked by assets. The top 15 banks are selected on a yearly basis, but, interestingly, the list of top 15 banks stayed remarkably stable throughout the period 1925 - 1965.[20]

I gather hand-collected data from two sources: yearly editions of Moody's Bank and Finance Manuals (for *new equity issuance*, *dividends*, *stock splits and dilutions*, *net income* both before and after changes to loan-loss reserves, and annual bank balance sheet data, such as *cash*, *loans*, *deposits*, and *book equity*) and Wall Street Journal archives (for *bank equity prices* on a quarterly basis). I also collect from Moody's Manuals specific data on how the new capital was raised, whether through secondary

[19]Studies of the Great Depression using individual bank-level data have had similarly to overcome these data limitations. One prominent example is Calomiris and Mason (2004), which hand-collects Fed call reports data of over 8,000 banks between 1929-1931 to study the causes of bank failures in the Great Depression.

[20]Only a few banks declined in size and dropped out of the sample (e.g., Irving Trust), while only a few banks entered the sample (e.g., Mellon). The three largest national banks just before the Great Depression were Chase, City (now Citi), and Bank of America, which are still the largest today.

offerings or subscriptions rights (which includes information on how many shares were offered and at what price). Using the above information, I calculate effective stock dilutions for each bank and year, which may result from stock dividends, rights issues, mergers, and actual splits.

5.4 Summary statistics

Table 1 presents summary statistics related to equity issuance, payouts, and other key variables for large U.S. banks. Panel A presents cross-sectional statistics for the 20 largest banks during 2005, around the peak of the most recent credit cycle. Summary statistics are reported for *bank assets*, *book equity to assets*, *annualized growth of assets*, and *annualized growth of book equity*. This panel demonstrates substantial heterogeneity across large banks, even at the peak of the credit cycle. For example, there is substantial variation in leverage across large banks, from 6.5% to 16.0% book equity to assets. There is also substantial heterogeneity in individual banks' balance sheet expansion, with growth of assets ranging from 0.6% to 72.8% and growth of book equity ranging from -10% to 17.2%. Thus, even though all firms are increasing their assets (with a mean growth rate of 18.7%), some banks are decreasing their book equity.[21]

Panel A also shows substantial heterogeneity in equity issuance and payouts in the cross-section of large banks in 2005. Summary statistics are reported for *net equity issuance minus dividends*, *issuance*, *repurchases*, *dividends*, and *retained income*, all expressed as a percentage of book equity. *Net equity issuance minus dividends* has a mean of -9.4%, suggesting that banks choose to pay out high levels of capital during credit expansions. *New issuance* and *share repurchases* are 0% of book equity for about half of banks, but range as high as 4.6% (for issuance) and 17.6% (for repurchases). *Dividend payouts* range from 1.2% to 11.0%.

[21]This decrease in book equity observed for some banks is not due to losses, since return on equity (after changes to loan-loss reserves) is positive for all large banks (results not shown).

Panel B presents aggregated time-series statistics over the period 1980-2012. The first row looks at *credit expansion* (the year-over-year change in bank credit to GDP). *Credit expansion* is also plotted in Figure 1, Panel A, along with alternative measures of the credit cycle such as change in total credit to households and non-financials (from the BIS) and change in bank loans to GDP (from the Fed flow of funds), which are shown to be similar. As shown in Table 1, Panel B, credit expansion can range as high as 2.4% of GDP (i.e. a credit expansion) and as low as -4.3% (i.e. a credit contraction).

Looking at equity issuance and payouts over time for the 20 largest banks aggregated together, Table 1, Panel B, reports that *net equity issuance minus dividends*, its components, and *retained income* are highly time-varying over credit cycles, ranging from annualized 6.9% to 7.0% for equity net issuance, from -12.9% to 3.2% for equity net issuance minus dividends, and from -4.4% to 14.5% for retained income. Figure 1, Panel B, plots *net equity issuance*, *dividends*, and *retained income* over time for the 20 largest banks aggregated. As shown later in the paper, banks' net equity issuance and retained earnings are countercyclical over the credit cycle, with net payouts during credit expansions, followed by net recapitalization during credit contractions.

6 Empirical results

In this section, I first present the main result of this paper, that net equity issuance by large banks over the period 1980-2012 is countercyclical. Then, I present evidence related to patterns of equity issuance across types of firms, historical time periods, and countries, followed by results of the difference-in-differences tests related to the adoption of the Euro. Lastly, I discuss the time-varying cost of bank capital and show that banks do not seem to be timing the market in aggregate.

6.1 Countercyclical equity issuance for large U.S. banks

I present the main result of this paper, that equity issuance by large banks over the period 1980-2012 is countercyclical.

To analyze the cyclicality of equity issuance, payouts, and retained earnings, I estimate Equation 8 with Newey-West standard errors using the dependent variable net equity issuance minus dividends aggregated over all large banks over the time period 1980-2012. Estimates are reported in Table 2, while Figure 2, Panel A, plots the corresponding scatter plot for large banks' net equity issuance over credit cycles, 1980-2012. Column 1 of Table 2 shows that a one standard deviation increase in credit expansion is associated with a statistically significant 2.4 percentage point decrease in net equity issuance (t = -3.383). I also estimate Equation 1 with various controls, such as return on equity (past 4-quarter smoothed), book-to-market, past stock returns (indicator variables of whether bank stocks increased more than 50% or decreased more than 25% in the last two years), term spread, and non-residential investment to capital (i/k). These controls account for other factors that might influence bank issuance: bank profitability, equity valuations, and investment opportunities. In Column 2, I find that the coefficient on credit expansion is relatively unchanged in magnitude and significance in the presence of these controls. Lastly, in Columns 3 and 4, I estimate the regression conditional on positive credit expansion (i.e. on the subsample where Δ(bank credit / GDP) ¿ 0), with and without controls. The fact that the estimates are about twice as large in magnitude and still significant indicates that the results for the full sample are not just driven by recapitalization of banks in bad times. Banks actually issue less equity in good times.

Equity net issuance minus dividends is the standard dependent variable in my analysis, but I also look at other measures of equity inflows and outflows of large commercial banks (1980-2012), such as new issuance, repurchases, dividends, and retained income. In Table 3, I repeat my analysis by estimating Equation 8 for these other dependent

variables, all of which are normalized by book equity. As in Table 2, I also estimate regressions with and without controls, both on the entire sample and conditional on credit expansion being positive. The estimates in Table 3 demonstrate that all of these measures of capital inflow and outflow are consistent with countercyclical bank equity.

One potential concern might be that equity issuance and payouts may simply be driven by cash flow considerations: during credit expansions, banks might be highly profitable and generate more profit than they have investment opportunities to re-invest those profits, in which case banks might just pay out the excess profits. If this were the case, we would see payouts increase proportionately to income (or some smoothed measure of income). However, I show that, since retained income is countercyclical, banks increase their payouts faster than their income increases.[22] In particular, the countercyclicality of retained income is not driven by increased provisioning during credit expansions, because this measure of retained income is before changes to loan-loss reserves. Thus, banks are actually increasing capital payouts during expansions and funding new investment opportunities with debt rather than retained income or new issuance.

I perform various robustness checks (see the Online Appendix) and find that equity issuance for large banks is robustly countercyclical. For example, firm level regressions with fixed effects (without and with firm-level controls, such as log assets, leverage, return on equity, book-to-market, and past equity returns), estimated with Thompson's (2012) dually clustered standard errors, also yield significantly negative coefficients. I also repeat the analysis but normalize issuance and payouts by market value of equity rather than book value. I also show that results are robust to calculating net equity issuance using shares outstanding from other data sources (CRSP, Bankscope, Datastream). Stephens and Weisbach (1998) also suggest alternative approaches for measuring equity issuances and repurchases, for example, using Compustat's Purchase

[22]Retained income is income minus payouts, so countercyclical retained income would imply that banks increase their payouts faster than their income increases.

of Common and Preferred Stock and Sale of Common and Preferred Stock. Results are robust to these alternative measures of equity issuance and payouts; however, these two series only extend back to the late-1990s or early 2000s, making them unsuitable for studying longer time periods. Lastly, I show that countercyclical bank equity issuance is robust to other measures of the credit cycle, such as change in loans from commercial banks to GDP (from Fed flow of funds, 1947-2012) and change in total credit to GDP (all credit, not just from banks, to households and non-financial firms, from BIS).

6.2 Equity issuance across types of firms, historical time periods, and countries

I present evidence related to patterns of equity issuance across types of firms, historical time periods, and countries.

6.2.1 Across types of firms

I look at differences in government guarantees across types of firms. Recall that Gandhi and Lustig (2013) find a tail-risk subsidy that is greatest for large commercial banks and GSEs, less for large investment banks, and negative for small commercial banks, while non-financials in the U.S. are generally not perceived to have government guarantees protecting creditors. Do we see different patterns of equity issuance corresponding to these different levels of government guarantees?

Table 4 reports estimates for Equation 8 for large banks, GSEs, investment banks, small banks, and non-financial firms over the sample period 1980-2012. Figure 2, Panels B-D, plot the corresponding scatter plots. Consistent with the hypothesis, large banks have the most countercyclical net equity issuance with a coefficient of -0.024 (t-stat = -3.383), followed by GSEs with a coefficient of -0.018 (t-stat = -1.927), followed by investment banks with a coefficient of -0.013 (t-stat = -1.082) and small banks with

a coefficient of -0.003 (t-stat = -0.731); non-financials have procyclical issuance with a positive coefficient of 0.008 (t-stat = 1.972). Coefficients and their corresponding t-values are relatively unchanged in the presence of controls. Thus, we see cyclicality of issuance consistent with different levels of government guarantees.

6.2.2 Across historical time periods

Next, I analyze equity issuance of large U.S. commercial banks across different historical periods with different levels of government guarantees. Few, if any, perceptions of government guarantees (with the exception of Fed discount window assistance) existed before the creation of the FDIC in 1933. Perceptions of TBTF government guarantees are generally regarded to have increased in the U.S. in the 1980s — for example, after large-scale interventions, such as the bailouts of Continental Illinois in 1984. I show that, consistent with different levels of government guarantees, equity issuance is strongly procyclical before the creation of the FDIC (pre-1935), moderately procyclical until around the start of the 1980s bailouts (1935-1979), and countercyclical subsequently (1980-2012).

Figure 3 plots net equity issuance minus dividends as a function of Δ(bank credit / GDP) over the sample periods 1926-1934 (Panel A) and 1935-1979 (Panel B). Figure 4 plots a rolling regression of Equation 8 with a past-10-year moving window to illustrate how the cyclicality of equity issuance changes over time: a dramatic shift occurs around 1935, as issuance changes from highly to moderately procyclical, followed by another in the 1980s, from procyclical to countercyclical issuance.

To formally test these shifts, I estimate Equation 8 over the three time subsamples using net equity issuance minus dividends (relative to book equity) as the dependent variable. Results are reported in Table 5. Equity issuance is highly procyclical over the period 1926-1934: a one standard deviation in credit expansion is associated with a 3.6 percentage point increase in net equity issuance minus dividends ($\delta = 0.036$, t-stat

= 0.889, though the sample is necessarily limited to N = 9 years). Equity issuance is moderately procyclical over the period 1935-1979 (δ = 0.006, t-stat = 2.544) and countercyclical over the period 1980-2012 (δ = -0.024, t-stat = -3.383). According to Table 5, results are robust to adding the standard controls and to estimating the regression only during positive credit expansions (i.e. conditional on Δ(bank credit / GDP) ¿ 0).

6.2.3 TBTF banks subsequent to bailouts in the 1980s

I show that in the shift in the 1980s from procyclical to countercyclical equity issuance is mainly driven by banks that equity markets perceive as having large implicit guarantees. I construct a measure of increased perceptions of implicit guarantees (see Section 4.2 for details) based on the wealth effects that are conferred to the shareholders of the largest banks following the announcements of large bailouts (e.g., Continental Illinois in 1984). Using these wealth effects as proxies for changed market perceptions of implicit guarantees, I find that for U.S. banks, greater market-perceived implicit guarantees predict increased countercyclical issuance, compared to banks that do not have greater market-perceived implicit guarantees.

A long literature (e.g., O'Hara and Shaw, 1990; Kho, Lee, and Stulz, 2000; Gandhi and Lustig, 2013) argues that the subsequent wealth effects to shareholders of large banks subsequent to announcements of bailout and regulatory changes are due to changing perceptions by shareholders of implicit guarantees provided to TBTF banks.[23] Take, for example, the failure and bailout of Continental Illinois in 1984. The Comptroller of the Currency announced on September 20, 1984 during Senate testimony that the U.S. would not let the 11 largest banks fail. Subsequent to that announcement, the shareholders of the 11 largest banks, and only those banks, immediately received a

[23]Similarly, Kelly, Lustig, and Nieuwerburgh (2011) show that during the recent crisis the out-of-the-money index put options of bank stocks were relatively cheap, due to government absorbing sector-wide tail risk.

wealth boost in the form of cumulative abnormal stock returns (O'Hara and Shaw, 1990). Taking these wealth effects as proxies for market perceptions of increased implicit guarantees, for each bank I average the abnormal announcement returns over two 1980s bailouts to get a implicit guarantees measure for each era.

I compare large banks that gained greater implicit guarantees to large banks that did not. I argue that magnitude of the wealth subsidy to large bank shareholders reflects the interaction of two effects: the perceived increase in probability of a future government intervention in the banking sector, along with the perceived systemic importance of each bank in the cross-section. This change over time, interacted with cross-sectional variation across banks in response to the announcement, reflects the differential effect that some announcements of implicit guarantees have on some large banks and not others.

To analyze whether the implicit guarantees measure predicts the cyclicality of equity issuance going forward, I plot in Figure 5 the change in cyclicality of equity issuance as a function of the TBTF measure. The cyclicality of equity issuance is estimated for each firm individually from the following time-series regression:

$$\begin{aligned} Net\,Equity\,Issuance\,minus\,Dividends_t &= \alpha + (\delta_0 + \delta^{S\&L} \cdot 1_{year>1984}) \\ &\quad \cdot CreditExpansion_t + \epsilon_t \end{aligned} \tag{12}$$

Consistent with the hypothesis, it is evident from Figure 5 that there is an inverse relationship: firms that are perceived by equity markets to be conferred greater implicit guarantees are subsequently found to pay out more equity during credit expansions.

To formally test whether bailouts and the resulting increased implicit guarantees predict increased countercyclicality, the 50 largest commercial banks are divided into two categories: the treatment group is firms with increased implicit guarantees (measure ¿ 0.01) and the control group is firms without increased implicit guarantees (measure

≤ 0.01). To analyze the change in the cyclicality of bank equity issuance post-1984, I estimate Equation 11 using Thompson's (2011) dually clustered standard errors. The estimation approach is conceptually similar to estimating $\delta^{S\&L}$ in Equation 12 firm-by-firm with a time-series regression and then testing the difference in the average $\delta^{S\&L}$ between the two groups.

Estimates corresponding to the difference-in-differences regression in Equation 11 are reported in Table 6. The treatment group is banks with market-perceived increases in implicit guarantees, and the $Post_t$ indicator denotes observations after 1984. The difference-in-differences coefficient is a statistically significant -0.017 (Column 2, which includes firm fixed effects; t = -2.151). The coefficient is relatively unchanged in magnitude and significance in the presence of various sets of controls (Columns 3 and 4). Thus, banks that are perceived by the market to have gained implicit guarantees subsequent to bailouts have even stronger countercyclical equity issuance, relative to banks that did not gain implicit guarantees.

6.2.4 Across countries

I look at differences across countries in the level of perceived implicit guarantees. Equity issuance for large banks might be less countercyclical, perhaps even procyclical, in countries with large relative banking sectors (e.g., Switzerland, Hong Kong, the Netherlands), since investors may perceive that some banks are "too-big-to-save." If governments have limited fiscal capacity to backstop their largest banks, implicit guarantees may be smaller.

To analyze whether size of a country is associated with the cyclicality of bank equity issuance, I estimate Equation 8 individually for 26 developed countries over the sample period 1980-2012. Results are plotted in Figure 6, where the coefficient of cyclicality, estimated from Equation 8, is plotted for each country as a function of bank assets

to GDP in the year 2000.[24] It appears in this plot that countries with bank assets greater than 200% of GDP (i.e. "small" countries) generally have procyclical bank equity issuance, while countries with bank assets less than 200% of GDP (i.e. "large" countries) have countercyclical bank equity issuance. Thus, I use the 200% cut-off to group countries into "large" (e.g., Canada, Japan, US) versus "small" (e.g., Switzerland, UK).[25] (I put these terms in quotation marks since I'm referring to the size of the countries relative to its banking sector.)

To formally test differences in equity issuance between "small" versus "large" countries, I estimate the panel regression in Equation 9 and report results in Table 7. For "large" countries, a one standard deviation increase in credit expansion is associated with a -0.8 (Column 1; t = -3.920) percentage point change in net equity issuance, corresponding to countercyclical equity issuance. In contrast, for "small" countries, a one standard deviation increase is associated with a 0.6 (Column 1; t = 3.925) percentage point increase in net equity issuance, corresponding to procyclical equity issuance for countries with large relative banking sectors. The difference coefficient is relatively unchanged in magnitude and significance in the presence of various sets of controls (Columns 2 through 4). Thus, consistent with the hypothesis, equity issuance for large banks is procyclical in "small" countries where implicit guarantees to the largest banks may be perceived as weaker.

[24] Results are robust to classifying banks according to bank assets to GDP in other years: 1980, 1990, 2005, and an average across 1980-2012.

[25] A threshold approach based on 200% bank assets to GDP is preferable to a regression approach based on continuous variation in bank assets to GDP. The reason is that there might simply be a mechanical relation between bank assets to GDP and credit expansion = Δ(bank credit / GDP), in the sense that a larger banking sector leads to more variation in Δ(bank credit / GDP), which reduces the magnitude of the cyclicality coefficient. The binary approach based on the 200% threshold is better suited for my purpose, because what I really want to test is the sign of the cyclicality coefficient: whether countries with large relative banking sectors have procyclical issuance, while countries with small relative banking sectors have countercyclical issuance.

6.3 Change in bank equity issuance upon adoption of the Euro

I look at the change in countercyclical bank equity issuance in different Eurozone economies upon the introduction of the Euro. Using interest rate convergence as a proxy for change in implicit guarantees, I show that countries with high interest rate convergence have increased countercyclical issuance when joining the Euro, compared to countries with low interest rate convergence.

For countries joining the Euro, there may be a perception among creditors that banks in Eurozone periphery countries (e.g., Italy, Portugal, Spain) gained implicit protections in joining the Euro, perhaps from the European Central Bank or implicit fiscal guarantees from other Euro countries. Even if the implicit guarantees are conferred mostly upon these countries' sovereign debt rather than banks, implicit guarantees for sovereign debt can indirectly support banks, because they provide the fiscal backing for individual governments to support their banks in the event of a crisis. The fact that the EU lent money to Ireland, Spain, and Portugal in 2011-2012 to jointly help these governments recapitalize their banks and to forestall sovereign debt crises illustrates how implicit guarantees for sovereign debt can indirectly confer protections to a country's banking sector.

I divide the initial group of countries to adopt the Euro into two groups based on whether interest rate convergence is low or high. For my interest rate convergence measure, I use the change between 1996 and 1999 in the real interest rate spread relative to Germany. I classify countries as having low interest rate convergence if the change spread is less than -0.01 (e.g., Germany, the Netherlands, and Finland) or high if the change in spread is greater than -0.01 (e.g., Ireland, Italy, and Portugal).[26] Figure 7 plots the change in cyclicality of issuance, which is estimated for each country individually from the following time-series regression:

[26]I specifically exclude Spain because Spain adopted in 2000 a stringent regulatory regime often referred to as "dynamic provisioning," which required provisioning in good times over and above standard loan loss provisions. See JimÃ©nez et al. (2012) for details.

$$Equity\,Issuance_t = \alpha + (\delta_0 + \delta_1 \cdot 1_{year>1998}) \cdot CreditExpansion_t + \epsilon_t \qquad (13)$$

Using the above classification, I estimate the difference-in-differences regression corresponding to Equation 11 using Thompson's (2011) dually-clustered standard errors to test whether countries with high interest rate convergence have increased countercyclical issuance when joining the Euro, compared to countries with low interest rate convergence. Estimates are reported in Table 8. The difference-in-differences coefficient is a statistically significant -0.026 (Column 2, which includes country fixed effects; t = -2.561). The coefficient is relatively unchanged in magnitude and significance in the presence of various sets of controls (Columns 2 through 4). Thus, countries with high interest rate convergence have increased countercyclical issuance when joining the Euro, compared to countries with low interest rate convergence, consistent with my hypothesis.

6.4 Cost of bank equity capital and market timing

In this section, I find that the *cost* of equity capital is countercyclical over the credit cycle for both non-financials and banks: lower cost-of-capital (i.e. high stock prices) during expansions and higher cost-of-capital (i.e. low stock prices) during contractions. Therefore, if bank managers take advantage of low costs of equity capital, then equity issuance should increase during credit expansions alongside the lower cost of capital, and issuance should be procyclical. While procyclical issuance for non-financials is consistent with market timing, countercyclical issuance by banks indicates that banks do not appear to be simply timing the market.

I estimate the time-varying cost of capital over the credit cycle with the following time-series regression with Newey-West standard errors:

$$(r_{t+K} - r^f_{t+K}) = \alpha + \beta \cdot CreditExpansion_t + \gamma' \cdot Controls_t + \epsilon_t \tag{14}$$

which predicts K = 4- or 8-quarter ahead excess return of an equity index ($r_{t+K} - r^f_{t+K}$) using a set of predictor variables including *credit expansion*. Although I employ Newey-West standard errors, I take a deliberately conservative approach with standard errors by using non-overlapping returns. That is, in calculating 4- or 8-quarter ahead returns, I drop the intervening observations from our data set. Non-overlapping returns thus make the estimation results robust to many potential econometric issues involved in estimating standard errors of overlapping returns.

Table 9 estimates the panel regression model specified in Equation 14 for 4-, and 8-quarter-ahead excess returns, without and with controls (dividend yield, book to market, T-Bill yield, investment to capital, and the corporate spread). Panel A reports coefficients for large U.S. commercial banks over the sample period 1925-2012, while Panel B reports coefficients for various time subsets and types of firms. Panel A reports that a one standard deviation increase in credit expansion predicts 6.4 (t = -2.241) and 13.3 (t = -2.874) percentage point *decreases* in subsequent returns for large U.S. commercial banks over 4-, and 8-quarter horizons, respectively. When controls are included, the coefficients are slightly higher in magnitude and have similar statistical significance. Panel B reports similar cost-of-capital coefficients for two time subsets (1925-1980 and 1980-2012), in addition to coefficients for large U.S. investment banks, small U.S. commercial banks, and non-financial U.S. firms (all over the period 1980-2012). To save space, coefficients on control variables are not reported. In particular, for non-financial firms, a one standard deviation increase in bank credit expansion predicts 3.1 and 12.4 percentage point *decreases* for 4-, and 8-quarter-ahead excess returns, respectively.[27]

[27]In estimating Equation 14, the cost of equity *relative* to debt is equity returns minus the three-month T-Bill yield. Since the goal is to understand why banks may choose equity over debt, it is important to use the correct funding cost for banks. To address concerns that the T-Bill rate might

This result on time-varying cost of capital ties into the ongoing debate on the cost of bank capital (see Admati, DeMarzo, Hellwig, and Pfleiderer, 2013b). While previous work (e.g., Baker and Wurgler, 2013; Kisin and Manela, 2014) estimates an average cost of bank capital over time, this result suggests that, due to the time-varying cost of capital, the cost of increased regulatory capital may be substantially lower than previously estimated, if regulators are careful to implement new capital requirements during expansions. This result may further lend support for the use of countercyclical regulatory capital buffers.

To the extent that bank managers do not seem to be timing the market, market timing theory (Baker and Wurgler, 2002) is only relevant when there is differential optimism between bank managers and equity markets. Previous work has indicated that bank equity investors are over-optimistic during credit booms (Baron and Xiong, 2014), as were professionals in the securitization industry during the recent housing boom (Cheng, Raina, and Xiong, 2014). To the extent that I do not detect market timing for large banks and broker-dealers in the data, the evidence is consistent with similar levels of optimism between bank managers and equity markets but no differential optimism driving equity issuance.

7 Conclusion

Bank equity issuance and retained income are countercyclical across credit cycles over the sample period 1980-2012. In contrast, equity issuance for non-financial firms is procyclical. Thus, during credit expansions, even though more capital might help banks better absorb shocks, banks raise equity the least.

not be the appropriate funding cost, I also check that the results are robust to other measures of bank funding costs, such as deposit rates (from Fed data), interbank lending benchmark rates, repo rates, and implied firm-specific deposit or borrowing rates from Compustat or Bankscope. See the Online Appendix. The reason why the exact funding cost does not seem to matter is that the variance in alternative funding costs (relative to the T-Bill yield) is small compared to the variance in equity returns.

I develop an illustrative model of moral hazard to understand why bailout distortions and government guarantees may lead banks to choose countercyclical equity issuance. In accordance with the model, evidence across types of firms, countries, and historical time periods strongly suggest that bank equity issuance is most countercyclical when and where government guarantees, especially implicit TBTF guarantees, are strongest. In particular, a shift from procyclical to countercyclical bank equity issuance shift in the 1980s is driven by banks that equity markets perceive as TBTF. Analyzing countries adopting the Euro and using interest-rate convergence as a proxy for increased implicit guarantees, I find that bank equity issuance becomes more strongly countercyclical in Eurozone countries that gain increased implicit guarantees upon adopting the Euro, compared to other Eurozone countries. Lastly, I find that the cost of equity capital (for both banks and non-financials) is countercyclical over the credit cycle. While procyclical issuance for non-financials is consistent with market timing, countercyclical issuance by banks indicates that banks do not appear to be simply timing the market. My results highlight the role of bailout distortions and government guarantees, which allow bankers to raise less and pay out more equity during credit expansions.

Appendix

A. Data construction

In this section, I provide addition information related to data construction.

New equity issuance and share repurchases. Following Jagannathan, Stephens, & Weisbach (2000), I separate out new equity issuance and share repurchases as *max(net equity issuance, 0)* and *min(net equity issuance, 0)*, respectively, where net equity issuance here is from monthly CRSP data, which Jagannathan, Stephens, & Weisbach (2000) show is generally accurate compared to other methods of constructing issuances

and repurchases.[28]

Credit Expansion. The bank credit data comes from the BIS Long Credit Series (data are generally from 1950 onwards for a variety of countries), merged with historical bank credit data from Schularick-Taylor (2012), which extends the dataset back to 1870 for 14 developed economies. I standardized credit expansion within each country.

Control variables. The *term spread* is the difference between *long-term interest rates* (10-year government bond yields, taken mostly from Global Financial Data and OECD) and *short-term interest rates* (3-month government t-bill rates taken from Global Financial Data, the IMF, OECD). *Investment to capital* is private non-residential fixed investment divided by the outstanding private non-residential fixed capital stock, which comes from the Kiel Institute's database on investment and capital stock. The *corporate spread* (just used for the U.S.) comes from *FRED*. For the predicted cost-of-capital regressions (U.S. data only), *dividend yield* (d/p) and *book-to-market* for the S&P 500 index come from Global Financial Data.

Announcement Returns. Announcement returns, defined as the one-day abnormal returns (relative to the S&P 500 index), are reported for each bank and each announcement in Table A1. The two announcement events are:

1. Sept. 20, 1984: The OCC announcement related to Continental Illinois. See Section 5.2 and O'Hara and Shaw (1990) for details.

2. Feb. 7, 1989: President George H. W. Bush unveils his proposal for the $166 billion Savings and Loan bailout, guaranteeing that *all* depositors of *all* banks (not just of S&Ls) would be protected, while also promising that the FDIC insurance premium would not rise (except for S&Ls). While S&L stocks rose, stocks of large money-center banks were the biggest beneficiaries, due to the blanket guarantees

[28]They show, in general, these measures tends to slightly understate issuance and repurchases, since they are netted out on a monthly basis.

for all deposits. See Mansur and Elyasuani (1994) for details.

B. Proof of Proposition 1

Substituting the creditors' breakeven constraint (Equation 7) into the shareholders' expected value (Equation 6) results in:

$$\begin{aligned} \text{Expected Shareholder Value} \;=\; & \frac{\frac{1}{2}(\overline{A}-L+(1+a)\cdot n)^2}{\overline{A}-\underline{A}} \\ & +V\cdot\frac{(\overline{A}-L+(1+a)\cdot n)}{\overline{A}-\underline{A}} \\ & -n\;-(1-\gamma)\,\frac{\frac{1}{2}(\underline{A}-L+(1+a)\cdot n)^2}{\overline{A}-\underline{A}} \end{aligned} \tag{15}$$

Equation 15 is an upward-facing quadratic equation with respect to n, and therefore no interior solution exists. Thus, a bank manager who is maximizing shareholder value will choose either n = +N (issuing the maximum amount) or n = -N (paying out the maximum amount). Mathematically, the optimum n = +N or -N depends on whether the vertex of Equation 15 is greater or less than n = 0. To find the vertex n^*, I differentiate Equation 15 with respect to n and set it equal to 0:

$$\begin{aligned} 0 \;&=\; \frac{d\,E[Shareholder\ Value]}{dn} \\ &=\; \frac{(1+a)}{(\overline{A}-\underline{A})}\left[V+\left(\overline{A}-L+(1+a)\cdot n\right)-(1-\gamma)\left(\underline{A}-L+(1+a)\cdot n\right)\right]-1 \end{aligned} \tag{16}$$

Solving for n yields:

$$-(1+a)n=\frac{1}{\gamma}\left[V+(\overline{A}-\underline{A})\cdot\left(\frac{a}{1+a}\right)\right]+(\underline{A}-L). \tag{17}$$

Substituting in $\sigma=\frac{\overline{A}-\underline{A}}{A_0}$ and $\sigma_D=\frac{A_0-\underline{A}}{A_0}$, dividing both sides by A_0, and labeling

both sides Ψ yields:

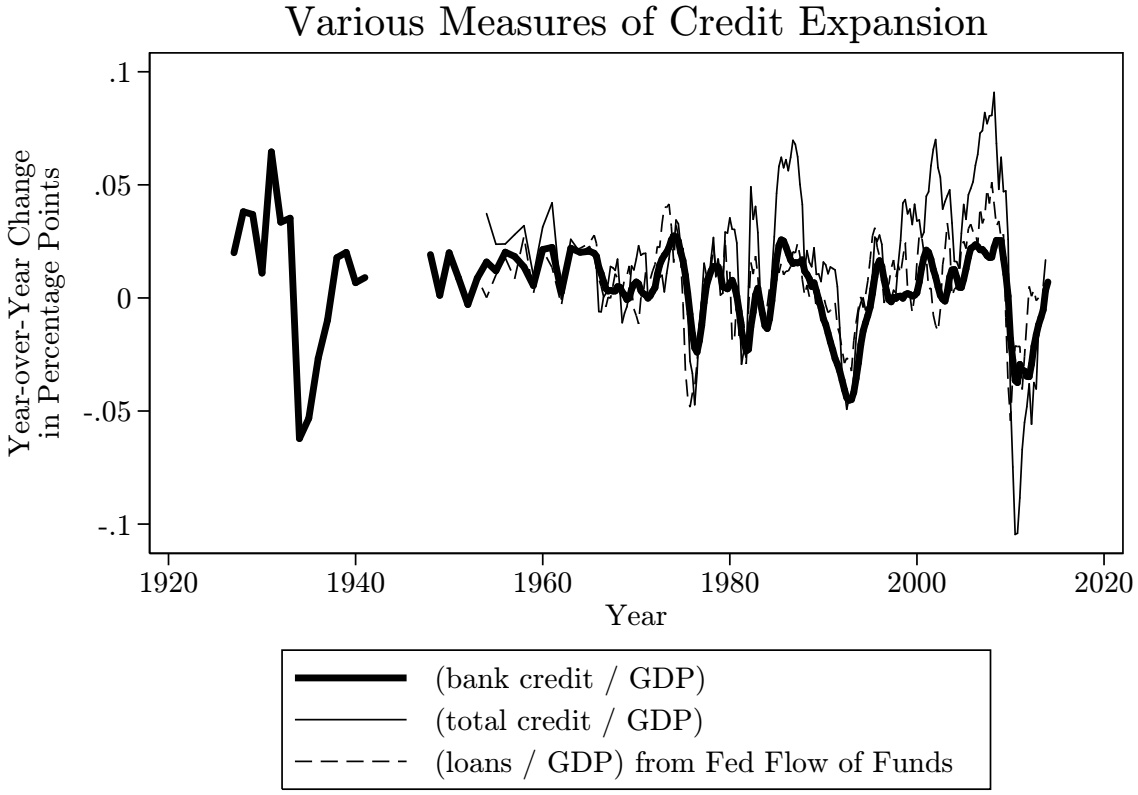
$$\Psi = -(1+a)n = \frac{1}{\gamma}\left[\frac{V}{A_0} + \sigma \cdot \frac{a}{1+a}\right] + (1-\sigma_D) - \frac{L}{A_0}. \tag{18}$$

From Equation 18, we can see that if $\Psi > 0$, then the vertex is $n^* < 0$, and therefore the bank manager will choose n = +N (issuing the maximum amount). If $\Psi < 0$, then the vertex is $n^* > 0$, and therefore the bank manager will choose n = -N (paying out the maximum amount). If $\Psi = 0$, then the vertex is $n^* = 0$, and therefore the bank manager will be indifferent between the two corner solutions.

Figure 1: Credit expansion and bank equity

Panel A plots credit expansion (i.e. $\Delta(bank\,credit/GDP)$) over time, along with alternative measures of the credit cycle, such as change in total credit to GDP and change in bank loans to GDP. Panel B plots net equity issuance, dividends, and retained earnings (normalized by book equity) over time for the 20 largest U.S. commercial banks, aggregated.

Panel A



Panel B

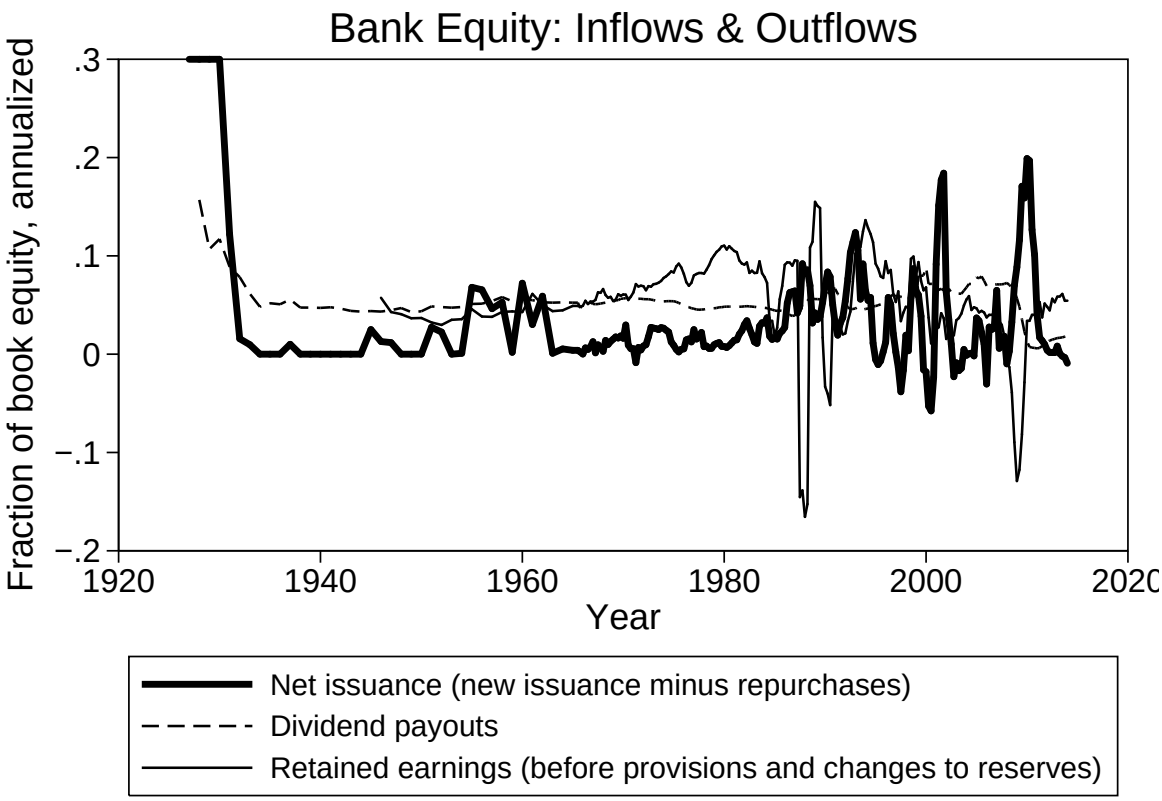

Bank Equity: Inflows & Outflows
Fraction of book equity, annualized
.3
.2
.1
0
−.1
−.2
1920
1940
1960
1980
2000
2020
Year
Net issuance (new issuance minus repurchases)
Dividend payouts
Retained earnings (before provisions and changes to reserves)

Figure 2: Equity issuance for different types of U.S. firms

Figure 2 plots net equity issuance (normalized by book value) as a function of $\Delta(bank\,credit/GDP)$. Net equity issuance is aggregated over all U.S. firms in each category. Panel A is for large commercial banks, Panel B is for large investment banks, Panel C is for small commercial banks, and Panel D is for non-financial firms. Observations are quarterly, 1980-2012.

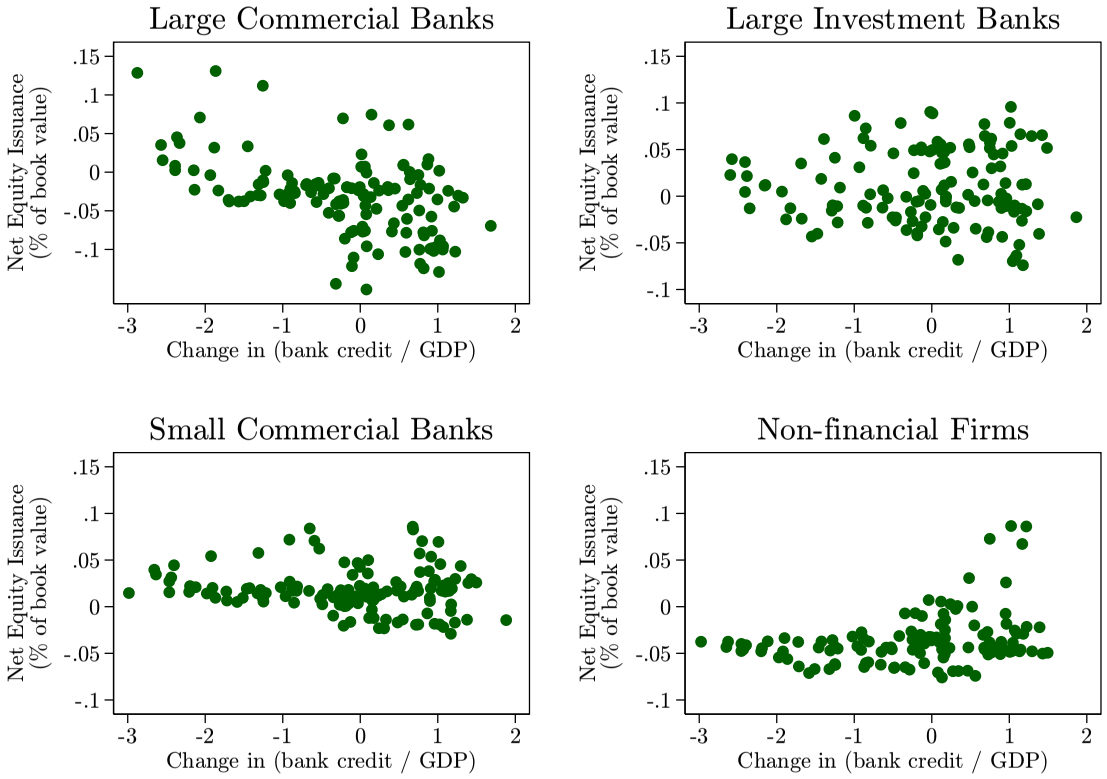
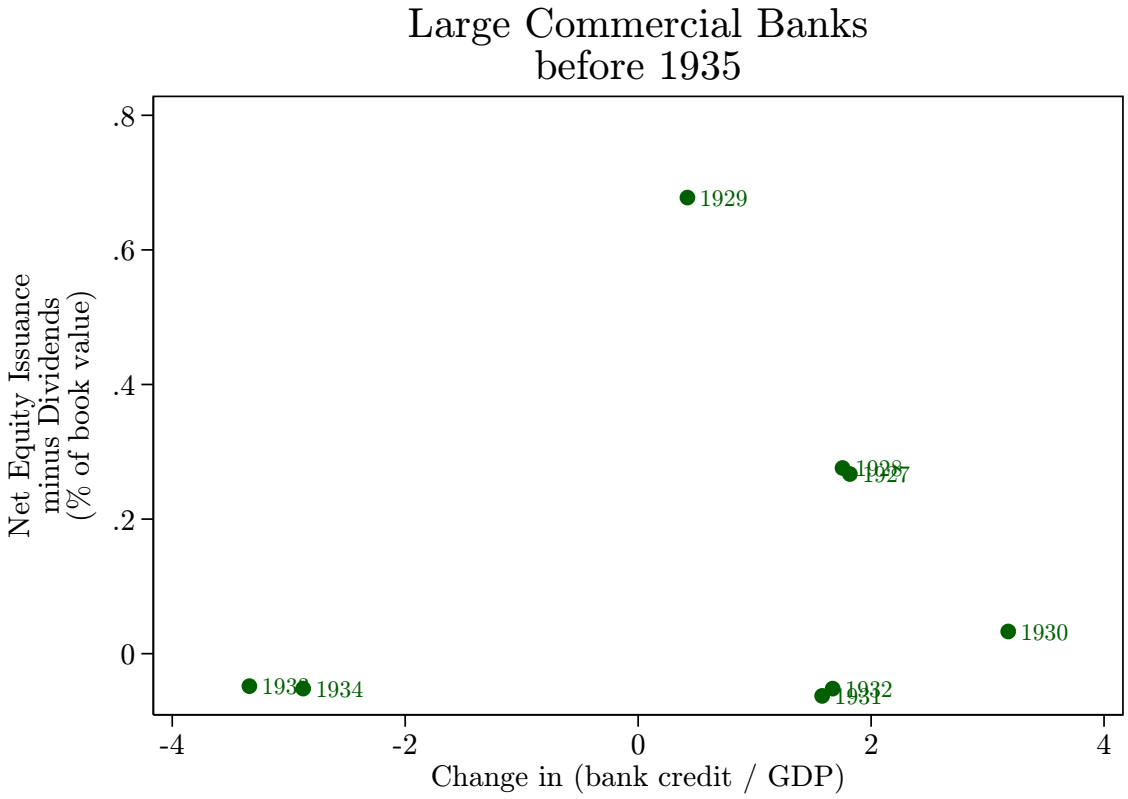


Figure 3: U.S. bank equity issuance over two historical time periods

Figure 3 plots net equity issuance minus dividends (normalized by book value, aggregated over all large U.S. commercial banks) as a function of $\Delta(bank\,credit/GDP)$. Panel A is for the period 1926-1935, while Panel B is for the period 1935-1980.

