Bank Capital and Risk Taking: A Loan Level Analysis*

Michael Ohlrogge †, ‡

August 28, 2017

Abstract

Does high leverage incentivize banks to systematically originate and hold riskier loans? I construct a novel data set consisting of 1.8 million small business and home mortgage loans, matched to the specific banks that originated them and verified to be held on bank portfolios, rather than sold. I measure the capital ratio (the inverse of the leverage ratio, defined as equity divided by asset value) for each bank at the time of each loan's origination. After controlling for both bank and time fixed effects, a one point increase in Tier 1 capital ratios (e.g. from 12% to 13%) is associated with a 4% decrease in the default risk of mortgage loans held on portfolio (from a mean foreclosure rate of 2.5% to 2.4% for loans originated between 2003 and 2012). When considering the average capital of banks in US counties between 2003 and 2006, a one point increase in Tier 1 capital ratios is associated with a 4.4% reduction in foreclosures between 2007 and 2012. These results are robust to an instrumental variables strategy for predicting bank capital, a wide range of measures of bank capital, different types of banks, types of loans, and time periods.

Keywords: Bank Capital, Asset Substitution, Portfolio Risk, Mortgages, Small Business Loans
JEL Classification: G11, G21, G28, G32, G38

*I would like to thank the Center for Financial and Risk Analytics (CFRA) and the John M. Olin Program in Law and Economics at Stanford University for providing financial support for this project. I am particularly indebted to Kris Kasianovitz and the Stanford Green Library for funding and assistance with acquiring data for this project, and to Paul Lomio, the Stanford Law School Library, and the John M. Olin Program for providing additional support for data acquisition. Anat Admati, Kay Giesecke, Markus Pelger, and Amit Seru have provided many helpful comments on drafts of this paper. I also thank the participants in the internal finance seminar at the Department of Management Science and Engineering at Stanford for feedback. Any errors contained in this work are entirely my own.

†Stanford University, Department of Management Science & Engineering
‡Stanford Law School

1 Introduction

Economists have long recognized that the amount of debt a firm uses to finance its operations can have a significant impact on the investment decisions that firm makes, the risk of those investments, and whether such investments are optimal from a private or social perspective (Jensen and Meckling, 1976; Myers, 1977). Banks are amongst the most heavily debt-funded firms in the economy, and enjoy unique government guarantees of much of that debt, through both implicit and explicit channels (Kelly et al., 2016; Gandhi and Lustig, 2015). Furthermore, as has been shown for instance in the recent financial crisis, banks have a significant impact on other sectors of the economy based on the investments they make and the consequences of those investments. Unsurprisingly then, theorists have paid considerable attention to modeling how a bank's levels of debt relative to equity, and government guarantees of that debt, will impact the risk of the investments a bank makes (Merton, 1977). The conclusions of these models have been mixed, however, with some predicting that an increase in debt financing by banks will lead to an increase in portfolio risk (Buser et al., 1981; Hellmann et al., 2000) with others predicting the opposite (Kim and Santomero, 1988; Chen et al., 2017).

The goal of this paper is to empirically test for a causal relationship between high leverage and high portfolio risk in banks, using a novel data set that connects millions of individual small business and home mortgage loans to the banks that originated them. Understanding the relation between leverage and risk-taking has implications for both the stability of the banking system as well as its investing impacts on the rest of the economy. Regulations govern bank capital ratios (equity divided by assets, or the inverse of leverage ratios). Frequently, capital is seen as important because of its role in absorbing losses and thus preventing insolvency. Yet, if shifts in bank capital also produce a dynamic change in the portfolio risk of banks, then, for instance, a reduction in bank capital requirements may increase the risk of bank failure both by reducing the loss-absorbing buffer of capital and by increasing the likelihood of large bank losses in the first place due to increased portfolio risk.

The relation between bank capital and risk is also important for understanding the economic efficiency of bank investing decisions. The asset substitution problem explored by Jensen and Meckling

(1976) predicts that decreases in capital can lead a bank to invest in assets that simultaneously have higher risk but lower net present value. Because of the crucial role that banks play in allocating trillions of dollars of credit in the global economy, understanding how bank capital impacts the efficiency of these investments is important for understanding the efficiency of the economy as a whole.

Prior empirical efforts to investigate the relation between bank capital and portfolio risk have been hampered by a lack of detailed data on the portfolio investments banks make. As private, competitive corporations, banks are understandably not keen to publicly release information on the specific assets they invest in or the risk of those assets. Instead, economists studying this question have been forced to rely on broad regulatory disclosures which give only high-level summaries of banks' current total portfolios and rough proxies for portfolio-wide risk, such as the percentage of assets in different regulatory risk weight categories (commercial loans, Treasury securities, etc.). This regulatory data, however, contains no information on specific assets, when such assets were acquired, and what levels of capital banks had at the time of acquisition. As such, a need for more precise information on bank assets in order to make progress on this question has been recognized in the literature (Gorton and Rosen, 1995).

This paper contributes to this debate by analyzing data with an unprecedented level of detail on the assets of banks and the risks of those assets. In particular, I analyze roughly 600,000 small business loans and 1.2 million home mortgage loans made by banks in the United States between the years of 2002 and 2013. These are loans held on the balance sheets of banks, rather than those sold via securitization or other channels. For each loan, I observe the identity of the bank that originated the loan, the level of the bank's capital at the time of the loan origination, loan-level features such as the interest rate, and the loan's outcome - whether it went into default or was repaid. I obtain the data on the small business loans through a Freedom of Information Act (FOIA) inquiry to the US Small Business Administration which runs a program allowing banks to purchase partial default insurance from the government on certain small business loans. I obtain the mortgage loan information from public filings made with county registrar of deeds offices when mortgages are taken out on properties (and when foreclosures or shortsales subsequently occur), and then match these records with the Home Mortgage

Disclosure Act (HMDA) data to isolate loans held on bank portfolios.

Using this data, I show that banks with higher levels of debt relative to equity make systematically riskier loans as measured across a wide array of risk metrics. For instance, when considering small business loans made by large banks, an increase in the capital ratio (capital divided by asset value) of one percentage point (e.g. from 12% to 13%) is associated with a reduction in loan default risk of approximately 1.5 percentage points (from a mean default rate of 18.5% to a reduced rate of 17.0%, a roughly 8% drop in total risk). I show also that when these loans default, the severity of losses (based on how much of the loan's principal can be recovered) are also greater for more lightly capitalized banks, suggesting an additional dimension to the risk they are taking on. For instance, moving from the 50th to 60th percentile of capital ratios is associated with a 10% reduction in average losses suffered upon default of small business loans.