Panel A



Panel B

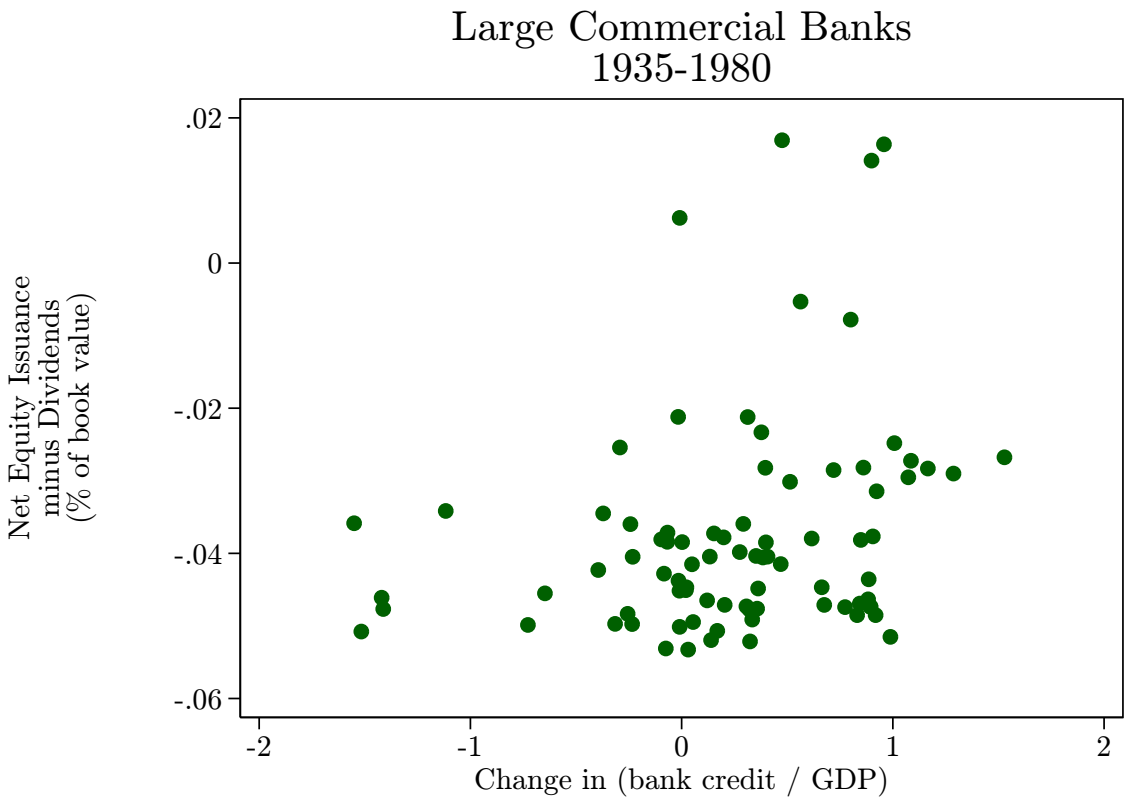

Large Commercial Banks
1935-1980
Net Equity Issuance
minus Dividends
(% of book value)
.02
0
-.02
-.04
-.06
-2
-1
0
1
2
Change in (bank credit / GDP)

Figure 4: Rolling regression of the cyclicality of bank equity issuance

Figure 5 plots a rolling regression of Equation 8 with a past 10-year moving window to illustrate how the cyclicality of bank net equity issuance minus dividends changes over time. The gray shading is the 95% confidence interval.

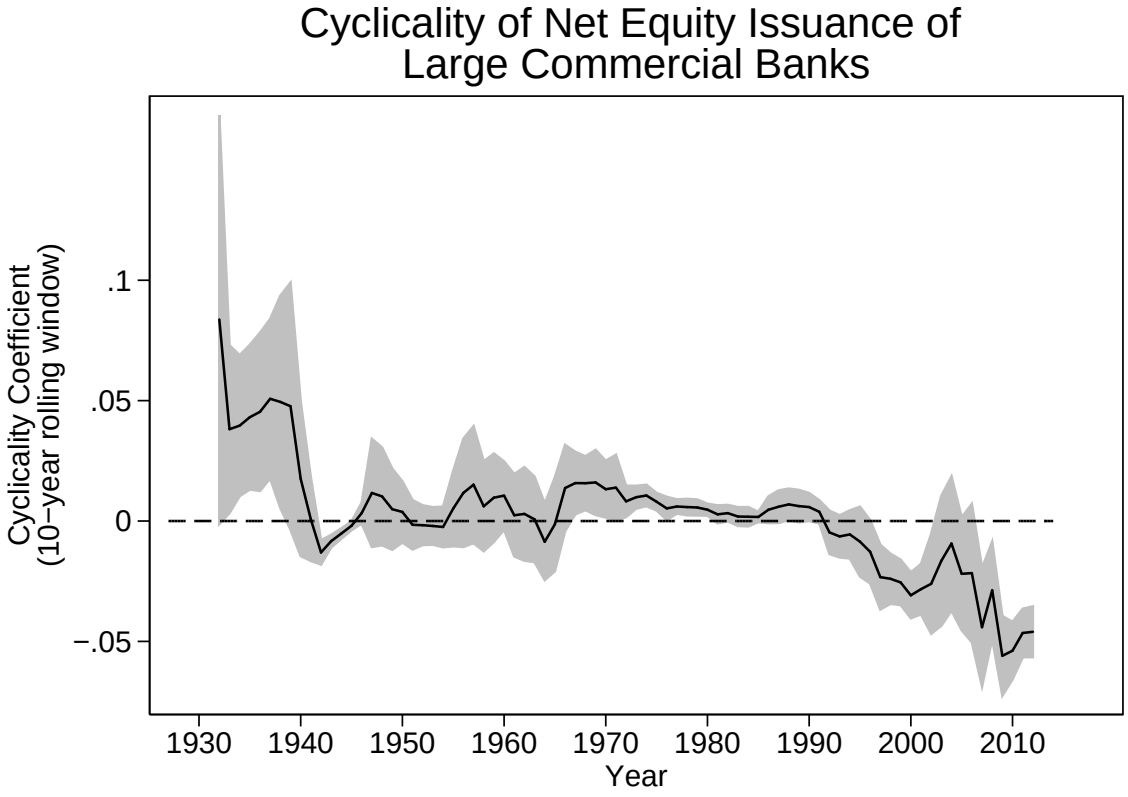


Figure 5: Change in bank equity issuance after 1980s bailouts

Figure 5 plots the change in cyclicality of net equity issuance minus dividends as a function of the implicit guarantees measure constructed from equity announcement returns. The cyclicality of net equity issuance minus dividends is estimated for each firm individually from Equation 8.

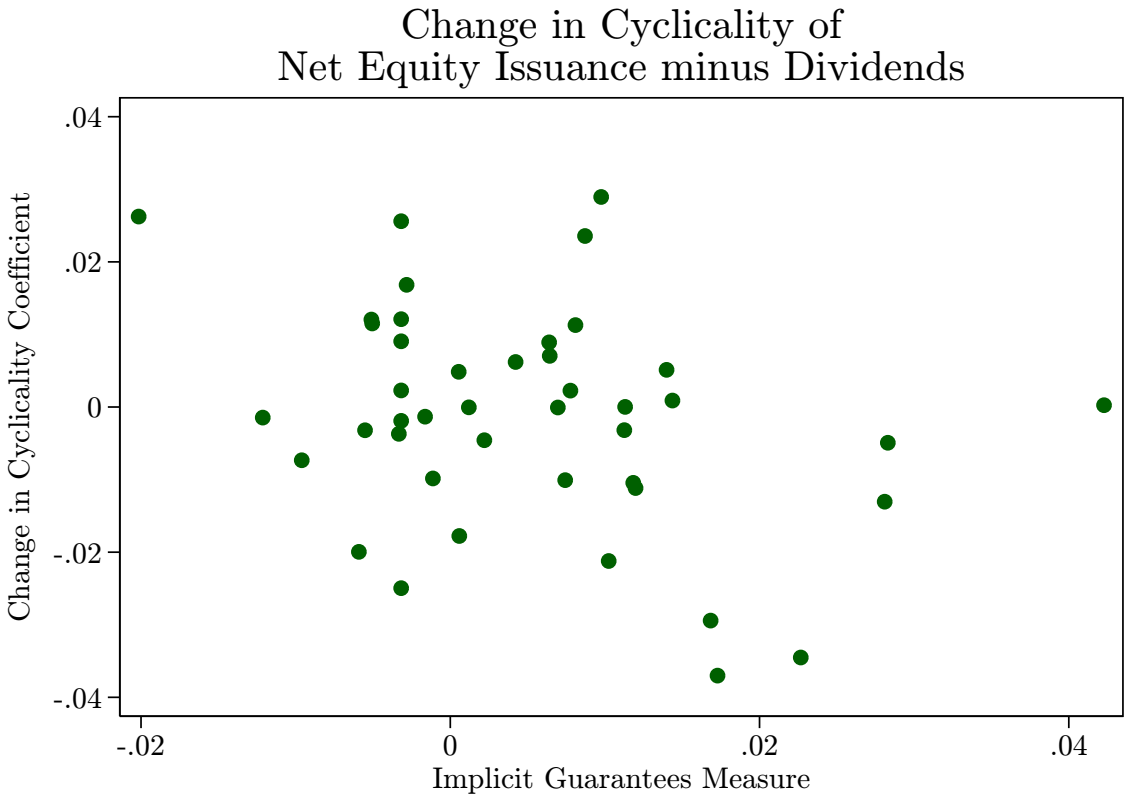


Figure 6: Bank equity issuance across countries

Figure 6 plots, for each country, the coefficient of cyclicality for bank equity issuance (estimated from Equation 8 over the period 1980 to 2012) as a function of bank assets to GDP.

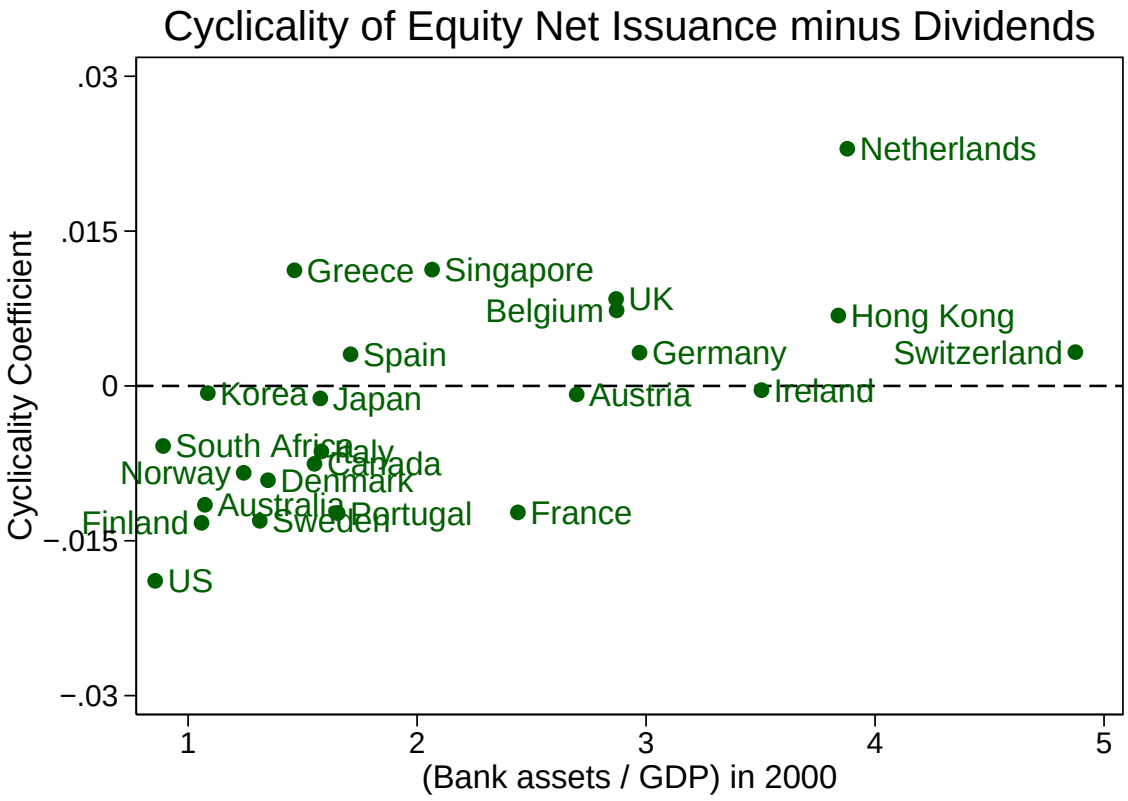


Figure 7: Change in bank equity issuance upon adoption of the Euro

Figure 7 plots the change in cyclicality of bank equity issuance (estimated for each country individually from Equation 8) as a function of the real interest rate convergence (change in real interest rate spread from 1995 to 1999, relative to Germany). Luxembourg, which has interest rate convegence of -0.015, is off the graph with a positive coefficient change of 0.1. Spain is omitted, as discussed in Section 5.3, due to its bank regulatory regime of "dynamic provisioning."

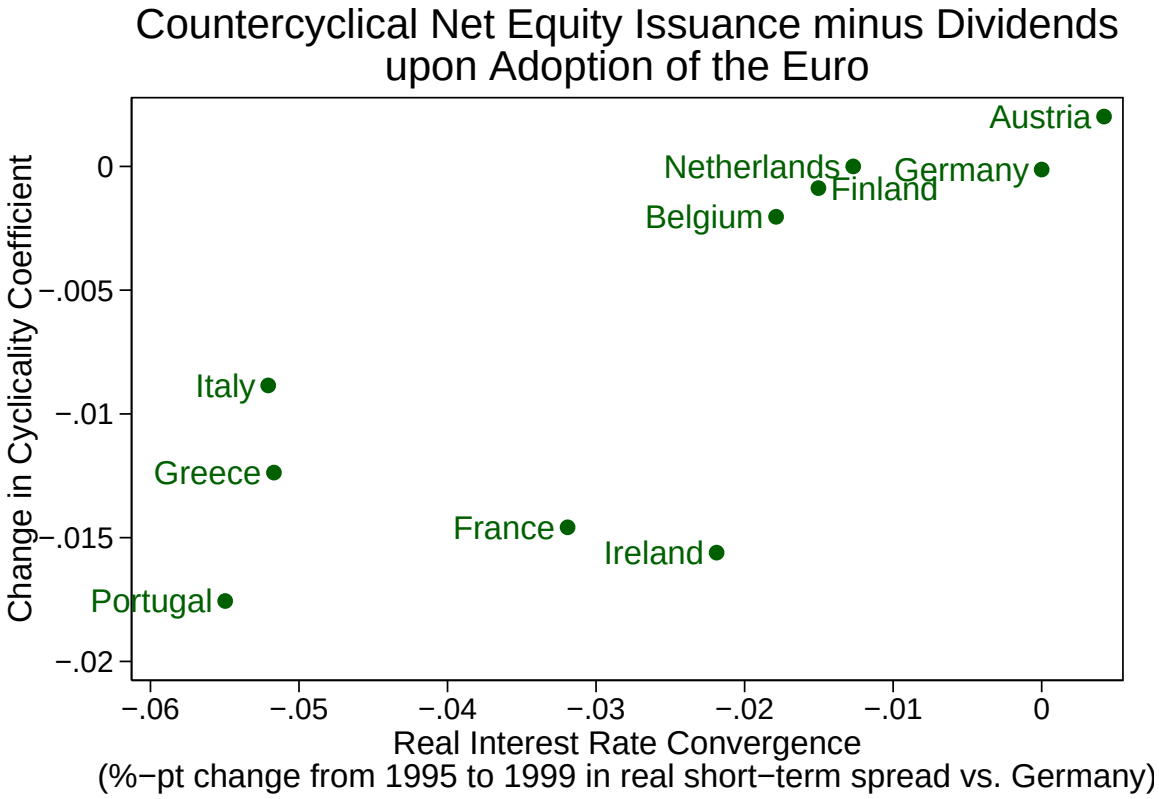


Table 1: Summary Statistics

Table 1 presents summary statistics for equity issuance, payouts, and other key variables for large U.S. banks. Panel A presents cross-sectional statistics for the 20 largest banks during 2005, around the peak of the most recent credit cycle. Panel B presents aggregated time series statistics over the period 1980-2012.

Panel A: Cross-sectional statistics, 20 largest banks in 2005

	Mean	Stdev.	5%	10%	25%	50%	75%	90%	95%
Bank balance sheet statistics									
Assets ($ millions)	336,449	437,320	54,841	60,386	91,015	109,170	434,981	1,212,239	1,472,793
Book equity / Assets	9.6%	3.0%	6.5%	6.8%	8.1%	9.0%	9.8%	12.6%	16.0%
Growth of assets (annualized)	18.7%	28.4%	0.6%	2.2%	4.9%	9.4%	21.0%	43.2%	72.8%
Growth of book equity (annualized)	3.4%	8.6%	-10.0%	-7.2%	-1.2%	3.2%	8.2%	14.4%	17.2%
Equity issuance and payouts (as percent of book equity, annualized)									
Net equity issuance minus dividends	-9.4%	6.9%	-24.1%	-19.5%	-13.8%	-7.9%	-5.5%	-2.6%	0.9%
Equity Issuance	1.1%	2.6%	0.0%	0.0%	0.0%	0.0%	1.1%	2.6%	4.6%
Repurchases	3.6%	4.6%	0.0%	0.0%	0.0%	2.0%	6.0%	10.5%	17.6%
Dividends	6.5%	2.3%	1.2%	3.7%	4.8%	7.0%	7.6%	8.8%	11.0%
Retained income	5.1%	5.7%	-5.5%	-0.7%	2.1%	6.0%	7.9%	10.3%	15.4%

Panel B: Time series statistics, quarterly observations (annualized) over 1980-2012

	Mean	Stdev.	5%	10%	25%	50%	75%	90%	95%
Credit expansion									
Δ (bank credit / GDP)	-0.1%	2.0%	-4.3%	-3.3%	-1.4%	0.3%	1.5%	2.2%	2.4%
Equity issuance and payouts of large banks	*(aggregated, as percent of book equity)*								
Net equity issuance minus dividends	-4.4%	4.7%	-12.9%	-10.9%	-7.2%	-3.8%	-1.8%	1.4%	3.2%
Equity issuance	2.9%	2.6%	0.2%	0.2%	0.9%	2.0%	3.9%	7.3%	8.6%
Repurchases	2.1%	3.1%	0.0%	0.0%	0.0%	0.9%	3.2%	5.2%	7.7%
Dividends	5.1%	2.1%	0.9%	1.7%	4.4%	5.2%	6.3%	7.3%	7.9%
Retained income	4.8%	10.5%	-4.4%	-2.1%	2.8%	6.0%	9.7%	12.3%	14.5%

Table 2: Net Equity Issuance minus Dividends over the Credit Cycle

Table 2 demonstrates that equity issuance for large U.S. commercial banks is counter-cyclical over credit cycles over the period 1980-2012. The table reports estimates from Equation 6 with the dependent variable net equity issuance minus dividends. The regression is estimated for the full sample in Columns 1 and 2 and conditional on positive credit expansion in Columns 3 and 4. Newey-West t-statistics are in parentheses. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Large U.S. Commercial Banks, 1980-2012

	Net Equity Issuance minus Dividends			
	Full sample		Conditional on Δ(bank credit/GDP) > 0	
	(1)	(2)	(3)	(4)
Δ (bank credit / GDP)$_t$	-0.024***	-0.016**	-0.028	-0.034**
	[-3.383]	[-2.907]	[-1.897]	[-2.747]
(Income / book equity)$_t$		-0.118*		-0.169*
		[-1.997]		[-2.100]
Log(book / market)$_t$		0.014		0.036
		[1.119]		[1.901]
(Bank stock prices increased)$_t$		-0.004		-0.006
		[-0.297]		[-0.345]
(Bank stock prices decrease)$_t$		0.083***		0.112***
		[6.465]		[4.825]
(Term spread)$_t$		0.531		0.040
		[1.375]		[0.065]
Log(i / k)$_t$		0.005		0.212
		[0.010]		[0.274]
Constant	-0.036***	-0.044	-0.029**	-0.057
	[-4.510]	[-0.714]	[-2.851]	[-0.615]
Adj. R^2	0.24	0.46	0.04	0.25
N	132	132	66	66

Table 3: Issuance, payouts, and retained earnings over the credit cycle

Table 3 repeats the analysis from Table 2 but for other dependent variables, such as new issuance, repurchases, dividends, and retained income. The table only reports estimates for coefficients on $\Delta(bank\,credit/GDP)$; coefficient estimates for control variables are omitted to save space. The regression is estimated for the full sample in Columns 1 and 2 and conditional on positive credit expansion in Columns 3 and 4. Newey-West t-statistics are in parentheses. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Large U.S. Commercial Banks, 1980 - 2012

	Full sample		Conditional on Δ(bank credit/GDP) > 0	
	No controls	Controls	No controls	Controls
	(1)	(2)	(3)	(4)
Net Equity Issuance minus Dividends	-0.024***	-0.016**	-0.028	-0.034**
	[-3.383]	[-2.907]	[-1.897]	[-2.747]
Equity Issuance	-0.010**	-0.008**	-0.016	-0.019*
	[-2.703]	[-3.036]	[-1.861]	[-2.404]
(negative) Repurchases	-0.008**	-0.004*	-0.001	-0.003
	[-2.963]	[-2.165]	[-0.083]	[-0.484]
(negative) Dividends	-0.009	-0.004	-0.003	-0.003
	[-1.939]	[-1.541]	[-0.488]	[-0.936]
Retained Earnings	-0.009	-0.010**	-0.040**	-0.010
(before changes in reserves)	[-1.619]	[-3.304]	[-3.276]	[-0.982]

Table 4: Net Equity Issuance for Various Types of Firms

Table 4 reports estimates for Equation 3 for various types of firms in the U.S. over the sample period 1980-2012. The dependent variable is net equity issuance minus dividends. Panel A compares various types of financial firms to non-financial firms, and Panel B compares different types of non-financial firms. Newey-West t-statistics are in parentheses. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Panel A: Various types of firms

	Large Commercial Banks	Fannie Mae Freddie Mac	Large Investment Banks	Small Commercial Banks	Non-Financial Firms (Manufacturers)
	(1)	(2)	(3)	(4)	(5)
Without control variables					
Δ (bank credit / GDP)$_t$	-0.024***	-0.018	-0.013	-0.003	0.008*
T-stat	[-3.383]	[-1.927]	[-1.082]	[-0.731]	[1.972]
Adj. R^2	0.24	0.06	0.05	0.01	0.09
N	132	132	132	132	132
With control variables (coefficients on controls not shown)					
Δ (bank credit / GDP)$_t$	-0.016**	-0.007	-0.012	-0.005	0.007*
T-stat	[-2.907]	[-1.071]	[-1.222]	[-1.629]	[2.046]
Adj. R^2	0.46	0.40	0.29	0.13	0.29
N	132	132	132	132	132

Panel B: Various types of non-financial firms

	Airlines	Autos (Big 3)	Defense	Business Services	Computers & Software	Retail	Wholesale
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Without control variables							
Δ (bank credit / GDP)$_t$	-0.064	-0.017	-0.014*	0.019	0.021	0.002	0.002
(T-stat)	[-1.700]	[-1.567]	[-1.994]	[0.809]	[0.815]	[0.153]	[0.293]
Adj. R^2	0.06	0.05	0.12	0.02	0.02	-0.01	-0.00
N	132	132	132	132	132	132	132
With control variables (coefficients on controls not shown)							
Δ (bank credit / GDP)$_t$	-0.057	-0.025**	-0.015**	0.002	0.001	0.005	0.002
(T-stat)	[-1.865]	[-2.912]	[-2.691]	[0.115]	[0.049]	[0.411]	[0.413]
Adj. R^2	0.08	0.15	0.36	0.28	0.27	0.15	0.31
N	132	132	132	132	132	132	132

Table 5: Cyclicality of bank equity issuance over several time periods

Table 5 reports estimates for Equation 6 for large U.S. commercial banks over three historical sample periods: 1926-1934, 1935-1979, and 1980-2012. Observations are quarterly starting in 1965 and annual beforehand. The dependent variable is net equity issuance minus dividends. The regression is estimated for the full sample in Columns 1 and 2 and conditional on positive credit expansion in Columns 3 and 4. Newey-West t-statistics are in parentheses. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Large U.S. Commercial Banks

		Full sample		Conditional on Δ(bank credit / GDP) > 0	
		No controls	Controls	No controls	Controls
		(1)	(2)	(3)	(4)
1926-1934	Δ (bank credit / GDP)$_t$	0.036	-	0.054	-
	T-stat	[0.889]	-	[0.756]	-
	Adj. R^2	0.02	-	0.08	-
	N	9	-	6	-
1935-1979	Δ (bank credit / GDP)$_t$	0.006*	0.006	0.014***	0.013*
	T-stat	[2.544]	[1.862]	[3.791]	[2.206]
	Adj. R^2	0.06	0.01	0.08	0.03
	N	83	83	58	58
1980-2012	Δ (bank credit / GDP)$_t$	-0.024***	-0.016**	-0.028	-0.034**
	T-stat	[-3.383]	[-2.907]	[-1.897]	[-2.747]
	Adj. R^2	0.24	0.46	0.04	0.25
	N	132	132	66	66

Table 6: Change in cyclicality of bank equity issuance after 1980s bailouts

Table 6 reports estimates from Equation 8 on changes in cyclicality of bank equity issuance after 1980s bailouts. The regression is estimated on a panel of the 50 largest commercial banks. The treatment group is banks with increased implicit guarantees (implicit guarantee measure ¿ 0.01), versus the control group without increased implicit guarantees (implicit guarantee measure $\leq$ 0.01). The Postt indicator denotes observations after 1984. Dually-clustered t-statistics are in parentheses. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Δ (bank credit / GDP)$_t$	-0.003*	-0.003*	-0.002	-0.002*	-0.007
	[-2.334]	[-2.655]	[-1.877]	[-2.022]	[-1.746]
Δ (bank credit / GDP)$_t$ · Post$_t$	-0.000	0.000	-0.000	0.001	0.001
	[-0.059]	[0.022]	[-0.098]	[0.201]	[0.594]
Δ (bank credit / GDP)$_t$ · Treat$_i$	0.008	0.009	0.009	0.006	0.008
	[1.637]	[1.780]	[1.725]	[1.484]	[1.495]
Δ (bank credit / GDP)$_t$ · Post$_t$ · Treat$_i$	-0.017*	-0.017*	-0.017*	-0.021**	-0.018*
	[-2.151]	[-2.237]	[-2.268]	[-2.758]	[-2.035]
Post$_t$		-0.007	-0.003	-0.000	-0.007
		[-1.558]	[-0.849]	[-0.070]	[-1.712]
Post$_t$ · Treat$_i$		0.009	0.009	0.003	0.013
		[0.655]	[0.630]	[0.213]	[0.871]
Constant	-0.040***	-0.038***	-0.037***	-0.035***	-0.037***
	[-404.78]	[-25.849]	[-11.749]	[-43.398]	[-25.051]
Firm FE	x	x	x	x	x
Controls$_{i,t}$			x		
Controls$_{i,t}$ · Firm				x	
Controls$_{i,t}$ · Δ (bank credit / GDP)$_t$					x
Adj. R^2	0.07	0.07	0.09	0.13	0.08
N	4984	4984	4438	4438	4438

Table 7: Cyclicality of bank equity issuance across developed economies

Table 7 reports estimates from Equation 7. The table tests the difference in the cyclicality of equity issuance between "small" versus "large" countries (these terms refer to the size of the countries relative to its banking sector). Dually-clustered t-statistics are in parentheses. Observations are quarterly over the sample period 1980-2012. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

	(1)	(2)	(3)	(4)
Δ (bank credit / GDP)$_t$ · LargeCountry$_t$	-0.008***	-0.007***	-0.003	-0.006**
	[-3.920]	[-3.616]	[-1.395]	[-2.595]
Δ (bank credit / GDP)$_t$ · SmallCountry$_t$	0.006***	0.006***	0.007***	0.008***
	[3.925]	[3.945]	[3.896]	[3.466]
Constant	-0.035***	-0.038***	-0.057***	-0.035***
	[-19.899]	[-13.114]	[-13.518]	[-20.630]
Country FE	x	x	x	x
Controls$_{i,t}$		x		
Controls$_{i,t}$ · Firm			x	
Controls$_{i,t}$ · Δ (bank credit / GDP)$_t$				x
R-squared	0.35	0.40	0.54	0.37
N	2299	2231	2231	2231

Table 8: Cyclicality of bank equity issuance upon adoption of the Euro

Table 8 reports estimates from Equation 8 on changes in cyclicality of bank equity issuance upon the adoption of the Euro. The treatment group is high interest rate convergence countries (change in real interest rate spread greater than -0.01), versus the control group of low interest rate convergence countries (change in real interest rate spread less than -0.01). The Postt indicator denotes observations after the adoption of the Euro in 1999. Dually-clustered t-statistics are in parentheses. Observations are quarterly across countries over the sample period 1980-2012. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)
Δ (bank credit / GDP)$_t$	-0.014	-0.014	-0.002	-0.002	-0.017*
	[-1.770]	[-1.886]	[-0.655]	[-0.613]	[-2.312]
Δ (bank credit / GDP)$_t$ · Post$_t$	0.016	0.018*	0.007	0.006	0.011*
	[2.036]	[2.606]	[1.602]	[1.413]	[2.370]
Δ (bank credit / GDP)$_t$ · Treat$_i$	0.020	0.019	0.009	0.012*	0.008
	[1.929]	[1.835]	[1.656]	[2.108]	[1.330]
Δ (bank credit / GDP)$_t$ · Post$_t$ · Treat$_i$	-0.026*	-0.023	-0.013*	-0.013	-0.015*
	[-2.561]	[-1.904]	[-2.000]	[-1.957]	[-2.186]
Post$_t$		-0.001*	0.000	-0.000	-0.000*
		[-2.478]	[0.586]	[-0.045]	[-1.993]
Post$_t$ · Treat$_i$		0.000	-0.000	-0.000	-0.000
		[0.556]	[-0.650]	[-0.408]	[-0.386]
Constant	-0.003***	-0.002***	-0.004***	-0.003***	-0.002***
	[-45.657]	[-12.101]	[-10.264]	[-6.553]	[-19.298]
Country FE	x	x	x	x	x
Controls$_{i,t}$			x		
Controls$_{i,t}$ · Firm				x	
Controls$_{i,t}$ · Δ (bank credit / GDP)$_t$					x
Adj. R^2	0.21	0.22	0.31	0.39	0.26
N	803	803	712	712	712

Table 9: Predicted Cost of Equity Capital

Table 9 reports estimates from Equation 12 and demonstrates that the cost of equity capital is countercyclical over the credit cycle for both non-financials and banks. The table estimates 4- and 8-quarter ahead (non-overlapping) excess return of an equity index using a set of predictor variables including credit expansion (which is standardized). Panel A is for large U.S. commercial banks over the sample period 1925-2012 (excluding the years 1942-1947, for which there is no data for $\Delta(bank\,credit/GDP)$), and Panel B is for various time subsets and types of firms. Newey-West t-statistics are in parentheses. *, **, and *** denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Panel A: Large U.S. commercial banks, 1925-2012

	4 quarters ahead			8 quarters ahead		
	(1)	(2)	(3)	(4)	(5)	(6)
Δ (bank credit / GDP)	-0.064*	-0.072*	-0.065*	-0.133**	-0.153**	-0.161*
	[-2.241]	[-2.607]	[-2.200]	[-2.874]	[-2.859]	[-2.607]
Log(d / p)		0.043	0.027		0.139	0.134
		[0.746]	[0.425]		[1.452]	[1.280]
Log(book / market)		0.027	0.018		0.089	0.064
		[0.471]	[0.294]		[0.919]	[0.645]
T-bill yield			-0.009			-0.055
			[-0.428]			[-0.971]
Log(i / k)			-0.020			0.023
			[-1.058]			[0.539]
Corporate spread			-0.010			0.021
			[-0.519]			[0.407]
Constant	0.018	0.024	0.026	0.046	0.071	0.069
	[0.713]	[0.856]	[0.916]	[1.086]	[1.454]	[1.249]
R^2	0.069	0.081	0.096	0.127	0.181	0.214
N	82	82	82	41	41	41

Panel B: Various time subsets and types of firms

	4 quarters ahead			8 quarters ahead		
	(1)	(2)	(3)	(4)	(5)	(6)
Large commercial banks, 1925-2012	-0.064*	-0.072*	-0.065*	-0.133**	-0.153**	-0.161*
	[-2.241]	[-2.607]	[-2.200]	[-2.874]	[-2.859]	[-2.607]
Large commercial banks, 1925-1980	-0.087*	-0.090*	-0.098*	-0.175**	-0.174*	-0.188*
	[-2.315]	[-2.150]	[-2.444]	[-3.065]	[-2.361]	[-2.693]
Large commercial banks, 1980-2012	-0.045	-0.055	-0.053	-0.095	-0.119	-0.131
	[-1.292]	[-1.714]	[-1.050]	[-1.286]	[-1.506]	[-0.869]
Large investment banks, 1980-2012	-0.082	-0.074	-0.140	-0.233*	-0.210*	-0.356*
	[-1.130]	[-1.165]	[-1.733]	[-2.560]	[-2.548]	[-2.209]
Small commercial banks, 1980-2012	-0.042	-0.039	-0.074	-0.141	-0.138	-0.213
	[-1.170]	[-1.080]	[-1.686]	[-1.776]	[-1.563]	[-1.728]
Non-financials, 1980-2012	-0.031	-0.036	-0.028	-0.124*	-0.137*	-0.145*
	[-1.165]	[-1.427]	[-1.024]	[-2.347]	[-2.410]	[-2.282]

Chapter II

Credit Expansion and Neglected Crash Risk

(with Wei Xiong, Princeton)

1 Introduction

Economists have long argued that credit expansion by banks and other intermediaries can lead to instability of the financial system and the economy, e.g., Fisher (1933), Minsky (1977), and Kindleberger (1978). The recent literature, e.g., Borio and Lowe (2002) and Schularick and Taylor (2012), provides empirical evidence showing that bank credit expansion predicts subsequent banking crises. In this paper, we examine financial instability associated with bank credit expansion through the lens of equity prices. As equity prices aggregate expectations and preferences of equity investors, the joint dynamics of equity prices, especially of bank stocks, with credit expansion provide a channel to analyze the expectations and preferences of equity investors regarding the financial instability associated with credit expansion.

We focus on three questions regarding credit expansion from the perspectives of equity investors: First, does credit expansion predict an increase in the crash risk of bank stocks and the equity market index in subsequent quarters? As equity prices tend to crash in advance of banking crises, the predictability of bank credit expansion for banking crises does not necessarily imply predictability for equity crashes. Our second question is concerned with whether increased equity crash risk is compensated by a higher equity premium. This question is not only a natural continuation of the first, but also sheds light on several views regarding the origin of credit expansion.

In the aftermath of the recent financial crisis, an influential view argues that credit

expansion may reflect active risk seeking by bankers as a result of their misaligned incentives with their shareholders, e.g., Allen and Gale (2000) and Bebchuk, Cohen, and Spamann (2010). A second view posits that credit expansion may also reflect largely increased risk appetite of financial intermediaries due to relaxed Value-at-Risk constraints (Danielsson, Shin and Zigrand, 2012; Adrian, Moench and Shin, 2013). This view belongs to a large literature that emphasizes the limited capital of financial intermediaries as an important factor driving financial market dynamics.[1] Lastly, credit expansion may be driven by widespread optimism shared by financial intermediaries and other agents in the economy. This view can be traced back to Minsky (1977) and Kindleberger (1978), who emphasize that prolonged periods of economic booms tend to breed optimism, which in turn leads to credit expansions that can eventually destabilize the financial system and the economy. The recent literature has proposed various mechanisms that can lead to such optimism, such as neglected risk (Gennaioli, Shleifer and Vishny, 2012, 2013), group think (Benabou, 2013), extrapolative expectations (Barberis, 2012), and this-time-is-different syndrome (Reinhart and Rogoff, 2009).

If credit expansion is simply caused by bankers acting against the will of their shareholders, we expect the shareholders to demand a higher equity premium as compensation for the increased crash risk they have to bear. On the other hand, credit expansion may also reflect over-optimism or elevated risk appetite of bankers and their shareholders, in which case there may not be a higher equity premium to accompany the increased crash risk.

Our last question focuses on separately measuring the equity premium during periods of large credit expansions and contractions. The risk-appetite view and the beliefs view differ in how much the equity premium can be reduced during credit expansions. As the beliefs view emphasizes the overvaluation of equity during expansions, the equity

[1]See, for example, Shleifer and Vishny (1997), Xiong (2001), Kyle and Xiong (2001), Gromb and Vayanos (2002), Brunnermeier and Pedersen (2009), He and Krishnamurthy (2012, 2013), and Brunnermeier and Sannikov (2014).

premium can become negative, in contrast to the low but positive premium implied by the risk-appetite view.

Our data set consists of 20 developed economies with data from 1920 to 2012. We measure credit expansion as the past three-year change in bank credit to GDP ratio in each country. In contrast to the perception that credit expansions are often global, bank credit expansion actually exhibits only a small cross-country correlation outside the two most prominent credit expansions, the boom of the late 1920s leading up to the Great Depression and the boom of the 2000s.

To analyze the first question, we test whether credit expansion predicts a significant increase in the crash risk of subsequent returns of the bank equity index and broad equity market index by estimating a probit panel regression. This estimation shows that credit expansion significantly predicts a higher likelihood of equity crashes for both indices in subsequent quarters. In addition to the probit specification, we also use quantile regressions to forecast an alternative measure of negative skewness in stock returns: the distance from the median to the lower tail (5nd quantile) minus the distance to the upper tail (95th quantile). This alternative measure also confirms the same finding that bank credit expansion predicts a significant increase in the equity crash risk, reflecting financial instability associated with bank credit expansions.

Next, we address the second question regarding whether the increased crash risk associated with credit expansion is compensated by a higher equity premium. We find that one to eight quarters after bank credit expansions, despite increased crash risk, the mean excess returns of the bank equity index and broad equity index are significantly lower rather than higher. One might argue that the lower returns predicted by bank credit expansion may be caused by a correlation of bank credit expansion with a time-varying equity premium, which is indeed present in the data. However, even after controlling for a host of variables known to be predictors of the equity premium, including dividend yield, book to market, inflation, the term spread, nonresidential

investment to capital, and other variables, bank credit expansion remains strong in predicting lower mean returns of the bank equity index and equity market index.

Taken together, our analysis shows that bank credit expansion predicts increased crash risk in the bank equity index and broad equity index, and the increased crash risk is accompanied by a lower, rather than higher, equity premium. To the extent that shareholders do not demand a higher equity premium to compensate themselves for the increased crash risk, there does not appear to be an outright tension between bankers and shareholders during credit expansions. The lack of such a tension suggests that the agency view alone is insufficient for fully explaining financial instability associated with bank credit expansions. Instead, there is a need to account for optimism and risk taking by shareholders during credit expansions.

We also find that conditional on credit expansions exceeding a 95th percentile threshold, the mean excess return for the bank equity index in the subsequent eight quarters is substantially negative at -25.8%. It is difficult to explain this substantially negative equity premium simply based on elevated risk appetite of intermediaries and shareholders. Instead, it points to the presence of over-optimism or neglect of crash risk by shareholders during credit expansions.

We further compare the mean and median excess returns of the bank equity index and broad equity index predicted by bank credit expansions. Interestingly, when there is a bank credit expansion, the median excess return predicted by bank credit expansion is also lower, which suggests that equity returns subsequent to credit expansions are decreased even in the absence of the occurrence of equity crashes. This decrease in the predicted median returns likely reflects the correction of shareholders' overly optimistic beliefs over time or increased risk appetite of shareholders and contributes to about two thirds of the decrease in the predicted mean returns. The remaining third reflects the contribution of the occurrence of subsequent crash events. If shareholders have rational expectations, they would fully anticipate the frequency and severity of the crash events

subsequent to credit expansion and thus demand a higher equity premium ex ante to offset the subsequent crashes. To the extent that the median return predicted by credit expansion is lower rather than higher, shareholders do not demand an increased premium to protect themselves against subsequent crash risk.

It is important to note that our findings by no means exclude the presence of distorted incentives of bankers and elevated risk appetite of shareholders in driving unstable credit expansions. To the contrary, it is likely that these factors are jointly present. While our study serves to highlight the presence of over-optimism and neglect of crash risk by shareholders, it is likely that in this environment bankers have even greater incentives to underwrite poor quality loans and seek risk in order to cater or take advantage of their shareholders, e.g., Stein (1996), Bolton, Scheinkman and Xiong (2006) and Cheng, Hong and Scheinkman (2013).

Following Rietz (1998) and Barro (2006), a quickly growing literature, e.g., Gabaix (2012) and Wachter (2013), highlights rare disasters as a compelling resolution of the equity premium puzzle. Gandhi and Lustig (2013) argue that greater exposure of small banks to bank-specific tail risk explains the higher equity premium of small banks. Furthermore, Gandhi (2011) presents evidence that in the U.S. data, aggregate bank credit expansion predicts lower bank returns and argues that this finding is driven by reduced tail risk during credit expansion. In contrast to this argument, we find increased rather than decreased crash risks subsequent to bank credit expansions in a sample of 20 countries. This finding suggests that shareholders neglect imminent crash risk during credit expansions, as pointed out by Gennaioli, Shleifer and Vishny (2012, 2013). Our analysis does not contradict the importance of tail risk in driving equity premium. Instead, it highlights that shareholders' perceived tail risk may or may not be consistent with the actual tail risk, as suggested by Weitzman (2007). In this regard, our analysis also reinforces the concern expressed by Chen, Dou and Kogan (2013) regarding a common practice of attributing puzzles in asset prices to "dark matter,"

such as tail risk, that is difficult to measure in the data.

In the aftermath of the recent global financial crisis, there is increasing recognition by policy makers across the world of the importance of developing early warning systems of future financial crises. Prices of equity and other securities are often considered potential indicators. To the extent that equity investors tend to be optimistic and neglect crash risk during credit expansions, our analysis suggests that the predictability of equity prices for future financial crises is limited and that quantity variables such as growth of bank credit to GDP are more promising.

This chapter is structured as follows. Section 2 discussed the related literature. Section 3 presents our hypotheses and empirical methodology. Section 4 describes the data and presents summary statistics. Section 5 presents the empirical results. Section 6 concludes.

2 Related Literature

The literature has recognized that bank credit expansion can predict banking crises. By using a sample of 34 countries between 1960 and 1999, Borio and Lowe (2002) compare a set of variables, including what they call gaps in equity prices, bank credit and investment (periods in which the variables deviate from their historic trends), to predict banking crises and find that the bank credit gap performs the best. Schularick and Taylor (2012) construct a historical data set of bank credit for 14 developed countries over a long sample period of 1870-2008 and confirm that a high growth rate of bank credit predicts banking crises. We expand the data sample of Schularick and Taylor to a larger set of countries and show that the growth rate of bank credit is a powerful predictor of equity crashes. More importantly, our analysis further finds that the increased crash risk is not compensated by a higher equity premium.

Our finding of bank credit expansion predicting an increased equity crash risk reflects reduced credit quality during credit expansions, which complements several recent studies. Mian and Sufi (2009) and Keys, et al. (2010) show that the credit boom of the U.S. in the 2000's allowed households with poor credit to obtain unwarranted mortgage loans, which led to the subsequent subprime mortgage default crisis. Using U.S. data back to 1920, Greenwood and Hanson (2013) find that during credit booms the credit quality of corporate debt borrowers deteriorates and that this deterioration forecasts lower excess returns to corporate bondholders. These findings suggest the presence of over-optimism by corporate bondholders during credit booms. Our study complements the analysis of Greenwood and Hanson by documenting the presence of not only over-optimism but also neglect of crash risk by equity holders during credit expansions.

In this regard, our analysis also echoes some earlier studies regarding the beliefs of financial intermediaries during the housing boom that preceded the recent global financial crisis. Foote, Gerardi, and Willen (2012) argue that before the crisis top investment banks were fully aware of the possibility of a housing market crash but "irrationally" assigned a small probability to this possibility. Cheng, Raina and Xiong (2013) provide direct evidence that employees in the securitization finance industry were more aggressive in buying second homes for their personal accounts than some control groups during the housing bubble and, as a result, performed worse.

Our study is also related to the growing literature that analyzes asset pricing implications of balance sheet quantities of financial intermediaries. Adrian, Moench and Shin (2013) and Adrian, Etula and Muir (2013) provide theory and empirical evidence for intermediary book leverage as a relevant pricing factor for both the time-series and cross-section of asset prices. Muir (2014) documents that risk premia for stocks and bonds increase substantially during financial crises after financial intermediaries suffer large losses. Different from these studies, our analysis builds on quantity of bank credit to GDP rather than intermediary leverage or capital. By examining equity returns sub-

sequent to both bank credit expansions and contractions, our analysis systematically summarizes the time-varying equity premium across credit cycles: the equity premium tends to be high during credit contractions (which are often crisis periods) and low or even negative during credit expansions.

A broader literature investigates real and financial effects of credit expansion from both domestic macroeconomic and international finance perspectives, highlighting various consequences of credit expansion such as bank runs, output losses, capital outflows, and currency crashes.[2] In the aftermath of the recent global financial crisis, the literature has made an effort to integrate financial instability and systemic risk originating from the financial sector into mainstream macroeconomic models, e.g., Gertler and Kiyotaki (2012), He and Krishnamurthy (2012, 2013), and Brunnermeier and Sannikov (2014). Our paper contributes to this literature by highlighting the need to incorporate the role of beliefs by intermediaries and shareholders during credit expansions, which may lead to financial crises.

3 Empirical Hypotheses and Methodology

This section introduces three empirical hypotheses that anchor our analysis, together with the regression methodology we use to analyze these hypotheses.

[2]Bernanke and Gertler (1989), Kashyap, Stein and Wilcox (1993), Kiyotaki and Moore (1997), and Holmstrom and Tirole (1997) show that credit frictions can have significant and persistent effects on the real economy. Mishkin (1978), Bernanke (1983), and Eichengreen and Mitchener (2003) study the role of credit in the propagation of the Great Depression in the U.S. Demirgüç-Kunt and Detragiache (1998), Kaminsky and Reinhart (1999), Eichengreen and Arteta (2002), Borio and Lowe (2002), Laeven and Valencia (2008), and Mendoza and Terrones (2008) analyze the role of credit in international financial crises.

3.1 Crash risk

We first examine financial instability associated with bank credit expansions by analyzing crash risk in equity prices. When there is a large expansion of bank credit, credit may flow to borrowers with poor credit quality, either households or non-financial firms. Reduced borrower quality exposes banks to increased defaults, which may be realized only after a substantial deterioration in the economy.[3] When default risk becomes imminent, banks' equity prices may crash. Given the critical role played by banks in channeling credit to the economy, investors' anticipation of the large losses suffered by banks spilling over to the rest of the economy will also cause the broad equity index to crash along with the bank index.

Motivated by these considerations, we hypothesize that bank credit expansion predicts greater crash risk in the bank equity index and the equity market index, as summarized below.

Hypothesis I: *Bank credit expansion predicts subsequent equity price crashes in both the bank equity index and the equity market index.*

To examine this hypothesis, we estimate probit regressions with an equity crash indicator as the dependent variable to see if credit expansion predicts increased crash risk. Specifically, we estimate the following probit model, which predicts future equity crashes using bank credit expansion and various controls:

$$Pr[Y = 1 \,|\, predictor\, variables_{i,t}] = \Phi[\alpha + \beta'_q \cdot predictor\, variables_{i.t}] \qquad (1)$$

[3]Biais, Rochet, and Woolley (2014) provide a model in which learning about the likelihood of negative shocks in an innovative industry can lead to increased crash risk. After some periods with no shock, confidence about the industry rises, causing inefficient managers exerting low risk-prevention effort to enter the industry and raise its crash risk.

Φ is the CDF of the standard normal distribution, and Y is a future crash indicator ($Y = 1_{crash}$), which takes on a value of 1 if there is an equity crash in the next K quarters (K = 1, 4, and 8) and 0 otherwise.[4] We define an equity crash when the log excess return of the underlying equity index is less than -20% in one quarter or less than -30% in two consecutive quarters. With this definition, the equity crash indicator takes on the value of 1 every 5.4% of quarters, or one quarter every 4.6 years on average. Given that an increased crash probability may be driven by increased volatility rather than increased crash risk on the down side, we also estimate equation (1) with ($Y = 1_{boom}$), where 1_{boom} is a symmetrically defined positive tail event (with respect to the mean), and compute the difference in the marginal effects between the two probit regressions (probability of a crash minus probability of a boom).

3.2 Equity premium

If bank credit expansion is indeed accompanied by increased equity crash risk, the equity premium during credit expansion offers a lens to uncover shareholders' expectations and preferences regarding the increased crash risk. The literature has highlighted that option-like compensation contracts incentivize bankers to underwrite poor quality loans and seek risk at the expense of their shareholders and creditors (e.g., Allen and Gale, 2000; Bebchuk, Cohen, and Spamann, 2010). Suppose that during bank credit expansions, shareholders anticipate bankers seeking excessive risk or underwriting poor quality loans against the shareholders' will. While they may not be able to discipline bankers from taking the risk, they can always choose to vote on their feet by selling their shares. As a result, we expect the equity prices to offer a higher equity premium

[4] Another potential way to measure tail risk (or perception of tail risk) is to use options data. However, such data is limited to recent years in most countries.

as compensation for the increased crash risk.[5]

Another view of credit expansion focuses on the role of beliefs. Bank credit expansion may be accompanied by widespread optimism in the economy, a view emphasized by Minsky (1977) and Kindleberger (1978), which would lead to a lower equity premium or even predictable losses for equity investors. During prolonged economic booms, both bankers and their shareholders may become overly optimistic about the economy due to neglected risk (Gennaioli, Shleifer and Vishny, 2012, 2013), group think (Benabou, 2013), extrapolative expectations (Barberis, 2013), or this-time-is-different syndrome (Reinhart and Rogoff, 2009). Such over-optimism may cause bankers to excessively expand credit to households and non-financial firms and at the same time induce shareholders to ignore increased crash risk.

It is worth mentioning that the agency view and the beliefs view are not mutually exclusive, as risk-seeking incentives of bankers and over-optimism of shareholders may be jointly present in driving bank credit expansions. In the presence of overly optimistic shareholders, even rational bankers may underwrite poor quality loans and seek risk to cater or take advantage of their shareholders' optimism (e.g., Stein, 1996; Bolton, Scheinkman and Xiong, 2006; Cheng, Hong and Scheinkman, 2013). Separately, incentives together with cognitive dissonance may also lead both bankers and shareholders to turn a blind eye to warning signs about potential tail risk they face in credit markets (e.g., Benabou, 2013; Cole, Kanz, and Klapper, 2013).

These different views motivate us to examine the following hypothesis regarding the relationship between the equity premium and credit expansion.

[5]The literature has also pointed out that implicit guarantees from governments create a "too big to fail" problem and may lead banks to excessively expand credit to the economy (e.g., Rajan, 2006, 2010; Acharya, et al., 2010). Note that excessive credit expansion induced by implicit government guarantees might even benefit shareholders. If bankers expand credit to take advantage of implicit government guarantees and if the guarantees provide sufficient protection to equity holders, then there would not be any increased equity crash risk associated with bank credit expansion and equity holders would then earn a reasonable expected return on their equity holdings.

Hypothesis II: *Bank credit expansion predicts a higher equity premium in both the bank equity index and the equity market index.*

To examine Hypothesis II, we use an OLS panel regression with country fixed effects:

$$r_{i,t+K} - r^f_{i,t+K} = \alpha_i + \beta' \cdot predictor\, variables_{i,t} + \epsilon_{i,t} \tag{2}$$

which predicts the K quarter ahead excess return of either the bank equity index or equity market index, conditional on a set of predictor variables including bank credit expansion. We test whether β, the coefficient of credit expansion in equation (2), is different from zero. By using a fixed effects model, we test Hypothesis II by focusing on the time series dimension within countries. As the predictor variables come from different sources for different countries, direct comparison of the level of the predictor variables across countries is not feasible.

From an empirical perspective, it is useful to note that bank credit expansion may also be correlated with a time-varying equity premium caused by forces independent of the financial sector, such as by habit formation of representative investors (Campbell and Cochrane, 1999) and time-varying long-run consumption risk (Bansal and Yaron, 2004). A host of variables are known to predict the time variation in the equity premium, such as dividend yield, inflation, book to market, the term spread, investment to capital, and consumption to wealth. See Lettau and Ludvigson (2010) for a review of this literature. It is thus important in our analysis to control for these variables to isolate effects associated with bank credit expansion.

When estimating regressions with bank equity returns, we do not control for market returns. This is because we hypothesize that financial instability associated with bank credit expansion starts in the banking sector before spilling over to the rest of the economy, rather than bank returns simply responding to market-wide fluctuations.

Lastly, in estimating coefficients for equation (2), we test for the possible presence of small-sample bias, which may produce biased estimates of coefficients and standard errors in small samples when a predictor variable is persistent and its innovations are highly correlated with returns, e.g., Stambaugh (1999). In Section 5.5.2, we use the methodology of Campbell and Yogo (2006) to show that small-sample bias is unlikely a concern for our estimates.

3.3 Magnitude of equity premium

Another view of credit expansion highlights the role of risk appetite of the financial sector. According to this view, bank credit expansion can be caused by relaxed risk constraints or an elevated risk appetite of bankers and financial intermediaries. Danielsson, Shin and Zigrand (2012) and Adrian, Moench and Shin (2013) develop models to show that falling asset price volatility (which tends to happen during economic booms) relaxes Value-at-Risk constraints faced by financial intermediaries and allows them to expand more credit to the economy. In their framework, the elevated risk appetite leads not only to credit expansions but also to a reduced equity premium as financial intermediaries are also the marginal investors in stock markets.

It is challenging to fully separate the effects caused by over-optimism and elevated risk appetite. We explore a quantitative difference between these views. An elevated risk appetite can reduce the equity premium down to zero but not below zero in standard settings,[6] while over-optimism can cause equity prices to be substantially overvalued and thus cause the equity premium to be negative. This quantitative difference motivates us to examine the magnitude of the equity premium during credit expansions, as stated in the following hypothesis.

[6]A caveat is that a sufficiently strong hedging motive by equity holders together with a certain correlation between equity returns and endowment risk faced by equity holders may turn the equity premium to negative.

Hypothesis III: *Predicted equity excess returns subsequent to credit expansions are negative for both the bank equity index and the equity market index, reflecting the over-optimism of shareholders during credit expansions.*

Generally speaking, theories of the effects of intermediary capital on financial markets, such as those referenced in Footnote 1, typically imply a negative relationship between risk premia in asset prices and intermediary capital and put a particular emphasis on the largely increased risk premia after financial intermediaries suffer large losses. In contrast, Hypothesis III is concerned with risk premia during credit expansions, which tend to occur during periods when financial intermediaries are well capitalized.

To examine Hypothesis III, we estimate a non-linear model of the predicted equity excess return subsequent to a large credit expansion:

$$r_{i,t+K} - r^f_{i,t+K} = \alpha_i + \beta \cdot 1_{creditexpansion>x} + k \cdot controls + \epsilon_{i,t} \tag{3}$$

where $x \geq 50\%$ is a threshold for credit expansion, expressed in percentiles of credit expansion in a country. In the absence of controls, this model is equivalent to computing a simple average conditional on credit expansion exceeding the given percentile threshold x. The advantage of this formal estimation technique over simple averaging is that it allows us both to add control variables and also to compute dually-clustered standard errors for hypothesis testing, since the error term $\epsilon_{i,t}$ is possibly correlated both across time and across countries. Adding control variables shows that, with the additional information from the controls available to shareholders at the time, the returns subsequent to credit expansions may be even more predictably negative. This model specification is non-linear with respect to credit expansion and thus also serves to ensure that our analysis is robust to the linear regression model in equation

(2). After estimating this model, we report a t-statistic to test whether the predicted equity premium $E[r_{i,t+K} - r_{i,t+K} \,|\, \cdot\,]$ is significantly different from zero.

Furthermore, to examine the predicted equity excess return subsequent to large credit contractions, we also estimate a similar model by conditioning on credit contraction, i.e., credit expansion lower than a percentile threshold $y \geq 50\%$:

$$r_{i,t+K} - r^f_{i,t+K} = \alpha_i + \beta \cdot 1_{creditexpansion<y} + k \cdot controls + \epsilon_{i,t} \tag{4}$$

3.4 Quantile regressions

In addition to the regressions described above, we also adopt quantile regressions for additional analysis. First, we compare the mean and median excess returns predicted by bank credit expansions to isolate the effect on mean returns of the occurrence of crash events from other effects such as elevated risk appetite and over-optimism. Second, we construct an alternative measure of crash risk.

A quantile regression estimates the best linear predictor of the qth quantile of future equity excess returns conditional on the predictor variables:

$$Quantile_q[r_{i,t+K} - r^f_{i,t+K} \,|\, predictor\, variables_{i,t}] = \alpha_{i,q} + \beta'_q \cdot predictor\, variables_{i,t} \tag{5}$$

This quantile regression allows one to study how predictor variables forecast the entire shape of the distribution of subsequent excess returns.

Specifically, we analyze a median regression (50th quantile regression) and compare the mean and median excess returns predicted by bank credit expansions. β_{median} estimated from equation (5) measures how much equity returns decrease "most of the time" during a credit expansion. A negative β_{median} indicates that equity excess returns

subsequent to credit expansions are likely to be decreased even in the absence of the occurrence of crash events. Such a negative coefficient reflects gradual correction of equity overvaluation induced by shareholders' over-optimism or elevated risk appetite of shareholders during credit expansions. The difference between β_{mean} estimated from equation (2) and β_{median} measures, subsequent to credit expansion, how much mean returns decrease due to the subsequent occurrence of crash events in the sample, which in turn reflects the degree to which crash risk contributes to the mean returns.

Furthermore, to assess the robustness of crash risk coefficients estimated from probit regressions, we adopt a quantile-based approach to study crash risk without relying on a particular choice of thresholds for crash indicator variables.[7] Specifically, we employ jointly estimated quantile regressions to compute the following negative skewness statistic to ask whether credit expansion predicts increased crash risk:

$$\beta_{negative\,skew} = (\beta_{q=50} - \beta_{q=5}) - (\beta_{q=95} - \beta_{q=50}) \tag{6}$$

where $\beta_{q=x}$ denotes the coefficient estimated for the x quantile. This statistic $\beta_{negative\,skew}$ equals the distance from the median to the lower tail minus the distance to the upper tail. As with the probit regressions, we do not measure just $(\beta_{q=50} - \beta_{q=5})$, the distance between the median and the left tail, because a larger number could simply be indicative of increased conditional variance. Instead, we measure the asymmetry of the return distribution, the increase in the lower tail minus the increase in the upper tail.[8]

[7] Quantile regression estimates have a slightly different interpretation from the probit estimates: the probits analyze the likelihood of tail events, while quantile regressions indicate the severity of tail events. It is possible, for example, for the frequency of crash events to stay constant, while the severity of such events to increase.

[8] In the statistics literature, this measure is called the quantile-based measure of skewness. We use the 2nd and 98th quantiles to represent tail events, though the results from the quantile regressions are qualitatively similar for various other quantiles (for example, 1st/99th or 5th/95th quantiles) but with slightly less statistical significance. There is a trade-off with statistical power in using increasingly extreme quantiles, since the number of extreme events gets smaller while the magnitude of the skewness coefficient gets larger.

3.5 Standard errors

Special care must be taken to estimate these aforementioned predictive return regressions in a financial panel data setting. An important concern is that both outcome variables (e.g. non-overlapping n-quarter-ahead excess returns, n = 1, 4, and 8) and explanatory variables (e.g. bank credit expansion and controls) are correlated across countries (due to common global shocks) and over time (due to persistent country-specific shocks). If these concerns are not appropriately accounted for, the standard errors of the regression coefficients can be biased downward. Therefore, we estimate standard errors that are dually clustered on time and country, following Thompson (2011), to account for both correlations across countries and over time.

We also take a deliberately conservative approach by using non-overlapping returns. That is, in calculating 4-, 8- or 12-quarter ahead returns, we drop the intervening observations from our data set, in effect estimating the regressions on annual, biennial or triennial data. As a result, we can assume that auto-correlation in the dependent variables (excess returns) is likely to be minimal. Using non-overlapping returns thus makes our estimation robust to many potential econometric issues involved in estimating standard errors of overlapping returns.

For panel linear regression models with fixed effects, we implement dually-clustered standard errors by using White standard errors adjusted for clustering on time and country separately, and then combined into a single standard error estimate as explicitly derived in Thompson (2011). For probit and quantile regressions (including median regressions), we estimate dually-clustered standard errors by block bootstrapping, drawing blocks that preserve the correlation structure both across time and country. In the case of testing linear restrictions of coefficients, multiple regressions are estimated simultaneously to account for correlations in the joint estimates of the coefficients. For example, in testing the null H_0: $\beta_{negative\,skew} = (\beta_{q=50} - \beta_{q=5}) - (\beta_{q=95} - \beta_{q=50}) = 0$,

standard errors are generated by block bootstrapping simultaneous estimates of the q=2, 50, and 98 quantile regressions. Similarly, the difference between the mean and median coefficients, H_0: $(\beta_{mean} - \beta_{median})$, is tested by simultaneously bootstrapping mean and median coefficients; the resulting Wald statistic is then used to compute a p-value.

4 Data and Summary Statistics

We construct a panel data set of 20 countries from 1920 to 2012 using quarterly data. The data set is mostly complete for most countries from around 1950 onwards, and for half of the countries from around 1920 onwards. The sample length of each variable for each country can be found in Table A1 in the appendix.

4.1 Key variables

The main predictor variable is the three-year change in bank credit to GDP, expressed as an annualized percentage point difference. Bank credit refers to credit extended from banks to domestic households and private non-financial corporations. It excludes interbank lending and only includes non-public end users of credit. Our time series on bank credit to GDP is derived from two sources: "bank credit" from the BIS's "long series on credit to private non-financial sectors," which covers a large range of countries but generally only covers the postwar era, and from the data of Schularick and Taylor (2012) on "bank loans," which extend back over a century but only for a subset of the countries.

Throughout the paper, we refer to the three-year change in bank credit to GDP

as credit expansion or credit growth (or credit contraction when the change is negative). We often denote this predictor variable as $\Delta(bank\,credit/GDP)$, which we define to equal $[(bank\,credit/GDP)_t - (bank\,credit/GDP)_{t-3}]$. We look at changes in bank credit to GDP, rather than levels, for the following reasons. First, as shown later in Figure 2, the change in bank credit is positive during booms and falling during crises, while the level of bank credit may still be high after the crisis. Thus, the change of credit, not the level, is more indicative of the expansion or contraction phase and separates before versus after the start of banking crises. Second, bank credit as a percentage of GDP exhibits long-term trends presumably related to structural and regulatory factors. Differencing bank credit removes the secular trend and emphasizes the cyclical movements corresponding to credit expansions and contractions.[9] The three-year horizon for differencing bank credit to GDP is roughly consistent with the frequency of credit cycles.[10] Finally, when estimating regressions, we standardize the three year change in bank credit to GDP by its mean and standard deviation within each country.[11]

Our main outcome variables are future returns for both the equity market index and the bank equity index for each country. We consistently use log excess total returns as our measure of returns throughout the paper.[12]

Our main source for the price series of both indices is Global Financial Data (GFD). We choose well-known broadly-focused, market-cap-weighted indices for each country. We construct bank equity excess returns and equity excess returns for all countries

[9] As an alternative approach, we repeat our analysis with the detrended level of bank credit, using a one-sided Hodrick-Prescott (HP) filter (λ=100,000) to avoid look-ahead bias; results were qualitatively similar.

[10] In the online appendix, we provide additional analysis in Table S4 to show the strongest predictive power occurs using the three-year horizon. Specifically, we repeat our main analysis but with decomposing the three-year change into various lags of one-year changes in bank credit to GDP. The greatest predictive power comes from the 2 and 3 year lags, with the magnitude of the coefficients strongly dropping off at longer lags.

[11] Standardization is based on the in-sample distribution of each country. In Table S7 of the online appendix, we also show that results are robust to standardizing with past data only.

[12] We also repeat all the main results in the online appendix (Table S5) with arithmetic returns as a robustness check. The results are significant, albeit slightly less in magnitude for the probit and quantile regressions.

by subtracting the short-term interest rate from the equity returns. Total returns are constructed by adding dividend yield: the dividend yield of the equity index is taken mainly from GFD, and a dividend yield for the bank index for each country is constructed from individual banks' dividend yields using Compustat, Datastream and hand-collected data from Moody's Bank and Finance Manuals.[13] For forecasting purposes, we construct one-quarter-ahead excess returns by applying a lead operator to the excess returns. We also construct 4-, 8- and 12-quarter-ahead excess returns in a non-overlapping fashion.[14]

We also employ several financial and macroeconomic variables known to predict the equity premium as controls. The main control variables are dividend yield, book-to-market, inflation, non-residential investment to capital, and term spread. The variable household consumption to wealth is only reliably available for several countries and, while used in some of the analysis, is generally not included as the "main" control variables due to limited data availability. We also employ various other measures of aggregate credit and leverage of the household, corporate and financial sectors, and measures of international credit. Further information on data sources and variable construction for all variables can be found in the Appendix.

Finally, we also define a crash indicator, which takes on the value of 1 if the log excess return of the underlying equity index is less than -20% in one quarter or less than -30% in two consecutive quarters, and 0 otherwise.

[13]See the appendix and online appendix (Table S9) for details on constructing the price and dividend yield indices for bank stocks in each country. Due to the difficulty in obtaining historical data, the bank dividend yield index for each country does not necessarily contain exactly the same banks as the bank price index.

[14]Throughout the paper, we specifically exclude quarters from our analysis when inflation within ± 1 year of the given quarter is greater than 30%, because returns and interest rates become unreliable on the quarterly level. Inflation over 30% rarely occurs in our sample.

4.2 Summary statistics

Table 1 presents summary statistics for equity index returns, bank equity index returns and credit growth. Observations in Table 1 are pooled across all time periods and countries. Table 1 reports summary statistics for annualized equity excess returns without and with dividends, equity real total returns (index returns + dividends - inflation), bank equity excess returns without and with dividends, and bank equity real total returns (defined as above but for the bank equity index). The returns and standard deviations are all expressed as annualized log returns. The label Δ(bank credit / GDP) is the annualized three-year change in bank credit to GDP.

In Table 1, the mean equity log excess return is 6.0% (2.3% without including dividends). The mean equity log real total return is 7.6%. Bank stocks have slightly lower mean excess returns (5.7% with dividends, 2.2% excluding dividends, and 6.9% real total returns). We also report the median returns for all variables. The standard deviations of returns are around 21% for equity index returns and around 27% for bank stock returns.

Given that we study crash indicators and negative skewness statistics from quantile regressions based on left tail events, it is useful to get a sense of what magnitude drops these percentiles correspond to. From Table 1, we see that a 5th percentile return, which occurs on average once every 5 years, corresponds to a -63.3% annualized log return, and a 1st percentile return corresponds to an annualized log return of -106.7%. Table 1 also gives a sense of the magnitudes and variability of credit expansion. On average, bank credit to GDP expanded by 1.2% per year. In terms of the variability of credit expansion, bank credit expansion grew as rapidly as 11.3% of GDP per year (99th percentile) and contracted as rapidly as -6.3% of GDP per year (1st percentile). Total credit to GDP, which includes both bank credit and credit from other sources extended to households and non-financial corporations, grew at twice the rate of bank

credit on average, 2.4%, and is similarly volatile, with total credit expansion growing as rapidly as 17.4% of GDP per year (99th percentile) and contracting as rapidly as -8.8% of GDP per year (1st percentile). In Table 2, we show that there is a 76.8% time-series correlation ($\sqrt{R^2}$)=0.768) between bank credit growth and total credit growth.

The variability of bank credit expansion can be seen visually in Figure 1, which plots Δ(bank credit / GDP) over time. The time series for all countries appear mean-reverting and cyclical, with periods of rapid credit expansion often followed by periods of credit contraction.

Table 2 provides additional characteristics of bank credit expansions. Panel A summarizes several variables that predict future credit expansion based on an OLS panel regression with fixed effects for the three-year change of bank credit to GDP (normalized within each country) against the three-year lagged value of each of the following variables: daily equity market volatility, real GDP growth, the corporate spread, and the sovereign yield spread. Consistent with our expectations, bank credit expansions tend to follow good economic states. More specifically, lower daily equity market volatility, higher real GDP growth, smaller corporate yield spreads, and lower sovereign yield spreads in the past three years tend to precede larger bank credit expansions in the subsequent three years.

Panel B shows that bank credit expansion is correlated to changes in other aggregate credit variables (total credit, total credit to households, total credit to non-financial corporations, bank assets to GDP, and growth of household housing assets), leverage (of the household, corporate, and banking sectors), and with change in international credit (current account deficits to GDP and change in gross external liabilities to GDP). All variables here are normalized within each country. In particular, R^2 is high for the total credit, household and corporate credit, and bank assets and modest for change in gross external liabilities and household and corporate leverage, demonstrating correlation between different measures of credit.

In Figure 2, we see that historical banking crises, based on data from Reinhart and Rogoff (2009), are accompanied by large drops in equity markets, and especially in bank stocks. On average, the equity market drop starts in the year leading up to the start of the banking crisis and continues until two to three years after the start of the crisis. The fact that equity prices drop before the actual banking crises confirms a common wisdom that equity prices tend to move in advance of events that might affect the firms and the economy. This also makes it non-trivial for credit expansion to predict equity price crashes. In addition, credit peaks at the start of the crisis, with credit gradually contracting during the subsequent two years.[15]

Table 3 presents cross-country correlations of a set of variables. To economize on space, Table 3 only presents the cross-country correlations of other countries with the U.S. In general, quarterly equity excess returns are moderately correlated across countries (average correlation = 0.48) and bank equity excess returns are even less so (0.37). Bank credit expansions have historically been relatively idiosyncratic in nature with an average correlation of 0.22. This is rather modest, considering that the two most prominent credit expansions, those leading up to the Great Depression and the recent crisis, were global in nature. In fact, the average correlations of bank credit expansions in 1950-2005 (outside of these two episodes) is only 0.11. The relatively idiosyncratic nature of historical credit expansions helps our analysis, as their associations with equity returns and crashes may be attributed directly to local conditions and not indirectly through spillover from crises in other countries.

[15]The gradual contraction process may be due to credit lines pre-committed by banks, which, as documented by Ivashina and Scharfstein (2010), prevented banks from quickly reducing outstanding bank loans during the recent financial crisis.

5 Empirical Results

In this section, we report our empirical findings. We first demonstrate that credit expansion predicts an increased equity crash risk in subsequent quarters and then that credit expansion predicts a decrease in mean equity excess returns. Next, we report mean equity excess returns, conditional on bank credit expansion either exceeding a high percentile threshold (a large credit expansion) or falling below a low percentile threshold (a large credit contraction). We also use a quantile regression approach to isolate the contribution of subsequent crash events to the mean returns predicted by credit expansion from other effects. Finally, we provide a set of robustness checks of our results.

5.1 Predicting crash risk

To test Hypothesis I, we estimate the probit regression model specified in equation (1) to examine whether bank credit expansion (normalized within each country) predicts an increased probability of equity crashes, both in the bank equity index and the market index, in subsequent 1, 4, and 8 quarters. Table 4 reports marginal effects estimated from the probit model, with the dependent variable being the crash indicator ($Y = 1_{crash}$), which as defined in Section 3 takes on a value of 1 if there is a future equity crash in the next K quarters (K = 1, 4, and 8) and 0 otherwise. Given that an increased crash probability may be driven by increased volatility rather than increased negative skewness, we also estimate equation (1) with ($Y = 1_{boom}$) as the dependent variable, where 1_{boom} is a symmetrically defined positive tail event, and then compute and test the difference in the marginal effects between the two probit regressions (i.e. we calculate the increased probability of a crash minus the increased probability of a boom).

Table 4 reports the marginal effects corresponding to crashes in the bank equity

index (panel A) and in the equity market index (panel B) conditional on a one standard deviation increase in bank credit expansion. Regressions are estimated with and without the control variables. The blocks of columns in Table 4 correspond to 1-, 4-, and 8-quarter-ahead excess returns. Each regression is estimated with three sets of controls: the first block of rows (rows 1-3) reports marginal effects conditional on credit expansion with no controls, the second block of rows (rows 4-8) adding dividend yield, and the third block of rows (rows 9-21) uses all five main control variables (dividend yield, book to market, term spread, investment to capital, and inflation).

Table 4 demonstrates that bank credit expansion predicts an increased probability of equity crashes. The interpretation of the reported marginal effects is as follows: using the estimates for 1-, 4-, and 8-quarter horizons without controls, a one standard deviation rise in Δ(bank credit / GDP) is associated with a subsequent increase in the probability of a crash in the bank equity index by 2.2, 4.3, and 5.1 percentage points, respectively, and a crash in the market equity index by 1.5, 3.2, and 4.3 percentage points, respectively, all statistically significant at the 5% level. The marginal effects are slightly reduced but still significant after adding controls: after adding in all five controls, a one standard deviation rise in Δ(bank credit / GDP) is associated with a subsequent increase in the probability of a crash in the bank equity index by 1.5 (not significant), 3.6, and 3.9 percentage points, (for 1-, 4-, and 8-quarter horizons, respectively), and a crash in the market equity index by 1.2, 2.5, and 2.9 (not significant) percentage points, respectively, all statistically significant at the 5% level except the two marked. The control variables are often statistically significant too: lower dividend yield, term spread, and book to market all predict increased crash risk.

To distinguish increased crash risk from the possibility of increased volatility of returns conditional on credit expansion, we subtract out the marginal effects estimated for a symmetrically defined positive tail event (i.e. using $Y = 1_{boom}$ as the dependent variable). After doing so, the marginal effects stay about the same or actually increase

slightly: the probability of a boom conditional on credit expansion tends to decrease, while the probability of a crash increases, suggesting that the probability of an equity crash subsequent to credit expansion is driven primarily by increased negative skewness rather than increased volatility of returns. Also, as a robustness check, we adopt an alternative measure of crash risk in Section 5.4 using a quantile-regression-based approach, which studies crash risk without relying on a particular choice of thresholds for crash indicator variables.

In summary, consistent with Hypothesis I, we find that bank credit expansion predicts an increase in the crash risk of the bank equity index and equity market index in the subsequent 1 to 8 quarters. This predictability is particularly strong for the bank equity index. This result expands the findings of Borio and Lowe (2002) and Schularick and Taylor (2012) by showing that bank credit expansion not only predicts banking crises but also equity crashes, which tend to precede banking crises.

5.2 Predicting the equity premium

We now turn to testing Hypothesis II. Table 5 estimates the panel regression model specified in equation (2) of Section 3.2 (the standard OLS fixed effects model), which predicts future equity excess returns conditional on a one standard deviation increase in credit expansion. Various columns in Table 5 report estimates of regressions on credit expansion without controls, with adding dividend yield as a control, with all five main controls (dividend yield, book to market, term spread, investment to capital, and inflation), and with an additional sixth control (consumption to wealth) for which there is limited data availability.[16]

[16]For all six control variables, missing values are imputed using each country's mean. Since the data set for the five main control variables is mostly complete, mean imputation has minimal effect on the regression results, which we verify. However, for consumption-to-wealth, for which there is limited data availability, mean imputation is necessary to retain a large enough sample size to analyze the coefficients on credit expansion. The robustness checks in Table S6 in the online appendix compare regressions

Panel A reports coefficients for the bank equity index as the dependent variable, and panel B reports coefficients for the equity market index. Groups of columns correspond to 1- 4-, and 8-quarter-ahead excess returns. Coefficients and t-statistics are reported, along with the (within-country) R^2 and adjusted R^2 for the mean regressions. The coefficients from the mean regression measure the change in the equity premium associated with normalized credit expansion. For the bank equity index, a one standard deviation increase in bank credit expansion predicts 1.0, 4.9, and 8.0 percentage point decreases in subsequent returns over the 1-, 4-, and 8-quarter, respectively, all significant at the 5% level. When the controls are included, the coefficients generally are slightly lower but have similar statistical significance.[17] For the equity market index, the coefficients are smaller: a one standard deviation increase in bank credit expansion predicts 0.7, 3.3, and 5.3 percentage point decreases (all significant at the 5% level) for 1-, 4-, and 8-quarter-ahead excess returns, respectively.[18]

In general, for both the equity market index and the bank equity index, coefficients for mean regressions are roughly proportional to the number of quarters, meaning that the predictability is persistent and roughly constant per quarter for each quarter up to 2 years.[19]

---

with consumption to wealth with and without imputed missing values, along with regressions without consumption-to-wealth but on the smaller sample size for which there is consumption-to-wealth data, to show that consumption-to-wealth, while a powerful predictor of the equity premium, does not kill off the effect from bank credit expansion.

[17]For all six control variables, missing values are imputed using each country's mean. Since the data set for the five main control variables is mostly complete, mean imputation has minimal effect on the regression results, which we verify. However, for consumption-to-wealth, for which there is limited data availability, mean imputation is necessary to retain a large enough sample size to analyze the coefficients on credit expansion. The robustness checks in Table S6 in the online appendix compare regressions with consumption to wealth with and without imputed missing values, along with regressions without consumption-to-wealth but on the smaller sample size for which there is consumption-to-wealth data, to show that consumption-to-wealth, while a powerful predictor of the equity premium, does not kill off the effect from bank credit expansion.

[18]We verify that the coefficients for the bank equity index are not higher due to bank stocks having a high market beta. The bank equity index has an average market beta of about 1. Even after estimating a time-varying beta for the bank stock index using daily returns, the idiosyncratic component of bank returns also exhibits increased crash risk and lower mean returns subsequent to credit expansion.

[19]The coefficients level off after about 3 years (in unreported results), implying that the predictability is mostly incorporated into returns within 3 years.

Coefficient estimates remain similar in magnitudes after including the controls. For the equity market index, higher dividend yield, book to market, term spread, and consumption to wealth are all associated with a higher equity premium, while higher inflation and investment to capital are both associated with a lower equity premium. The signs of these coefficients are in line with prior work on equity premium predictability. These control variables tend to have stronger predictability for the equity market returns than for the bank equity returns.[20] Most importantly, the coefficient for bank credit expansion remains approximately the same magnitude and significance, despite the controls that are added. Thus, bank credit expansion adds new predictive power beyond these other variables and is not simply reflecting another known predictor of the equity premium.[21]

Table 5 also reports $R^2$and adjusted R^2 (as both have been variously reported in the equity premium predictability literature). In the univariate framework with just credit expansion as a predictor, the R^2 is 0.8%, 3.3%, and 7.3% for bank returns and 0.6%, 2.6%, and 4.5% for equity index returns for 1-, 4-, and 8-quarters ahead, respectively. Adding the five standard controls (column 3) increases the R^2 to 1.5%, 5.3%, and 8.2% for bank returns and 2.8%, 8.5%, and 12.9% for equity index returns for the same horizons. These values are within the range of values previously reported in the literature for the various control variables.[22] As a robustness check, in Section 5.5.1

[20]We use the same set of control variables for predicting returns of the banking index and equity market index because these controls are chosen to capture the time-varying environment in the broad equity market. It is thus not surprising that these controls have greater predictability for the equity market index than for the banking index.

[21]Table S2 in the online appendix provides various robustness checks to show that the results are not driven by changes in the denominator (GDP) but by changes in the numerator (bank credit) of the main predictor variable. Table S2 also shows that the findings are robust to using log change in credit or log change in credit/GDP as the predictor variable, and also to controlling for change or log change in GDP on the right hand side of the regression.

[22]There is a large range of R^2 and adjusted R^2 values reported in the literature for common predictors of the equity premium in U.S. data. For example, Campbell, Lo, and MacKinlay (1996) report R^2 for dividend yield: 0.015, 0.068, 0.144 (1, 4, 8 quarter overlapping horizons, 1927-1994); Lettau and Ludvigsson (2010) report adjusted R^2 for dividend yield: 0.00, 0.01, 0.02, and for cay: 0.08, 0.20, 0.28 (1, 4, 8 quarter overlapping horizons, respectively, 1952-2000); Cochrane (2012) reports R^2 for dividend yield: 0.10, for cay and dividend yield together: 0.16, and for i/k and dividend yield together: 0.11 (for 4 quarter horizons, 1947-2009); Goyal and Welch (2008) report adjusted R^2 of 0.0271, -0.0099,

we re-analyze the probit and mean regressions but on various geographical subsets and various subsets in time. In general, the results are robust to the subsamples analyzed.

Taken together, the results in these two subsections show that despite the increased crash risk associated with bank credit expansion, the predicted equity excess return falls rather than increases.[23] It is important to note that bank credit expansions are directly observable to the public. Thus, it is rather surprising that bank shareholders and stock investors do not demand a higher equity premium from their stock holdings to compensate themselves for the increased crash risk. In this sense, there does not appear to be an outright tension between shareholders and bankers during periods of bank credit expansions. Our finding thus challenges the narrowly-focused agency view that bank credit expansions are simply caused by bankers acting against the will of shareholders. Instead, there is a need to account for the presence of over-optimism and elevated risk appetite of shareholders.