For home mortgage loans, a one percentage point increase in capital ratios is associated with a reduction of foreclosure rates from a mean of 2.5% down to 2.4%, or a 4% drop in foreclosure risk. This implies, for instance, that an increase of five percentage points in capital ratios could have reduced mortgage foreclosures by approximately 20%. A change in capital ratios of this magnitude is roughly on par with what actually occurred between the lowest points of bank capitalization in recent history in 2007 and the high points of 2012. Thus, it is a plausible thought experiment to consider the implications of an increase in capital levels of this magnitude.

I document very similar figures when I move from considering individual loans on bank balance sheets to investigating more macro-level impacts of bank capital. In particular, I calculate average capital levels of all banks across each US county during the 2003 to 2006 period, and compare these to the total mortgage foreclosures each county experienced from 2007 to 2012. In these analyses, an increase of 5 percentage points in the average capital ratios of banks in a county is associated with a drop of at least 14.5% in aggregate foreclosures, even after controlling for a host of demographic and other factors. Bank capital levels, and the risk-taking that they influence amongst bank investments, can thus have a profound impact on the real economy.

There are several empirical challenges in determining whether these associations indeed reflect a causal relationship. One challenge is that the management of banks may have pre-existing risk preferences that drive both the risk of the assets banks invests in and the amount of debt banks finance those investments with. A risk-loving CEO, for instance, might choose a corporate strategy with both high loan risk and high debt (leverage).[1] Reducing the leverage of such a bank via capital regulation, for instance, would not necessarily lead to a reduction in its loan risk. If anything, the reaction could be the opposite. Another problem in drawing causal inferences is that both bank capital and other aspects of bank investing are influenced by regulations which can change in concert with one another. For instance, regulators may simultaneously increase bank capital requirements and increase other supervisory or regulatory requirements that constrain risk-taking. This too would lead to a spurious relation between bank capital and risk taking.[2]

The empirical specifications in this investigation are designed to meet these identification challenges in several ways. To control for the possibility of management risk preferences driving both capital and risk levels of banks, I use bank fixed effects in my empirical specifications. Thus, to the extent that such management characteristics are relatively stable within a given bank over time, these will be controlled for. If management preferences change over the course of the study period, however, these fixed effects will not be able to resolve the identification problem. Accordingly, I also fit my main model specifications over a series of five-year sub-intervals of my total study period, again using bank fixed effects. Across these set of model specifications, I show that even considering a given bank, that bank will tend to make riskier loans as its capital levels decline.

In order to confront the challenges that bank regulations governing capital ratios may change at the same time as other shifts in the regulation or supervision of banks that also impact bank risk taking, my models also include time fixed effects. Thus, these show that even at a given point in time, banks with lower levels of capital make riskier loans. Despite these controls, if bank regulation were to shift differently for different types of banks (for instance, small local banks versus large multi-national institutions), then these time fixed effects may not adequately control for changes in regulatory envi-

[1] See, e.g. Bernile et al. (2017). A similar effect could be induced by compensation policies that reward returns without adequately adjusting for risk (Becker and Ivashina, 2015).

[2] Or, conversely, regulators may substitute higher capital requirements for other types of regulation and supervision, leading potentially to an opposite spurious conclusion, that high levels of capital lead to greater bank risk taking.

ronment. Accordingly, I also fit my models separately for each of three categories of bank based on size. Additionally, in order to address concerns that capital measurements might be distorted from a more pure theoretical concept, I consider a wide array of different measures of capital, both those based on regulatory definitions (such as risk weighted assets), those based on accounting definitions (such as net assets), and those based on market valuations (such as market capitalization). I also consider both absolute measures of bank capital ratios as well as percentile rankings against other banks at a given time. Although I document some differences in the predictive ability of these metrics, the core results are robust across a wide range of definitions of bank capital.

In order to further ensure that the results here identify a causal relationship between capital and loan risk, I construct an instrumental variables strategy that is based on the instrument developed by Granja et al. (2017) to study the impact of bank capital levels on the ability of banks to purchase the assets of failed banks in government run auctions. This is a Bartik-type instrument which exploits the fact that banks are strikingly stable in the geographic distribution of where they do business. Different banks were founded in different areas of the country and generally continue to do business in the areas where they traditionally have operated. As such, and as I explicitly document, a bank's geographic area of operation reflects more upon its institutional history than upon factors like its managers' current risk taking preferences or strategies.

During the recent recession, different regions experienced dramatically different patterns in the timing and magnitude of house price declines. I show that these declines are strongly predictive of changes in the capitalization of banks operating in the given regions. Based on this, I instrument for a bank's capital level by creating a weighted house price change index for each bank. This index takes the changes in house prices in each of the metropolitan areas in which a bank operates and weights them by the percentage of a bank's total deposits which it has in that area. Since my risk models are at the loan-level, rather than the bank-level, I create a separate prediction of a bank's capital for each metropolitan area it operates in. This enables me, when calculating the weighted house price change index, to exclude the metropolitan area in which a loan itself is made. Thus, by design, this instrument will be insulated from local economic conditions which might impact both bank asset values and the risk characteristics of loans being made. The results from this IV analysis confirm those results of my main analyses.

This paper most closely relates to the empirical investigations on the relationship between bank capital and risk-taking by Gorton and Rosen (1995) and Gan (2004).[3] A key advantage of this study is that it is the first to use detailed information that identifies specific assets on bank balance sheets, the realized default risk of those specific assets, and the time those assets were acquired by banks. A second important contribution of this study is that it examines a much more recent time period, and a larger and more fully representative sample of banks as compared to earlier works. This study also contributes to the literature by linking the risk-taking influences of bank capital to the macro-economic effects these have on the broader economy.

Another strain of related literature uses bank capital levels to predict stock returns, as in Beltratti and Stulz (2012) , Haldane (2012), and Demirguc-Kunt et al. (2013). The empirical specifications here also relate to the theoretical models developed in work such as in Buser et al. (1981), Marcus (1984), Kim and Santomero (1988), Keeley (1990), Flannery (1989), Gennotte and Pyle (1991), Hellmann et al. (2000), VanHoose (2007), Chen et al. (2017), Admati et al. (2017).

The remainder of this paper proceeds as follows. Section 2 discusses the theoretical motivation of the analyses I perform and Section 3 discusses data. Section 4 presents the empirical strategies I use and Section 5 presents the main results. Section 6 presents robustness checks and Section 7 offers concluding remarks. The online appendix presents additional robustness checks, supplemental analyses, discussion of the regulatory background, and additional data descriptions.

2 Theoretical Background and Motivation

When a corporation funds a significant portion of its assets with debt relative to equity, the incentives of that firm's owners may change in ways that lead them to either to reduce the risk it takes due to debt overhang (Myers, 1977), or to increase the risk that it takes due to the asset substitution prob-

[3] Another related study is Bidder et al. (2017), which finds that when banks suffered losses and reduced capital due to oil price drops, they shifted more of their assets to categories with lower regulatory risk weights (e.g. trading mortgages held on books for mortgage backed securities). Becker and Ivashina (2015) investigate the incentives of insurance companies to invest in corporate bonds with different risk and return properties.

lem (Jensen and Meckling, 1976). Furthermore, the more highly leveraged an institution is, the less incentive its equity holders will have to reduce these agency costs of debt (Admati et al., 2017), a phenomenon that is particularly relevant to banks due to their high leverage relative to other corporations. The potential incentives for banks to increase risk as equity funding declines are also intensified due to the implicit and explicit insurance they receive on their debt, insurance whose price is rarely able to fully reflect its actuarial costs (Merton, 1977). Given then that banks may have multiple sources of incentive to increase risk taking beyond socially optimal levels, researchers have long pointed to capital regulation as a method to reduce the investment incentive distortions that banks face due to their high levels of debt (Buser et al., 1981; Hellmann et al., 2000; Admati and Hellwig, 2013).

Yet, some scholars have argued that the effects of higher capital and higher capital requirements might not be a reduction in risk taking but actually an increase. If banks are targeting a given net return on assets (and presuming their shareholders are relatively insensitive to the risks that accompany those expected returns), then an increase in bank capital, which would tend to reduce expected returns on assets (again, not considering the risk of those returns), could lead banks to compensate by increasing the riskiness of the assets a bank invests in (Kim and Santomero, 1988). Furthermore, non-deposit debt funding, which most large banks also use, must be rolled over. In the absence of implicit government guarantees of this debt, costs of debt rollover will increase both with asset risk and with overall debt levels, meaning that higher debt levels may incentivize a bank to reduce asset risk as a way of constraining debt rollover costs (see, e.g., Chen et al., 2017). Overall then, depending on which of these theoretical considerations proves dominant, the relationship between bank capital and loan risk could potentially run in either direction. A major objective of this paper therefore is to empirically test these competing theoretical predictions.

Banks with more debt might be led not just to make riskier investments, but also ones with lower net present value (NPV). This may be due either to agency conflicts between equity and debt holders of the firms (Jensen and Meckling, 1976) or because of “risk externalities” - bank incentives to take advantage of mis-priced government insurance, either explicit or implicit (Buser et al., 1981). Because of the large role that banks play in allocating credit throughout the economy, lower NPV loans by banks can also have a serious consequence for the efficiency of the economy as a whole. One particularly salient example of this can be seen in what are now widely recognized as the excesses in mortgage risk taken during the lead up to the recent financial crisis and the widespread externalities that defaulting mortgages then imposed on the broader economy (see, for instance Mian and Sufi, 2015). Determining the actual NPV of loans made by banks is extremely difficult, requiring, at a very minimum, some reliable way to measure the costs to banks of making loans and the precise revenues banks receive from those loans. Data of this detail goes beyond what is available for this study, which bars any definitive investigation of the impact of debt funding on the NPV of loans made by banks.

Nevertheless, this study does aim to shed some light on the topic of the relation between bank debt funding and the NPV of loans in two ways. First, at a minimum, in the absence of agency conflicts or risk externalities, one might more naturally expect that banks with higher levels of capital would specialize in making riskier loans (larger capital buffers making such banks better able to withstand losses on such loans), whereas banks with less capital would pursue strategies of safer, lower-margin loans whose profitability is amplified through additional leverage. Indeed, venture capital firms, which make some of the riskiest investments in startup companies, tend to be funded almost exclusively with equity. Consideration of debt rollover costs would also suggest that banks with less leverage would be better positioned and more incentivized to make riskier loans. If, therefore, banks with more leverage are actually those that make riskier loans, it raises the question of what theoretical considerations apart from agency conflicts or risk externalities could account for such behavior.

Secondly, given the losses sustained by financial institutions on home mortgage loans, and the negative externalities the default of those loans imposed on the larger economy (Mian and Sufi, 2015), it is clear that at least a large number of home mortgage loans made during the first decade of the twentieth century imposed costs (both private and public) that far exceeded their benefits.[4] To the extent that the issuance these damaging mortgage loans is attributable in part to the capital levels of banks making them, it lends an additional suggestive piece of evidence that the levels of bank capital may impact the efficiency of the economy as a whole. As such, the empirical tests presented in Section 4 address ways

[4]Though again, a full determination of these loans as positive or negative NPV *ex ante* is very challenging and one that goes beyond the scope of this study.

to investigate not just the risk of individual loans that banks invest in, but also, the association between those risks and impacts on the broader economy.

Several other important theoretical considerations bear on the specific methodological approach this investigation follows. While the presence of deposit insurance is relatively consistent across a wide range of banks, in recent decades, the phenomenon of certain banks as being “too big to fail” (TBTF) has become increasingly apparent, with the governments of the United States and other nations becoming unwilling to allow banks beyond a certain size to fail, or even for their creditors to take substantial losses. TBTF status then offers an additional form of insurance for particularly large banks (O’hara and Shaw, 1990; Kelly et al., 2016; Gandhi and Lustig, 2015). As such, there is reason to believe that if capital does indeed have an incentive effect to increase bank risk-taking, that effect may be particularly pronounced for large banks. Large banks might also be more or less susceptible to the incentive effects of capital based on differences in the regulatory treatment they receive. For instance, large banks may be able to more effectively conceal risks from regulators due to their increased use of complex risk weighting procedures in calculating capital requirements (Behn et al., 2016). Motivated by this, the empirical specifications described in Section 4 also investigate disparate effects of bank capital based on the size of banks.