5.3 Excess returns subsequent to credit expansions and contractions

Next, we test Hypothesis III by examining excess returns subsequent to large credit expansions and contractions. We find that predicted excess returns subsequent to large credit expansion are significantly negative and large in magnitude.

We estimate the magnitude of equity excess returns subsequent to credit expansions and contractions using non-linear regression models (3) and (4) discussed in Section 3.3. These regressions estimate 4-, 8-, and 12-quarter-ahead excess returns on an indicator

-0.0094, 0.0414, 0.0663, 0.1572 (annual returns, 1927-2005) for dividend yield, inflation, term spread, book to market, i/k, and cay, respectively.

[23] Gandhi (2011) also shows that in the U.S. data aggregate bank credit expansion negatively predicts the mean return of bank stocks, although he does not examine the joint presence of increased crash risk subsequent to bank credit expansions.

for credit expansion exceeding a high percentile threshold (an event we call a large credit expansion) or falling below a low percentile threshold (a large credit contraction), along with the five standard control variables. These regressions allow us to test whether the predicted excess return is negative subsequent to a credit expansion and positive subsequent to a credit contraction without relying on the linear specifications used in our earlier analysis.

The predicted excess returns conditional on credit expansion exceeding or falling below given percentile thresholds are plotted in Figure 3 and reported in Table 6. Specifically, Figure 3 plots the predicted 8- and 12-quarter-ahead excess returns conditional on credit expansion exceeding various high percentile thresholds varying from the 50th to 98th percentiles (i.e., large credit expansions), and on credit expansion below various low percentile thresholds from the 2nd to 50th percentiles (i.e., large credit contractions). Panel A is for the bank equity index, and panel B is for the equity market index. A 95% confidence interval is plotted for each of the returns.

Figure 3 shows that the predicted excess returns for both the bank equity index and the equity market index are decreasing with the threshold and remain negative across the upper percentile thresholds. Table 6 reports the same information but in tabular form. The predicted negative returns are weaker for the 4-quarter horizon but get increasingly stronger for 8- and 12- quarter horizons. For example, at the 95th percentile threshold and without controls, the predicted negative returns are -11.5%, -25.8%, and -43.7% for the 4-, 8-, and 12-quarter ahead horizons, with t-statistics of -0.875, -2.301, and -2.544, respectively. After controlling for the five standard controls, both the magnitude and t-statistic of the predicted negative returns at the 8- and 12-quarter ahead horizons remain similarly strong. Also note that there are a reasonably large number of observations satisfying the 95th percentile threshold. According to Table 6, there are 68, 28, and 21 observations for 4-, 8- and 12-quarter ahead horizons, respectively.

The predicted negative returns for the broad equity market index, while weaker in both magnitude and t-statistic than those for the bank index, are nevertheless substantial. Panel B of Table 6 shows that at the 95th percentile threshold, the predicted returns are -4.9%, -11.8%, and -18.4% for the 4-, 8-, and 12-quarter ahead horizons without controls, with t-statistics of -0.631, -1.603, and -2.241, respectively.

The large and significantly negative excess returns predicted by credit expansion confirm Hypothesis III and present a challenge for models that use only elevated risk appetite to explain the joint presence of increased crash risk and decreased mean return subsequent to credit expansion. Instead, this finding suggests that shareholders are overly optimistic and neglect crash risk during credit expansions.

Finally, Figure 3 and Table 6 also show that subsequent to credit contractions, the excess returns are positive. When credit contraction is less than the 5th percentile threshold, the predicted excess return in the subsequent 8 quarters is 24.9% for the bank index and 20.4% for the equity market index, both significant at the 5% level. As bank credit tends to contract during banking crises, the positive equity premium subsequent to a credit contraction is consistent with the finding of Muir (2014) that risk premia tend to be large during financial crises.

We provide various robustness checks in Section 5.5.3 to show that predicted excess returns subsequent to large credit expansions are robustly negative: 1) even after grouping observations of large credit expansions into distinct episodes (clusters) and then averaging across these episodes (addressing the concern that concurrent credit expansions in multiple countries during the same global episode ought to be treated as a single observation rather than separate observations), and 2) defining the percentile thresholds for each country strictly based on past observations for that country.

Figure 3 and Table 6 document a full picture of the dramatic, time-varying equity premium across credit cycles. During large bank credit expansions the expected excess returns of both the bank equity index and broad equity market index are substantially

negative, while during large bank credit contractions the expected excess returns are substantially higher than the long-run level of the equity premium.

5.4 Quantile regression-based analysis

In this subsection, we use a quantile regression approach to examine the median equity excess returns predicted by credit expansion and analyze the difference between the predicted mean and median returns. Furthermore, we adopt an alternative measure of crash risk to assess the robustness of increased crash risk subsequent to credit expansions.

Recall the quantile regression model specified in equation (5) of Section 3.4, which examines the predictability of bank credit expansion (normalized within each country) for the full distribution of subsequent equity returns. Table 7 reports estimates from the quantile regressions. The columns correspond to 1-, 4-, and 8- quarter-ahead excess returns, first for the bank equity index and then for the equity market index. The top half reports estimates for quantile regressions on credit expansion with no controls, while the bottom half reports estimates on credit expansion with the standard set of five controls (dividend yield, inflation, book to market, term spread, and investment to capital). To save space, coefficients on control variables are not reported in Table 7.

β_{median} measures how much quarterly equity returns decrease "most of the time" when there is credit expansion, while $(\beta_{mean} - \beta_{median})$ measures how much the mean return is reduced due to the occurrence of tail events in the sample. Table 7 reports the coefficients and t-statistics for the estimates of β_{mean} and β_{median} , as well as their difference and its associated p-value. The estimates for β_{median} are -0.007, -0.030, and -0.066 for the bank equity index and -0.005, -0.024, and -0.052 for the equity market index (1-, 4- and 8- quarter horizons, respectively); all coefficient estimates are significant at

the 5% level. After including the controls, the estimates remain at similar values. Taken together, these estimates show that subsequent to credit expansion, the median excess return is decreased across different horizons. As this decrease in the median excess return is not related to the occurrence of crash events, it reflects either the gradual correction of shareholders' over-optimism over time or the elevated risk appetite of shareholders.

In general, the median coefficients are about two-thirds of the level of corresponding mean coefficients. The remaining third of the decrease (i.e., $\beta_{mean} - \beta_{median}$) reflects the contribution of the occurrence of crash events in the sample to the change in the mean return associated with credit expansion. If shareholders have rational expectations, they would fully anticipate the frequency and severity of the crash events subsequent to credit expansion and thus demand a higher equity premium ex ante to offset the subsequent crashes. To the extent that the median return predicted by credit expansion is lower rather than higher, shareholders do not demand an increased premium to protect them against subsequent crash risk.

Table 7 also reports the coefficients and t-statistics for credit expansion from the three quantile regressions, $\beta_{q=5}$, $\beta_{q=50}$, and $\beta_{q=95}$, followed by the alternative crash risk measure — the conditional negative skewness coefficient $\beta_{negative\,skew} = (\beta_{q=50} - \beta_{q=5}) - (\beta_{q=95} - \beta_{q=50})$ — and its associated t-statistic. For bank equity index returns without control variables, the coefficient for negative skewness, $\beta_{negative\,skew}$, is estimated to be 0.023, 0.086, and 0.101 (all significant at the 5% level) for 1-, 4- and 8-quarter horizons, respectively. Once the controls are included, the coefficient for 1-quarter horizon remains roughly the same and significant at the 5% level, while for 4- and 8-quarter horizons becomes smaller and insignificant. As expected, tail risk for equity market index returns has a similar although smaller association with bank credit expansion because the tail risk in the equity market index originates indirectly from the financial instability of banks. Overall, the alternative quantile measure of crash risk confirms

our earlier finding from probit regressions of increased crash risk associated with credit expansion.

5.5 Further analysis

In this subsection, we perform several robustness checks. First, we check the robustness of the probit and mean regressions in geographical subsets and subsets in time. We also verify that small-sample bias due to persistent predictor variables is not likely to be a concern. Finally, we examine alternative ways to cluster standard errors and classify observations in our analysis of the large negative and positive returns subsequent to large credit expansions and contractions.

5.5.1 Robustness in subsamples

As a robustness check, we re-estimate the probit and mean regressions in various subsamples. In general, the coefficients have similar magnitudes regardless of the subsamples analyzed, reflecting the fact that our results are not driven predominantly by particular countries or historical time periods.

Table 8 reports mean and probit coefficients for Δ(bank credit / GDP) on future equity excess returns for various subsets of countries and time periods. Using a 4-quarter forecasting horizon, the regressions are the same as those reported in Tables 4 and 5. In Panel A, the data is subdivided into geographical regions, and separate regressions are run for each of the regions. In Panel B, we change the time period: one set of regressions is run on the full sample (1920-2013), another is run excluding the most recent crisis (1920-2005), and a third is run excluding both the recent crisis and the Great Depression (1950-2005).

In Panel A, for both the bank equity index and the equity market index, we see that the coefficients for the mean and probit regressions are roughly similar for each of the geographical subsets as they are for the full sample of developed countries. The mean coefficients are slightly larger for some regions (South Europe, Western Europe, Scandinavia) and slightly lower for other regions (the U.S. and English-speaking countries). The statistical power is reduced for several regions, though that is probably due to the smaller sample size in these subsets. The probit coefficients for both the bank equity index and equity market index are similar across regions, and with somewhat less statistical power again due to the smaller sample size.

Panel B shows the estimated mean and probit coefficients for different sample periods. In general, the coefficients have almost the same magnitude and statistical significance regardless of the sample period we use, implying that our results are not driven simply by the Great Depression or the recent financial crisis.

5.5.2 Test for small-sample bias

Tests of predictability in equity returns may produce biased estimates of coefficients and standard errors in small samples when a predictor variable is persistent and its innovations are highly correlated with returns, e.g., Stambaugh (1999). This small-sample bias could potentially pose a problem for estimating coefficients in our study, because the main predictor variable, three-year change in bank credit to GDP, is highly persistent on a quarterly level. In this subsection, we test for the possibility of small-sample bias using the methodology of Campbell and Yogo (2006) and find that small-sample bias is probably not a concern for our estimates.

The idea behind the methodology of Campbell and Yogo (2006) is that three conditions need to be jointly met for small-sample bias to be a concern: 1) the predictor variable needs to be persistent; 2) its innovations need to be highly correlated with

returns, and 3) the sample size needs to be small. Campbell and Yogo (2006) present Monte Carlo evidence to demonstrate when small-sample bias is or is not likely a concern, given parameter values for the sample size, persistence of the regressor, and the correlation of its innovations with returns.[24] Section B of the Appendix discusses the methodology in detail to test for small-sample bias along the lines of Campbell and Yogo (2006).

Table 9 reports parameter values corresponding to the sample size (N), persistence of bank credit expansion (ρ), and the correlation of its innovations with returns (δ). Table 9 shows that our data correspond to parameter values well outside the region for which small-sample bias is likely to be a concern. The key reason is that δ is small in our data.

To test for small-sample bias in multivariate regressions that use the five standard control variables, the same parameter values are also computed after replacing returns with the return residuals after controlling for the five standard control variables: see Section B of the Appendix for details. Because our data set is a panel and because fixed effects may also cause biased estimates in small samples, as an extra and overly-conservative robustness check, we also obtain tables of parameter estimates for each of the 20 countries individually (results reported in the online appendix, Table S8) and find that individual countries' parameters, with only rare exceptions, also fall into the region for which small-sample bias is probably not a concern.[25]

[24]Specifically, the Monte Carlo simulations report regions of the parameter space for which the actual size of the nominal 5% t-statistic (generated when testing the estimated β against the true β_0 with null hypothesis $\beta = \beta_0$ and alternative $\beta > \beta_0$) is greater than 7.5%.

[25]The few cases in which parameters fall into the region for which small-sample bias may still be a concern: Ireland (4, 8-quarters only, bank returns only), Portugal (8-quarters only, equity index returns only), and Spain (1-quarter only, both bank and equity index returns), since these countries had unusually large and persistent credit expansions in the 2000s. See Table S8 in the online appendix for details.

5.5.3 Robustness of negative returns subsequent to large credit expansions

Clustering observations by historical episodes. Recall Table 6, which analyzes equity excess returns subsequent to large credit expansions and large credit contractions. One might argue that concurrent observations of large credit expansions across multiple countries might reflect a single global episode rather than various local events. Accordingly, the episode may have correlated effects across countries and over the duration of the episode in ways not captured by dually-clustered standard errors. Here we demonstrate that the predicted excess returns subsequent to large credit expansions are robustly negative, even after grouping observations of large credit expansions into distinct episodes (clusters) and then averaging across these episodes.

By our count, there are 20 distinct historical episodes of credit expansion that satisfy the 95th percentile threshold for the 8-quarter ahead predictive regression analysis. These 20 historical episodes are widely dispersed throughout our sample period. Some of these 20 distinct historical episodes are well-known (like the Japanese boom of the late 80s, the boom preceding the 1997-8 East Asian Crisis, and the boom preceding the Scandinavian financial crises of the late 80s and early 90s, to name three examples), while other historical episodes are less well-known. This robustness check thus averages large credit expansion observations across multiple countries and years that are part of the same historical episode, and then considers each historical episode (cluster) as a single data point.

Specifically, we do as follows. Since countries undergoing large credit expansions (or contractions) may remain over the 95th (or under the 5th) percentile thresholds for multiple years, to collapse observations across time, we select only the returns subsequent to the first year in which credit expansion first crosses the 95th (or 5th) percentile thresholds.[26] Then we group concurrent observations across countries into distinct his-

[26] A list of all large credit expansions based on the 95thpercentile threshold and large credit contractions based on the 5th percentile thresholds, with their subsequent equity and bank equity returns,

torical episodes. Finally, returns from the resulting 20 historical episodes in the sample are averaged together to generate Table 10, taking each such historical episode as a single, independent observation.

Even after grouping observations into distinct historical episodes and averaging across these historical episodes, the subsequent returns are robustly negative. Table 10 reports the average returns in the 4, 8, and 12 quarters following the start of historical episodes of large credit expansions (Panel A) and large credit contractions (Panel B). For large credit expansions, subsequent 4-, 8-, and 12-quarter excess returns for the bank equity index are: -12.1%, -22.8%, and -17.6% (t-stats of -2.12, -2.82, and -2.06, respectively) For large credit contractions, subsequent 4-, 8-, and 12-quarter excess returns for the bank equity index are: 17.8%, 29.4%, and 43.50% (t-stats of 3.90, 3.27, and 3.61, respectively).

Classifying large credit expansions based strictly on past information. One may worry that the percentile thresholds for classifying large credit expansions and contraction use future information, since the percentiles are calculated for each country with the full in-sample distribution of credit expansion. To address this concern, we repeat the analysis used to predict negative returns conditional on large credit expansions but this time calculate the percentile thresholds for each quarter based only on past observations (percentile thresholds are only calculated when there is at least 5 years of past data for that country). For example, for credit expansion to be above the 95% threshold, credit expansion in that quarter must be greater than 95% of all previous observations for that country.

Table 11 demonstrates that predicted excess returns subsequent to large credit expansions are robustly negative, even when conditioning returns strictly on past information. The predicted negative returns are similar to those reported in Table 6 for the 95th and 98th percentiles, though slightly weaker in terms of both magnitude and

grouped together into historical episodes, can be found in Table S1 in the online appendix.

t-statistics of the coefficients. Thus, the predicted negative returns are robust, even conditioning strictly on only past information.

6 Conclusion

In a set of developed economies, we find that bank credit expansion predicts significantly increased crash risk of the bank equity index and equity market index in subsequent one to eight quarters. Despite the increased crash risk, credit expansion predicts both lower mean and median returns of these indices in the subsequent quarters, even after controlling for a host of variables known to predict the equity premium. The predicted excess return of the bank equity index in the eight quarters after credit expansion exceeding the 95th percentile for that country is significantly negative with a magnitude of -25.8%. It is difficult to explain the joint appearance of increased crash risk and decreased excess return subsequent to credit expansions simply by bankers acting against the will of shareholders or by elevated risk appetite of bankers and intermediaries. Instead, our findings suggest a need to account for the role of over-optimism and neglect of crash risk by shareholders.

Appendix

A. Data construction

This appendix contains additional information related to data sources and variable construction. The sample length for each country and variable is reported in Table A1. All older historical data was extensively examined country-by-country for each variable

to ensure accuracy and was compared across multiple sources whenever possible.

Bank credit expansion. The main explanatory variable is bank credit to GDP. As explained in Section III, bank credit refers to credit extended from banks to private end users of credit: domestic households and private non-financial corporations. The data for this variable are derived from two sources: "bank credit" from the BIS's "long series on credit to private non-financial sectors" and from the data of Schularick and Taylor (2012) on "bank loans." In merging the two series, we scale the level of "bank loans" to avoid breaks in the series. Still, there are slight discrepancies between the two data sources, most likely coming from differing types of institutions defined as banks, differing types of credit instruments considered "credit," and differing original sources used to compile the data. However, the BIS and Schularick-Taylor data match qualitatively, as their overlap is highly correlated.

Market and bank index excess total returns. We chose well-known broadly-focused, market cap weighted indices for each country. Our main data source for equity returns was Global Financial Data (GFD), though in a few cases we took data directly from stock exchanges' websites. In countries with several internationally-known equity indices (for example, the S&P 500, DJIA and NASDAQ in the U.S.), we favor the index with the broadest scope and the longest time series (the S&P 500 in the U.S.). For bank equity indices, we similarly choose market cap weighted indices of banking stocks, or when a bank-specific index was not available, an index of the financial sector (see Table S9a in the online appendix for details on bank price index construction). Total returns are constructed by adding dividend yield: To get total returns, the dividend yield of the equity index is taken from GFD (occasionally supplemented by Compustat and Datastream), and a dividend yield for the bank index for each country was constructed from individual bank's dividend yields using Compustat and Datastream (1973 onwards) and from hand-collected price and dividend data (1920–1978) of the largest publicly-listed banks in each country from Moody's Bank and Finance Manuals (see Table S9b in the

online appendix for details on bank dividend yield index construction). Due to the difficulty in obtaining historical data, the bank dividend yield index for each country does not necessarily contain exactly the same banks as the bank price index.

Controls. Dividend yield comes from GFD, supplemented by data from Thompson Reuters Datastream. *Book-to-market* comes from Datastream. *Inflation* is calculated from CPI data from GFD. *Long-term interest rates* are the yields on 10-year government bonds taken mostly from GFD and OECD. *Short-term interest rates* are almost always the 3-month government t-bill rates taken from GFD, the IMF, OECD, Schularick-Taylor (2012), and other sources. Occasionally, for older data, the short-term interest rate was taken to be the yield on central bank notes, high-grade commercial paper, deposits, or overnight interbank lending; since some of these rates can rise in times of market distress and also historically have been regulated, care was taken to make sure these alternative rates, when used, were representative of the market short-term interest rate. The term spread is the long-term interest rate minus the short-term interest rate.

Household *consumption to wealth* is private consumption expenditure from national accounts taken from GFD divided by aggregate financial assets held by the household sector from Piketty and Zucman (2014). *Investment to capital* is private non-residential fixed investment divided by the outstanding private non-residential fixed capital stock, which comes from the Kiel Institute's database on investment and capital stock. *Daily stock volatility* is computed for each country and quarter as the standard deviation of daily returns by using daily stock returns from GFD of the equity market index. The *corporate yield spread* is the yield spread between the AAA-rated 10-year-maturity corporate bond index from GFD and the 10-year government bond. The *sovereign spread* is the yield on the 10-year government bond minus the yield on the U.S. 10-year Treasury. *Real GDP growth* (year-over-year) is calculated from nominal GDP and the GDP deflator taken from GFD.

Other measures of credit and leverage. The data on bank credit is compared with

several other measures of credit: total credit refers to credit extended from all sources to domestic households and private non-financial corporations. The variables *total credit to households* and *total credit to nonfinancial corporations* are the same as *total credit* but decomposed into household and corporate components. All variables are normalized by GDP. Like bank credit, these credit aggregates are taken from the BIS's "long series on credit to private non-financial sectors" and cover credit extended to end users (domestic households and/or private non-financial corporations) and excludes interbank lending.

Other indirect measures of credit: *bank assets to GDP*, which comes mainly from Schularick and Taylor (2012), and *household housing asset growth*, which is the real growth in housing assets owned by the household sector, from Piketty and Zucman (2014). We also looked at leverage of the household, non-financial corporate, and banking sectors: specifically, *household debt to assets* (which is aggregate household debt to aggregate household assets from Piketty and Zucman (2014)) and *non-financial equity to assets* and *bank equity to assets* (using book values taken from Thompson Reuters Datastream). Lastly, we also examined international credit flows and aggregates using *current account to GDP* (gathered from the IMF's external debt database and OECD) and *gross external liabilities to GDP* (both public and private liabilities, from Lane and Milesi-Ferretti's (2007) database on countries' external assets and liabilities).

Backfilling/forward-filling. This paper performs all analysis on quarterly data. When data comes only in annual time series, as some of the older historical data does, the annual data (assuming it is an explanatory variable, not an outcome variable) is filled *forward* for the three subsequent quarters. We fill explanatory variables forward to avoid look-ahead bias in forecasting, since forward filled information for each quarter would already be known.

B. Methodology and results for small-sample bias test

Following the Campbell and Yogo (2006) methodology, we estimate the following regressions:

$$r_{i,t+K} - r^f_{i,t+K} = \alpha_i + \beta \cdot credit\,expansion_{i,t} + u_{i,t} \tag{7}$$

$$credit\,expansion_{i,t+K} = \gamma_i + \rho \cdot credit\,expansion_{i,t} + \epsilon_{i,t} \tag{8}$$

Table 9 reports parameter values corresponding to the sample size (N), persistence of the main predictor variable, bank credit expansion (ρ and $c = N \cdot (\rho - 1)$), and the correlation of its innovations with returns ($\delta = corr(u_{i,t}, \epsilon_{i,t})$). In addition, to test for small-sample bias in multivariate regressions that use the five standard control variables, we estimate the following additional regression:

$$r_{i,t+K} - r^f_{i,t+K} = \alpha_i + k' \cdot controls_{i,t} + z_{i,t} \tag{9}$$

and replace the left-hand side variable in equation (7) with the residual, $z_{i,t}$, taken from equation (9). Parameters obtained in the presence of control variables are also reported in Table 9.

From Table 9, we can see that all the values of δ are less than 0.125 (meaning there is minimal correlation between innovations in credit expansion with equity returns), the critical threshold reported in Campbell and Yogo (2006) for which small-sample bias is likely not to be a problem regardless of the value of c. In addition, because of the large sample size of our data, $c = N \cdot (\rho - 1)$ is universally larger than the threshold for which small-sample bias is likely not to be a problem regardless of the value of δ. Thus, our data correspond to parameter values well outside the region for which small-sample bias may be a concern.

Figure 1: Credit expansion

Credit expansion, measured as the past three-year change of bank credit to GDP, is plotted over time for the 20 countries in the sample. Observations are quarterly, 1920-2012. Bank credit refers to credit issued by banks to domestic households and domestic private non-financial corporations.

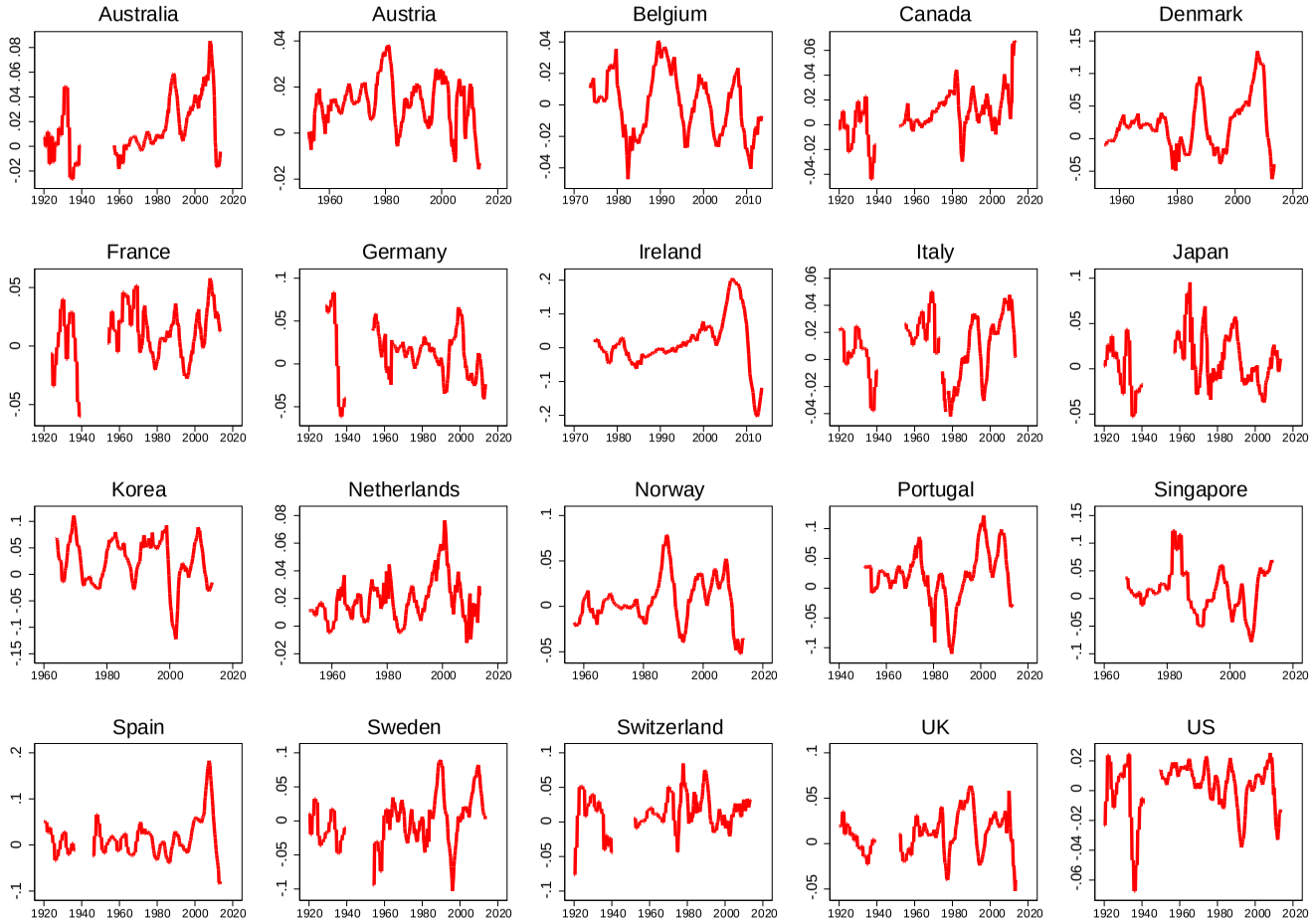


Figure 2: Credit and equity prices before and after banking crises

Bank credit to GDP (relative to each country's historical mean) and the bank equity and equity market cumulative total return indices (relative to their pre-crisis peaks) are plotted over time before and after the start of banking crises, where the start of banking crises is based on data from Reinhart and Rogoff (2009). Bank credit to GDP and the two equity indices are pooled averages across time and countries, conditional on the given number of years before or after the start of a banking crisis. Data are from the 20 countries shown in Figure 1 for 1920-2012.

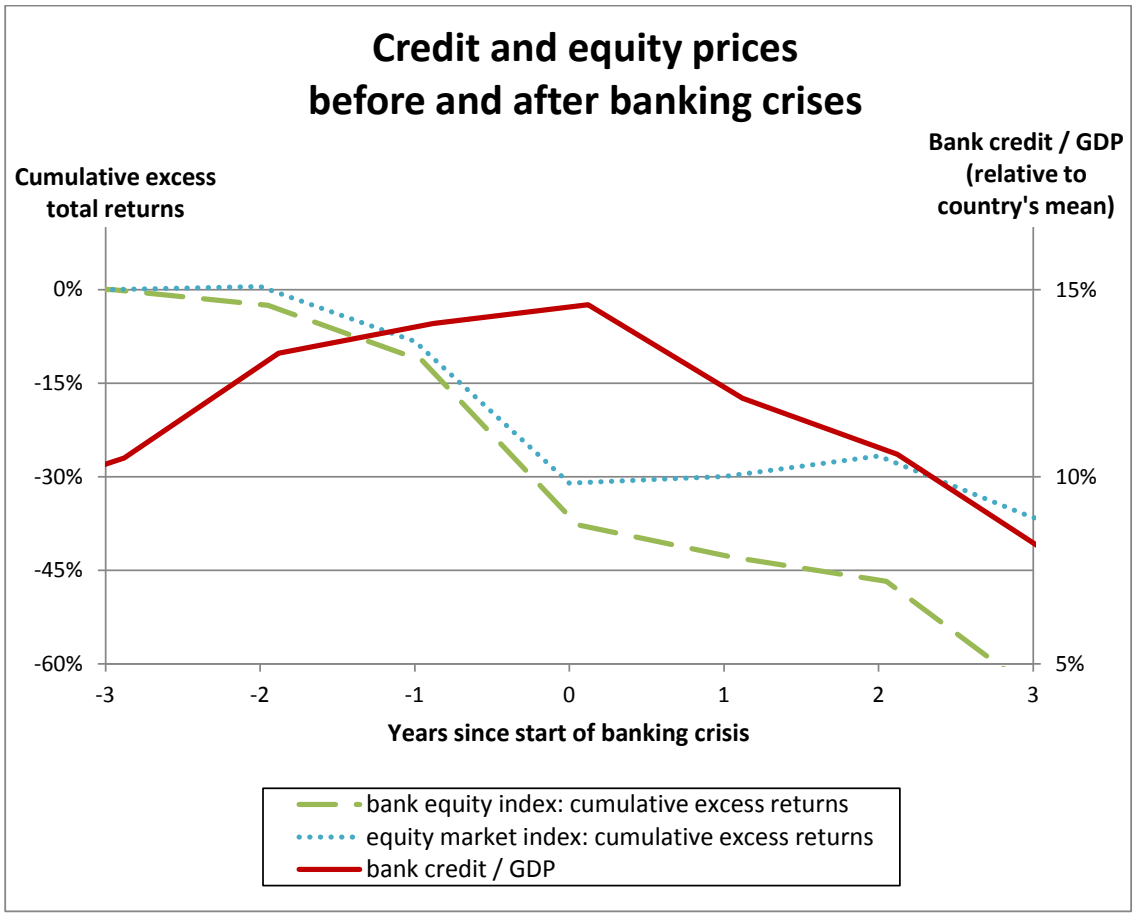
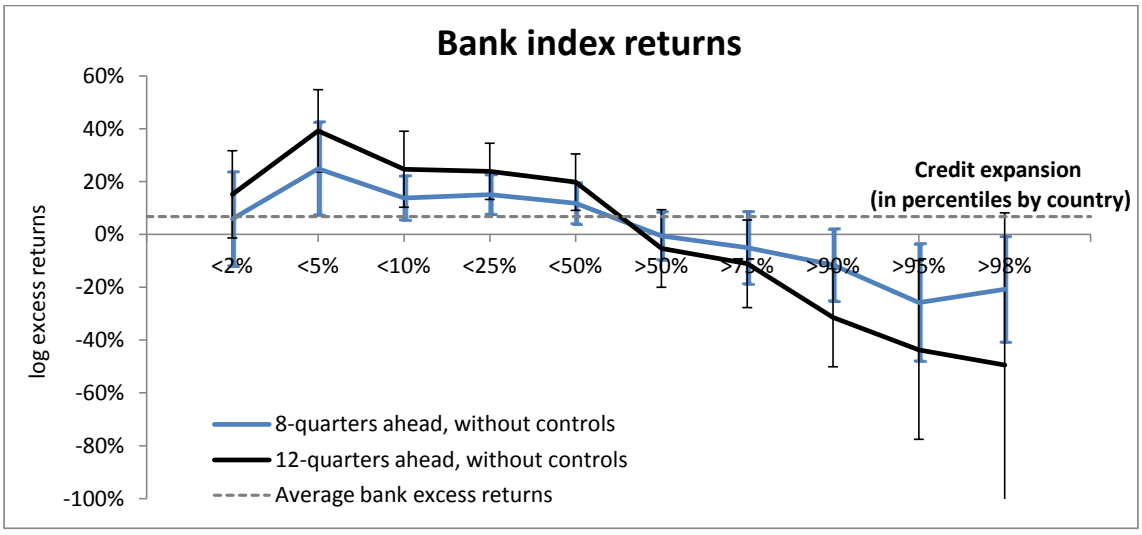


Figure 3: Negative predicted returns subsequent to large credit expansions and contractions

Panel A (for bank index returns) and Panel B (for equity market index returns) plot estimates and confidence intervals reported in Table 6. The plot shows the magnitude of equity excess returns 8- and 12-quarters subsequent to large credit expansions (when credit expansion exceeds a given percentile threshold), in addition to average returns subsequent to large credit contractions (when credit expansion falls below a given percentile threshold). Average returns conditional on the thresholds are computed using regression models (3) and (4) with non-overlapping returns. 95% confidence intervals are computed using dually-clustered standard errors.

Panel A: Bank index returns



Panel B: Equity index returns

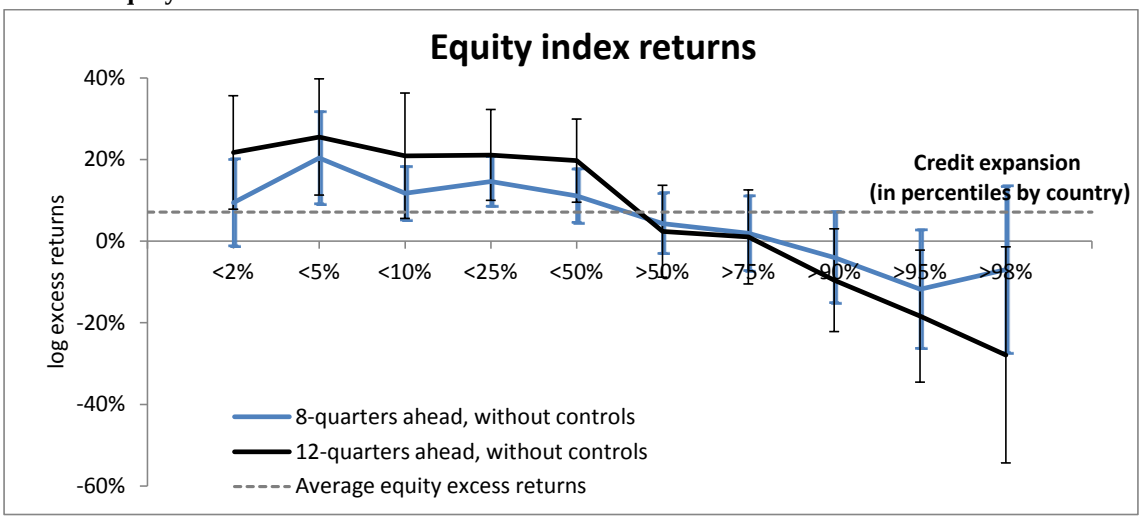


Table 1: Summary statistics

Summary statistics are reported for equity log excess returns (with and without dividends) and real returns for both the bank equity and equity market indices. Summary statistics are also reported for the three-year change in bank credit to GDP (credit issued by banks to domestic households and domestic private non-financial corporations) and three-year change in total credit to GDP (credit issued by all sources to domestic households and domestic private non-financial corporations). All statistics are pooled across countries and time.

	N	Mean	Median	Stdev.	1%	5%	10%	90%	95%	99%
Quarterly log returns, annualized										
Equity excess returns (w/out dividends)	5556	0.023	0.024	0.215	-1.101	-0.644	-0.456	0.472	0.661	1.196
Equity excess returns (incl. dividends)	4733	0.06	0.064	0.2145	-1.067	-0.633	-0.44	0.52	0.71	1.161
Equity real returns (incl. dividends)	5084	0.076	0.083	0.2145	-1.054	-0.600	-0.421	0.537	0.719	1.206
Bank stocks excess returns (w/out dividends)	4846	0.022	0.008	0.278	-1.378	-0.755	-0.509	0.529	0.770	1.677
Bank stocks excess returns (incl. dividends)	4454	0.057	0.048	0.275	-1.362	-0.743	-0.486	0.568	0.807	1.690
Bank stocks real returns (incl. dividends)	4665	0.069	0.055	0.269	-1.301	-0.707	-0.457	0.581	0.808	1.649
Credit to private households and non-financial corporations, 3 year percentage point change										
Δ (Bank credit / GDP)	5253	1.2%	1.1%	3.1%	-6.3%	-3.3%	-2.2%	4.8%	6.2%	11.3%
Δ (Total credit / GDP)	3811	2.4%	2.1%	5.0%	-8.8%	-4.6%	-2.8%	8.2%	10.3%	17.4%

Table 2: Time series correlations

This table reports estimates from panel regressions with fixed effects and the dependent variable being three-year change of bank credit to GDP (standardized within each country). Panel A shows the coefficient and R^2 from regressing future three-year change of bank credit to GDP on the past-three-years equity market daily volatility, real GDP growth, corporate yield spreads, and sovereign yield spreads. Panel B presents the coefficient and R^2 from regressing three-year change of bank credit to GDP on contemporaneous changes in other similar credit measures, including other aggregate credit variables (total credit to households [HHs], non-financial corporations [NFC]), leverage (of the household, corporate, and banking sectors), and changes in international credit. All variables are standardized within each country. The *, **, and *** marked on the regression coefficients denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Panel A: Variables that predict future credit expansion

		RHS variable:			
LHS variable:		Daily volatility	Real GDP growth	Corporate yield spread	Sovereign yield spread
Future 3-year change in (bank credit / GDP)	β	-.253**	.153*	-.227**	-0.13*
	R^2	0.09	0.02	0.13	0.02
	N	282	435	200	396

Panel B: Contemporaneous variation with other credit variables

		RHS variable:									
LHS variable:		Δ (total credit)	Δ (total credit to HHs)	Δ (total credit to private NFCs)	Δ (Bank assets / GDP)	Growth of household housing assets	HH debt / assets	NFC equity / assets	Bank equity / assets	Δ (gross external liabilities / GDP)	Current account deficit / GDP
Current 3-year change in (bank credit / GDP)	β	.752***	.653***	.651***	.626***	.232*	0.19	-0.249	-.25*	.332**	.16*
	R^2	0.59	0.45	0.41	0.37	0.12	0.14	0.12	0.13	0.14	0.05
	N	333	221	217	324	126	117	189	188	235	328

Table 3: Cross-country correlations

The table presents cross-country correlations of several variables between other countries and the U.S.

Correlation with U.S.

	Countries	Quarterly equity excess total returns	Quarterly bank equity excess total returns	Equity Crash indicator	Bank Equity Crash indicator	Δ (Bank credit / GDP)	D/P	Inflation	Term Spread	Book / Market	I/K	C/W
	US	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Commonwealth	Australia	0.43	0.33	0.39	0.33	0.40	0.57	0.70	0.37	0.88	0.68	0.13
	Canada	0.79	0.56	0.42	0.21	0.24	0.92	0.87	0.49	0.82	0.72	0.77
	UK	0.50	0.42	0.31	0.32	0.25	0.63	0.66	0.41	0.94	0.82	0.68
W. Europe	Austria	0.39	0.42	0.31	0.60	-0.09	0.68	0.23	-0.22	0.40	0.61	
	Belgium	0.61	0.47	0.66	0.52	-0.03	0.59	0.48	0.26	0.14	0.51	
	France	0.41	0.41	0.39	0.38	0.35	0.33	0.52	0.19	0.80	0.54	0.61
	Germany	0.37	0.34	0.43	0.41	0.49	0.36	0.61	0.15	0.82	0.53	0.41
	Ireland	0.57	0.49	0.61	0.42	0.56	0.88	0.62	0.19	0.54	0.64	
	Netherlands	0.63	0.42	0.48	0.34	0.00	0.90	0.62	0.20	0.91	0.68	
	Switzerland	0.52	0.49	0.24	0.51	0.22	0.71	0.68	0.09	0.85	0.39	
Scandinavia	Denmark	0.37	0.30	0.35	0.39	0.48	0.68	0.50	-0.12	0.56	0.69	
	Norway	0.47	0.28	0.56	0.40	0.47	0.59	0.43	0.26	0.00	0.32	
	Sweden	0.48	0.33	0.24	0.26	0.21	0.67	0.62	0.00	0.73	0.69	
S. Europe	Italy	0.38	0.34	0.29	0.28	0.25	0.32	0.43	-0.03	0.50	0.53	0.41
	Portugal	0.41	0.44	0.53	0.47	0.11	0.10	0.56	0.03	0.16	0.65	
	Spain	0.55	0.44	0.47	0.34	0.24	0.42	0.53	0.23	0.69	0.23	
Asia	Japan	0.28	0.14	0.12	0.23	0.46	0.61	0.33	0.19	0.05	-0.26	0.19
	Korea	0.38	0.22	0.26	-0.06	-0.25	0.62	0.47	-0.14	-0.30		
	Singapore	0.61	0.29	0.52	0.35	-0.17	0.27	0.59	0.40	0.48		
	Average	0.48	0.37	0.40	0.35	0.22	0.57	0.55	0.15	0.52	0.53	0.46

Table 4: Predictive probit regressions using crash indicators.