Another consideration is the role that bank managers may play in influencing a bank’s risk taking. If bank manager compensation is based on short-term measures of returns and does not accurately account for risk, then managers may have incentives to increase risk taking even apart from those of firm shareholders as a whole (Becker and Ivashina, 2015). Or, even apart from specific manager incentives, key managers, such as CEOs, may have existing risk preferences that influence the investments they guide their firms to take (Bernile et al., 2017). Because these effects are conceptually distinct from the risk-shifting channel of Jensen and Meckling (1976), their presence could interfere with causal inference on the relation between bank capital and risk taking. For instance, the presence of a risk-loving CEO or management compensation that overly encourages risk-taking could lead a bank’s management to pursue a strategy that involves both high leverage and high loan risk, leading to a spurious correlation between low bank capital and high loan risk. Accordingly, Section 4 discusses various empirical tools designed to overcome this identification challenge.

3 Data

3.1 Small Business Administration Data

The small business loan data analyzed in this study consists of all Section 7a loans backed by the US Small Business Administration (SBA). The loan information was obtained from the SBA via their internal response system to Freedom of Information Act (FOIA) inquiries submitted to the agency. The Section 7a program allows banks to make loans to small businesses and to purchase partial default insurance on those loans from the SBA.[5] In order to be approved as an SBA 7a loan, the loan must meet certain criteria, such as maximums on loan amount, interest rate, and so on. Any loan meeting the basic criteria, and initiated by a lender accredited by the SBA to make such loans, is automatically approved by the SBA. Once approved, the bank can choose to purchase default insurance from the SBA to cover up to 75% or 85% of the loan's value (depending on loan size). The SBA charges a fixed portion of the guaranteed amount for this insurance. While the SBA does not directly assess loan risk when either approving a loan to guarantee, or when deciding the price to charge for its guarantee, if a lender consistently has default levels that far exceed program guidelines then their ability to further participate in the program may be curtailed. Small businesses may use 7a loans for a variety of operational expenses.[6]

There are a total of 1.4 million loans in the SBA data set. For the purposes of this study, however, I consider only loans made in the fourth quarter of 2002 or later, as that is the point at which the capital variables that I study become publicly available from US regulators. I observe the outcomes of these loans up through the third quarter of 2016. For my sample, I consider all loans made prior to 2014, as very few loans made towards the end of the sample period yet had any opportunity to default. Naturally, I consider only loans made by banks for which capital information is reported. After imposing these

[5] `https://www.sba.gov/category/lender-navigation/sba-loan-programs/7a-loan-programs`

[6] Section A.2.5 of the online appendix explicitly examine the ways that this partial insurance influences the relationship between bank capital and loan risk.

restrictions I am left with just under 600,000 of these loans to analyze. Figure 1 plots a number of descriptive statistics for the small business loan data, and Figure 2 plots the number of loans originated each quarter along with the portion of those loans which eventually default. Table 1 gives information on the twenty banks with the greatest number of loans represented in this sample.

There is an active market for the securitization of SBA loans (de Andrade and Lucas, 2009; Craig et al., 2006). Initially, this might raise concerns about a tenuous relationship between bank capital and loan risk, if banks securitize most of the SBA loans they originate. In fact, however, this securitization market consists almost entirely of the guaranteed portions of the SBA loans, since it is these which are a far more standardized commodity and less sensitive to information about how carefully a given bank screened loan applicants.(Colomer, 2012; de Rugy, 2007). Thus, when a bank originates an SBA backed loan, there is a very high probability that the loan will remain as a source of risk on the bank's balance sheet.

3.2 County Registrar of Deeds Home Mortgage Data

The second set of loan data that I analyze consists of approximately 100 million home mortgage loans, using data gathered from nearly 3000 registrars of deeds in counties across the United States. Keeping track of ownership and rights to real property is critical to any functioning economy. As part of this system in the United States, whenever a piece of property changes ownership or becomes encumbered by a lien of any type, documentation must be filed on this with the Registrar of Deeds for the county in which the property is located. This information is then publicly available to anyone who wishes to look it up with the Registrar of Deeds. The data firm CoreLogic gathers this data and licenses it out to end users. By examining the history of documents recorded on each property, I am able to construct histories of specific loans and their outcomes. Section A.4 of the online appendix describes the details of how I create these records.

There are two critical advantages that constructing mortgage data from deed records offer for this study. First of all, because the deed filings include the name of the institution extending the loan, I am able to gather comprehensive information on the financial institution that originates every loan in the data set, something that is omitted from most of the leading sources of mortgage data used in the academic literature. Second, deed records are filed for loans regardless of whether those loans are subsequently sold, either to a securitization vehicle or another financial party. A great many of the mortgage data sets that are frequently analyzed in the academic literature consist only of loans that were securitized. Since the focus of this investigation is the impact of capital on the riskiness of assets that banks hold in their portfolios, a pool of exclusively securitized loans would be of significantly less use. To the best of my knowledge, I am aware of no other set of mortgage data that 1. is available to be licensed by any member of the public willing to pay the license fee, 2. gives the name of the financial institution that originated the loan, 3. includes loans that were held on the books of financial institutions as well as those that were securitized and 4. gives the outcome of the loan in terms of whether it defaulted, was pre-paid, etc. The construction of this unique, extensive, and rich data set is central to the success of the empirical strategy of this paper. Section A.2.1 gives descriptive plots for the home mortgage data, equivalent to those presented for the small business loan data above. Table 2 lists the top twenty lenders represented in the home mortgage data.

3.3 Home Mortgage Disclosure Act Data

Although the data that I gather from the deed records covers loans that were held on the books of financial institutions as well as those that were sold, it does not directly identify which is which. In order to obtain this information for the loans, I therefore match the deed mortgage records that I construct with those from the Home Mortgage Disclosure Act (HMDA), a database which does contain information on whether a loan is sold or retained on portfolio. HMDA is a federal law which requires that lenders report information to the government on approximately 90% of all loan applications and originations made in the United States (Williams, 2015). This information is then subsequently made publicly available. Section A.4 of the online appendix describes in detail the process by which I match the mortgage data created from the deed records and the HMDA data. In total, I am able to match 68% of the mortgages in the deed records with those in HMDA through this procedure, giving me a total of 65 million mortgages that combine data from the two sources. This success rate closely matches that of other researchers who have performed similar such matching.[7]

Although the HMDA data gives information on whether a loan was sold or retained on the books of the institution that originated it, this information is only reported as of the end of each calendar year. As such, for instance, if a loan is originated in November and sold the following January, then it will not be marked as being sold in the HMDA data. Because my focus is on the risk that banks choose to take for the portfolios on their books, I therefore restrict my sample to the loans made during the first half of each calendar year in order to ensure that a larger portion of the loans marked as retained are actually retained. The deed records that I have extend through the second quarter of 2014. Similar to with the SBA loans, therefore, I restrict my sample to loans originated prior to 2013, as loans originated just before the end of my available data have had very little time in which to default and be foreclosed upon.[8] I similarly begin my sample after the fourth quarter of 2002, when the bank capital information that I use becomes available, and restrict myself to only those lenders that I have capital information on. After making these restrictions I am left with approximately 1.2 million mortgage loans for the purpose of my analyses.