This table reports estimates from the probit regression model specified in equation (1) for the bank equity index (Panel A) and the market index (Panel B) in subsequent 1, 4, and 8 quarters. The dependent variable is the crash indicator ($Y = 1_{crash}$), which takes on a value of 1 if there is a future equity crash, as defined in Section 4, in the next K quarters (K = 1, 4, and 8) and 0 otherwise, which is regressed on the three-year change in bank credit to GDP (standardized within each country) and several subsets of control variables known to predict the equity premium. All reported estimates are marginal effects, so that a coefficient of 0.024 means that a one-standard deviation increase in Δ(bank credit / GDP) predicts a 2.4 percentage point increase in the likelihood of a future crash. This table also reports estimates from equation (1) with ($Y = 1_{boom}$), a symmetrically defined right tail event, and the difference in the marginal effects between the two probit regressions (the probability of a crash minus probability of a boom). Standard errors are dually-clustered on country and time. The *, **, and *** marked on the regression coefficients denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Panel A: Crash in bank index

		1			4			8		
		Crash	Boom	Difference	Crash	Boom	Difference	Crash	Boom	Difference
No controls	Δ (bank credit / GDP)	0.022**	-0.007	0.029*	0.043**	-0.024	0.067*	0.051***	-0.033	0.084**
		(2.71)	(-0.95)	(2.53)	(2.79)	(-1.45)	(2.34)	(3.90)	(-1.81)	(2.85)
	N	4172	4172	4172	1057	1057	1057	540	540	540
With two controls	Δ (bank credit / GDP)	0.019*	-0.008	0.027*	0.038*	-0.024	0.062*	0.043**	-0.030	0.072*
		(2.17)	(-1.09)	(2.20)	(2.56)	(-1.43)	(2.21)	(2.94)	(-1.48)	(2.19)
	log(D/P)	-0.060*	0.014	-0.074*	-0.111**	0.052	-0.163*	-0.092**	0.024	-0.116
		(-2.24)	(0.65)	(-1.99)	(-2.64)	(1.33)	(-2.57)	(-2.61)	(0.72)	(-1.79)
	N	3866	3866	3866	996	996	996	507	507	507
With all controls	Δ (bank credit / GDP)	0.015	-0.011	0.026*	0.036**	-0.029	0.065*	0.039*	-0.029	0.068
		(1.86)	(-1.45)	(2.15)	(2.59)	(-1.76)	(2.46)	(2.22)	(-1.36)	(1.83)
	log(D/P)	-0.042	-0.003	-0.039	-0.096	0.035	-0.131	-0.072	0.017	-0.089
		(-1.21)	(-0.14)	(-0.94)	(-1.80)	(0.81)	(-1.65)	(-1.64)	(0.41)	(-1.11)
	inflation	0.011	-0.076	0.087	-0.146	0.272	-0.418	-0.266	0.487	-0.752
		(0.05)	(-0.31)	(0.31)	(-0.42)	(0.68)	(-0.72)	(-0.63)	(1.01)	(-0.88)
	term spread	0.200	0.128	0.072	-0.358	-0.222	-0.136	-0.645	0.151	-0.796
		(0.31)	(0.24)	(0.07)	(-0.31)	(-0.21)	(-0.06)	(-0.50)	(0.13)	(-0.33)
	log(book / market)	-0.015	0.052	-0.067	0.008	0.046	-0.038	-0.023	-0.007	-0.015
		(-0.29)	(1.77)	(-1.17)	(0.14)	(0.77)	(-0.36)	(-0.34)	(-0.11)	(-0.12)
	log(I/K)	0.092	0.096	-0.004	0.088	0.086	0.001	0.076	0.007	0.069
		(1.14)	(1.37)	(-0.04)	(0.85)	(0.58)	(0.01)	(0.67)	(0.05)	(0.31)
	N	3659	3659	3659	943	943	943	479	479	479

Panel B: Crash in equity index

		1			4			8		
		Crash	Boom	Difference	Crash	Boom	Difference	Crash	Boom	Difference
No controls	Δ (bank credit / GDP)	0.015**	-0.002	0.017*	0.032*	-0.012	0.045	0.043**	-0.020	0.063*
		(2.76)	(-0.30)	(2.23)	(2.35)	(-0.85)	(1.82)	(3.11)	(-1.48)	(2.57)
	N	4370	4370	4370	1128	1128	1128	574	574	574
With two controls	Δ (bank credit / GDP)	0.015**	-0.001	0.016*	0.028*	-0.010	0.038	0.034*	-0.012	0.046
		(2.66)	(-0.18)	(1.98)	(2.23)	(-0.70)	(1.64)	(2.45)	(-0.79)	(1.71)
	log(D/P)	-0.062**	0.020	-0.083***	-0.154***	0.060	-0.214***	-0.194***	0.088**	-0.282***
		(-3.20)	(1.44)	(-3.56)	(-3.72)	(1.81)	(-3.55)	(-4.20)	(2.60)	(-3.88)
	N	4322	4322	4322	1118	1118	1118	564	564	564
With all controls	Δ (bank credit / GDP)	0.012*	-0.001	0.013	0.025*	-0.006	0.031	0.029	-0.006	0.035
		(2.03)	(-0.25)	(1.66)	(2.06)	(-0.44)	(1.45)	(1.78)	(-0.37)	(1.11)
	log(D/P)	-0.037	0.013	-0.051*	-0.132***	0.034	-0.166**	-0.156**	0.054	-0.209**
		(-1.81)	(0.82)	(-2.26)	(-3.35)	(0.90)	(-2.71)	(-3.22)	(1.37)	(-2.61)
	inflation	0.219	0.003	0.216	0.187	0.089	0.098	0.316	0.071	0.244
		(1.25)	(0.02)	(1.10)	(0.58)	(0.31)	(0.25)	(1.13)	(0.21)	(0.48)
	term spread	-0.613*	0.440	-1.053*	-2.015**	0.256	-2.272	-1.673*	-0.093	-1.581
		(-2.32)	(1.29)	(-2.40)	(-2.92)	(0.36)	(-1.94)	(-2.31)	(-0.12)	(-1.17)
	log(book / market)	-0.041	0.042	-0.083*	-0.047	0.086	-0.133	-0.074	0.045	-0.118
		(-1.39)	(1.82)	(-2.24)	(-1.23)	(1.53)	(-1.61)	(-1.87)	(0.96)	(-1.47)
	log(I/K)	0.068	0.030	0.038	0.075	-0.007	0.082	0.152	-0.093	0.245
		(0.96)	(0.57)	(0.42)	(0.57)	(-0.05)	(0.36)	(1.27)	(-0.72)	(1.11)
	N	3995	3995	3995	1035	1035	1035	522	522	522

Table 5: Equity premium predictability regressions

This table reports estimates from the panel regression with fixed effects model specified in equation (2) for the bank equity index (Panel A) and the market equity index (Panel B), in subsequent 1, 4, and 8 quarters. Returns are non-overlapping, and standard errors are dually-clustered on country and time. A coefficient of -0.011 means that a one-standard deviation increase in Δ(bank credit / GDP) predicts a 1.1 percentage point decrease in subsequent returns. The dependent variable is log excess total returns, which is regressed on the three-year change in bank credit to GDP (standardized within each country) and several subsets of control variables thought to predict the equity premium. The *, **, and *** marked on the regression coefficients denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

Panel A: Bank index

	1 quarter horizon				4 quarter horizon				8 quarter horizon			
Δ (bank credit / GDP)	-0.010*	-0.009*	-0.009*	-0.008*	-0.049*	-0.045*	-0.047*	-0.045*	-0.086**	-0.080**	-0.079**	-0.078**
	(-2.554)	(-2.151)	(-2.298)	(-2.201)	(-2.161)	(-2.014)	(-2.435)	(-2.386)	(-3.268)	(-3.010)	(-2.833)	(-2.808)
log(D/P)		0.015	0.010	0.009		0.093	0.085	0.076		0.095	0.067	0.061
		(1.166)	(0.669)	(0.602)		(1.921)	(1.346)	(1.209)		(1.584)	(0.853)	(0.794)
inflation			-0.112	-0.115			-0.289	-0.320			-0.119	-0.147
			(-0.987)	(-1.025)			(-0.666)	(-0.737)			(-0.177)	(-0.211)
term spread			0.447	0.404			1.533	1.320			1.787	1.643
			(1.214)	(1.091)			(1.037)	(0.885)			(0.935)	(0.865)
log(book / market)			0.020	0.016			0.035	0.019			0.095	0.082
			(0.728)	(0.574)			(0.322)	(0.175)			(0.600)	(0.516)
log(I/K)			0.016	0.017			0.081	0.082			0.036	0.039
			(0.317)	(0.342)			(0.438)	(0.445)			(0.180)	(0.198)
consumption / wealth				0.259**				0.985**				0.829
				(3.192)				(2.645)				(1.342)
R^2	0.008	0.010	0.015	0.017	0.033	0.047	0.053	0.059	0.073	0.080	0.082	0.084
Adj. R^2	0.003	0.004	0.008	0.011	0.014	0.027	0.029	0.034	0.036	0.039	0.034	0.035
N	4150	3849	3643	3643	1037	979	927	927	521	493	467	467

Panel B: Equity index

	1 quarter horizon				4 quarter horizon				8 quarter horizon			
Δ (bank credit / GDP)	-0.007**	-0.007**	-0.006*	-0.006*	-0.033*	-0.032*	-0.032*	-0.031*	-0.053**	-0.050**	-0.049*	-0.047*
	(-2.968)	(-2.769)	(-2.159)	(-2.059)	(-2.309)	(-2.349)	(-2.362)	(-2.302)	(-2.691)	(-2.661)	(-2.235)	(-2.173)
log(D/P)		0.014	0.009	0.007		0.080*	0.068*	0.059		0.123*	0.082	0.071
		(1.852)	(1.146)	(1.020)		(2.547)	(2.144)	(1.853)		(2.241)	(1.553)	(1.391)
inflation			-0.182*	-0.185*			-0.520	-0.536			-0.867*	-0.887*
			(-2.175)	(-2.210)			(-1.453)	(-1.506)			(-2.001)	(-2.029)
term spread			0.454*	0.414*			1.569*	1.360			1.637	1.423
			(2.532)	(2.291)			(1.981)	(1.719)			(1.364)	(1.202)
log(book / market)			0.031*	0.027*			0.078	0.062			0.168*	0.147*
			(2.271)	(1.988)			(1.638)	(1.322)			(2.357)	(2.098)
log(I/K)			-0.006	-0.005			-0.013	-0.012			-0.051	-0.046
			(-0.200)	(-0.162)			(-0.112)	(-0.104)			(-0.439)	(-0.411)
consumption / wealth				0.255***				1.006***				1.291**
				(4.056)				(3.655)				(3.229)
R^2	0.006	0.010	0.028	0.032	0.026	0.055	0.085	0.098	0.045	0.087	0.129	0.141
Adj. R^2	0.002	0.005	0.022	0.026	0.007	0.036	0.063	0.075	0.008	0.050	0.086	0.098
N	4324	4276	3950	3950	1073	1070	987	987	538	536	494	494

Table 6: Negative predicted returns subsequent to large credit expansion

Panel A (for the bank equity index) and Panel B (for the equity market index) report average log excess returns in 4-, 8- and 12-quarters subsequent to large credit expansions (when credit expansion exceeds a given percentile threshold) and subsequent to large credit contractions (when credit expansion falls below a given percentile threshold). Average returns conditional on the thresholds (and corresponding t-statistics and adjusted R^2) are computed using regression models (3) and (4) with non-overlapping 4-, 8-, and 12-quarter ahead returns. T-statistics are computed using dually-clustered standard errors.

Panel A: Bank index

Threshold in percentiles:			**<2%**	**<5%**	**<10%**	**<25%**	**<50%**	**>50%**	**>75%**	**>90%**	**>95%**	**>98%**
4-quarter ahead returns	no controls	E[r]	0.006	0.144	0.098	0.080	0.060	-0.005	-0.033	-0.077	-0.115	-0.206
		(t-stat)	(0.12)	(3.04)	(2.452)	(3.513)	(2.708)	(-0.137)	(-0.67)	(-0.973)	(-0.875)	(-1.684)
		Adj. R^2	0.01	0.018	0.016	0.019	0.019	0.02	0.024	0.024	0.024	0.023
		N	24	56	110	252	507	555	301	124	68	25
	with controls	E[r]	0.031	0.141	0.098	0.079	0.060	-0.005	-0.035	-0.078	-0.114	-0.208
		(t-stat)	(0.523)	(2.469)	(2.11)	(3.106)	(2.509)	(-0.143)	(-0.706)	(-0.972)	(-0.901)	(-1.774)
		Adj. R^2	0.029	0.039	0.034	0.045	0.048	0.052	0.055	0.051	0.059	0.039
		N	22	50	99	228	442	509	289	120	68	25
8-quarter ahead returns	no controls	E[r]	0.059	0.249	0.138	0.151	0.118	-0.006	-0.050	-0.116	-0.258	-0.208
		(t-stat)	(0.655)	(2.806)	(3.285)	(3.996)	(3.013)	(-0.132)	(-0.721)	(-1.691)	(-2.301)	(-2.053)
		Adj. R^2	0.049	0.072	0.064	0.082	0.101	0.107	0.102	0.126	0.109	0.077
		N	10	25	56	125	251	291	163	64	28	14
	with controls	E[r]	0.080	0.267	0.146	0.150	0.114	-0.001	-0.054	-0.131	-0.253	-0.209
		(t-stat)	(0.788)	(2.58)	(3.185)	(3.739)	(2.768)	(-0.023)	(-0.775)	(-1.806)	(-2.348)	(-2.347)
		Adj. R^2	0.034	0.039	0.038	0.041	0.041	0.04	0.044	0.045	0.046	0.046
		N	9	23	52	115	220	267	155	61	28	14
12-quarter ahead returns	no controls	E[r]	0.152	0.392	0.247	0.239	0.198	-0.053	-0.111	-0.315	-0.437	-0.495
		(t-stat)	(1.814)	(4.946)	(3.381)	(4.423)	(3.653)	(-0.711)	(-1.319)	(-3.34)	(-2.544)	(-1.69)
		Adj. R^2	0.047	0.055	0.05	0.059	0.059	0.058	0.066	0.066	0.071	0.055
		N	10	21	43	88	174	188	108	43	21	8
	with controls	E[r]	0.202	0.378	0.263	0.245	0.205	-0.051	-0.114	-0.337	-0.427	-0.492
		(t-stat)	(1.924)	(3.672)	(3.336)	(4.329)	(3.468)	(-0.667)	(-1.29)	(-3.872)	(-2.793)	(-1.869)
		Adj. R^2	0.105	0.116	0.114	0.126	0.14	0.139	0.139	0.173	0.156	0.125
		N	9	18	38	80	153	171	101	40	21	8

Panel B: Equity index

	Threshold in percentiles:		<2%	<5%	<10%	<25%	<50%	>50%	>75%	>90%	>95%	>98%
4-quarter ahead returns	no controls	E[r]	0.070	0.125	0.094	0.079	0.063	0.015	0.001	-0.029	-0.049	-0.087
		(t-stat)	(1.956)	(5.234)	(4.185)	(3.968)	(3.207)	(0.544)	(0.035)	(-0.572)	(-0.631)	(-1.03)
		Adj. R^2	0.005	0.011	0.01	0.012	0.012	0.016	0.015	0.016	0.015	0.012
		N	28	58	113	260	508	594	335	142	80	30
	with controls	E[r]	0.098	0.131	0.103	0.080	0.064	0.011	-0.005	-0.037	-0.054	-0.092
		(t-stat)	(2.469)	(4.221)	(4.247)	(3.966)	(3.402)	(0.42)	(-0.137)	(-0.707)	(-0.717)	(-1.157)
		Adj. R^2	0.011	0.018	0.012	0.024	0.019	0.026	0.027	0.033	0.037	0.018
		N	25	52	101	238	461	551	314	133	76	28
8-quarter ahead returns	no controls	E[r]	0.094	0.204	0.117	0.146	0.111	0.043	0.019	-0.040	-0.118	-0.070
		(t-stat)	(1.75)	(3.563)	(3.524)	(4.726)	(3.346)	(1.154)	(0.41)	(-0.71)	(-1.603)	(-0.678)
		Adj. R^2	0.018	0.024	0.024	0.036	0.051	0.073	0.048	0.06	0.056	0.039
		N	13	26	58	131	254	309	180	75	35	17
	with controls	E[r]	0.117	0.214	0.128	0.143	0.110	0.037	0.007	-0.063	-0.128	-0.109
		(t-stat)	(1.651)	(2.768)	(3.841)	(4.074)	(3.311)	(0.959)	(0.158)	(-1.13)	(-2)	(-1.249)
		Adj. R^2	0.064	0.069	0.068	0.069	0.069	0.073	0.071	0.074	0.072	0.069
		N	11	23	52	119	230	286	168	69	32	15
12-quarter ahead returns	no controls	E[r]	0.217	0.255	0.209	0.211	0.197	0.024	0.010	-0.096	-0.184	-0.279
		(t-stat)	(3.075)	(3.529)	(2.68)	(3.734)	(3.808)	(0.42)	(0.171)	(-1.502)	(-2.241)	(-2.076)
		Adj. R^2	0.094	0.099	0.095	0.103	0.099	0.104	0.106	0.116	0.113	0.101
		N	12	22	44	92	172	201	117	49	24	8
	with controls	E[r]	0.251	0.255	0.250	0.225	0.200	0.007	-0.002	-0.113	-0.184	-0.290
		(t-stat)	(2.178)	(2.894)	(3.391)	(4.004)	(4.112)	(0.14)	(-0.036)	(-2.206)	(-2.482)	(-3.161)
		Adj. R^2	0.172	0.176	0.18	0.186	0.198	0.216	0.187	0.205	0.201	0.184
		N	11	20	39	83	158	184	108	44	23	8

Table 7: Quantile regressions

The first block of rows reports β_{mean}, the coefficient from estimating linear regression model (2), β_{median}, the coefficient from a median regression (50th quantile regression), and the difference $(\beta_{median} - \beta_{mean})$, together with their associated t-statistics and p-value. The dependent variable is subsequent non-overlapping 4-, 8-, or 12-quarter ahead returns of the bank equity index or the market equity index, which is regressed on credit expansion and other controls. The second block of rows reports the coefficients and t-statistics for the three quantile regressions, $\beta_{q=5}$, $\beta_{q=50}$, $\beta_{q=95}$, followed by the conditional negative skewness coefficient $\beta_{negative\,skew} = (\beta_{q=50} - \beta_{q=5}) - (\beta_{q=95} - \beta_{q=50})$. Standard errors are dually-clustered on country and time. The *, **, and *** marked on the regression coefficients denote statistical significance at 5%, 1%, and 0.1% levels, respectively.

		Bank index			Equity index		
Explanatory variables:		1	4	8	1	4	8
Δ (bank credit / GDP)	Mean	-.01*	-.049*	-.086**	-.007**	-.033*	-.053**
	(t stat)	(-2.55)	(-2.16)	(-3.27)	(-2.97)	(-2.31)	(-2.69)
	Median	-.007***	-.03**	-.066***	-.005**	-.024*	-.052
	(t stat)	(-4.45)	(-2.88)	(-4.06)	(-2.79)	(-2.29)	(-2.31)
	Difference	.003*	.019	.02	.002*	.009*	.001
	(p-value)	(.049)	(.059)	(.195)	(.044)	(.039)	(.928)
	Q5	-.042	-.119***	-.093***	-.034***	-.077***	-.106***
	(t stat)	(-4.07)	(-4.02)	(-3.73)	(-3.31)	(-3.97)	(-4)
	Q50 (median)	-.007***	-.03**	-.066***	-.005**	-.024*	-.052
	(t stat)	(-4.45)	(-2.88)	(-4.06)	(-2.79)	(-2.29)	(-2.31)
	Q95	.006	-.028	-.14***	-.003	-.028	-.027
	(t stat)	(.65)	(-1.2)	(-5.06)	(-.55)	(-1.55)	(-1.49)
	Negative skew	.023**	.086**	.101*	.026*	.056	.029
	(t stat)	(2.8)	(2.62)	(2.11)	(2.15)	(1.5)	(1.01)
	N	4150	1037	521	4324	1073	538
Δ (bank credit / GDP), with D/P, inflation, book to market, term spread, and i/k as controls (coefficients on controls not shown)	Mean	-.009*	-.047*	-.079**	-.006*	-.032*	-.049*
	(t stat)	(-2.3)	(-2.43)	(-2.83)	(-2.16)	(-2.36)	(-2.23)
	Median	-.006***	-.037**	-.034	-.005**	-.03*	-.055***
	(t stat)	(-8.16)	(-2.91)	(-1.2)	(-2.67)	(-2.51)	(-3.72)
	Difference	.004*	.015	.043*	.002	.004	-.003
	(p-value)	(.013)	(.177)	(.039)	(.26)	(.523)	(.690)
	Q5	-.034*	-.085*	-.03	-.024**	-.045	-.018
	(t stat)	(-2.21)	(-2.22)	(-.810)	(-2.66)	(-1.82)	(-.700)
	Q50 (median)	-.006***	-.037**	-.034	-.005**	-.03*	-.055***
	(t stat)	(-8.16)	(-2.91)	(-1.2)	(-2.67)	(-2.51)	(-3.72)
	Q95	-.005	-.003	-.073*	-.002	-.02	-.019
	(t stat)	(-.54)	(-.14)	(-2.06)	(-.66)	(-.76)	(-.44)
	Negative skew	.028***	.015	.035	.016	.005	-.073*
	(t stat)	(4.52)	(.33)	(.33)	(1.91)	(.13)	(-2.05)
	N	3643	927	467	3950	987	494

Table 8: Robustness in geographical and time subsamples

This table demonstrates that the estimates reported in Tables 4 and 5 for the mean and probit regression models are robust within various geographical and time subsets. Panel A analyzes various geographical subsets, while Panel B analyzes various time subsets: 1920-2012 (the full sample), 1920-2005 (excluding the recent crisis), and 1950-2005 (excluding both the recent crisis and the Great Depression). The table reports estimates of mean and probit coefficients (using the same methodology as in Tables 4 and 5) of non-overlapping 4-quarter-ahead log excess returns of either the bank equity index or the equity market index regressed on credit expansion with or without the five standard controls. Coefficients reported in this table are always on Δ(bank credit / GDP); coefficients on control variables are omitted.

Panel A: Robustness by geographical region (4 quarter non-overlapping forecast horizon)

			All	Largest Eight	U.S.	English speaking	Western Europe	Southern Europe	Scandinavia
Bank Index	**probit - without controls**	Δ (bank credit / GDP)	0.043**	0.041**	0.052	0.048**	0.053**	0.098***	0.057*
		(t-stat)	(2.79)	(2.58)	(1.56)	(2.59)	(2.91)	(3.98)	(1.98)
		N	1057	524	76	286	741	100	185
	probit - with controls	Δ (bank credit / GDP)	0.032*	0.035*	0.040	0.046**	0.043*	0.048*	0.046
		(t-stat)	(2.35)	(2.53)	(0.82)	(2.93)	(2.39)	(2.21)	(1.88)
		N	1128	587	84	294	752	135	166
	mean - without controls	Δ (bank credit / GDP)	-0.049*	-0.033	-0.038	-0.023	-0.065*	-0.084*	-0.066*
		(t-stat)	(-2.161)	(-1.944)	(-1.489)	(-1.465)	(-2.405)	(-2.151)	(-2.195)
		N	1037	518	76	283	725	97	179
	mean - with controls	Δ (bank credit / GDP)	-0.047*	-0.031	-0.004	-0.008	-0.061**	-0.149*	-0.069***
		(t-stat)	(-2.435)	(-1.610)	(-0.126)	(-0.569)	(-2.719)	(-2.481)	(-3.673)
		N	927	517	76	282	667	97	155
Equity Index	**probit - without controls**	Δ (bank credit / GDP)	0.036**	0.035**	0.017	0.033*	0.041*	0.134*	0.054
		(t-stat)	(2.59)	(2.68)	(0.45)	(2.57)	(2.39)	(2.57)	(2.00)
		N	943	522	76	284	681	100	159
	probit - with controls	Δ (bank credit / GDP)	0.025*	0.029*	0.010	0.038*	0.027	0.057	0.019
		(t-stat)	(2.06)	(2.24)	(0.26)	(2.28)	(1.80)	(1.69)	(0.69)
		N	1035	584	84	292	744	133	164
	mean - without controls	Δ (bank credit / GDP)	-0.033*	-0.027*	-0.015	-0.019	-0.047**	-0.038	-0.029
		(t-stat)	(-2.309)	(-2.049)	(-0.715)	(-1.356)	(-2.719)	(-1.569)	(-0.928)
		N	1073	580	84	290	702	132	144
	mean - with controls	Δ (bank credit / GDP)	-0.032*	-0.025	0.003	-0.004	-0.047**	-0.072**	-0.008
		(t-stat)	(-2.362)	(-1.615)	(0.129)	(-0.270)	(-3.262)	(-2.914)	(-0.275)
		N	987	579	84	290	699	130	144

Panel B: Robustness by time period (4 quarter non-overlapping forecast horizon)

			1920-2012	1920-2005	1950-2005
Bank Index	**probit - without controls**	Δ (bank credit / GDP)	0.043**	0.041***	0.041**
		(t-stat)	(2.79)	(3.30)	(3.00)
		N	1057	943	842
	probit - with controls	Δ (bank credit / GDP)	0.032*	0.026*	0.026*
		(t-stat)	(2.35)	(2.32)	(2.04)
		N	1128	1014	855
	mean - without controls	Δ (bank credit / GDP)	-0.049*	-0.038**	-0.042**
		(t-stat)	(-2.161)	(-3.182)	(-3.049)
		N	1037	924	828
	mean - with controls	Δ (bank credit / GDP)	-0.047*	-0.035**	-0.045**
		(t-stat)	(-2.435)	(-2.691)	(-2.916)
		N	927	825	730
Equity Index	**probit - without controls**	Δ (bank credit / GDP)	0.036**	0.032**	0.033*
		(t-stat)	(2.59)	(2.78)	(2.47)
		N	943	841	742
	probit - with controls	Δ (bank credit / GDP)	0.025*	0.018	0.022
		(t-stat)	(2.06)	(1.57)	(1.38)
		N	1035	933	777
	mean - without controls	Δ (bank credit / GDP)	-0.033*	-0.028*	-0.031*
		(t-stat)	(-2.309)	(-2.516)	(-2.537)
		N	1073	959	845
	mean - with controls	Δ (bank credit / GDP)	-0.032*	-0.028*	-0.037*
		(t-stat)	(-2.362)	(-2.141)	(-2.431)
		N	987	885	772

Table 9: Test for possible small-sample bias

This table tests for the possibility of small-sample bias using the methodology of Campbell and Yogo (2006). Equations (8) and (9) are estimated, and parameter values corresponding to the sample size (N), persistence of bank credit expansion (ρ), and the correlation of its innovations with returns ($\delta = corr(u_{i,t}, \epsilon_{i,t})$) are reported. Panel A corresponds to bank equity index returns, and Panel B corresponds to equity market index returns.

Panel A: Bank stock returns

Quarters ahead	Controls?	ρ	δ	N	N * (ρ - 1)
1	N	0.964	0.031	4117	-148.21
1	Y	0.964	0.038	3614	-130.10
4	N	0.796	0.037	1018	-207.67
4	Y	0.796	0.058	888	-181.15
8	N	0.497	0.013	497	-249.99
8	Y	0.497	0.022	434	-218.30

Panel B: Index returns

Quarters ahead	Controls?	ρ	δ	N	N * (ρ - 1)
1	N	0.964	0.023	4285	-154.26
1	Y	0.964	0.037	3913	-140.87
4	N	0.796	0.024	1046	-213.38
4	Y	0.796	0.049	966	-197.06
8	N	0.497	-0.007	512	-257.54
8	Y	0.497	0.010	472	-237.42

Table 10: Robustness of negative returns: Clustering observations by historical episodes

This table groups concurrent observations of large credit expansions (exceeding the 95th percentile threshold) and large credit contractions (below the 5th percentile threshold) across countries into distinct historical episodes. A list of all large credit expansions and large credit contractions, grouped together by historical episode, can be found in Table S1 in the online appendix. Returns from the resulting historical episodes in the sample are averaged together, taking each such historical episode as a single, independent observation.

Panel A: Returns subsequent to large credit expansions (observations grouped by episodes)

	returns on bank equity			returns on market index		
Quarters ahead:	4	8	12	4	8	12
Average over episodes:	-0.121	-0.228	-0.176	-0.036	-0.085	-0.033
T-STAT	-2.119	-2.821	-2.059	-0.723	-1.586	-0.370
S.E.	0.057	0.081	0.086	0.050	0.053	0.089
N (episodes)	18	18	18	20	20	20

Panel B: Returns subsequent to large credit contractions (observations grouped by episodes)

	returns on bank equity			returns on market index		
Quarters ahead:	4	8	12	4	8	12
Average over episodes:	0.178	0.294	0.435	0.136	0.244	0.321
T-STAT	3.898	3.268	3.608	3.890	4.588	4.264
S.E.	0.046	0.090	0.121	0.035	0.053	0.075
N (episodes)	21	21	21	21	21	21

Table 11: Robustness of negative predicted returns: out-of-sample predictability

This table is similar to Table 6, but the percentile threshold for each quarter is calculated only using previous information for that country (given at least 5 years of past data for that country). For example, for credit growth to be above the ¿95% threshold, credit growth must be greater than 95% of all previous observations for that country.

Panel A: Bank index

	Threshold in percentiles:		**<2%**	**<5%**	**<10%**	**<25%**	**<50%**	**>50%**	**>75%**	**>90%**	**>95%**	**>98%**
4-quarter ahead returns	no controls	E[r]	7.3%	9.3%	9.2%	6.7%	6.9%	-0.7%	-2.4%	-4.4%	-5.3%	-8.2%
		(t-stat)	(1.566)	(2.428)	(2.93)	(3.078)	(3.175)	(-0.202)	(-0.528)	(-0.705)	(-0.7)	(-1.245)
		Adj. R^2	0.012	0.015	0.016	0.015	0.022	0.023	0.022	0.02	0.018	0.018
		N	69	98	141	278	462	595	367	194	122	73
	with controls	E[r]	6.9%	9.5%	9.8%	6.9%	6.7%	-0.6%	-2.7%	-4.5%	-4.8%	-7.8%
		(t-stat)	(0.896)	(1.738)	(2.353)	(2.664)	(2.766)	(-0.167)	(-0.567)	(-0.707)	(-0.634)	(-1.183)
		Adj. R^2	0.034	0.033	0.036	0.04	0.048	0.053	0.053	0.051	0.051	0.043
		N	43	72	111	240	401	545	342	184	119	72
8-quarter ahead returns	no controls	E[r]	16.6%	13.8%	14.6%	13.2%	12.5%	-0.8%	-3.7%	-8.8%	-14.2%	-14.0%
		(t-stat)	(3.717)	(1.926)	(4.037)	(3.531)	(3.39)	(-0.157)	(-0.566)	(-1.277)	(-1.747)	(-1.814)
		Adj. R^2	0.058	0.065	0.065	0.078	0.097	0.103	0.095	0.101	0.09	0.097
		N	34	47	70	145	235	304	196	94	54	35
	with controls	E[r]	19.3%	14.5%	15.6%	13.3%	12.2%	-0.3%	-3.8%	-8.3%	-13.5%	-13.1%
		(t-stat)	(2.517)	(1.474)	(3.511)	(3.08)	(3.104)	(-0.065)	(-0.563)	(-1.187)	(-1.778)	(-1.859)
		Adj. R^2	0.034	0.036	0.038	0.039	0.043	0.041	0.044	0.042	0.039	0.04
		N	21	34	57	126	204	280	183	89	53	35
12-quarter ahead returns	no controls	E[r]	25.2%	27.2%	23.8%	22.7%	21.0%	-4.3%	-8.3%	-18.7%	-23.7%	-37.3%
		(t-stat)	(6.527)	(4.789)	(3.846)	(4.669)	(4.111)	(-0.565)	(-1.016)	(-1.871)	(-2.145)	(-2.851)
		Adj. R^2	0.048	0.048	0.05	0.054	0.058	0.059	0.063	0.06	0.062	0.056
		N	26	36	50	95	157	203	131	66	40	23
	with controls	E[r]	28.4%	30.7%	26.5%	23.2%	21.0%	-3.9%	-8.7%	-18.7%	-21.9%	-35.7%
		(t-stat)	(2.505)	(3.068)	(3.337)	(3.834)	(3.552)	(-0.471)	(-0.951)	(-1.741)	(-2.212)	(-3.103)
		Adj. R^2	0.108	0.112	0.115	0.12	0.131	0.131	0.135	0.136	0.127	0.138
		N	16	26	39	83	139	183	119	62	39	23

Panel B: Equity index

Threshold in percentiles:			<2%	<5%	<10%	<25%	<50%	>50%	>75%	>90%	>95%	>98%
4-quarter ahead returns	no controls	E[r]	8.2%	8.4%	8.8%	6.5%	6.8%	1.4%	0.6%	-1.0%	-0.9%	-2.7%
		(t-stat)	(2.678)	(3.905)	(4.625)	(3.643)	(3.372)	(0.532)	(0.197)	(-0.237)	(-0.192)	(-0.626)
		Adj. R^2	0.006	0.007	0.01	0.008	0.013	0.017	0.014	0.014	0.01	0.011
		N	59	91	142	283	465	631	399	220	143	90
	with controls	E[r]	8.4%	8.3%	9.2%	7.0%	6.9%	1.1%	-0.1%	-1.8%	-1.2%	-3.0%
		(t-stat)	(2.124)	(3.237)	(3.762)	(4.163)	(3.469)	(0.43)	(-0.023)	(-0.443)	(-0.269)	(-0.706)
		Adj. R^2	0.011	0.011	0.014	0.019	0.018	0.026	0.024	0.032	0.028	0.023
		N	46	78	123	254	420	588	375	208	140	88
8-quarter ahead returns	no controls	E[r]	10.8%	10.7%	12.9%	12.9%	11.5%	4.3%	2.8%	-1.9%	-3.7%	-4.1%
		(t-stat)	(2.299)	(2.233)	(4.446)	(3.937)	(3.394)	(1.1)	(0.624)	(-0.307)	(-0.556)	(-0.609)
		Adj. R^2	0.016	0.02	0.017	0.03	0.051	0.069	0.052	0.051	0.057	0.063
		N	27	44	72	148	234	325	212	108	67	45
	with controls	E[r]	11.3%	10.4%	12.9%	12.8%	11.4%	3.6%	1.8%	-3.2%	-5.6%	-5.1%
		(t-stat)	(1.718)	(1.865)	(3.824)	(3.47)	(3.318)	(0.926)	(0.404)	(-0.539)	(-0.993)	(-0.929)
		Adj. R^2	0.063	0.064	0.066	0.066	0.069	0.074	0.072	0.072	0.067	0.068
		N	21	38	64	132	211	302	199	101	65	44
12-quarter ahead returns	no controls	E[r]	12.8%	19.7%	15.2%	20.0%	20.7%	3.0%	1.2%	-3.9%	-10.2%	-19.7%
		(t-stat)	(1.599)	(3.273)	(2.777)	(4.117)	(4.564)	(0.535)	(0.214)	(-0.581)	(-1.186)	(-2.017)
		Adj. R^2	0.094	0.094	0.094	0.099	0.099	0.105	0.103	0.111	0.11	0.105
		N	22	33	49	96	156	214	138	74	46	27
	with controls	E[r]	16.1%	23.2%	18.0%	20.6%	20.5%	2.0%	0.5%	-4.6%	-10.9%	-20.2%
		(t-stat)	(1.564)	(2.968)	(3.4)	(3.956)	(4.243)	(0.375)	(0.082)	(-0.639)	(-1.418)	(-2.269)
		Adj. R^2	0.173	0.174	0.173	0.181	0.191	0.204	0.194	0.189	0.196	0.206
		N	17	28	42	87	143	198	129	69	45	27

Table A1 - Data and sample length

This table shows the sample length for each variable by reporting the first year of data for each variable within each country.

Country	Bank credit / gdp	equity return	bank equity return	D / P	bank D / P	first year of banking crisis	exchange rate	inflation	three mo tbill yield	govt 10yr yield	longterm corpbond yield	E / P	stock daily volatility	book / market	i / k	stock market turnover	currentaccount / gdp	gross external debt / gdp	c / w	real gdp growth	central govt debt / gdp	totalcredit / gdp	Hhdebt / totalassets	totalcreditHH / gdp	growth in housing assets	NFC equity / assets	totalcreditNFC / gdp	BANK equity / assets
Australia	1920	1920	1920	1920	1924	1920	1920	1920	1928	1920	1983	1973	1958	1980	1960	1973	1960	1970	1978	1920	1920	1954	1978	1977	1961	1981	1977	1983
Austria	1949	1922	1986	1925	1986	1920	1920	1920	1960	1923	1970	1973	1975	1980	1960	1989	1960	1970		1949	1924	1949		1995		1993	1995	1987
Belgium	1970	1920	1934	1927	1965	1920	1920	1920	1948	1920	1970	1969	1973	1980	1960	1980	1960	1970		1935	1920	1970		1980		1981	1980	1981
Canada	1920	1920	1920	1934	1923	1920	1920	1920	1934	1920	1970	1956	1973	1980	1960	1995	1961	1970	1970	1920	1920	1954	1970	1969	1971	1981	1969	1981
Denmark	1951	1921	1921	1969	1952	1920	1920	1920	1921	1920	1994	1969	1979	1980	1960	1991	1960	1970		1922	1920	1951		1994		1981	1994	1979
France	1920	1920	1920	1920	1924	1920	1920	1920	1922	1920	1970	1971	1968	1980	1960	1991	1960	1970	1970	1921	1920	1969	1970	1977	1971	1981	1977	1988
Germany	1925	1924	1928	1920	1928	1920	1920	1920	1920	1920	1970	1969	1970	1980	1960	1968	1960	1970	1970	1926	1925	1950	1950	1970	1951	1981	1970	1979
Ireland	1971	1934	1973	1973	1973	1920	1946	1922	1960	1928		1973	1973	1981	1960	1997	1960	1970		1949	1924	1971		2002		1982	2002	1985
Italy	1920	1920	1973	1925	1973	1920	1920	1920	1922	1920	1970	1984	1957	1981	1960	1993	1960	1970	1980	1920	1920	1950	1966	1950	1967	1982	1950	1983
Japan	1920	1920	1946	1920	1958	1920	1920	1920	1920	1920	1970	1956	1948	1980	1960	1972	1960	1970	1980	1920	1920	1964	1970	1964	1971	1980	1964	1980
Korea	1960	1962	1975	1963	1987	1920	1920	1948	1969	1973	1972	1988	1962	1986		1978	1970	1971		1954	1920	1962		1962		1987	1962	1990
Netherlands	1948	1920	1928	1969	1928	1920	1920	1920	1920	1920	1970	1969	1973	1980	1960	1986	1960	1970		1949	1920	1961		1990		1981	1990	1979
Norway	1953	1945	1988	1969	1986	1945	1945	1945	1959	1945	1970	1969	1980	1984	1960	1987	1960	1970		1945	1946	1953		1975		1981	1975	1979
Portugal	1947	1934	1938	1988	1989	1920	1920	1930	1981	1920		1988	1986	1986	1960	1988	1960	1972		1954	1920	1947		1979		1990	1979	1996
Singapore	1963	1965	1970	1972	1986	1920	1920	1948	1972	1998		1972	1965	1980		1983		1970		1958	1969	1991		1991		1993	1991	1982
Spain	1920	1920	1940	1920	1966	1920	1920	1920	1924	1920		1979	1971	1990	1960	1990	1960	1970		1920	1920	1970		1980		1993	1980	1979
Sweden	1920	1920	1920	1920	1926	1920	1920	1920	1920	1920	1974	1969	1980	1982	1960	1995	1960	1970		1920	1920	1961		1981		1982	1981	1979
Switzerland	1920	1920	1930	1920	1930	1920	1920	1920	1920	1920	2000	1969	1969	1980	1960	1989	1960	1970		1930	1924	1975		1999		1993	1999	1979
UK	1920	1920	1920	1923	1923	1920	1920	1920	1920	1920	1970	1927	1969	1980	1960	1966	1960	1970	1971	1920	1920	1962	1971	1962	1972	1981	1976	1981
US	1920	1920	1920	1920	1929	1920	1920	1920	1920	1920	1920	1920	1928	1980	1960	1934	1960	1970	1960	1920	1920	1952	1946	1952	1947	1952	1952	1980

Chapter III

Risk and Return in High Frequency Trading

(with Jonathan Brogaard, University of Washington, and Andrei Kirilenko, MIT)

1 Introduction

Driven in part by their seeming ability to profit in all circumstances, high-frequency trading (HFT) firms have recently attracted a great deal of attention. Still, many basic questions about their strategies and profitability remain unanswered. For example, how do HFT firms make money? How do they compete with one another? And how useful is speed in driving revenue?

This paper seeks to address these questions, motivated by concerns raised in the academic literature and popular press concerning the incentives of individual HFTs to take liquidity and to compete over small increases in speed. We use a proprietary transaction-level data set from the Commodity Futures Trading Commission (CFTC) to study the these concerns through the lens of risk and return of individual HFT firms in the E-mini S&P 500 futures contract.[1]

Consistent with these concerns, we show that HFT firms who specialize in liquidity-taking (aggressive) strategies generate substantially more revenue than those who spe-

[1]We identify "HFT" firms by using activity-based selection criteria introduced in Kirilenko, Kyle, Samadi, and Tuzun (2014). For brevity, we use the term "firm" to be synonymous with a specific trading account, even though several HFT firms in our data set have more than one trading desk, as it is our understanding that, due to regulatory and clearing requirements, most (but not all) trading desks operate as semi-independent trading entities.

cialize in liquidity-providing (passive) strategies. Moreover, revenue persistently and disproportionally accumulates to the top performing HFTs, suggesting winner-takes-all market structure. We further show that speed is an important determinant of revenue generation, and the relation is strongest for HFTs with liquidity-taking (aggressive) strategies. Our results imply that, consistent with many of the concerns highlighted in the theoretical literature, HFT firms have strong incentives to take liquidity and to compete over small increases in speed.

We start by establishing some basic empirical facts regarding the risk and return characteristics of individual HFT firms. The median HFT firm demonstrates unusually high and persistent risk-adjusted performance with an annualized Sharpe ratio of 4.3 and a four-factor (Fama-French plus momentum) annualized alpha of 22.02%.[2] Unlike for many non-HFT investment strategies, firm-level performance is strongly persistent over both days and months. Risks are kept low by strict inventory management and rapid turnover of contracts. Despite the strong outperformance on the firm level, effective HFT trading costs paid by non-HFT investors are only 0.22 basis points.

Our focus on distinguishing liquidity-demanding versus liquidity-providing trading strategies of HFTs is motivated by the idea that speed can be helpful in different ways. The theoretical literature puts forth several ideas about the ways that HFTs use speed to generate profits. Some theories (e.g., Martinez and Rosu, 2013; Foucault, Hombert and Rosu, 2013; and Biais, Foucault, and Moinas, 2014) view HFTs as aggressive (liquidity-demanding) traders who use speed and aggressive orders to trade an instant before others — whether in reaction to news, order flow, or latency arbitrage — and pick off stale limit orders or trade ahead of others' information. In this view, HFTs increase adverse selection and trading costs on other investors. Other theories, in contrast, (e.g., Jovanovic and Menkveld, 2012; Ait-Sahalia and Seglam, 2013; and various viewpoints

[2]As explained further in Section 3, these returns assume fully-capitalized positions in the futures contract and should thus be interpreted as a highly conservative lower bound on returns. Given that margin requirements in the E-mini are about 10% of the notional value of the contract, actual HFT returns on capital are most likely several times higher than reported here.

presented in the mainstream media[3]), view HFTs as passive market makers who use speed to cancel or modify limit orders in response to informed trading, thus mitigating adverse selection and providing tighter bids and asks.

In distinguishing these two different views of HFTs, we find firm-level specialization: a majority of HFTs consistently specialize either in liquidity taking (whom we label Aggressive HFTs) or liquidity-provision (Passive HFT). More importantly, Aggressive HFTs earn substantially higher returns than Passive HFTs — the average Aggressive HFTs earns an annualized alpha of 90.67%, while the average Passive firm earns 23.22% — suggesting that there is a strong profit motive for liquidity taking rather than liquidity providing.

We further distinguish Aggressive and Passive HFTs in terms of their risk and return characteristics and trading behavior. Spectral analysis (following Hasbrouck and Sofianos, 1993) shows that Aggressive HFTs as a whole lose money on shorter time scales (presumably from the bid-ask spread and price impact) but gain money by predicting price movements on longer (but still intraday) time scales. In contrast, Passive HFTs show the opposite, making money at short horizons and losing money over longer intervals. We also decompose from whom (i.e. from which other trader types) HFT firms earn their trading revenue, and show that Passive and Aggressive HFTs win and lose from different trader types. In particular, Aggressive HFTs make about 45% of their revenue from adversely selecting the other HFT subtypes.

A second motivation for studying trading revenue and competition among firms is that the competitive trading structure of HFT firms can lead to a winner-takes-all environment, whereby the trader who is first able to identify and respond to a profitable opportunity will capture all the gains (see, for example, Budish, Cramton and Shim, 2013; Jones, 2013; Weller, 2013)[4]. Other firms who are even milliseconds late will miss

[3]See, for example, http://www.bloombergview.com/articles/2014-03-20/why-do-high-frequency-traders-never-lose-money

[4]Short-lived profit opportunities may derive from trading on news (for example, using direct data feeds) or on order flow (using information obtained from "pinging", flash quotes, or exploiting delays

out: the magnitude of the profit may be sharply reduced or the trading opportunity may have disappeared completely. A winner-takes-all environment leads to socially inefficient investment in faster technology (Budish, Cramton and Shim, 2013; Biais, Foucault, and Moinas, 2014), as small increases in trading speed lead to large payouts, driving an arms race for seemingly small reductions in latency. This type of environment may further lead to incentives to exploit speed advantages via liquidity-taking trading strategies: picking off limit orders of slower traders and liquidity-providers, which may reduce liquidity or force effective trading costs upon other investors.

According to the winner-takes-all hypothesis, we expect to see a highly right-skewed cross-sectional distribution of revenue and a high concentration of revenue (measured by the Herfindahl index) that persists over time. Several other consequences may follow from a winner-takes-all environment. For example, Budish, Cramton and Shim (2013) theorize that that if speed advantages are relative, then increased competition won't drive profit opportunities to zero, since HFTs can always one-up the competition with an ever-smaller increase in speed. Thus, aggregate trading revenue and the concentration of revenue would not decrease over time. Additionally, we expect entrants to earn substantially lower returns than established HFT firms and be more likely to exit, as presumably incumbents have advantages due to their experience allowing them to capture most of the profits. Finally, we also expect speed to be a strong determinant of profitability, and the relation is strongest for HFTs with liquidity-taking (aggressive) strategies.

In looking at the data, we find evidence consistent the above predictions: the cross-sectional distribution of returns is highly right-skewed, with revenue disproportionally and persistently accumulating to top-performing firms. Revenue is concentrated among the top performers, as measured by the Herfindahl index. Both results are consistent with a winner-takes-all environment. Additionally, we find that aggregate revenue and

in public order book updates), taking advantage of mispriced orders or others' trading mistakes, or using other predictive, momentum or signal trading strategies.

concentration of revenue among top-performers does not decrease over our two-year sample, after adjusting for volatility and non-HFT trading volume. New entrants earn substantially fewer profits and are more likely to exit. Finally, we study a measure of relative (rank-order) speed developed in Weller (2013), which measures latency in terms of reaction time to incoming order flow. While Weller (2013) previously demonstrated that relative speed is correlated with returns, we show that the relation is strongest Aggressive HFTs. Each of the above findings is consistent with a winner-takes-all environment with speed and aggressiveness being key components of success. Overall, our analysis thus reveals an industry dominated by a small number of increasingly-fast, liquidity-taking incumbents with high and persistent returns.

The rest of the paper is as follows. Section 2 discusses the related literature, Section 3 describes the data and methods, Section 4 examines the risk and return performance of HFTs, Section 5 analyzes competition, market concentration, and entry/exit of firms within the HFT sector, Section 6 studies speed, and Section 7 concludes.

2 Related Literature

This paper contributes mainly to two literatures: the growing body of work on HFT and the study of investment performance of different groups of traders. Given that there is no publicly available data set on HFT firms, several papers study HFT activity despite being unable to directly observe individual high frequency traders (e.g., Hasbrouck and Saar, 2013). Other papers make use of limited or aggregated proprietary data sets. For example, Jovanovic and Menkveld (2012) and Menkveld (2013) study the July 2007 entry into Dutch stocks of a single high-frequency market maker, and Brogaard, Hendershott, and Riordan (2013) study aggregated HFT activity on NASDAQ. In contrast to these papers, this paper uses a data set that allows us to identify

the trades of individual HFT firms.

Most empirical papers on HFT and algorithmic trading assess the potential costs and benefits of HFT using natural experiments, such as analyzing technological upgrades to trading venues (e.g., Hendershott, Jones, and Menkveld, 2011; Boehmer, Fong, and Wu, 2014; Riordan and Storkenmaier, 2012; and Gai, Yao, and Ye, 2013) or changes in fee structures only affecting HFT firms (e.g., Malinova, Park, and Riordan, 2013). These papers generally show that HFT activity improves liquidity (in terms of bid-ask spreads, depth, and price impact), lowers adverse selection (for example, more price discovery taking place via quotes rather than trades), and lowers transaction costs for institutional and retail traders. In contrast to these papers, this study looks at HFT from a risk and returns perspective and analyzes the incentives and competitive forces that shape HFT activity.

Finally, a number of previous studies have evaluated the investment performance of different types of traders. For example, Harris and Schultz (1998) study the profitability of SOES bandits, a group of individual traders in the 1990s who would quickly enter and exit trades and who were thought by some to have unfair advantages. Hasbrouck and Sofianos (1993) study the profitability of NYSE specialists. Like Ackermann, McEnally, and Ravenscraft (1999), who study the profitability of different hedge fund strategies, this paper studies different trading strategies of HFTs. Similar to studies look into factors that induce different traders to trade (e.g., Grinblatt and Keloharju, 2001), this paper studies incentives for entry and exit, speed, and aggressiveness.

3 Data and Methods

3.1 The E-mini S&P 500 futures market

We use transaction-level data with trader identifiers for the E-mini S&P 500 stock index futures contract (E-mini). Our data set spans over two years, from the start of August 2010 to the end of August 2012.

The E-mini is a favorable setting for studying HFT for the following reasons.

First, it is an important and highly liquid market with several different types of market participants regularly trading, including a high number of HFT firms. The E-mini is the second-most traded futures contract in the world with a notional trading volume of approximately $200 billion per day in August 2010. Hasbrouck (2003) shows that the E-mini futures contract is the largest contributor to the price discovery process of the S&P 500 index.

Second, the contract is in zero net supply and buying and selling are symmetrical, so there are no short-selling constraints. Trading in the E-mini is a zero-sum game: one trader's profits come directly at the expense of the opposite trader.

Third, because the contract trades only electronically and only on the Chicago Mercantile Exchange (CME), there is no concern about unobserved trades occurring on other exchanges or on the floor.

Lastly, the E-mini has no designated market makers, no maker-taker fees or liquidity rebates for the front-month contract, and no obligations for certain market participants (such as quoting two sides or making prices continuous). There is no institutionalized class of intermediaries in this market. HFT trading activity is presumably in explicit pursuit of profits, undistorted by other requirements or competing incentives.

Our data set is trade-by-trade and contains common fields such as price, number of contracts traded, and time of the trade in units of seconds (and in milliseconds

for a few months). The CME's Globex matching engine stamps a unique matching ID on each regular transaction, which enables us to construct the exact ordering of transactions. Cancelled and other irregular transactions are filtered out. In addition, the data set contains unique identifiers for the ultimate buyer and the ultimate seller (not just their brokers), an identifier for which side initiated the trade (passive for the side with the resting order, aggressive, otherwise), and an identifier for each executed order that allows us to group multiple transactions into a single underlying order (since large executable orders may be executed against several different resting limit orders).

Each E-mini contract is $50 times the value of the underlying S&P 500 index; as a result, the notional contract is valued at approximately $50,000. The tick size is $12.50. The contract is cash-settled against the value of the underlying S&P 500 index. Initial margins for speculators and hedgers (members) are $5,625 and $4,500, respectively, in August 2010; maintenance margins for all traders are $4,500. We exclude months in which the leading contract expires (March, June, September, December) in order to exclude rollover effects and multiple expirations trading simultaneously. Outside of the rollover months, the front-month contract usually has well over 99% of the trading volume, although we do analyze trades of all expirations. While the E-mini futures contract trades nearly around-the-clock, we only use data during normal trading hours: 8:30 a.m. - 3:15 p.m. Central Standard Time (CST).

One limitation of the paper is that our profit calculations do not account for all the costs of an HFT firm. While we know such direct costs of trading as trading fees ($0.15 per contract), the cost of direct data feeds, and the cost of co-location, we cannot adequately calculate other costs such as computer systems, labor, and risk management systems. We report gross trading revenues throughout to limit speculative assumptions from influencing our findings and because our focus is on trading performance.

3.2 Categorizing Traders

Following Kirilenko, Kyle, Samadi, and Tuzun (2014), we define different trader types based on two selection criteria: inventory and trading volume. HFTs are identified as those firms with extremely high volume, low intraday inventory and low overnight inventory. As shown in Kirilenko, Kyle, Samadi, and Tuzun (2014) and confirmed by us, by analyzing all market participants along the dimensions of volume and inventories, HFT firms stand out as a distinct cluster, with daily trading volume orders of magnitude higher than other traders. In addition, we use the aggressiveness indicator (assigned by the matching engine) to group HFT firms into three different categories based on their liquidity impact.

The precise selection criteria for HFTs are as follows. For all individual market participants, we calculate four metrics: median trading volume (median trading volume across days which the firm is active, where active is defined as an HFT trading more than 1000 contracts per day), end-of-day inventory ratio (daily end-of-day inventory position divided by total volume, median across days which the firm is active), intraday inventory ratio (maximum daily position minus minimum daily scaled by daily total volume, median across days which the firm is active), and aggressiveness ratio (ratio of aggressive trading volume to total trading volume). For each month, there are three categories a potential trader must satisfy to be considered a HFT: (1) median trading volume greater than 5,000 contracts per day (about $250 million in notional daily trading volume)[5], (2) end-of-day inventory ratio less than 5%, and (3) intraday inventory ratio less than 10%. As a final step, we drop from our sample firms that are active fewer than 10 trading days per month. We find 85 firms satisfy these HFT criteria, though because of entry and exit, not all these firms necessarily trade in the same months.

We find that a majority of HFT firms (or HFTs) consistently specialize either in

[5]This is a small (overly-inclusive) minimum requirement for trading volume compared to what most HFTs in our sample actually trade. The average daily trading volume per firm in our sample is 36,000 contracts (about $1.8 billion).

liquidity taking (Aggressive HFTs) or liquidity-provision (Passive HFTs) over the two-year sample we study. Thus, we categorize HFTs as follows: Aggressive HFTs have an aggressiveness ratio greater than 60%, while Passive HFTs have an aggressiveness ratio less than 20%. The Aggressive HFT designation is meant to capture HFTs who are liquidity takers and generate trading profits by predicting future market movements, while the Passive HFT designation is meant to capture algorithmic market makers and liquidity providers whose revenue is partly generated by the bid-ask spread. We later show that these two classes of HFTs have markedly different characteristics in terms of trading behavior and profitability. We call firms that are not in either category Mixed HFTs and later show that they have characteristics that resemble a mix the two main categories. There are 18 Aggressive, 39 Mixed, and 28 Passive HFTs in our sample, though again not all these firms necessarily trade in the same months.

For each month we classify the rest of the universe of non-HFT firms into four different subcategories: Non-HFT Market Maker, Fundamental, Small, and Opportunistic traders. The non-HFT market maker category captures traditional market makers and liquidity providers, such as those examined by Hasbrouck and Sofianos (1993) and Coughenour and Saad (2004). Specifically, Non-HFT market makers are non-HFT market participants who have a (1) aggressiveness ratio less than 20%, (2) end-of-day inventory ratio less than 15%, and (3) median trading volume of at least 20 contracts (about $1 million) per day. The fundamental trader category is meant to capture large institutional traders who adopt buy-and-hold strategies, such as those studied by Anand, Irvine, Puckett, and Venkataraman (2012) and Puckett and Yan (2011). Specifically, Fundamental traders have a (1) median trading volume of 1000 contracts per day (about $50 million) and (2) end-of-day inventory ratio greater than 30%. The Small trader category most likely captures retail traders (e.g. Kaniel, Saar, and Titman, 2008; Seasholes and Zhu, 2010). Small traders have a median daily trading volume of less than 20 contracts (or about $1 million). Note that HFTs and the

above three categories are all mutually exclusive by definition. The remaining firms are designated Opportunistic. The Opportunistic category captures either medium-sized buy-and-hold investors (firms not big enough to be considered Fundamental traders) or large traders who unlike Fundamental traders move in and out of positions throughout the day but with significantly larger fluctuations and lower frequency than HFTs. This group likely captures arbitrageurs, small asset managers, hedge funds, and other hard-to-classify traders. In August 2010, we identify 737 Non-HFT Market Maker, 346 fundamental traders, 21,761 small traders and 8,494 opportunistic traders.

Table 1 presents a summary of trading behavior for these different trader types for August 2010. For each trader type, we report two statistics: the daily percent of market volume traded and the daily aggressiveness ratios (aggressive contracts by trader category divided by total contracts by trader category). Table 1 Row 8 shows that, on average, 3.2 million contracts are traded in the E-mini market per day. The E-mini is an extremely liquid market with an average of about 70 contracts traded every second. HFTs as a whole trade 54.4% of the double-counted trading volume ($\frac{Buys_{HFT}+Sells_{HFT}}{2 \cdot MktVolume}$), or 1.73 million contracts daily (Row 1+2+3). The next largest category is Opportunistic, with 31.93% of contract volume by its 8,494 participants (Row 7).

The variation within the HFT categories over different days is considerable. For example, Aggressive HFTs range from 9.5% to 17.6% of market volume across days; they have the largest variation of the three HFT categories (Row 1). Passive HFTs make up a significantly smaller portion of HFT volume (8.87%, Row 3) than Aggressive HFTs (15.22%, Row 1). Looking at liquidity taking, Aggressive HFTs take liquidity in 84.22% (Row 9) of the contracts they trade, while Passive HFTs only take liquidity in 12.35% (Row 11). Fundamental traders make up 8.42% (Row 4) of trade quantity, while taking liquidity about half of the time (57.68%, Row 10). Small traders and non-HFT market makers are both a small share of the market, with 1.04% (Row 5) and 4.24%

(Row 6) of the trade quantity, respectively, while taking liquidity 57.82% (Row 11) and 12.98% (Row 12) of the time, respectively. Opportunistic traders make up the largest share of trading volume at 31.93% (Row 7), while taking liquidity 59.08% (Row 13) of the time.

Figure 1 examines trading volume of the different trader types across the two-year span. Panel A shows the average daily HFT volume of the HFT subgroups for each month over the two year span. The total HFT market volume decreases over the two-year span from a peak of 3,187,011 contracts in August 2010 to 2,322,787 in August 2012. About half of that decline was from a reduction in Mixed HFT trading from 960,643 contracts in August 2010 to 564,200 contracts in August 2012, while Aggressive and Passive HFT volume was relatively constant.

Panel B looks at the percent of market volume traded by each trader type over the two year span. Although average daily market volume fluctuates considerably over the two-year span, the percent traded by each type is relatively stable. For example, HFT percentage volume starts in August 2010 at 54.37% and ends in August 2012 at 55.53%; However, Aggressive HFT volume increases over the two years from 15.22% to 22.63%, while Mixed HFT volume decreases from 30.28% to 24.59%. Overall, Figure 1 highlights the strong stability in HFT trading volume over a two year span in the E-mini S&P 500 futures contract.

3.3 Returns and Profitability

Throughout the paper, daily profits for each firm, i, are calculated for each trading day, t, according to marked-to-market accounting, assuming that each trader starts each day with a zero inventory position. More precisely, for each trader, we calculate the end-of-day profits as the cumulative cash received from selling short positions minus

the cash paid from buying long positions, plus the value of any outstanding positions at the end-of-day, marked to the market price at close: $\pi_{i,t} = \sum_{n=1}^{N_{i,t}} p_n y_{i.n} + p_N y_{i,N}$, where p_n is the price of the nth trade (out of N total trades on day t) and $y_{i,n}$ is the signed quantity of the nth trade by trader i. As Table 5 shows, HFTs usually end the day flat, so marking-to-market at the end of the trading day is relatively innocuous.

Daily returns are calculated by dividing the daily profits of each firm by the assumed capitalization of each firm. Capitalization is calculated for each firm on a monthly basis by looking at the maximum position taken in the month and multiplying by the notional value of the contract. Since HFTs inventory generally exhibit sharp, well-defined maximum and minimum positions they're willing to take (for example, bounds of 100 or 200 contracts exactly), this number represents the maximum amount of capital that an HFT would need to execute its specific strategy.[6]

Finally, we evaluate the Sharpe ratios of HFTs (Sharpe, 1966). Four-factor (Fama-French plus momentum) alphas are computed for each HFT using the standard Fama-French model (Fama and French, 1993) plus the Carhart (1997) momentum factor, using daily portfolio returns data from Kenneth French's website. The Sharpe ratio for each trader is calculated as $SR_i = \frac{r_i - r_f}{\sigma_i} \cdot \sqrt{252}$, where r_i is the average daily return, r^f is the risk-free rate, which is zero throughout our sample, and σ_i is the standard deviation of trader i's returns over the calculated sample period (usually one month).

[6]In practice, HFTs may use much less capital than the amount we calculate, since firms only have to satisfy margin requirements, which are generally about one-tenth the notional value of the contract. We decided to use the notional value of the contract rather than the margin requirement per contract to calculate the firm's capitalization because a firm that loses money on any given day will have to post new margins. The maximum possible capital that a firm would need is the full notional value of the contract. In practice, HFT returns on capital are presumably several times higher than we report, and thus these numbers should be interpreted as conservative lower bounds.

4 Risk and Return Performance of HFTs

In this section we present empirical facts about the investment performance of HFTs. First, HFTs consistently outperform the market, with the median HFT firm achieving an annualized Sharpe ratio of 4.3 and a four-factor annualized alpha of 22.02%. Aggressive HFTs earn substantially higher returns than Passive HFTs — the average Aggressive HFTs achieves an annualized alpha of 90.67%, while the average Passive firm earns 23.22% — suggesting that there is a strong profit motive for liquidity taking rather than liquidity providing. Second, the cross-sectional distribution of returns is highly skewed towards top earners whose Sharpe ratios and alphas are many times those of the median, and firm-level performance is strongly persistent over both days and months. We interpret this as showing profits persistently and disproportionally accumulate to the top performing HFTs, suggesting winner-takes-all competition. The winner-takes-all nature of returns is further confirmed in Section 5 when we look at industry concentration of returns and report high Herfindahl concentration indices.

Looking further into how HFTs minimize their risk, we show that high Sharpe ratios are achieved by minimizing positional risk through strict inventory control and rapid turnover of contracts. Finally, to help differentiate Aggressive versus Passive HFTs' strategies, we analyze returns to different types of HFTs by performing spectral analysis (following Hasbrouck and Sofianos, 1993) and by decomposing from whom (i.e. from which other trader types) HFTs earn their profits. We show that Aggressive HFTs as a whole lose money on shorter time scales (presumably from the bid-ask spread and price impact) but gain money by predicting price movements on longer (but still intraday) time scales; in contrast, Passive HFTs show the opposite, making money at short horizons and losing money over longer intervals. In particular, Aggressive HFTs make a large part of their profits from adversely selecting Passive HFTs.

Finally, the decomposition of profits also shows that despite strong outperformance

on the individual firm level, effective HFT trading costs imposed on non-HFT investors are only 0.22 basis points.

4.1 Distribution of Returns

To get a general sense of HFT returns, we first present the distribution of daily returns pooled across all firms within each subtype. We highlight the higher profitability of Aggressive HFTs, suggesting a stronger profit motive for liquidity taking than liquidity providing. In the next subsection, we look more specifically at the cross-sectional distribution across firms.

Table 2 presents the distribution of HFT returns and profits per contract. Panel A reports annualized returns, determined from the daily returns of each HFT firm and aggregated into groups. Panel B reports profit per contract. Mean, median, skew, and kurtosis are reported for both returns and profits per contract.

Table 2 shows that Aggressive HFTs earn substantially higher returns than Passive HFTs suggesting that there is a strong profit motive for liquidity taking rather than liquidity providing. For example, Aggressive HFTs earn a mean annualized return of 122.10% (Column 2), while Mixed and Passive HFTs earn significantly less: only 32.29% and 25.69%, respectively.[7] The p-values on the statistical significance of profits (Column 7) show that these values differ significantly from zero. While the averages are all positive, there is a sizeable distribution of profitability across observations: the standard deviation of returns (Column 4) has Aggressive HFTs realizing the highest variation in profits (28.39%) and Passive HFTs the lowest (8.89%). The daily skewness and kurtosis (Columns 5 and 6) statistics show that the distribution of profits is non-normal: there is excess weight in the tails, in the upper tail especially (positive skew),

[7]We also calculate the returns for 24-hour continuous trading (as opposed to looking only at regular trading hours) and find qualitatively similar results for the three subtypes of HFTs.

for all three types of HFTs.

Panel B reports profits per contract and similarly confirms that Aggressive HFTs are considerably more profitable than Mixed and Passive HFTs. We also find, in unt-abulated results, that in dollar terms the average Aggressive HFT firm earns a daily profit of $74,000 in gross trading revenues, while average Mixed and Passive firms earn $27,910 and $12,600, respectively. Annualized, these numbers corresponds to trading revenues of over $18.6, $7.0, and $3.2 million for Aggressive, Mixed, and Passive HFTs, respectively, in the E-mini alone. Overall, Table 2 highlights the higher profitability of Aggressive HFTs, suggesting a stronger profit motive for liquidity taking than liquidity providing.

4.2 HFT Risk and Return in the Cross-Section

Next, we look at the cross-section of HFTs to study the winner-takes-all nature of competition and find wide variation in returns across HFT firms in the cross-section. In particular, we analyze returns, abnormal returns (alphas), profits per contract, and Sharpe ratios. Other than profits per contract, the reported statistics are annualized. The cross-sectional distribution of returns is highly skewed towards top earners whose Sharpe ratios and alphas are many times those of the median.

Table 3 reports the mean and standard deviation across firms, in addition to the 10th, 25th, 50th, 75th and 90th percentiles. The results are broken down for Aggressive, Mixed, and Passive HFTs. The cross-sectional distribution of annualized returns (Rows 1-4) shows that there is a wide distribution of returns across firms.

We focus on alphas (abnormal returns), which provide a better sense of risk-adjusted performance than absolute returns. Alphas are reported for various specifications, though we emphasize the four-factor (Fama-French + Carhart momentum) alphas. The

average HFT firm earns abnormal annualized returns of 39.92%. Comparing this number to absolute returns of 39.49% shows that the returns of HFTs are unrelated to market returns. The breakdown among the subcategories is striking. Aggressive HFTs earn alphas of 90.67% while Mixed and Passive earn 28.18% and 23.22%, again suggesting a stronger profit motive for liquidity taking than liquidity providing.

The fact about the distribution we wish to emphasize is the winner-takes-all aspect: that firms in the top decile strongly outperform those at the median. Firms in the bottom ten percent earn small negative returns, while firms above the median earn large positive profits. The distribution is right skewed, reflecting a few high earners at the top. For example, 25% of HFTs have annualized four-factor alphas greater than 50.92%, and 10% of HFTs have alphas greater than 70.51%. Similarly, the fact that the median firm's alpha (22.02%) is lower than mean firm's alpha (39.78%), especially for Aggressive HFTs (median = 37.06%, mean = 90.67%), suggests that the mean is driven by top performing firms who earn particularly high returns, providing further evidence for positive skewness in the cross-section of firms.

Table 3 also reports the distribution across firms of average daily profit per contract (rows 11-14). Similar to the raw returns and alphas, profit per contract is widely dispersed around the median with a pronounced positive skew: profit per contract in the top decile of firms is $2.01, compared with 0.46% for the median.

Lastly, we examine annualized Sharpe ratios, which show that while HFTs bear some risk, their risk-adjusted returns are unusually high. Menkveld (2013) has previously calculated the annualized Sharpe ratio of an HFT firm trading in equities markets to be 9.35. Passive HFTs generate the highest Sharpe ratio of 5.85, implying that they keep risk low, as their annualized returns are the lowest at 23.13%. Mixed and Aggressive HFTs earn the Sharpe ratios of 5.26 and 4.29, respectively. To get a sense of the magnitude of these numbers, the performance of the median HFT firm Sharpe ratio (4.30) is over 13 times higher than the Sharpe ratio of the S&P 500 (0.31) (Fama

and French, 2002) and 18 times higher than the mean Passive HFT firm. Thus, while HFTs bear some risk, their risk-adjusted returns are high.

Yet even within these different types of HFTs, the Sharpe ratio for firms varies widely, especially with the top earners, consistent with the winner-takes-all idea. For example, 25% of HFTs have a Sharpe ratio greater than 9.10, and 10% of HFTs have a Sharpe ratio greater than 12.68. Again, firms in the top decile strongly outperform those at the median, suggesting winner-takes-all.

4.3 Persistence of Returns

Next we test for persistence of returns, both in terms of absolute returns and also relative to other HFT firms. Persistent profits over time suggest that something other than luck is driving a firm's performance. There is an extant literature showing that performance for actively-managed mutual funds in period t generally does not predict performance in period t+1 (e.g. Carhart, 1997). However, there is evidence of persistence by some investors (e.g. Jagannathan, Malakhov, and Novikov (2010) for hedge funds and Kaplan and Schoar (2004) for private equity). Nonetheless, the expectation for most types of investors and funds is that ability has little influence in investing returns. The null hypothesis here is thus that HFT firms do not exhibit persistence in returns.

To formally analyze persistence of returns, we look at the role that lagged returns (or rank ordering versus other firms) have in predicting today's returns (or rank order), both using daily and monthly firm-level returns. We estimate the following ordinary least squares (OLS) regression:

$$r_{i,t} = \alpha + \beta \cdot r_{i,t-1} + Time\,FEs + \epsilon_{i,t} \tag{1}$$

where $r_{i,t}$ is daily or monthly returns. While there may be higher idiosyncratic risk in

daily returns, resulting in moderate persistence at the daily level, we find that returns are more strongly persistent on the monthly level. The regression is estimated for each group, Aggressive, Mixed, and Passive, separately. We also estimate Equation (1) with other dependent variables and their lags, including normalized returns, normalized profits, normalized profits per contract, Sharpe ratios, and the rank order of these metrics.[8]

The results are reported in Table 4. Results based on daily observations are reported in Panel A, those from monthly observations are reported in Panel B, and those based on rank ordering in Panel C. Regressions considering normalized dependent variables are performed without time fixed effects as the normalization removes any average time variation. With normalization, the estimate of the constant, which is always zero, is not reported.

Estimates for each type of HFT show that one-day lagged performance is a strong predictor of current performance. HFTs exhibit persistence. For daily returns, Aggressive HFTs have coefficients of 0.421 (for daily returns) and 0.723 (for monthly returns). Mixed and Passive HFTs have mild or no persistence on the daily level (0.109 and -0.73, respectively), at least before normalizing returns, but strong persistence on the monthly level (0.407 and 0.725, respectively). Normalized daily returns are qualitatively similar, though we see slightly weaker persistence for Aggressive HFTs on the daily level than before (0.276) and stronger persistence for Mixed and Passive HFTs on the daily level than before (0.169 and 0.205, respectively). The coefficients reported above are all statistically significant with p¡0.001.

The adjusted R^2 estimates imply strong explanatory power of past returns: 25.5%, 7.9% and 12.4% for daily returns and 52.2%, 22.4% and 33.5% for monthly returns for Aggressive, Mixed and Passive HFTs, respectively. Persistence estimates using normalized profits and profits per contract are qualitatively similar, though slightly weaker,

[8]By normalized, we mean subtracting out the mean and dividing by the standard deviation across firms for that day. We normalize to control for the time-varying mean and volatility of HFT returns.

than those using returns. However, persistence estimates using monthly normalized Sharpe ratios, however, are particularly strong with coefficients of 0.639, 0.659 and 0.452 for Aggressive, Mixed and Passive HFTs, respectively, with also high adjusted R^2 estimates.

The above findings show that the level of performance in the past is predictive of the level of performance in the future. We perform a similar analysis in Panel C but instead consider whether relative rank among HFTs within their subcategories in the past is predictive of current rank. Rank is calculated using daily returns, monthly returns, monthly profits, monthly profits per contract, and Sharpe ratios. For each measure separately, we rank the relative performance of the HFT firms in each day or month, with 1 representing the best performing firm, 2 the second-best performing firm, etc. We then repeat the analysis in Equation (8). A positive coefficient means that if a firm performed well in the past, it is likely to perform well in the future. All measures and all subcategories suggest persistence in relative performance.[9]

In thinking about the drivers of persistence, it may be that human skill and experience allow firms to distinguish themselves. Alternatively, it could be purely technological advantages. A combination of the both is also probable. Regardless of the precise mechanism driving the persistence, persistent performance may lead to industry concentration as strong performing firms will continue to perform well. Less successful firms may exit the industry. In Section 4 we observe both and analyze the implications of competition among HFT given strong persistence among firms.

[9]As the dependent variables are both discrete and restricted to non-negative values there are additional econometric techniques that apply other than OLS. Two techniques include Poisson and negative binomial models. See Long (1997) for a thorough description these models and how to analyze count data. In untabulated results, we repeat the Table 4, Panel C analysis using both the Poisson and negative binomial models. The results are qualitatively the same.

4.4 Risk and Inventory Management

We've shown that HFTs earn unusually high risk-adjusted returns with a distribution that is strongly positively skewed: top earners achieve alphas and Sharpe ratios many times the median. In addition, firm-level performance is highly persistent over the sample. The next three subsections on inventory management, spectral analysis, and HFTs profits by counterparties provide further evidence of how HFTs consistently produce high risk-adjusted returns.

First, we show that high risk-adjusted returns (alphas and Sharpe ratios) are partly achieved by minimizing positional risk through strict inventory control and rapid turnover of contracts. As shown earlier, HFT risk is low, both in terms of systematic risk (as evidenced by the high alphas) and total risk (as evidenced by Sharpe ratios). Minimizing intraday positional risk not only minimizes HFT exposure to market risk but also puts a bound on maximum possible losses.

Table 5 provides evidence on the inventories of HFTs in the cross-section. The first row reports the Average Daily Volume, which captures the average number of contracts an HFT firm trades each day. The remainder of the table focuses on inventory. The average end-of-day inventory is the average number of contracts (across days) held at the end of the regular trading day, 3:15 CST, by the median firm. The end of day inventory is not always zero, but is generally very low. End-of-day inventory of the median Aggressive HFT is 49.3 contracts, while end-of-day inventory of the median Passive HFT is only 10.5 contracts. The small standard deviations (across firms) show almost all firms keep their average end-of-day inventories very low. To get a sense of how large a position this is for a given firm, we normalize the average end-of-day inventory for each firm by that firm's average daily trading volume. For Aggressive HFTs, this reveals that only 0.4% of trading volume is held at the end of the day by the median firm, while for Passive HFTs, it is only 0.1%.

The second measure of inventory is the average intraday inventory range, which represents the average (across days) of the maximum intraday inventory variation (daily maximum position minus minimum position). This is an estimate of how much exposure a firm is willing to bear. Aggressive HFT take a maximum variation in intraday positions as large as 525.9 contracts on average, while Passive HFTs take a maximum variation of 204.2 contracts. But a fractions of each firm's average daily trading volume though, these numbers are only 7.3% and 2.2%, respectively.

While we select HFTs based on low intraday and end-of-day inventories, Table 5 shows that intraday and end-of-day inventories for the majority of firms are substantially lower than the bounds in our selection criteria, implying that tight inventory management is not just an artifact of our selection criteria but an active goal of HFTs. The results show that minimizing positional risk, and elimination the potential large loss associated with a sizeable direction position, is an important component of HFTs maximizing risk-adjusted returns.

4.5 Spectral Analysis

To further understand the investment horizon of HFTs, we follow Hasbrouck and Sofianos (1993) and decompose HFT profits over different time horizons using spectral analysis. The time horizon over which HFTs make their profits provides insight into their trading strategies and allows us to further examine the differences between the types of HFTs. We find that Aggressive HFTs as a whole lose money on shorter time scales (presumably from the bid-ask spread and price impact) but gain money by predicting price movements on longer (but still intraday) time scales (after 1,000 market transactions). In contrast, both Mixed and Passive HFTs show the opposite, tending to make money at short horizons (1,000 transactions or fewer) and lose money over longer

intervals.

Spectral analysis decomposes profits for each firm into different time horizons. Profits are calculated from two time series, prices and inventories, which are analyzed as sums of Fourier modes of varying frequencies. Intuitively, for any given frequency, if prices and inventories are waves that move in phase (traders buy before the price is going up), profits are generated, while if they are out of phase (traders buy before the price goes down), traders incur losses.

Mathematically, we start with marked-to-market profit for any individual trader at transaction time τ[10]:

$$\pi_\tau = \sum_{t=0}^{\tau} x_t \cdot (p_t - p_{t-1}) = \sum_{t=0}^{\tau} x_t \cdot \Delta p_t \tag{2}$$

where x_t is the inventory holdings of that trader and p_t is the price at time t.[11] By defining the following functions:

$$\widehat{x}_\omega \;\; = \;\; \sum_{t=0}^{\tau} x_t \cdot e^{2\pi i t\omega/T} \tag{3}$$

$$\widehat{\Delta p}_\omega = \sum_{t=0}^{\tau} \Delta p_{t+1} \cdot e^{2\pi i t\omega/T} \tag{4}$$

where ω is interpreted as a wavelength having units of transaction time and $\widehat{x}_\omega$ and $\widehat{\Delta p}_\omega$ are the spectral densities of the x_t and p_t, we can re-express marked-to-markets profits using the following formula (see Hasbrouck and Sofianos, 1993, for mathematical

[10]For computational reasons, spectral analysis must be conducted in transaction time, as opposed to clock time. For reference, there are approximately 10 transactions per second.

[11]Spectral analysis requires us to assume that x_t and Δp_t are stationary processes, which is a valid assumption to make given that x_t, HFT firms' inventories, is a mean-reverting process and Δp_t, is the first difference of the price process.

details)[12]:

$$\pi_T = \frac{1}{T}\sum_{\omega=1}^{\infty} \widehat{x}_\omega \cdot \overline{\widehat{\Delta p}_\omega} = \frac{1}{T}\sum_{\omega=1}^{\infty} 2 \cdot Real(\widehat{x}_\omega \cdot \overline{\widehat{\Delta p}_\omega}) \tag{5}$$

The $2 \cdot Real(\widehat{x}_\omega \cdot \overline{\widehat{\Delta p}_\omega})$ term captures the component of the marked-to-market profits generated at trading wavelength ω. Thus, if the summation in Equation (5) is broken up into wavelength blocks, then it is possible to decompose profits into the following time horizons bins (expressed in units of market transactions): ω = 1-10, 11-100, 101-1,000, 1,001-10,000, 10,000-100,000, and 100,000+. We thus decompose profits into these time horizon bins each HFT firm. Then for each time horizon, we take the median profits across days and then across firms (along with the 25th and 75th percentiles across firms). Spectral analysis results are reported in Table 6.[13]

Table 6 shows that in Aggressive HFTs tend to make positive profits at medium time scales, in the 1,001-10,000 and 10,001-100,000 transaction range, with negative profits at short ranges (11-100 and 101-1,000 transaction intervals) and the longest time scale of 100,000+ transactions. In order for an aggressive trade to be profitable, an HFT must not only predict the direction of the price process but also overcome the bid-ask spread. We suspect that this is the reason Aggressive HFTs fail to make money at the shortest time intervals. The spectral analysis results are consistent with the notion that Aggressive HFTs make money by anticipating price movements in the 1,000-100,000 transactions range, while losing money at both very short and long time scales. In contrast, both Mixed and Passive HFTs tend to gain money at short time scales (1-10, 11-100, and 101-1,000 transactions) while losing money on longer time scales (1,001-10,000, 10,001-100,000, and 100,000+ transactions). These results are consistent with the idea that Mixed and Passive HFTs earn the bid-ask spread in the

[12]Real is the function that takes the real part of a complex number. The last equality follows because the imaginary part of $\widehat{x}_\omega \cdot \overline{\widehat{\Delta p}_\omega}$ sums to 0

[13]We only analyze one month (August 2010) rather than two years of HFT profits due to the computational challenges involved in doing spectral analysis. However, results from several other months were qualitatively similar.

short-run but are adversely selected on a longer time scale.

4.6 HFT Profits by Counterparty

We would like to know from whom HFTs earn their profits. We decompose HFT profits by counterparty, analyzing both dollar profits and profits-per-contract. Analyzing dollar profits tells us where the money is (which groups HFTs have strong incentives to target), and analyzing profits-per-contract tells us about adverse selection costs and which traders are informed traders or best able to evade HFT activity. Interestingly, Aggressive HFTs make about 45% of their profits from adversely selecting the other HFT subtypes.

To calculate from whom HFTs earn their profits, we divide the remaining universe of traders in the E-mini into four categories of traders, as discussed in Data and Methods: Fundamental traders (likely large institutional traders), Non-HFT market makers, Small traders (likely retail traders), and Opportunistic traders (likely arbitrageurs, small asset managers, hedge funds, and other hard-to-classify traders). We find that: 1) HFTs in aggregate make most of their largest dollar profits from Opportunistic traders, 2) on a per-contract basis, Fundamental traders incur the least cost to HFTs, while Small traders incur the most, 3) interestingly, Aggressive HFTs make a large fraction of their profits from Mixed and Passive HFTs, and 4) the effective HFT transaction cost on non-HFT trades is approximately 0.22 basis points.

Whether HFTs make more or fewer profits from different types of other traders is hard to predict using theory. For example, Fundamental traders may trade in a way that makes their order flow noticeable and leaves a detectable pattern in their trading activity (as in Hirschey, 2013; Heston, Korajczyk and Sadka, 2010), in which case HFTs would make more profits against them. On the other hand, due to their

size and sophistication (as in Badrinath, Kale, and Noe, 1995; Boehmer and Wu, 2008; Boulatov, Hendershott, and Livdan, 2013; Hendershott, Livdan, and Schurhoff, 2012), Fundamental traders may be more skilled at evading HFTs and minimizing price impact, in which case HFT would profits less against them.

Similarly, Small traders (retail investors) might be relatively less informed (e.g., Hvidkjaer, 2008; Kaniel, Saar, and Titman, 2008; Barber, Odean, and Zhu, 2009) and thus incur significant losses to HFTs. However, because retail traders are small and trade noisily, they may not leave patterns in the data and consequently HFTs may have a more difficult time detecting them and thus earn less from them. As noted earlier, the result that we find is, on a per-contract basis, Fundamental traders incur the least cost to HFTs while Small traders incur the most, demonstrating that size and sophistication (even at the expense of leaving detectable order-flow patterns) are likely important for minimizing HFT-related trading costs.