3.4 Bank Capital Data

I obtain information on the capital levels of banks from Call Reports - comprehensive regulatory disclosures made by banks, which are then made publicly available. Within these, I consider an array of different measures of capital. In particular, I consider the ratio of Tier 1 capital to risk weighted assets (often referred to as the Tier 1 capital ratio), the ratio of Tier 1 plus Tier 2 capital to risk weighted assets (often referred to as the total capital ratio) for each bank, and the ratio of Tier 1 capital to total assets, not risk weighted, (often referred to as the leverage ratio). Additionally, for those banks owned by a bank holding company, which cover the majority of loans in my sample, I also consider these ratios at the level of the ultimate holding company. In the event a bank is not owned by a parent holding company, the parent and bank level metrics would be identical. If a parent company is also a publicly

[7]See, e.g. work by Nancy Wallace: `https://bfi.uchicago.edu/sites/default/files/file_uploads/3_Wallace_MFM.pdf`

[8]In the deed data, I only observe when either a foreclosure or shortsale occurs, rather than the initial loan delinquency which is often used as a metric of loan default.

traded company, then I also consider the ratio of its market capitalization to its total assets and risk weighted assets. Figure 3 considers each of the small business loans in the data, the capital level of the bank that originated the loans at the time of origination, and then plot the distribution of these capital levels across all of the loans.[9] In order to get a sense of how capital ratios have changed across time, Figure 4 depicts the 25th, 50th, and 75th percentiles of each of the key capital ratios I consider, with percentiles taken of the set of the 100 largest banks operating in the United States.

A very small number of banks in my sample have either very low or very high capital ratios. I exclude from my sample all observations from banks that have capital ratios below the regulatory minimums. For these banks, regulators will be taking an extremely active and forceful role in supervising bank operations until the bank becomes more adequately capitalized. As such, it is not meaningful to examine the impact of capital on bank incentives for risk-taking. For these banks, it is the priorities of the regulators, not the incentives of the bank, that determine risk-taking. Similarly, a very small number of banks have extremely high capital levels, e.g. Tier 1 capital ratios of 30%, 50% or more (as compared to the 4% regulatory minimum and a 75th quantile of between 10% and 15%). I exclude these as well, as I consider them outliers that represent very unusual cases.[10] In total, these restrictions eliminate 2.1% of the small business loans and 3.5% of the home mortgage loans - a very small portion of the sample. Furthermore, I find little impact on my results by varying the cutoff levels for these exclusions.[11]

3.5 Additional Data Sources

In my analyses of total foreclosures per county during the 2007-2012 period, I use data from CoreLogic's Loan Level Market Analytics (LLMA) data source. This is collected from mortgage servicers and covers approximately 82% of the residential market in the United States. This data has the disadvantage of not identifying specific banks that make the loans, making it unusable for my main analyses. But, for

[9] Section A.2.1 in the online appendix depicts comparable information for the home mortgage loans.

[10] Furthermore, when it comes to evaluating the results of this study from a policy perspective, banks whose capital levels are within the main stream are by far the most relevant. This study aims, for instance, to lend insight to regulators considering lower capital levels to 3% or raising them to 15%, 20%, even 25%. For regulators considering changes beyond these levels, down to, for instance, 1% or up to 40%, this study simply does not have enough data to make a meaningful contribution to such a decision.

[11] Even if I include all of these outliers, my fundamental results are largely unchanged. Some coefficients become smaller and less significant, but not to the extent to impact any of the key conclusions of this study. The basic cause of this result is self-evident. While there is a robust relationship between higher bank capital and lower loan risk, that relationship starts tapering off once capital gets beyond a certain point.

calculating aggregate foreclosures over a geographic area, it is a reliable data source that is frequently used in the academic literature. I also use this to calculate mean FICO scores of borrowers by zip code, and the total number of open mortgages in each county as of the end of 2006, which I use in the analysis of total foreclosures. In my analyses of formal enforcement actions by bank supervisors, I use data posted on the websites of each of the Federal Reserve, Office of the Comptroller of the Currency, Federal Deposit Insurance Commission that list all formal actions by each of these agencies, as well as of the Office of Thrift Supervision, whose duties were largely subsumed by the Office of Comptroller of the Currency.

4 Empirical Strategy and Causal Identification

4.1 Core Empirical Specification

The core of the analytical framework for this paper lies in fitting conditional logistic regressions to model the probability that a given loan will default as a function of the capital levels of the bank that originated it and other predictors.[12] These regressions are of the general form:

$$\mathbf{P}\left[\text{Default}_{ijt}\right] = \Lambda\left(\beta \text{Capital}_{jt} + \Gamma X_t + \delta B_{jt} + \mu_j + \eta_t + \varepsilon_{ijt}\right) \tag{1}$$

Here $\Lambda()$ represents the conditional logit function,[13] i indexes individual loans, j indexes banks, and t indexes time, measured as the month in which a loan is originated. Default_{it} indicates whether loan i, originated at time t, eventually defaults.[14] Capital_{jt} represents the capital ratio of bank j at time t, with capital being measured according to a variety of specifications discussed below. X_t represents a set of macro-economic controls (interest rates, unemployment rates, and past stock market returns), all measured as of the month of loan origination. B_{jt} is an additional set of bank characteristics, μ_j is a bank fixed effect and η_t is a time fixed effect with a separate value for each origination year.

[12] All of these results are also robust to using a linear probability model.

[13] Because my models involve bank fixed effects, estimating each of these as a separate parameter, as in a regular logistic regression, would lead to the problem of incidental parameter bias, which the conditional logit avoids.

[14] The small business data does not explicitly identify the time that default occurs. Furthermore, the mortgage data indicates the time that a foreclosure or short sale occurs, but not the time of the underlying default, which can be months or years prior. Accordingly, I do not attempt to incorporate the timing of defaults into this analysis, simply whether they occur at any point.

In various specifications, I set by construction certain of the coefficients to be zero. In particular, for the majority of my analyses, bank capital and bank fixed effects are the only bank-specific variables that I use. But, I also consider several supplemental exercises in which I use fuller sets of bank-level predictors. I discuss these issues more fully in Section A.1.1.

Similarly, my base analyses contain no borrower or loan specific information. Suppose a loan characteristics, such as the interest rate, were highly correlated with loan risk (which it certainly is). By directly including such a variable in an analysis alongside the bank capital variables, I would be effectively conditioning on the outcome (loan risk) that I am trying to study. The primary goal of this investigation is to study the impact of capital on total loan risk, not loan risk conditional on a set of observable factors already known to be predictive of risk. Nevertheless, in a few of my supplemental analyses, I do include borrower and loan specific information, for particular reasons which I articulate below in the context of those analyses.

My preferred measure of goodness of fit for these analysis is the area under the receiver operator characteristic curve (AUC). The AUC is a standard metric for binary classifiers such as the logistic regression. It measures the accuracy of the classifier in terms of both false positive and false negative classifications. The AUC ranges from 0.5, which indicates a classifier no better than random, to 1, which represents perfect accuracy.

The baseline specification in Equation 1 is designed to address many of the potential identification concerns that lie in drawing a causal inference between bank capital and risk taking, with the instrumental variables specification, described further below, serving as an additional check on the validity of the inference. One possible identification challenge stems from the potential that banks may have managers with pre-existing risk preferences that lead them to adopt a portfolio that is both highly leveraged and comprised of high-risk loans.[15] This phenomenon could then lead to a negative bias on the β parameter of interest in Equation 1, in other words, a correlation between low capital and high loan risk, but one which is not causal in nature.

Another possibility that could bias results in the other direction could stem from banks' differing

[15] See, for instance, Bernile et al. (2017) documenting that early life experiences can impact CEO risk-taking preferences.

abilities or incentives to manipulate the calculations that go into the capital ratios they report. For instance, Behn et al. (2016) show that large German banks used the ability to calculate risk weights with complex internal models to make loans that were riskier than those made by banks using simpler risk-weighting procedures but that carried lower risk weights according to their internal models. To the extent that this phenomena is present in the banks that I study, it would bias the β parameter in Equation 1 upwards, artificially weakening the relationship between low capital ratios (measured using risk weighted assets) and high loan risk.[16]

In both of these cases, however, provided the given bank characteristic is relatively stable over time, it will be controlled for by the bank fixed effects that I use in specifications based on Equation 1. If, on the other hand, these bank characteristics shift more frequently, then the bank fixed effects will not be able to fully address the problems they raise for causal inference. Therefore, as a robustness check, I also fit versions of Equation 1 over five-year sub-intervals of my sample period. These smaller sample periods help to bolster the plausibility that unobservable bank characteristics such as these are stable over the sample period. Section 6, which deals with robustness checks, presents the results of these analyses.

Another identification concern lies with the ways that regulators influence not just the capital levels of banks but also other aspects of their operations that relate to the risks of the loans they make. If, for instance, regulators change the capital levels required of banks at the same time they change other aspects of bank regulation or supervision, that could create spurious correlations between capital levels and risk taking. If regulators increase capital requirements at the same time they increase other risk-constraining aspects of regulation, then it would lead to a spurious correlation between high capital and low risk. If instead, regulators saw higher capital requirements as a partial substitute for other forms of regulation or supervision, then it would lead to a spurious correlation in the opposite direction.

As an initial response to these challenges, the models based on Equation 1 include year fixed effects. As such, they are essentially measuring, for a given point in time, whether banks that have less capital make riskier loans than banks with more capital. To the extent therefore that these regulatory shifts

[16] To the extent that the non-equity instruments that can be counted as either Tier 1 or Tier 2 capital do not meaningfully function as loss absorbing capital, the numerator in capital measures could likewise be manipulated, and certain banks are likely better positioned to do so than others. See Section A.3 of the online appendix for more details on regulatory matters.

are common amongst banks, the year fixed effects will control for them. At the same time, however, it is possible that regulatory and supervisory treatment may be quite different for small local banks as compared to large multi-national institutions. In that case, simple year fixed effects on their own may not be adequate to address this identification concern.

As a response to the potential for different regulatory treatment for different sizes of banks, I also fit models separately for each of three groups of banks based on size of assets. The smallest such size group corresponds with bank holding companies with $17 billion or less in risk weighted assets. The mid-size group corresponds with bank holding companies with risk weighted assets between $17 billion and $345 billion, and the large group corresponds with bank holding companies with risk weighted assets between $345 billion and $1.7 trillion. I choose these sizes so as to have roughly comparable numbers of loans in each group as well as with an eye towards which institutions might most plausibly enjoy too big to fail insurance. By considering the main analyses of Equation 1, I am also able to investigate the hypotheses described in Section 2 that larger banks may have greater sensitivity to the incentive effects created by low capital levels due to the additional too big to fail insurance they enjoy as compared to other banks.

In Section 3.4 I discuss the various metrics that I use for the capital variable in Equation 1 and in Section A.3 of the online appendix I discuss the regulatory background for the calculation of those metrics. The use of these different metrics is designed to address concerns about the difficulty in accurately measuring bank capital.[17] For instance, in addition to the regulatory-defined metrics of the leverage ratio, Tier 1 ratio, and total capital ratio, I also use alternative capital metrics that use a bank's market capitalization, which has been suggested to at times more accurately reflect the value of a bank's equity (Haldane, 2012). Another way to control for measurement error in the capital variables is to replace absolute values of the variables in Equation 1 with the percentile rank of each bank, as compared to other banks during a given quarter. In this way, as long as distortions are relatively consistent across banks, the quantile measure will control for them. Even if distortions differ by bank type, as long as they are consistent within, for example, the size groupings of banks that I discuss above, then using quantile

[17]See, e.g. Haldane (2012), Admati et al. (2013), for discussions of limitations in regulatory-based measures of bank capital.

measures of capital for models constrained to banks of a particular type will control for those distortions. The tables in the body of this paper focus on several key regulatory and market based capital measures, with the online appendix giving a set of results over a wider range of capital measures.