Table 7 breaks down the trading profits by trading pairs in August 2010: the rows identify who receives the profits, whereas the different columns represent from whom the profits are derived. Table 7 is constructed by considering only the trades (in August 2010) between two groups and calculating the profit flows that result from those trades. To illustrate, we calculate the Aggressive HFT-Fundamental profit flows by removing all trades except those for which one party was an Aggressive HFT and the other party was a Fundamental trader. Since gains/losses from large, long-term positions can make profits look extremely volatile, we calculate short-term profits: defined here as gains/losses resulting 1 minute after each trade. Since the futures market is a zero-sum game, the resulting profits matrix is symmetrical and zero along the diagonals.[14]

Panel A shows that Aggressive and Mixed HFTs make positive profits from all other

[14]Note that our focus is on returns during a short horizon. A loss during this interval does not imply that the trader (or trader category) loses money overall. In addition, we only observe a market participant's activities in the E-mini, which may be one of multiple markets in which a trader participates. For instance, Fundamental traders may be using the E-mini contract as a hedge against real-world risks and Opportunistic traders may be doing cross-market arbitrage. Thus, a loss in the E-mini market does not imply the trading firm loses money overall.

types of traders, while Passive HFTs make positive profits from all other types of traders except from fundamental traders. In particular, Aggressive HFTs make about 45% of their profits (= (7,190,140 + 2,557,038) / 21,952,215 = 44.4%) from adversely selecting the other HFT subtypes. Panel A shows that all types of HFTs make the majority of their dollar profits from opportunistic traders. Panel B describes the profits and losses on a per contract basis. These results provide an estimate of the effective transaction costs involved in trading with certain groups. Since we compute profits on a 1 minute basis while resetting each trader's inventory to zero, their profits can be interpreted as short-term transaction costs extracted from the rest of the market, not gains from long-term directional positions. For example, Fundamental traders incur a loss of -$1.92 (Column 2, Row 5) when trading with Aggressive HFTs, while Small traders experience a much larger loss of -$3.49 (Column 2, Row 7). Interestingly, Small traders lose similar amounts per contract to non-HFT traders. Finally, the Total column in Panel A allows us to calculate the effective HFT transaction cost imposed on all other traders. Since HFTs make $35.23 million from all non-HFTs in August 2010 (= $21.95 million + $12.77 million + $0.50 million), while everyone else trades 31.96 million contracts (see Table 1), the effective HFT cost per trade on all other traders is $1.10 per contract. Scaling this per-contract transaction cost by $50,000 the approximate price of a contract yields an estimated HFT-imposed transaction cost of 0.0022% or 0.22 basis points.

5 Competition

We analyze competition in the HFT industry to show and understand why profits are concentrated among a small number of incumbents who realize high and persistent returns. According to the winner-takes-all idea, we expect to see a high concentration of profits (measured by the Herfindahl index) that persist over time. Combined with

the evidence from the previous section showing high firm-level persistence and strong outperformance of top-earning firms, this section provide evidence of an industry with profits concentrated among a small number of established firms.

In trying to understand the strong performance and persistence of returns to HFT firms, we also examine new entry. New entrants can potentially introduce competition and drive down both firm-level and industry profits. Given HFTs' high profitability, one might expect there to be strong incentives for entry, especially for firms that already have the technological capabilities and financial expertise to trade (market makers, hedge funds, broker-dealers, etc.), leading to increased competition and downward pressure on both firm-level and industry profits. However, in contrast to this expectation, we find that new entrants are substantially less profitable than incumbents and more likely to exit. Likely as a result, we further observe that industry concentration (as measured by the Herfindahl index) remains high and relatively constant over the two year sample. Also, aggregate profitability does not decrease, after adjusting for volatility and non-HFT trading volume. Combined with the evidence from the previous section showing high firm-level persistence and strong outperformance of top-earning firms, this section provides evidence of an industry dominated by a handful of established firms that are somehow able to keep the challengers at bay.

Our results are consistent with theoretical predictions regarding winner-takes-all competition based on speed. As mentioned previously, Budish, Cramton and Shim (2013) theorize that that if speed advantages are relative, then increased competition won't drive profit opportunities to zero, since HFTs can always one-up the competition with an ever-smaller increase in speed. Thus, aggregate profits and the concentration of profits would not decrease over time. Additionally, we expect entrants to earn substantially lower returns than established HFT firms and be more likely to exit, as presumably incumbents have advantages due to their experience allowing them to capture most of the profits.

5.1 HFT Market Concentration over Time

Here we show that the HFT industry concentration (as measured by the Herfindahl index of profits or trading volume) remains high and relatively constant over the two-year sample and also that aggregate profitability does not decrease, after adjusting for volatility and non-HFT trading volume. Our results indicate that the HFT industry exhibits concentration, in line with competition and new entry not decreasing the profitability of the most successful HFT firms. Combined with high firm-level persistence of returns, a high and steady Herfindahl index suggests that top performing incumbents maintain their position in the market.

We calculate Herfindahl indices, a commonly used measure of concentration of market share or earnings within an industry, for both profits and trading volume. The profit Herfindahl index is calculated as

$$Herfindahl_{i,t} = \sum_{i=1}^{N} \left[\frac{Profit_{i,t}}{HFT\,Total\,Profit_t} \right]^2 \tag{6}$$

where N is the number of firms in the HFT subgroup (Aggressive, Mixed, or Passive) in month t that earn non-negative profits, $Profit_{i,t}$ is firm i's total trading profits in month t, and $HFT\,Total\,Profit_t$ is the total trading profits of all HFT firms in the HFT subgroup. The volume Herfindahl index is calculated using the same formula but considering the number of contracts traded instead of trading profits. The Herfindahl index has a range of (0,1].

Table 8 reports the Herfindahl index for each subgroup of HFTs, first by calculating the index for each trading day and then averaging over days for every half-year block in the data set. Standard deviations are reported in parentheses, and stars indicate the significance from a t-test comparing the daily Herfindahl index in 2010 to the current half-year period.

A larger number implies a more concentrated industry, which is a proxy for the

level of competition in the industry. As a reference, Van Ness, Van Ness, and Warr (2005) find that market makers have an average Herfindahl index (created with trading volume, instead of profit) on NASDAQ of 0.14, with a range of 0.037 to 0.439. The results in Table 8 for both the Profit-based and Volume-based number are in line with this range. Perhaps more interesting, however, than the level is the direction. By performing a regression with time-based dummy variables, we show that Profit-based and Volume-based Herfindahl indices between 2010 and each subsequent half-year block are not decreasing, and in fact are somewhat increasing for Mixed and Passive HFTs.

For example, for Aggressive HFTs, the profit and volume Herfindahl indices are 0.362 and 0.2, respectively; these Herfindahl indices increase to 0.381 and decrease to 0.183 in the second half of 2011 and is not significant for other half-year blocks. In other words, there is no apparent directional trend. However, contrast those results with those of Passive HFTs, whose Profit Herfindahl index increases from 0.287 in 2010, trending steadily upward, to 0.545 in the second half of 2012; similarly the volume Herfindahl index of Passive HFTs increases from 0.129 in 2010, trending steadily upward, to 0.331 in the second half of 2012. We conclude that for Mixed and Passive HFTs, concentration of volume and profits actually increases between 2010 and 2012.

HFT is a rather new industry, and so one might expect concentration to decrease as new entrants join the market. We observe the opposite.

5.2 HFT Profitability over Time

In line with HFT industry concentration not decreasing over time, we find that HFT daily returns, profits and profits-per-contract have not fallen, and in some cases have actually increased over time, after taking into account market volatility and non-HFT trading volume.

Figure 2, Panel A shows that the aggregate monthly profits of HFTs vary substantially, though there is no clear directional trend. Aggressive HFTs consistently earn positive daily profit over the two-year span, with a mean of $395,875 per day over the two-year span, a high of $2,745,724 per day in August 2011 and a low of -$86,296 in January 2012. Mixed HFTs earn an even higher positive average daily profit over the two-year span (because there are more firms, even though each firm individually makes less), with a mean of $525,064 per day over the two-year span, a high of $890,828 per day in October 2010 and a low of $63,027 in January 2011. Passive HFTs earn the lowest average daily profit over the two-year span, with a mean of $107,239 per day over the two-year span, a high of $403,259 per day in January 2012 and a low of -$22,499 in February 2011.

The same is roughly true of profits per contract: Panel B shows the time-series of total monthly profits divided by total monthly contracts traded. Aggressive HFTs earn the highest average daily profits per contract over the two-year span, with a mean of $0.85 over the two-year span, a high of $2.01 in May 2011 and a low of $0.09 in January 2012. Mixed HFTs earn less with a mean of $0.52 per day over the two-year span, a high of $0.93 in May 2011 and a low of -$0.23 in January 2011. Passive HFTs earn the lowest with a mean of $0.46 over the two-year span, a high of $1.18 in January 2012 and a low of -$0.47 in February 2011.

More formally we examine the daily returns per firm, after controlling for daily volatility and non-HFT trading volume, which may affect returns on a daily or longer-term basis. To do this analysis, we estimated the following regression, where we use indicator variables to evaluate the average profits over half-year blocks:

$$\begin{aligned} r_{i,t} &= \alpha_{2010\,baseline} + \beta_1 1_{1st\,half\,2011} + ... + \beta_4 1_{2nd\,half\,2012} \\ &\quad + \beta_5 \log(volatility_t) + \beta_6 \log(nonHFT\,volume_t) + \epsilon_{i,t} \end{aligned} \tag{7}$$

where the indicator variables take the value of 1 during the corresponding half-year period, and controlling for daily log volatility (calculated in transaction time) and daily non-HFT trading volume.[15]

The coefficients are reported in Table 9 for average daily returns, daily profits and daily profit per contract for Aggressive, Mixed and Passive HFTs. The constant α is interpreted as the average HFT returns or profitability for 2010 (after controlling for volatility and non-HFT volume), and the coefficients β_1 through β_4 capture the cumulative increase or decrease in average returns or profits (after controlling for volatility and non-HFT volume) since 2010. According to, the baseline HFT returns for 2010 (α) are 0.219%, 0.105% and 0.0841% for Aggressive, Mixed and Passive HFTs, respectively; the baseline HFT average daily profits for 2010 (α) are \$68,100, \$58,800 and \$2,140 for Aggressive, Mixed and Passive HFTs, respectively; and the baseline HFT average daily profits per contract for 2010 (α) are \$1.74, \$0.872 and \$0.254 for Aggressive, Mixed and Passive HFTs, respectively.

Interestingly, the coefficients on log volatility are positive and statistically significant for Aggressive HFTs and negative or not significantly different from zero for Mixed and Passive HFTs, suggesting that Aggressive HFTs benefit from increased volatility (perhaps more opportunities to adversely select other traders), whereas Mixed and Passive HFTs are harmed by increased volatility (more opportunities to be adversely selected themselves). The coefficients on log non-HFT volume are positive but only sometimes statistically significant, suggesting that increasing that increased trading by non-HFT market participants increases returns and profits, especially for Aggressive HFTs, though may have no effect on returns and profits for Mixed and Passive HFTs.

The indicator coefficients, β_1 through β_4, test whether average daily returns and profitability decreases over the two year sample. In most cases, they are not statisti-

[15]Both controls (log volatility and log non-HFT volume) are reported as deviation from the average so that the constant in the regression can be interpreted as the average profitability in 2010 when volatility and volume are average.

cally significantly different from zero. When the coefficients are significant, they are positive for all but one coefficient, suggesting that profits per firm were constant or increasing from 2010 to 2012. The non-statistically significant coefficients are evenly divided between being positive and negative, suggesting that the issue isn't simply a lack of statistical power but that there is no downward trend on returns and profits over time. Taken together the evidence fails to support the hypothesis that HFT profits and returns have decreased over time. They appear to be highly variable day-to-day and month-to-month, but there is no trend either up or down in the aggregate or on the firm level.

5.3 Entry and Exit

We analyze entry and exit of HFTs into the E-mini market. Focusing on new entrants, we find that new entrants (especially Aggressive HFT entrants) underperform established HFTs and have a higher probability of exiting. We regress returns and profitability measures on indicator variables for length of time in the market, designating less than 1 month, 2 months or 3 months as indicators of new entry.[16] We exclude the observations in 2010 as this is when we first observe any firm. Formally, we estimate the following OLS regression equation:

$$r_{i,t} = \alpha + \beta_1 1_{one\,month\,i,t} + \beta_2 1_{two\,month\,i,t} + \beta_3 1_{three\,month\,i,t} + day\,FEs + \epsilon_{i,t} \qquad (8)$$

If new entrants are less competitive and profitable than established firms, the coef-

[16]Specifically, the 1 month old indicator variable takes the value 1 (otherwise 0) for firm i during days t to t + 30 if firm i began trading on day t-1. The 2 and 3 month old indicators likewise use t + 60 and t + 90 windows, respectively. In accounting for entry and exit, we look at the entire lifetime of each firm, not just the months for which the firm is classified as an HFT or as a specific type of HFT. Finally, we ignore gaps and just count the overall first and last trading days of a firm as entry and exit dates.

ficient on the one-month dummy should be negative; the two- and three-month dummy coefficients should also be negative but increasing, showing that experience in the market matters. Based on the regression specification, the constant is interpreted as average returns for established firms. In addition to daily returns, we also look at normalized returns, normalized profits and normalized profits per contract: coefficients in the normalized regressions are interpreted as how many standard deviations new entrants performance are above or below the mean.

The results are reported in Table 10, Panel A. Looking at daily returns (Columns 1-3), Aggressive HFTs have statistically significant negative coefficients corresponding to the new entry dummies, with the largest negative value for the 1-month dummy (-0.749%) and decreasing over time (-0.625% and -0.254%). By comparing these coefficients to the constant (0.665%), we can see that new entrants have negative returns for the first month and only slowly increase their profitability over the subsequent two months. Thus, new Aggressive HFTs show a large discrepancy in average daily returns between new entrants and established firms.

For Mixed HFTs, the coefficients are less negative and only significant for the 1-month dummy; comparing the 1-month coefficient (-0.065%) to the constant (0.144%), we see that new entrants who are Mixed HFTs still have positive returns, though less so than established firms. For Passive HFTs, the results are even weaker, suggesting that there is no disadvantage for new entrants who are Passive HFTs. Similar results hold when looking at normalized daily returns (Column 4-6), normalized daily profits (Columns 7-9) and normalized profits per contract (Columns 10-12): for Aggressive HFTs (and less so for Mixed HFTs) entrants are considerably less profitable than established firms, while for Passive HFTs entrants are associated with no such disadvantage.

For further evidence on the lack of competition, we examine whether new entrants are more likely to exit. We find that new Aggressive and Mixed HFTs have an approximately 1% per day increase in the probability of exit than more established HFTs. To

address the probability of exit, we perform a probit regression on each month separately to determine whether new entrants are more likely to exit:

$$Pr[Exit_{i,t} = 1] \; = \; \Phi[\alpha + \beta_1 1_{x-month\, i,t} + day\, FEs + \epsilon_{i,t}] \tag{9}$$

where $Exit_{i,t}$ takes the value 1 on day t for firm i if that is the last day firm i trades. The coefficients of interest are again those on the dummy variable for new entrants, x-month, which takes the value one, two, or three for either one-, two- or three-months after first appearing as an HFT, depending on the regression specification. We do not count exits in August 2012, due to it being the last month of the data set.

The marginal effects corresponding to the coefficients of the probit regression are reported in Table 10, Panel B. The Pseudo R^2 is reported below the coefficients. Column 1 shows that Aggressive HFTs who are new entrants (1 month) are much more likely to stop trading than their more experienced competitors, with a coefficient of 0.0186: a new entrant has an increased daily probability of exit of 1.86%, which is reasonably high considering that this is the increased probability of exit for *each day* (not the cumulative probability of exit at any future point). Aggressive probability of exit for new entry is statistically significant and positive for the first three months. Mixed HFT new entrants have heightened exit probabilities in months two and three. However, for Passive HFTs the marginal effects are not statistically significant.

In conclusion, new entrants are less profitable and more likely to exit than established HFTs. This finding helps explain the continued concentration of profits among a small subset of established firms: if new entrants are less profitable and tend to exit, then concentration of profits will continue or even increase among established firms. Furthermore, we speculate that new entrants' propensity to exit will limit the degree to which competitive forces reduce the profits of existing HFTs.

6 Speed

Why might HFT be dominated by a small number of incumbents whose past experience allows them to stay ahead of the competition and maintain their market share? In Section 4, we show that Aggressive HFTs outperform their Mixed and Passive counterparts. Here we focus on another important factor, relative (rank-order) speed, which we show to be an important determinant of profitability, especially for Aggressive HFTs.

We look at the relationship between speed and profits and show a positive association. Our measure of latency comes from Weller (2013): we measure *latency* of a firm as the 5th percentile of the duration between switching from a passive trade to an aggressive trade, measured in milliseconds. While this measure may not be ideal for every HFT, such as HFTs that trade all passively or all aggressively, it is a way to capture an active decision that is likely a response to market events. In contrast to other measures such as the duration between switching from an aggressive buy to an aggressive sell, which may suggest more about the timing of a firm's inventory management strategies, our latency measure aims to capture a firm's reaction speed to market events given its technological constraint.

We sort Aggressive, Mixed and Passive HFTs' firm-day observations into decile bins corresponding to *latency*. *Speed* is the reciprocal of the average latency (plus 1 millisecond, to ensure we do not divide by zero). Figure 3 plots the average of *Speed* versus average profits overall all firm-day observations within each decile bin.

Panel A shows *Speed* vs. average profits for Aggressive and Mixed HFTs (grouped together), and Panel B shows the same for Passive HFTs. The three lines trace the relationship between speed and profits at approximately one-year intervals (October 2010, July 2011, August 2012).[17] Figure 3 suggests a positive relationship between profits and *Speed*. In the months examined, there is a roughly increasing relationship

[17]Because millisecond time stamps were needed for this analysis, we are limited to which months we can analyze.

between profits and *Speed*.

The graphs, which trace the relationship between *Speed* and profits at approximately one-year intervals (October 2010, July 2011, August 2012), show that Speed is increasing over the two-year span. Between October 2010, July 2011 and August 2012, the average Speed of the top decile increases from approximately 0.4 to 0.75 to 1. To interpret, a value of 0.5 implies that an HFT switches from aggressive to passive, or vice versa, in one millisecond. A value of 1 implies that an HFT switches within the same millisecond. The speed of the second-to-top decile increases from approximately 0.27 to 0.32 to 1. There are similar increases in *Speed* for Passive HFTs. Although *Speed* increases over the two-year span, average profit fluctuates month-to-month with no clear trend over time.

For a formal analysis, we estimate the following regression from firm-day observations of HFT profitability:

$$\log(Profits_{i,t}) = \alpha + \beta_2 RelativeSpeed_{i,t} + time\,FEs + \epsilon_{i,t} \tag{10}$$

where $\log(Profits_{i,t})$ is a modified version of log profits, actually $sign(profits) \cdot \log(1 + |profits|)$, to allow for negative values, $RelativeSpeed_{i,t}$ is the firm's ranking that day in terms of speed among all firms of that subtype scaled by the total number of firms that day of that subtype. The results are reported in Table 11.

Table 11 shows that for all three subtypes of HFTs relative speed is important, with statistically significant coefficients of -6.72, -4.44 and -2.62 for Aggressive, Mixed, and Passive HFTs, respectively. The signs of the coefficients are as one would expect: profits are decreasing in relative speed (higher relative rank means higher profits). The coefficient is highest for Aggressive HFTs. Given that the left-hand side of the regression is in logs, this translates to a four order-of-magnitude higher change in profits for Aggressive HFTs than Passive HFTs, conditional in a one-rank improvement in speed. By showing that speed is an important determinant of profitability, our analysis suggests

that HFTs have strong incentives to compete over small increases in speed in an industry dominated by a small number of incumbents earning high and persistent returns.

7 Conclusion

We study the risk and return performance of HFT firms. We document several important descriptive statistics, many of which suggest superior investment performance of HFTs. HFTs earn Sharpe ratios that are several times higher than those for other asset classes or trader types. HFT returns are highly persistent, while risks are kept very low through tight inventory control and rapid turnover of contracts. HFT profits accumulate to the fastest and most aggressive liquidity-taking incumbents, while new entrants are less profitable and more likely to exit.

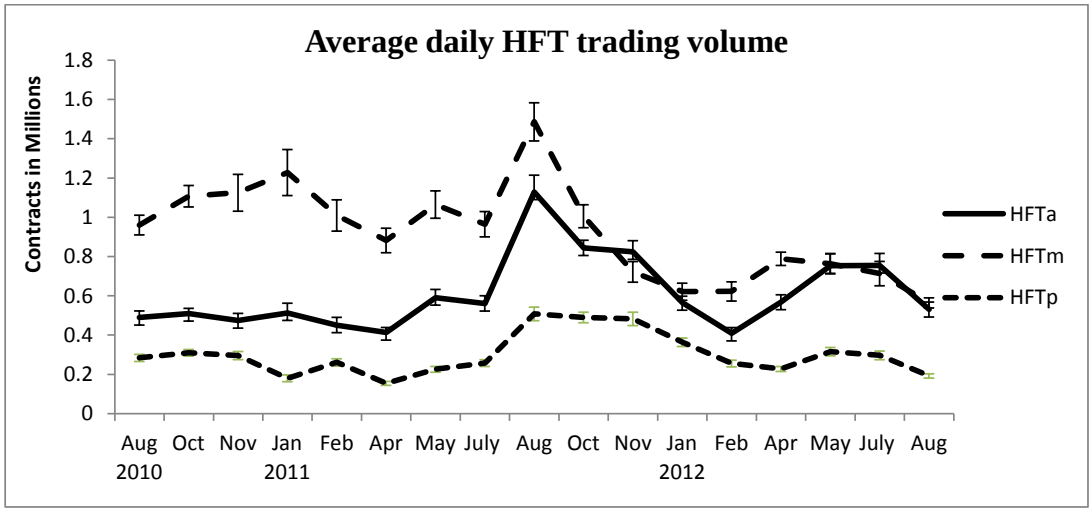
These facts highlight the importance of understanding the industrial organization of HFTs. Economists generally think that competition from new entrants will improve markets: there will be more liquidity, greater price efficiency, lower transaction costs for investors, and less potential for any one firm to influence markets. However, the cutthroat competitive environment in which HFTs interact may influence their impact on market quality. With limited competition from new entrants to engage incumbent HFTs, market quality may not improve as much as it would otherwise. Recent theoretical papers have highlighted concerns of faster traders adversely selecting slower traders and competition on speed leading to socially inefficient arms races for speed.

Our results suggest HFTs have strong incentives to take liquidity and compete over small increases in speed in an industry dominated by a small number of incumbents earning high and persistent returns. Understanding the industrial organization of HFTs allows researchers to think more comprehensively about the role of HFTs in financial markets.

Figure 1: Aggregate HFT trading activity

Panel A shows the average daily HFT volume of all HFT accounts combined (with standard error bars). Panel B shows the percent of market volume by trader type.

Panel A: HFT average daily trading volume



Panel B: Percent market volume by trader types

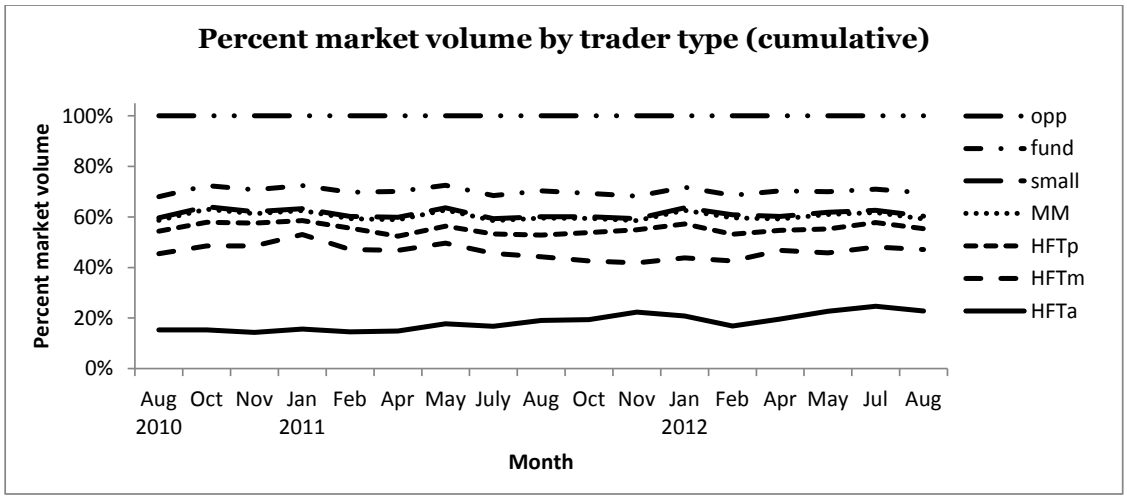
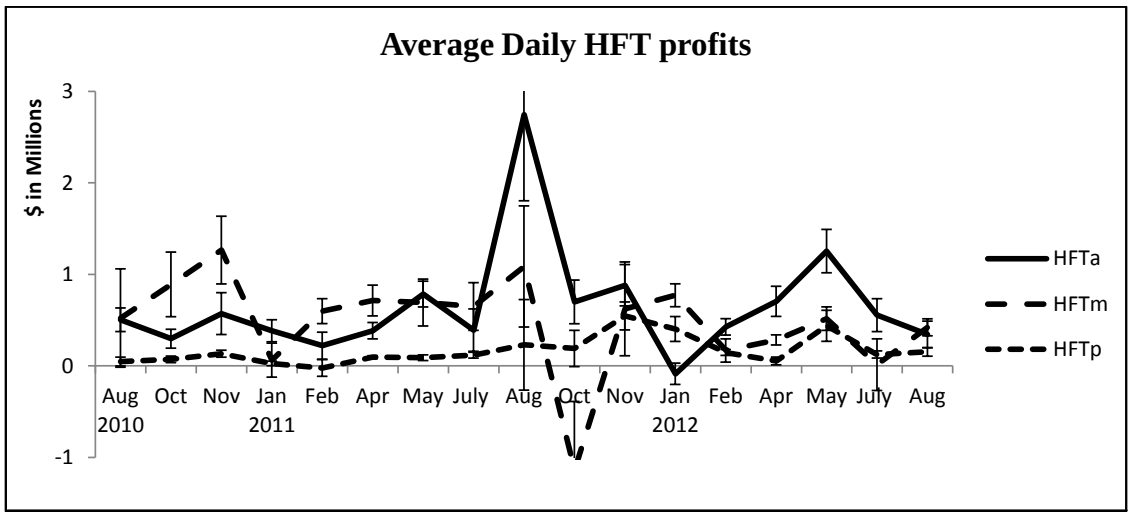


Figure 2: Aggregate HFT returns and profitability

Panel A shows the average daily HFT returns of all HFT accounts combined. Panel B shows average daily profits per contract, where the contract price is approximately $50,000. Standard error bars are calculated based on daily observations.

Panel A: Aggregate HFT profits



Panel B: Aggregate HFT profits per contract

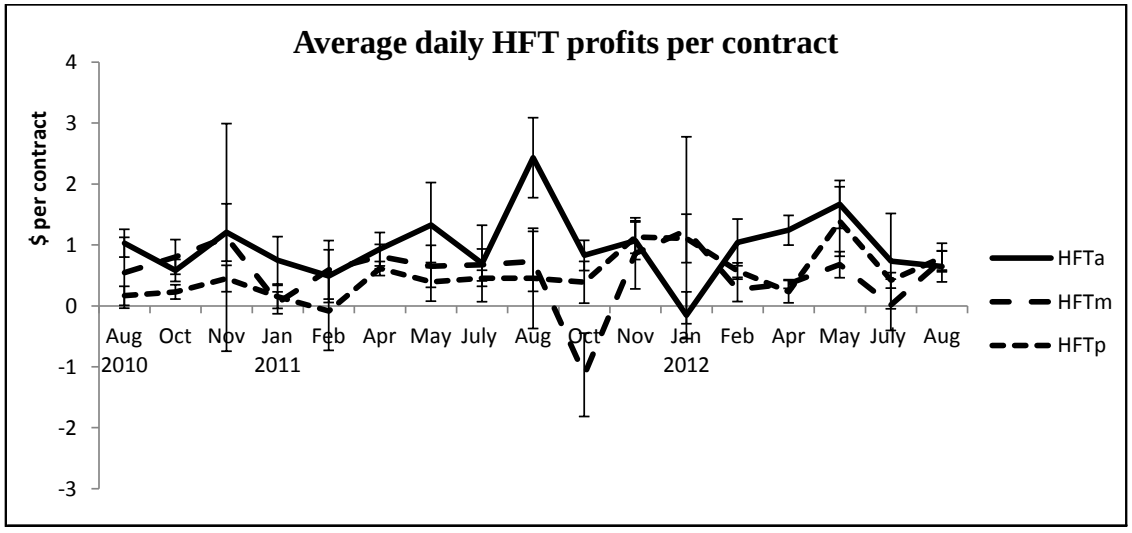
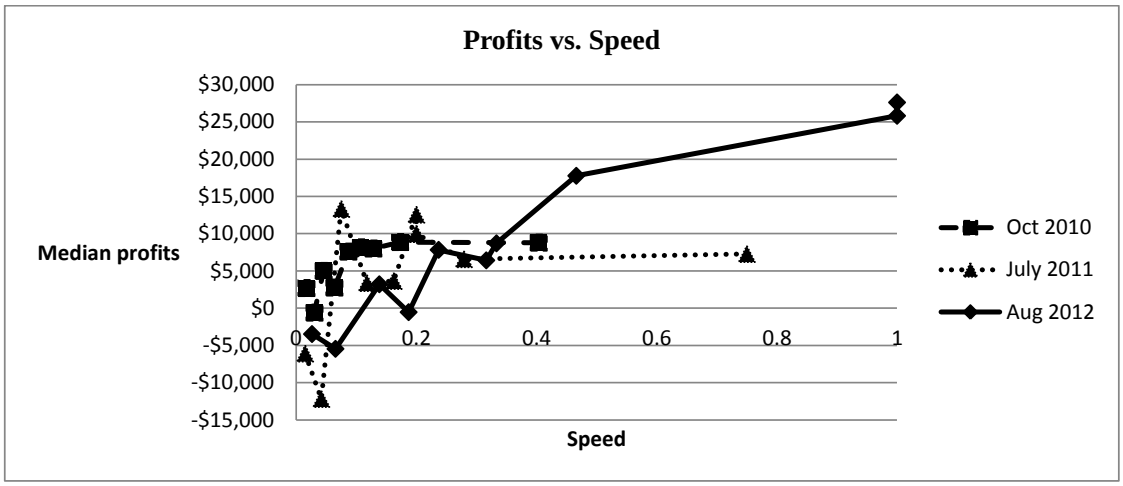


Figure 3: Speed and profitability

The figures shows the relationship between speed and profits. $Speed_{i,t} = (1+Latency_{i,t})^{-1}$, where $Latency_{i,t}$ is calculated as the 5th percentile (within each firm-day) of the time in milliseconds between successive trades that switch from a passive trade to an aggressive trade. Speed observations are grouped into decile bins, which are plotted against median profits within each bin. Panel A analyzes Aggressive and Mixed HFTs, and Panel B analyzes Passive HFTs. Within each panel, the relationship between speed and profits is plotted separately for three months for which we have millisecond data: October 2010, July 2011, and August 2012.

Panel A: Aggressive and Mixed HFT



Panel B: Passive HFT

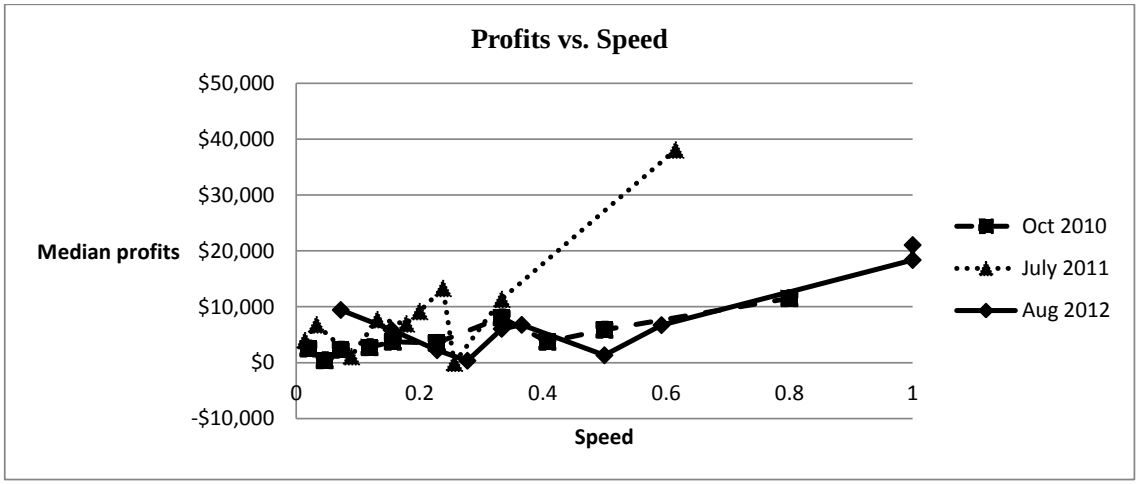


Table 1: Summary statistics of S&P 500 E-mini market, August 2010

The mean, standard deviation, minimum, median, and maximum of the daily trading activity of seven different trader types' trading activity are reported: Aggressive HFTs (HFT^A), Mixed HFTs (HFT^M), Passive HFTs (HFT^P), fundamental, small trader, non-HFT market maker, and opportunistic. Two different statistics are reported: *Daily % Market Volume* is the daily percent of market volume traded and *Daily Aggressive Ratio* is the daily fraction of contracts that are liquidity taking.

Daily % Market Volume	**Mean**	**Std. Dev.**	**Min.**	**Median**	**Max**
HFT^A (n=14)	15.22%	1.89%	9.50%	15.74%	17.57%
HFT^M (n=30)	30.28%	1.65%	27.46%	30.11%	33.94%
HFT^P(n=21)	8.87%	0.96%	6.36%	8.96%	10.93%
Fundamental (n=346)	8.42%	1.17%	6.97%	8.26%	12.69%
Small Trader (n=21761)	1.04%	0.11%	0.82%	1.02%	1.24%
Non-HFT Market Maker (n=737)	4.24%	0.37%	3.60%	4.31%	4.83%
Opportunistic (n=8494)	31.93%	1.40%	29.78%	31.77%	34.44%
Total (n=31403)	3,187,011	819,419	1,652,052	3,081,016	4,465,574
Daily Aggressive Ratio					
HFT^A	84.22%	1.90%	80.05%	84.26%	87.79%
HFT^M	37.08%	1.56%	34.66%	37.29%	40.08%
HFT^P	12.35%	1.45%	10.73%	11.83%	15.54%
Fundamental	57.68%	3.83%	48.68%	57.58%	64.90%
Small Trader	57.82%	1.25%	55.64%	57.53%	59.94%
Non-HFT Market Maker	12.98%	0.99%	11.45%	12.88%	15.12%
Opportunistic	59.08%	1.21%	56.80%	58.91%	61.89%

Table 2: HFT returns

The table reports statistics on the daily returns and profitability of HFT firms from August 2010 to August 2012. Each observation is a firm-day. Panel A reports returns and Panel B reports profits per contract (nominal contract value of approximately $50,000). The labels HFT^A, HFT^M, HFT^P correspond to Aggressive, Mixed, and Passive HFTs, respectively. Means, medians, and standard deviations are reported annualized, while skewness and kurtosis are based on daily observations. The p-value tests whether daily means are statistically significantly different from zero (tested over N=530 trading days).

Panel A: Annualized Returns (in %)

	N	Mean	Median	Std. Dev.	Skew.	Kurt.	P-Value
HFT (all)	13713	51.67	18.56	16.78	6.816	147.154	< 0.001
HFT^{A}	3274	122.10	25.15	28.39	5.679	65.525	< 0.001
HFT^{M}	6147	32.29	18.84	11.40	0.547	58.503	< 0.001
HFT^{P}	4292	25.69	15.35	8.89	6.402	249.666	< 0.001

Panel B: Profit Per Contract

	N	Mean	Median	Std. Dev.	Skew.	Kurt.	P-Value
HFT (all)	13713	$0.641	$0.519	$0.486	$0.125	$79.320	< .05
HFT^{A}	3274	$1.296	$0.792	$0.536	$5.556	$96.318	< .001
HFT^{M}	6147	$0.404	$0.466	$0.568	$-2.044	$46.798	0.15
HFT^{P}	4292	$0.481	$0.470	$0.268	$-3.986	$183.799	< 0.01

Table 3: Cross-sectional distribution of returns across firms

The table reports statistics on the cross-section of HFT firms who are active for at least three months. Returns, alphas, profits per contract (on a notional value of approximately $50,000), and Sharpe ratios (based on daily observations and reported in annualized terms) are first calculated for each firm; statistics on the cross-section of firms are reported below. The labels HFT^A, HFT^M, HFT^P correspond to Aggressive, Mixed, and Passive HFTs, respectively. Data-use restrictions prevent certain top and bottom percentiles from being reported.

	N	Mean	S.D.	10%	25%	50%	75%	90%
Returns (% annualized)								
HFT (all)	85	39.49	87.83	-0.35	8.07	20.40	51.14	74.13
HFT^A	18	91.08	170.55			46.64		
HFT^M	39	27.42	45.61			22.51		
HFT^P	28	23.13	22.24			15.02		
Alphas (% annualized)								
HFT (all), CAPM	85	39.92	88.81	-0.26	8.95	19.20	51.15	73.33
3-factor	85	39.74	89.43	-1.49	8.09	19.10	50.82	72.93
4-factor	85	39.78	89.21	-1.80	8.13	22.02	50.92	70.51
HFT^A 4-factor	18	90.67	174.74			37.06		
HFT^M 4-factor	39	28.18	44.90			23.15		
HFT^P 4-factor	28	23.22	22.24			14.88		
Profit per contract ($)								
HFT (all)	85	0.54	1.43	-0.16	0.28	0.46	0.82	2.01
HFT^A	18	0.65	1.56			0.65		
HFT^M	39	0.40	1.68			0.46		
HFT^P	28	0.66	0.91			0.41		
Sharpe Ratios (annualized)								
HFT (all)	85	5.25	5.01	-0.18	2.16	4.30	9.10	12.68
HFT^A	18	4.29	4.57			3.86		
HFT^M	39	5.26	5.75			4.38		
HFT^P	28	5.85	4.18			4.52		

Table 4: Persistence of returns

The table analyzes persistence in firms' returns and profits on both a daily and monthly level. The following regression is estimated: $r_{i,t} = \alpha + \beta \cdot r_{i,t-1} + Time\,FEs + \epsilon_{i,t}$. Regressions are also estimated in which the dependent variable (and lag) is normalized across firms for each day (or month) to account for the time-varying mean and variance of returns. With normalization, the regression is performed without time fixed effects as the normalization removes any average time trend, and the estimate of the constant, which is always zero, is not reported. Similar regressions are estimated with profits, profits-per-contract, and Sharpe ratios. Panel A analyzes persistence using daily observations, Panel B uses monthly observations, and Panel C considers persistence of rank ordering of firms on each day or month. *, ** and *** correspond to p-values ¡ 5%, 1%, 0.1%, respectively. Standard errors are in parentheses.

Panel A: Daily persistence

	Returns(i,t)			Normalized returns(i,t)			Normalized profits(i,t)			Normalized profits-per-contract(i,t)		
	HFT^A	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**
Dependent var (i,t-1)	.421*** (.0193)	.109*** (.0147)	-.073*** (.0167)	.276*** (.0195)	.169*** (.0147)	.205*** (.0174)	.16*** (.0199)	.141*** (.0148)	.227*** (.0173)	.111*** (.0202)	.093*** (.015)	.074*** (.0176)
Constant	.271*** (.0322)	.111*** (.0103)	.11*** (.0086)									
Adj-R^2	0.255	0.079	0.124	0.077	0.029	0.044	0.035	0.021	0.054	0.018	0.010	0.010
N	2703	4728	3458	2702	4724	3454	2702	4724	3454	2702	4724	3454

Panel B: Monthly persistence

	Returns(i,t)			Normalized returns(i,t)			Normalized profits(i,t)			Normalized profits-per-contract(i,t)			Normalized Sharpe ratio(i,t)		
	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P
Dependent var (i,t-1)	.723***	.407***	.725***	.809***	.449***	.699***	.416***	.33***	.616***	.418***	.296***	.52***	.639***	.659***	.452***
	(.0654)	(.0569)	(.0978)	(.0495)	(.0541)	(.0623)	(.0839)	(.0585)	(.0523)	(.0812)	(.0598)	(0.080)	(.0634)	(.0441)	(.0686)
Constant	.149*	.0779***	.0234												
	(.061)	(.0165)	(.0158)												
Adj-R^2	0.522	0.224	0.335	0.656	0.216	0.428	0.176	0.114	0.452	0.208	0.099	0.246	0.439	0.470	0.212
N	161	270	197	161	270	197	161	270	197	161	270	197	161	270	197

Panel C: Persistence of rank ordering

	Daily returns(i,t)			Monthly returns(i,t)			Monthly profits(i,t)			Monthly profits-per-contract(i,t)			Sharpe ratio(i,t)		
	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P	HFT^A	HFT^M	HFT^P
Dependent var (i,t-1)	0.376***	0.238***	0.278***	0.695***	0.474***	0.444***	0.679***	0.490***	0.402***	0.479***	0.328***	0.289***	0.629***	0.555***	0.479***
	(0.0152)	(0.0121)	(0.014)	(0.0504)	(0.0478)	(0.0551)	(0.0515)	(0.0474)	(0.0562)	(0.0616)	(0.0513)	(0.0588)	(0.0545)	(0.0452)	(0.0539)
Constant	3.868***	7.802***	5.678***	2.150***	5.775***	4.978***	2.265***	5.607***	5.347***	3.677***	7.384***	6.362***	2.614***	4.892***	4.666***
	(0.111)	(0.144)	(0.128)	(0.422)	(0.61)	(0.575)	(0.431)	(0.604)	(0.588)	(0.515)	(0.654)	(0.615)	(0.456)	(0.576)	(0.564)
Adj-R^2	0.141	0.056	0.077	0.481	0.223	0.194	0.458	0.238	0.159	R-sq	0.225	0.105	0.393	0.306	0.226
N	2703	4728	3458	161	270	197	161	270	197	161	270	197	161	270	197

Table 5: HFT inventories

Inventory statistics are reported for Aggressive HFTs (HFT^A), Mixed HFTs (HFT^M), and Passive HFTs (HFT^P). Median and standard deviations across firms are reported for three statistics: *Average Daily Volume* (in thousands), *Average end-of-day inventory* (both in number of contracts and as a percentage of the firm's average daily trading volume), and *Average intraday inventory range*. Medians (and s.d.) across firms are calculated by taking the average across days for each firm and then the median (and s.d.) across firms.