4.2 Instrumental Variables Specification

The inclusion of bank and year fixed effects, and variations on the basic model in Equation 1 to cover smaller time intervals and to fit separately over different types of banks go a long way to address the potential identification concerns. Nevertheless, it is always possible that remaining challenges to causal identification could persist, even accounting for those factors. For instance, if the phenomenon of banks manipulating risk weights described in Behn et al. (2016) is more dynamic, with banks changing their degree of manipulation frequently to match their changes in risk, then even with bank and year fixed effects, this could create upwards bias on the β parameter in Equation 1, leading to a false conclusion that there is a weaker relationship between low capital and high risk than there actually is, or even that the relationship runs in the opposite direction.[18] Likewise, it is conceivable that the three groupings of banks, designed in part to address concerns of different regulatory treatment, might be too coarse of a division or might not line up with a more ideal way of separating banks based on different regulatory treatment. For these reasons, I also implement an instrumental variables specification to add further robustness to the causal inferences based on the associations between bank capital and loan risk that this study documents.

The instrument I use was first proposed by Granja et al. (2017) in a study of the impact of bank capitalization on ability to purchase failed banks being resolved by the Federal Deposit Insurance Corporation (FDIC). The instrument exploits the fact that different regions of the country experienced different levels of decline in property values during the recent financial crisis. Declining property values in an area will tend to deplete the value of a bank's assets, both of real property as well as of financial instruments, such as mortgages and leases, backed by those assets. As asset values decline, capital will also decline, unless the bank acquires additional infusions of new capital investments or

[18] A similar bias but in the opposite direction would occur if risk preferences or compensation incentives of managers shift with sufficient frequency.

sells additional assets to compensate.

In order to construct the instrument, and following the procedure in Granja et al. (2017), I start by collecting for each bank in my sample the total value of deposits it holds in each Core Base Statistical Area (CBSA) in the United States. This information is available from the FDIC in their Statement of Deposits (SOD) disclosures. The idea behind the instrument is that banks with larger fractions of deposits in a given area are also more likely to be exposed to real estate price declines in that area. The instrument, however, considers nothing in terms of the actual investment decisions of banks - for instance, which types or how many assets they invest in for each CBSA, anything else about the portfolio allocations of the banks, or other factors that might relate to strategies or preferences that favor more or less risk. The identifying assumption, then, is that the distribution of the geographies where banks have historically accepted deposits is not connected to factors such as the risk preferences of managers of those banks, after controlling for bank and year fixed effects. After discussing the instrument in further depth, I present several explicit validation exercises relating to this assumption.

The specific details for constructing the instrument are as follows. I let p_{rt} represent the house price index, collected from the Zillow,[19] in CBSA r at time t.[20] The goal is to create an index that weights the changes in house prices in CBSAs where a bank operates by the fraction of its total deposits in each of those areas. This thus serves as a metric of how severe of losses a bank likely experienced due to declining property values in the areas in which it operates. My reference point for measuring declines is the first quarter of 2006, so I measure changes in the house price index of each CBSA compared to levels in Q1 2006. I consider house prices from 2008 onwards as proportional changes from this baseline.

Suppose that bank i makes a loan in CBSA j at time t. I represent the set of all CBSAs in which bank i has the deposits as Ω_i. I consider the universe of deposits that bank has in all CBSAs *except* the given CBSA ($\Omega_i - j$), in order that any calculations of this metric be independent of any local regional conditions in the CBSA that might affect the probability of loan default.[21] For each CBSA, n, other than

[19]Using instead the price index from the Federal Housing Administration, FHA, makes little difference. The Zillow index has a slightly broader coverage of CBSAs, which is why I prefer it.

[20]I also construct the same instrument using house prices from Zillow and get essentially identical results.

[21]Banks with deposits in only a single CBSA are thus by construction excluded from tests basted on this instrument.

j, I calculate the value of deposits that a bank has in that CBSA (d_{in}) as of Q1 2006. I then divide that by the total value of all deposits the bank has in CBSAs other than j, ($\sum_{m\in\Omega_i-n} d_{im}$). By design, therefore, a bank will have different values of this weighted house price index for each of the CBSAs that it makes loan in. Based on these specifications, then, the weighted house price index for bank i making a loan in CBSA j at time t would be:

$$\Delta p_{ijt} = \sum_{n\in\Omega_i-j} (p_{nt}/p_{n2006} - 1) \frac{d_{in}}{\sum_{m\in\Omega_i-j} d_{im}} \tag{2}$$

Having constructed this weighted house price index, I then calculate, for each bank, the change in its capital ratio (using each of the different metrics I consider here) from time t compared to the first quarter of 2006: $\Delta\text{Capital}_{ijt} := \text{Capital}_{ijt} - \text{Capital}_{ij2006}$. Finally, in an OLS regression, I use the weighted house price index, and other controls, to predict this change in capitalization for each region in which a bank operates:

$$\Delta\text{Capital}_{ijt} = \alpha + \beta\Delta p_{ijt} + \mu_i + \gamma_t + \eta X_t + \varepsilon_{ijt} \tag{3}$$

This is essentially the same set of predictors as in Equation 1 including time and bank fixed effects and a matrix X of macro controls. One difference is that whereas for my main analyses, I use a conditional logit formulation, for the instrumental variables analyses I switch to a linear probability model for a practical and tractable way of implementing the IV specification. Although my first stage is used to predict change in capitalization, rather than capitalization itself, because the change is with respect to a fixed reference point for each bank (Q1 2006), with the inclusion of bank fixed effects the use of this change metric is identical to the use of total bank capital. I find that the drops in this weighted house price index are a strong and reliable predictor for the regulatory-based capital measures for the years from 2008 to 2013.[22] For instance, when predicting the Tier 1 capital ratio, the F statistic on the instrument is 36.7 and a one standard deviation movement in the instrument is associated with a 0.58 point change in the capital ratio, compared to a mean Tier 1 capital ratio change of 2.17 over the IV

[22]For years 2006 and 2007, there is relatively little variation in the ΔHPI metric compared to the reference point of Q1 2006, only in 2008 did the larger declines in house prices begin to manifest. With little instrument variation I find little power in the IV for this time period and thus exclude it.

period. The R^2 on the first stage is 0.82.[23]