	Median (and s.d.) across firms		
	HFT^A	**HFT^M**	**HFT^P**
Average Daily Volume (thous.)	7,838	9,134	8,693
	(28,087)	(36,826)	(17,895)
Average end-of-day inventory	49.3	16.7	10.5
	(154.6)	(266.2)	(203.8)
- Normalized by firm's average	0.4%	0.2%	0.1%
daily trading volume	(4.6%)	(1.3%)	(1.3%)
Average intraday inventory range	525.9	313.6	204.2
	(861.1)	(1030.0)	(659.3)
- Normalized by firm's average	7.3%	4.0%	2.2%
daily trading volume	(5.8%)	(3.7%)	(3.4%)

Table 6: Spectral Analysis

The table analyzes trading profits in August 2010 over different time horizons using spectral analysis, following the methods of Hasbrouck and Sofianos (1993). We first compute the spectral decomposition of profits for each individual HFT firm and each trading day, decomposing profits over the following intervals: 1-10, 11-100, 101-1,000, 1,001-10,000, 10,000-100,000 and 100,000+ market transactions. For each HFT firm we calculate the median profit across days for each interval, and then report the median and the 25th and 75th percentiles (in brackets) across firms.

	Transaction interval					
	1-10	**11-100**	**101-1,000**	**1,001-10,000**	**10,001- 100,000**	**100,000+**
HFT^{A}	$870	-$678	-$5,348	$21,939	$22,108	-$8,637
	[$-2825, $8252]	[$-10887, $5997]	[$-45231, $12597]	[$-9633, $73428]	[$6213, $44481]	[$-19056, $4234]
HFT^{M}	$12,145	$23,171	$8,811	-$21,832	-$8,483	-$1,935
	[$7825, $19111]	[$12301, $35883]	[$-5835, $27894]	[$-36494, $-5288]	[$-13018, $2360]	[$-4179, $1452]
HFT^{P}	$5,236	$12,991	$11,408	-$7,917	-$9,774	-$3,428
	[$3840, $11170]	[$10174, $20701]	[$7920, $19186]	[$-14282, $-1512]	[$-20778, $-6990]	[$-9596, $-2509]

Table 7: Profit Breakdown

The table analyzes the decomposition of average daily short-term profits among different trader types in August 2010. The table is constructed by considering the trades between each pair-wise group and calculating the profit flows that result from those trades. We calculate each type's short-term profits (gains/losses realized on a trade one-minute after that trade). The rows identify who receives the profits, whereas the different columns represent from whom the profits are derived. Panel A analyzes average daily trading profits, and Panel B describes the profits on a per contract basis.

	Counterparty							
Profits to:	**HFT^{A}**	**HFT^{M}**	**HFT^{P}**	**Non-HFT MM**	**Fundamental**	**Opportunistic**	**Small**	**Total**
Panel A: Profits								
HFT^{A}	$0	$7,190,140	$2,557,038	$1,962,675	$1,691,738	$8,171,350	$379,275	$21,952,215
HFT^{M}	-$7,190,140	$0	$1,400,125	$1,607,400	$376,300	$15,358,325	$1,219,050	$12,771,060
HFT^{P}	-$2,557,038	-$1,400,125	$0	$84,300	-$373,438	$4,323,100	$425,538	$502,338
Non-HFT MM	-$1,962,675	-$1,607,400	-$84,300	$0	-$382,688	$1,060,000	$120,525	-$2,856,538
Fundamental	-$1,691,738	-$376,300	$373,438	$382,688	$0	$2,321,775	$119,550	$1,129,413
Opportunistic	-$8,171,350	-$15,358,325	-$4,323,100	-$1,060,000	-$2,321,775	$0	$339,150	-$30,895,400
Small Trader	-$379,275	-$1,219,050	-$425,538	-$120,525	-$119,550	-$339,150	$0	-$2,603,088
Total	-$21,952,215	-$12,771,060	-$502,338	$2,856,538	-$1,129,413	$30,895,400	$2,603,088	$0
Panel B: Profit/Loss Per Trade								
HFT^{A}	$0.00	$2.02	$1.93	$2.64	$1.92	$2.49	$3.49	$2.04
HFT^{M}	-$2.02	$0.00	$0.94	$1.95	$0.18	$1.92	$4.42	$0.60
HFT^{P}	-$1.93	-$0.94	$0.00	$0.75	-$0.62	$1.79	$5.05	$0.08
Non-HFT MM	-$2.64	-$1.95	-$0.75	$0.00	-$1.59	$1.11	$4.25	-$0.97
Fundamental	-$1.92	-$0.18	$0.62	$1.59	$0.00	$1.45	$2.87	$0.19
Opportunistic	-$2.49	-$1.92	-$1.79	-$1.11	-$1.45	$0.00	$2.05	-$1.38
Small Trader	-$3.49	-$4.42	-$5.05	-$4.25	-$2.87	-$2.05	$0.00	-$3.67
Total	-$2.04	-$0.60	-$0.08	$0.97	-$0.19	$1.38	$3.67	$0.00

Table 8: HFT Herfindahl Concentration

The table reports estimates on the Herfindahl index for each class of HFTs over successive six-month sample windows and analyzes whether market concentration is increasing or decreasing over time. We report a profit-based Herfindahl Index as well as a volume-based one. The profit Herfindahl index is calculated as $Herfindahl_{i,t} = \sum_{i=1}^{N} \left[\frac{Profit_{i,t}}{HFT\,Total\,Profit_t}\right]^2 \epsilon\,(0,1]$, and the volume Herfindahl is calculated similarly using trading volume instead of trading profits. A larger index implies a more concentrated industry. To assess whether market concentration is increasing or decreasing over time, the first-half of 2010 is used a baseline, and subsequence six-month sample periods are tested against the baseline. *, ** and *** correspond to p-values ¡ 5%, 1%, 0.1%, respectively. Standard errors are in parentheses.

	Profit Herfindahl Index			**Volume Herfindahl Index**		
	HFT^{A}	**HFT^{M}**	**HFT^{P}**	**HFT^{A}**	**HFT^{M}**	**HFT^{P}**
1st half 2010	0.362 (0.131)	0.287 (0.135)	0.287 (0.167)	0.2 (0.059)	0.239 (0.061)	0.129 (0.056)
1st half 2011	0.377 (0.202)	0.316 (0.164)	0.495*** (0.236)	0.195 (0.106)	0.26 (0.142)	0.329*** (0.104)
2nd half 2011	0.318* (0.155)	0.259 (0.135)	0.471*** (0.229)	0.183* (0.039)	0.174*** (0.136)	0.289*** (0.083)
1st half 2012	0.34 (0.178)	0.376*** (0.156)	0.513*** (0.286)	0.199 (0.054)	0.266 (0.132)	0.348*** (0.182)
2nd half 2012	0.347 (0.144)	0.412*** (0.141)	0.545*** (0.219)	0.199 (0.062)	0.248 (0.063)	0.331*** (0.129)

Table 9: Long-run profitability

The table analyzes long-run trends in HFT profitability by reporting estimates of the following OLS regression: $r_{i,t} = \alpha_{2010\,baseline} + \beta_1 1_{1st\,half\,2011} + ... + \beta_4 1_{2nd\,half\,2012} + \gamma' \cdot controls_t + \epsilon_{i,t}$. The notation $1_{six\,month\,window}$ corresponds to a series of indicators capturing six-month windows in 2011 and 2012; a positive coefficient on one of these indicators would correspond to an increasing trend in HFT firms' profits in comparison to the benchmark year 2010 (the first year in which we have data). We control for daily volatility and non-HFT-volume. *, ** and *** correspond to p-values ¡ 5%, 1%, 0.1%, respectively. Standard errors are in parentheses.

	Daily returns (%)			**Daily profits ($, thous.)**			**Daily profit-per-contract**		
	HFT^A	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**
Constant (2010 baseline)	.219***	.105***	.0841***	68.1***	58.8***	2.14	1.74***	.872*	.254
	(.0567)	(.0232)	(.0227)	(13.2)	(11.5)	(9.66)	(.303)	(.347)	(.195)
1st half of 2011 indicator	-.0233	.0258	-.032	.132	-17.8	-4.04	-.501	-.332	-.218
	(.0762)	(.0333)	(.0369)	(17.7)	(16.4)	(15.7)	(.407)	(.498)	(.317)
2nd half 2011 indicator	.0494	.0645*	.128***	7.36	-33.5*	35.4*	-.762	-.391	.627*
	(.0771)	(.0328)	(.0349)	(17.9)	(16.2)	(14.8)	(.412)	(.49)	(.299)
1st half 2012 indicator	.321***	.132***	.0632	21.8	-13.1	25.1	-.38	.016	.575
	(.0752)	(.0331)	(.0345)	(17.5)	(16.3)	(14.7)	(.402)	(.495)	(.296)
2nd half 2012 indicator	.353***	.0584	-.0719	10.3	-31.1	8.83	-.897	-.177	-.691
	(.0917)	(.0426)	(.0443)	(21.3)	(21)	(18.9)	(.49)	(.637)	(.38)
Log volatility$_t$	.209**	-.0331	-.0326	43.4**	-44.6**	-23.3	.892*	-.861	-.719*
(deviation from mean)	(.0669)	(.0334)	(.0344)	(15.6)	(16.5)	(14.6)	(.358)	(.5)	(.295)
Log non-HFT volume$_t$	.196*	.18***	.047	68.7***	39.7	3.49	.177	.813	.699
(deviation from mean)	(.0777)	(.0488)	(.0429)	(18.1)	(24)	(18.2)	(.415)	(.728)	(.367)
Adj-R^2	0.026	0.009	0.013	0.027	0.004	0.004	0.006	0.001	0.008
N	3195	4525	2679	3195	4525	2679	3195	4525	2679

Table 10: HFT Entry and Exit

The table analyzes the performance of new HFT entrants. Panel A reports the results of the following OLS regression: $r_{i,t} = \alpha + \beta_1 1_{one\,month\,i,t} + \beta_2 1_{two\,month\,i,t} + \beta_3 1_{three\,month\,i,t} + day\,FEs + \epsilon_{i,t}$. The notation $1_{one\,month\,i,t}$ takes the value 1 if firm i began trading in the last 30 days, otherwise 0; two and three month dummy variables are defined similarly. We exclude the observations in 2010 as this is when we first observe any firm. Panel B reports the results of the probit regression $Pr[Exit_{i,t} = 1] = \Phi[\alpha + \beta_1 1_{x-month\,i,t} + day\,FEs + \epsilon_{i,t}]$ where $Exit_{i,t}$ takes the value 1 on day t for firm i if that is the last day firm i trades. The coefficient of interest is again the dummy variable, $1_{x-month\,i,t}$, defined as above, but regressions are univariate (one indicator at a time, with x-month representing either one-, two-, or three- months) in contrast to Panel A. In addition to the excluded observations described in Panel A, this regression excludes observations in August 2012, the last month of the analysis, due to this being the last month of the data set. *, ** and *** correspond to p-values ¡ 5%, 1%, 0.1%, respectively. Standard errors are in parentheses.

Panel A: Profitability of New Entrants

	Daily returns(i,t)			Normalized daily returns(i,t)			Normalized daily profits(i,t)			Normalized profits-per-contract(i,t)		
	HFT^A	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**	**HFT^A**	**HFT^M**	**HFT^P**
< 1 month old indicator	-.749***	-.0653*	-.0169	-.479***	-.144***	-.0283	-.396***	-.173***	-.0394	-.144**	.0047	.003
	(.0858)	(.0276)	(.0224)	(.0511)	(.039)	(.0419)	(.0513)	(.039)	(.0419)	(.0517)	(.0391)	(.0419)
< 2 month old indicator	-.625***	-.0605	.0176	-.316***	-.132*	.128*	-.179	-.154*	-.0237	-.0356	-.0388	.0254
	(.16)	(.044)	(.0325)	(.0953)	(.0622)	(.0607)	(.0957)	(.0622)	(.0607)	(.0964)	(.0623)	(.0607)
< 3 month old indicator	-.254*	.006	-.0734*	-.103	.0418	-.0032	-.0357	-.121*	-.112*	-.0034	.0082	-.0886
	(.12)	(.0338)	(.029)	(.0715)	(.0478)	(.0542)	(.0718)	(.0478)	(.0542)	(.0724)	(.0479)	(.0542)
Constant	.665***	.144***	.113***	.138***	.0366*	-0.0063	.103***	.066***	0.0323	0.0345	0.001	0.0104
	(0.04)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)
Adj-R^2	0.134	0.074	0.113	0.026	0.003	0.001	0.018	0.004	0.001	0.002	0.000	0.001
N	3695	6433	4741	3694	6429	4737	3694	6429	4737	3694	6429	4737

Panel B: Probability of Exit, marginal effects

	HFT^A		HFT^M		HFT^P	
< 1 month old indicator	0.0186**	(0.009)	0.0046	(0.005)	0.0017	(0.005)
Pseudo R^2	0.057		0.0076		0.0032	
< 2 month old indicator	0.0204**	(0.008)	0.0123*	(0.007)	0.0003	(0.004)
Pseudo R^2	0.076		0.053		0.0026	
< 3 month old indicator	0.0184***	(0.007)	0.0082*	(0.005)	0.0024	(0.003)
Pseudo R^2	0.077		0.053		0.0055	

Table 11: Speed and Profits

The table reports coefficients from the regression: $\log(Profits_{i,t}) = \alpha + \beta_2 RelativeSpeed_{i,t} + time\ FEs + \epsilon_{i,t}$, where $\log(profits)_{i,t}$ is a modified version of log profits, namely $sign(profits) * \log(1 + |profits|)$ to allow for negative values. $RelativeSpeed_{i,t}$ is the firm's ranking that day in terms of speed among all firms of that sub-type scaled by the total number of firms that day of that sub-type. *, ** and *** correspond to p-values ¡ 5%, 1%, 0.1%, respectively. Standard errors are in parentheses.

	HFT^{A}	**HFT^{M}**	**HFT^{P}**
Relative Speed	-6.72***	-4.44***	-2.62***
	(1.3)	(0.73)	(0.77)
N	730	1693	1334
Adj-R^2	0.035	0.020	0.009

References

1. Acharya, Viral V., Deniz Anginer, and A. Joseph Warburton. "The End of Market Discipline? Investor Expectations of Implicit Government Guarantees." Working Paper (2014).

2. Acharya, Viral V., Hanh Le, and Hyun Song Shin. "Bank capital and dividend externalities." Working Paper (2013).

3. Acharya, Viral V., Irvind Gujral, Nirupama Kulkarni, and Hyun Song Shin. "Dividends and Bank Capital in the Financial Crisis of 2007-2009." Working Paper (2012).

4. Acharya, Viral V., Itamar Drechsler, and Philipp Schnabl. "A pyrrhic victory? Bank bailouts and sovereign credit risk." *Journal of Finance*, forthcoming (2014).

5. Acharya, Viral, Thomas Cooley, Matt Richardson and Ingo Walter. "Manufacturing tail risk: A perspective on the financial crisis of 2007-09." *Foundations and Trends in Finance* 4 (2010): 247-325.

6. Ackermann, C., R. McEnally, and D. Ravenscraft. "The performance of hedge funds: Risk, return, and incentives." *Journal of Finance* 54.3 (1999): 833-874.

7. Admati, Anat R., Peter M. DeMarzo, Martin F. Hellwig, Paul Pfleiderer. "Fallacies, Irrelevant Facts, and Myths in the Discussion of Capital Regulation: Why Bank Equity Is Not Expensive." Working paper (2013).

8. Admati, Anat R., Peter M. DeMarzo, Martin F. Hellwig, Paul Pfleiderer. "The Leverage Ratchet Effect." Working Paper (2013).

9. Adrian, Tobias, and Hyun Song Shin. "Procyclical leverage and value-at-risk." *Review of Financial Studies*, forthcoming (2013).

10. Adrian, Tobias, and Nina Boyarchenko. "Intermediary leverage cycles and financial stability." Staff Report, Federal Reserve Bank of New York (2012).

11. Adrian, Tobias, Emanuel Moench, and Hyun Song Shin. "Leverage asset pricing." Staff Report, Federal Reserve Bank of New York (2013).

12. Adrian, Tobias, Erkko Etula, and Tyler Muir. "Financial intermediaries and the cross section of asset returns." *Journal of Finance* 69.6 (2014): 2557-2596

13. Ait-Sahalia, Y. and M. Saglam. "High-Frequency Traders: Taking Advantage of Speed." Working Paper (2013).

14. Allen, Franklin and Douglas Gale. "Bubbles and crises." *Economic Journal* 110 (2000): 236-255.

15. Allen, Franklin, Elena Carletti, and Robert Marquez. "Credit market competition and capital regulation." *Review of Financial Studies* 24.4 (2011): 983-1018.

16. Anand, A., P. Irvine, A. Puckett, and K. Venkataraman. "Performance of institutional trading desks: An analysis of persistence in trading costs." *Review of Financial Studies* 25.2 (2012): 557-598.

17. Badrinath, S., J. Kale, and T. Noe. "Of shepherds, sheep, and the cross-autocorrelations in equity returns." *Review of Financial Studies* 8.2 (1995): 401-430.

18. Baker, Malcolm P., and Jeffrey Wurgler. "Would stricter capital requirements raise the cost of capital? Bank capital regulation and the low risk anomaly." Working Paper (2013).

19. Baker, Malcolm, and Jeffrey Wurgler. "Market timing and capital structure." *Journal of Finance* 57.1 (2002): 1-32.

20. Balasubramnian, Bhanu, and Ken B. Cyree. "Market discipline of banks: Why are yield spreads on bank-issued subordinated notes and debentures not sensitive to bank risks?" *Journal of Banking & Finance* 35.1 (2011): 21-35.

21. Bansal, Ravi and Amir Yaron. "Risks for the long run: A potential resolution of asset pricing puzzles." *Journal of Finance* 59 (2004): 1481-1509.

22. Barber, B., T. Odean, and N. Zhu. "Do retail trades move markets?" *Review of Financial Studies* 22.1 (2009): 151-186.

23. Barberis, Nicholas. "Psychology and the financial crisis of 2007-2008." in *Financial Innovation and Crisis*, M. Haliassos ed., MIT Press (2012).

24. Barro, Robert. "Rare disasters and asset markets in the twentieth century." *Quarterly Journal of Economics* 121 (2006): 823-866.

25. Bebchuk, Lucian, Alma Cohen and Holger Spamann. "The wages of failure: Executive compensation at Bear Stearns and Lehman 2000-2008." *Yale Journal on Regulation* 27 (2010): 257-282.

26. Benabou, Roland. "Groupthink: Collective delusions in organizations and markets." *Review of Economic Studies* 80 (2013): 429-462.

27. Bernanke, Ben and Mark Gertler. "Agency costs, net worth, and business fluctuations." *American Economic Review* 79 (1989): 14-31.

28. Bernanke, Ben S. "Nonmonetary Effects of the Financial Crisis in Propagation of the Great Depression." *American Economic Review* 73.3 (1983): 257–76.

29. Biais, B., T. Foucault, and S. Moinas. "Equilibrium Fast Trading." Working Paper (2014).

30. Biais, Bruno, Jean-Charles Rochet, and Paul Woolley. "The dynamics of innovation and risk." *Review of Financial Studies*, forthcoming (2014).

31. Boehmer, E. and J. Wu. "Order flow and prices." Working Paper (2008).

32. Boehmer, E., K. Fong, and J. Wu. "International Evidence on Algorithmic Trading." Working Paper (2014).

33. Bolton, Patrick, Jose Scheinkman and Wei Xiong. "Executive compensation and short-termist behavior in speculative markets." *Review of Economic Studies* 73 (2006): 577-610.

34. Borio, Claudio and Philip Lowe. "Asset prices, financial and monetary stability: exploring the nexus." Working paper (2002).

35. Boulatov, A., T. Hendershott, and D. Livdan. "Informed trading and portfolio returns." *Review of Economic Studies* 80.1 (2013): 35-72..

36. Brogaard, J., T. Hendershott, and R. Riordan. "High frequency trading and price discovery." *Review of Financial Studies*, forthcoming (2013).

37. Brown, Craig O., and I. Serdar Dinc. "Too many to fail? Evidence of regulatory forbearance when the banking sector is weak." *Review of Financial Studies* 24.4 (2011): 1378-1405.

38. Brunnermeier, Markus and Lasse Heje Pedersen. "Market Liquidity and Funding Liquidity." *Review of Financial Studies* 22 (2009): 2201-2238.

39. Brunnermeier, Markus and Yuliy Sannikov. "A Macroeconomic Model with a Financial Sector." *American Economic Review* 104 (2014): 379-421.

40. Budish, E., P. Cramton, and J. Shim. "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response." Working Paper (2013).

41. Calomiris, Charles W., and Charles M. Kahn. "The role of demandable debt in structuring optimal banking arrangements." *American Economic Review* (1991): 497-513.

42. Calomiris, Charles W., and Joseph R. Mason. "Consequences of bank distress during the Great Depression." *American Economic Review* (2003): 937-947.

43. Campbell, John and John Cochrane. "By force of habit: A consumption-based explanation of aggregate stock market behavior." *Journal of Political Economy* 107 (1999): 205-251.

44. Campbell, John and Motohiro Yogo. "Efficient tests of stock return predictability." *Journal of Financial Economics* 81 (2006): 27-60.

45. Campbell, John, Andrew Lo and Craig MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press (1996).

46. Carhart, M. "On persistence in mutual fund performance." *Journal of Finance* 52.1 (1997): 57-82.

47. Carow, Kenneth A., and Glen A. Larsen Jr. "The effect of FDICIA regulation on bank holding companies." *Journal of Financial Research* 20.2 (1997): 159-74.

48. Chen, Hui, Winston Dou, and Leonid Kogan. "Measuring the 'dark matter' in asset pricing models." Working paper (2013).

49. Cheng, Ing-Haw, Harrison Hong and Jose Scheinkman. "Yesterday's heroes: compensation and creative risk-taking." *Journal of Finance*, forthcoming (2013).

50. Cheng, Ing-Haw, Sahil Raina and Wei Xiong. "Wall Street and the housing bubble." *American Economic Review*, forthcoming (2013).

51. Cochrane, John. "Discount rates." *Journal of Finance* 66 (2011): 1047-1108.

52. Cole, S., M. Kanz, and L. Klapper. "Incentivizing calculated risk-taking: Evidence from an experiment with commercial bank loan officers." Working paper (2013).

53. Coughenour, J. and M. Saad. "Common market makers and commonality in liquidity." *Journal of Financial Economics*, 73.1 (2004): 37-69.

54. Danielsson, Jon, Hyun Song Shin and Jean-Pierre Zigrand. "Procyclical leverage and endogenous risk." Working paper (2012),.

55. DeAngelo, Harry, and Rene M. Stulz. "Why high leverage is optimal for banks." Working Paper (2013).

56. Demirguc-Kunt, A., and E. Detragiache. "Inside the crisis: An empirical analysis of banking systems in distress." *Journal of International Money and Finance* 25.5 (2006): 702-718.

57. Demirguc-Kunt, Asli, and Enrica Detragiache. "Does deposit insurance increase banking system stability? An empirical investigation." *Journal of Monetary Economics* 49.7 (2002): 1373-1406.

58. Demirguc-Kunt, Asli, and Harry Huizinga. "Bank activity and funding strategies: The impact on risk and returns." *Journal of Financial Economics* 98.3 (2010): 626-650.

59. Demirguc-Kunt, Asli, and Enrica Detragiache. "The Determinants of Banking Crises in Developing and Developed Countries." IMF Staff Papers 45.1 (1998): 81–109.

60. Diamond, Douglas W., and Raghuram G. Rajan. "A theory of bank capital." *Journal of Finance* 55.6 (2000): 2431-2465.

61. Diamond, Douglas W., and Raghuram G. Rajan. "Banks and liquidity." *American Economic Review* (2001): 422-425.

62. Dittmar, Amy, and Anjan Thakor. "Why do firms issue equity?" *Journal of Finance* 62.1 (2007): 1-54.

63. Duchin, Ran, and Denis Sosyura. "Safer ratios, riskier portfolios: Banks response to government aid." *Journal of Financial Economics* 113.1 (2014): 1-28.

64. Eichengreen, Barry, and Carlos Arteta. "Banking Crises in Emerging Markets: Presumptions and Evidence." In *Financial Policies in Emerging Markets* edited by Mario Blejer and Marko Skreb. Cambridge, Mass.: MIT Press (2002), 47–94.

65. Eichengreen, Barry, and Kris J. Mitchener. "The Great Depression as a Credit Boom Gone Wrong." Working Paper (2003).

66. Fahlenbrach, Rudiger, Robert Prilmeier, and Rene M. Stulz. "This time is the same: Using bank performance in 1998 to explain bank performance during the recent financial crisis." *Journal of Finance* 67.6 (2012): 2139-2185.

67. Fama, E. and K. French. "The equity premium." *Journal of Finance* 57.2 (2002): 637-659.

68. Fama, E., and French, K. "Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33.1 (1993): 3-56.

69. Fama, Eugene F., and Kenneth R. French. "Testing tradeoff and pecking order predictions about dividends and debt." *Review of Financial Studies* 15.1 (2002): 1-33.

70. Fisher, Irving. "The debt-deflation theory of great depressions." *Econometrica* 1 (1933): 337-357.

71. Flannery, Mark J., and Sorin M. Sorescu. "Evidence of bank market discipline in subordinated debenture yields: 1983-1991." *Journal of Finance* 51.4 (1996): 1347-1377.

72. Flannery, Mark, and Stanislava Nikolova. "Market discipline of US financial firms: Recent evidence and research issues." in *Market discipline across countries and industries* (2004): 87-100.

73. Foote, Christopher, Kristopher Gerardi and Paul Willen. "Why did so many people make so many ex-post bad decisions? The causes of the foreclosure crisis." Working paper (2012).

74. Foucault, T., J. Hombert, and I. Rosu. "News trading and speed." Working Paper (2013).

75. Gabaix, Xavier. "Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance." *Quarterly Journal of Economics* 127 (2012): 645-700.

76. Gai, J., M. Yao, and C. Ye. "The Externalities of High Frequency Trading." Working Paper (2013).

77. Gandhi, Priyank. "The relationship between credit growth and the expected returns of bank stocks." Working paper (2011).

78. Gandhi, Priyank, and Hanno Lustig. "Size Anomalies in US Bank Stock Returns." *Journal of Finance*, forthcoming (2013).

79. Geanakoplos, John. "The leverage cycle." in *NBER Macroeconomics Annual*, Volume 24. University of Chicago Press (2010): 1-65.

80. Gennaioli, Nicola, Andrei Shleifer and Robert Vishny. "Neglected risks, financial innovation, and financial fragility." *Journal of Financial Economics* 104 (2012): 452-468.

81. Gennaioli, Nicola, Andrei Shleifer and Robert Vishny. "A model of shadow banking." *Journal of Finance*, forthcoming (2013).

82. Gertler, Mark and Nobuhiro Kiyotaki. "Banking, Liquidity and Bank Runs in an Infinite Horizon Economy." Working Paper (2012)

83. Goyal, Amit and Ivo Welch. "A comprehensive look at the empirical performance of equity premium prediction." *Review of Financial Studies* 21 (2008): 1455-1508.

84. Graham, John R. "How big are the tax benefits of debt?" *Journal of Finance* 55.5 (2000): 1901-1941.

85. Greenwood, Robin and Samuel G. Hanson. "Issuer quality and corporate bond returns." *Review of Financial Studies* 26 (2013): 1483-1525.

86. Grinblatt, M. and M. Keloharju. "What makes investors trade?" *Journal of Finance* 56.2 (2001): 589-616.

87. Gromb, Denis and Dimitri Vayanos. "Equilibrium and welfare in markets with financially constrained arbitrageurs." *Journal of Financial Economics* 66 (2002): 361-407.

88. Hanson, Samuel G., Anil K. Kashyap, and Jeremy C. Stein. "A macroprudential approach to financial regulation." *Journal of Economic Perspectives* (2011): 3-28.

89. Harris, J. and P. Schultz. "The trading profits of SOES bandits." *Journal of Financial Economics* 50.1 (1998): 39-62.

90. Hasbrouck, J. and G. Saar. "Low-latency trading." *Journal of Financial Markets* 16.4 (2013): 646-679.

91. Hasbrouck, J. and G. Sofianos. "The trades of market makers: An empirical analysis of NYSE specialists." *Journal of Finance* 48.5 (1993): 1565-1593.

92. Hasbrouck, J. "Intraday price formation in US equity index markets." *Journal of Finance* 58.6 (2003): 2375-2400.

93. He, Zhiguo and Arvind Krishnamurthy, "A Model of Capital and Crises." *Review of Economic Studies* 79 (2012): 735-777.

94. He, Zhiguo and Arvind Krishnamurthy. "Intermediary Asset Pricing." *American Economic Review* 103 (2013): 732-70.

95. Hendershott, T., C. Jones, and A. Menkveld. "Does algorithmic trading increase liquidity?" *Journal of Finance* 66 (2011): 1-33.

96. Hendershott, T., D. Livdan, and N. Schurhoff. "Are institutions informed about news?" Working Paper (2012).

97. Heston, S., R. Korajczyk, and R. Sadka. "Intraday patterns in the cross-section of stock returns." *Journal of Finance* 65.4 (2010): 1369-1407.

98. Hirschey, N. "Do high-frequency traders anticipate buying and selling pressure?" Working Paper (2013).

99. Holmstrom, Bengt and Jean Tirole. "Financial intermediation, loanable funds and the real sector." *Quarterly Journal of Economics* 112 (1997): 663-91.

100. Hong, Harrison, Jiang Wang, and Jialin Yu. "Firms as buyers of last resort." *Journal of Financial Economics* 88.1 (2008): 119-145.

101. Huang, Sheng, and Anjan V. Thakor. "Investor heterogeneity, investor-management disagreement and share repurchases." *Review of Financial Studies* 26.10 (2013): 2453-2491.

102. Hvidkjaer, S. "Small trades and the cross-section of stock returns." *Review of Financial Studies* 21.3 (2008): 1123-1151.

103. Ivashina, Victoria and David Scharfstein. "Bank lending during the financial crisis of 2008." *Journal of Financial Economics* 97 (2010): 319-338.

104. Jagannathan, Murali, Clifford P. Stephens, and Michael S. Weisbach. "Financial flexibility and the choice between dividends and stock repurchases." *Journal of Financial Economics* 57.3 (2000): 355-384.

105. Jagannathan, R., A. Malakhov, and D. Novikov. "Do hot hands exist among hedge fund managers? an empirical evaluation." *Journal of Finance* 65.1 (2010): 217-255.

106. Jimenez, Gabriel, et al. "Macroprudential policy, countercyclical bank capital buffers and credit supply: Evidence from the Spanish dynamic provisioning experiments." Working Paper (2012).

107. Jones, C. "What do we know about high-frequency trading?" Working Paper (2013).

108. Jovanovic, B. and A. Menkveld. "Middlemen in limit-order markets." Working Paper (2012).

109. Kaminsky, Graciela L. and Carmen M. Reinhart. "The twin crises: The causes of banking and balance of payments problems." *American Economic Review* 89 (1999): 473-500.

110. Kaniel, R., G. Saar, and S. Titman. "Individual investor trading and stock returns." *Journal of Finance* 63.1 (2008): 273-310.

111. Kaplan, S. and A. Schoar. "Private equity performance: returns, persistence, and capital flows." *Journal of Finance* 60.4 (2004): 1791-1823.

112. Kashyap, Anil K., Raghuram Rajan, and Jeremy C. Stein. "Banks as liquidity providers: An explanation for the coexistence of lending and deposit taking." *Journal of Finance* 57.1 (2002): 33-73.

113. Kashyap, Anil, Jeremy Stein, and David Wilcox. "Monetary policy and credit conditions: Evidence from the composition of external finance." *American Economic Review* 83 (1993): 78-98.

114. Kashyap, Anil, Raghuram Rajan, and Jeremy Stein. "Rethinking capital regulation." in Maintaining stability in a changing financial system (2008): 431-471.

115. Keeley, Michael C. "Deposit insurance, risk, and market power in banking." *American Economic Review* (1990): 1183-1200.

116. Kelly, Bryan T., Hanno Lustig, and Stijn Van Nieuwerburgh. "Too-systemic-to-fail: What option markets imply about sector-wide government guarantees." Working Paper (2011).

117. Keys, Benjamin, Tanmoy Mukherjee, Amit Seru and Vikrant Vig. "Did securitization lead to lax screening? Evidence from subprime loans." *Quarterly Journal of Economics* 125 (2010): 307-362.

118. Kho, Bong-Chan, Dong Lee, and Rene M. Stulz. "U.S. Banks, Crises, and Bailouts: From Mexico to LTCM." *American Economic Review* 90.2 (2000): 28-31.

119. Kindleberger, Charles. *Manias, Panics, and Crashes: A History of Financial Crises*. New York: Basic Books (1978).

120. Kirilenko, A., A. Kyle, M. Samadi, and T. Tuzun. "The flash crash: the impact of high frequency trading on an electronic market." Working Paper (2014).

121. Kisin, Roni, and Asaf Manela. "The shadow cost of bank capital requirements." Working Paper (2013).

122. Kiyotaki, Nobu and John Moore. "Credit cycles." *Journal of Political Economy* 105 (1997): 211-48.

123. Kyle, Albert and Wei Xiong. "Contagion as a wealth effect." *Journal of Finance* 56 (2001): 1401-1440.

124. Laeven, Luc, and Fabian Valencia. "Systemic banking crises: A new database." Working paper (2008).

125. Lane, Philip R. and Gian Maria Milesi-Ferretti. "The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities," 1970–2004." *Journal of International Economics* 73 (2007): 223-250.

126. Lettau, Martin and Sydney Ludvigson. "Measuring and Modeling Variation in the Risk-Return Tradeoff." in *Handbook of Financial Econometrics*, edited by Yacine Ait-Sahalia and Lars-Peter Hansen (2010): 617-690.

127. Long, J. *Regression models for categorical and limited dependent variables*. Sage Publications, London (1997).

128. Malinova, K., A. Park, and R. Riordan. "Do retail traders suffer from high frequency traders?" Working Paper (2013).

129. Malmendier, Ulrike, and Geoffrey Tate. "CEO overconfidence and corporate investment." *Journal of Finance* 60.6 (2005): 2661-2700.

130. Mansur, Iobal, and Elyas Elyasuani. "An examination of the impact of the 1989 FIRREA on the market value of commercial banks and savings and loans." *Applied Financial Economics* 4.1 (1994): 11-22.

131. Marsh, Paul. "The choice between equity and debt: An empirical study." *Journal of Finance* 37.1 (1982): 121-144.

132. Martinez, V. and I. Rosu. "High frequency traders, news and volatility." Working Paper (2013).

133. Mehran, Hamid, and Anjan Thakor. "Bank capital and value in the cross-section." *Review of Financial Studies* 24.4 (2010): 1019-1067

134. Mendoza, Enrique and Marco Terrones. "An anatomy of credit booms: Evidence from macro aggregates and micro data." Working Paper (2008).

135. Menkveld, A. "High frequency trading and the new-market makers." *Journal of Financial Markets* 16.4 (2013): 712-740.

136. Mian, Atif and Amir Sufi. "The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis." *Quarterly Journal of Economics* 124 (2009): 1449-1496.

137. Miller, Merton H. "Debt and Taxes." *Journal of Finance* 32.2 (1977): 261-275.

138. Minsky, Hyman, "The financial instability hypothesis: An interpretation of Keynes and an alternative to 'standard' theory." *Nebraska Journal of Economics and Business* 16 (1977): 5-16.

139. Mishkin, Frederic. "The household balance sheet and the Great Depression." *Journal of Economic History* 38 (1978): 918-937.

140. Morgan, Donald P., and Kevin J. Stiroh. "Too big to fail after all these years." Staff Report, Federal Reserve Bank of New York (2005).

141. Muir, Tyler. "Financial crises and risk premia." Working paper (2014).

142. Myers, Stewart C. "Determinants of corporate borrowing." *Journal of Financial Economics* 5.2 (1977): 147-175.

143. Myers, Stewart C., and Nicholas S. Majluf. "Corporate financing and investment decisions when firms have information that investors do not have." *Journal of Financial Economics* 13.2 (1984): 187-221.

144. O'Hara, Maureen, and Wayne Shaw. "Deposit insurance and wealth effects: the value of being 'too big to fail'." *Journal of Finance* 45.5 (1990): 1587-1600.

145. Pagano, Marco, Fabio Panetta, and Luigi Zingales. "Why do companies go public? An empirical analysis." *Journal of Finance* 53.1 (1998): 27-64.

146. Panageas, Stavros. "Bailouts, the incentive to manage risk, and financial crises." *Journal of Financial Economics* 95.3 (2010): 296-311.

147. Petersen, Mitchell A. "Estimating standard errors in finance panel data sets: Comparing approaches." *Review of Financial Studies* 22.1 (2009): 435-480.

148. Piketty, Thomas and Gabriel Zucman. "Capital is back: Wealth-income ratios in rich countries, 1700-2010." *Quarterly Journal of Economics*, forthcoming (2014)

149. Puckett, A., and X. Yan. "The interim trading skills of institutional investors." *Journal of Finance* 66.2 (2011): 601-633.

150. Rajan, Raghuram. "Has finance made the world riskier?" *European Financial Management* 12 (2006): 499-533.

151. Rajan, Raghuram. *Fault Lines*. Princeton University Press (2010).

152. Reinhart, Carmen and Kenneth Rogoff. *This Time Is Different: Eight Centuries of Financial Folly*. Princeton University Press (2009).

153. Rietz, Thomas. "The equity risk premium: A solution." *Journal of Monetary Economics* 22 (1988): 117-131.

154. Riordan, R. and A. Storkenmaier. "Latency, liquidity and price discovery." *Journal of Financial Markets* 15.4 (2012): 416–437.

155. Schularick, Moritz and Alan Taylor. "Credit booms gone bust: monetary policy, leverage cycles and financial crises, 1870–2008." *American Economic Review* 102 (2012): 1029-1061.

156. Seasholes, M., and N. Zhu. "Individual investors and local bias." *Journal of Finance* 65.5 (2010): 1987-2010.

157. Sharpe, W. "Mutual fund performance." *Journal of Business* 39.1 (1966): 119-138.

158. Shleifer, Andrei and Robert Vishny. "The Limits of Arbitrage." *Journal of Finance* 52 (1997): 35-55.

159. Simsek, Alp. "Belief disagreements and collateral constraints." *Econometrica* 81.1 (2013): 1-53.

160. Sironi, Andrea. "Testing for market discipline in the European banking industry: Evidence from subordinated debt issues." *Journal of Money, Credit, and Banking* 35.3 (2003): 443-472.

161. Stambaugh, Robert. "Predictive regressions." *Journal of Financial Economics* 54 (1999): 375-421.

162. Stein, Jeremy. "Rational capital budgeting in an irrational world." *Journal of Business* 69 (1996): 429-455.

163. Stephens, Clifford P., and Michael S. Weisbach. "Actual share reacquisitions in open market repurchase programs." *Journal of Finance* 53.1 (1998): 313-333.

164. Strahan, Philip E. "Too Big to Fail: Causes, Consequences, and Policy Responses." *Annu. Rev. Financ. Econ.* 5.1 (2013): 43-61.

165. Thompson, Samuel B. "Simple formulas for standard errors that cluster by both firm and time." *Journal of Financial Economics* 99.1 (2011): 1-10.

166. Van Ness, B., R. Van Ness, and R. Warr. "The impact of market-maker concentration on adverse selection costs for NASDAQ stocks." *Journal of Financial Research* 28.3 (2005): 461-485.

167. Wachter, Jessica. "Can time-varying risk of rare disasters explain aggregate stock market volatility?" *Journal of Finance* 68 (2013): 987-1035.

168. Weitzman, Martin. "Subjective expectations and asset-return puzzles." A*merican Economic Review* 97 (2007): 1102-1130.

169. Weller, B. "Intermediation chains and specialization by speed: evidence from commodity futures markets." Working Paper (2013).

170. Xiong, Wei. "Convergence trading with wealth effects: An amplification mechanism in financial markets." *Journal of Financial Economics* 62 (2001): 247-292.