4.2.1 Validating the Instrument Specification

In order for this instrument to be valid, it must be exogenous to unobserved factors, such as bank managers' strategic preferences or banks' abilities to manipulate their capital measures, that could affect both capital levels and loan risk. As Granja et al. (2017) note, the instrument likely satisfies the exclusion restriction by construction. Nevertheless, I perform several additional tests in order to validate this assumption. The instrument is based on exploiting geographic variation in where banks have their branches located. If, for instance, banks were strategically adjusting their branch locations, with more risk-loving managers moving their banks to areas with higher risk and more risk-averse managers moving operations to geographic areas with lower risk (to the extent that such areas could be reliably identified), then it would present an identification problem for this strategy. Even this in and of itself would not necessarily violate the exclusion restriction on my IV. If bank manager risk preferences are stable over time, then they will be controlled for in the IV by the bank fixed effects that I employ. Thus, an actual violation of the exclusion restriction for this IV would require that bank managers shift their risk preferences along shorter time scales *and* that these shorter-term shifts in risk preferences lead them to also shift their geographic distributions of activity.

One way to test whether this is true is to look at how stable bank geographic distributions are over time. If banks tend to remain in largely the areas in which they have historically operated, then it suggests that they are not shifting their geographic footprints in response to strategic considerations correlated with their corporate strategies for asset risk. To examine this, I consider for each bank the set of all CBSAs in which it has a branch at any point between 1994 and 2016, the years for which the FDIC SOD information is available. For each year, I then calculate the fraction of the bank's total deposits that occur in each CBSA. Thus, if a bank entered a CBSA for the first time in 2002, it would have 0 recorded for its total deposits in that CBSA in years prior to CBSA. I then use a panel regression to model the fraction of total deposits that each bank has in each CBSA. I use in this regression only Bank-CBSA

[23] The instrument strength for the other regulatory capital measures are similar, with F stats ranging from 16 to 36 (t stats between 4 and 6) and first stage R^2 values around 0.75. The instrument does not pass the strong instrument test for the market capitalization based measures. These presumably have greater variability based on a wider range of factors. Thus, I restrict results in this section to the regulatory-based measures.

fixed effects, which by construction are constant over the entire 23 year period of this exercise. The adjusted R^2 on this regression is 0.93. In other words, almost all of the variation in where banks locate, even over nearly a quarter century long period, can be explained by time invariant fixed effects. This strongly suggests against the notion that banks are shifting their geographic locations in response to shorter term changes in strategy and risk preferences.

As an additional test to validate the exclusion restriction on this IV, I consider an alternative formulation of it. Instead of using bank geographic distributions as of 2006, as I do in the main specification, I consider a version in which I use geographic distributions as of 2000, several years before the beginning of my study period. The results of this alternative specification (which I present in Section A.1.2 of the online appendix) are very similar to those for my main IV, with only minor increases in standard errors in certain specifications. This provides further validation that bank geographic distributions are stable over time and that any correlation between geographic distribution and risk preferences of banks can be well controlled for by the bank fixed effects in my specifications.

4.3 Additional Metrics of Loan Risk

The primary analyses of this paper measure loan risk as the probability of default. I also consider three other methods of investigating links between loan risk and bank capital. First, I consider losses conditional on default (LGD). The small business loan data that I use contains detailed information on the dollar value of losses sustained in the event that loans default. I divide these by the initial principal of each loan to calculate a loss severity, expressed as a percentage of the loan's principal. These severities range from 0 to approximately 125%.[24] I model these using regressions that follow the basic form of Equation 1 but using loan the logarithm of one plus LGD as the dependent variable and using a linear rather than conditional logit specification.[25]

Another way to assess the loan risks that banks take on is based on the amount of default insurance banks choose to buy on the small business loans they fund as part of the SBA's 7a program. Since taking

[24]Severities greater than 100% are possible because banks may expend more resources trying to collect on an unpaid loan balance than they recover on that loan.

[25]A functional form such as a beta regression might seem more natural for the loss-given-default specification. Yet, as a non-linear model, the beta regression would suffer from an incidental parameter bias problem with the inclusion of bank fixed effects. For this reason, I use the simpler linear formulation.

out default insurance on a higher percentage of a loan's principal value reduces a bank's risk from that loan, bank risk preferences can also be discerned based on the amount of insurance they take out. I model these using another regression of the form of Equation 1 but with the dependent variable as the percentage of a loan's principal that is covered by default insurance purchased by a bank. For these analyses, I also use a full set of borrower characteristics available in the small business lending data, plus fixed effects at the CBSA level. The reason for these is to capture, as well as possible, the amount of insurance that banks purchase, given all observable information on loan risk.[26]

A final perspective on risk comes from analyzing the interest rates charged on loans. If banks are aware that they are extending loans to riskier borrowers, then they should rationally respond by increasing the interest rates on those loans. I therefore use an additional set of linear regressions based on the form of Equation 1 to investigate this dimension of loan risk. Section 5.3 summarizes the results of these analyses using additional metrics of loan risk, with Section A.2.2 of the online appendix contains providing full details.

4.4 Supplemental Analyses

As discussed in Section 2, capital levels are relevant not just for understanding the risk that banks take onto their asset sheets but also for understanding the impacts those investment decisions by banks have on the larger economy. In order to investigate this, I consider mortgage lending in the lead up to the recent financial crisis, a consequential recent example of the impact on the real economy that bank investment decisions can have. In particular, I investigate whether counties that had more lightly capitalized banks experienced appreciably riskier lending practices in the lead up to the financial crisis and thus higher rates of mortgage foreclosures during the crisis.

I start by calculating the average capital levels of banks in each county over the pre-crisis period of 2003 to 2006. To do this I use the FDIC's Statement of Deposits (SOD) data to calculate, for each county, the percentage of total deposits accounted for by each bank for which I have capital measures from the Call Reports data. I then take the level of capital for each of these banks, weight them by

[26]Banks by no means report the full set of information available to them on the small business loans to the SBA. The borrower features include basic information like interest rate, loan amount, loan term, type of business, and details on the type of loan.

the bank's market share in a county, and sum those to generate an average capital measure for each county-year combination from 2003 to 2006. I then take the average of these yearly capital measures to arrive at an average capital level for banks in each county over the entire period from 2003 to 2006. From here, I investigate whether these capital measures are predictive of total foreclosures that occur in each county from 2007 to 2012. The specific form of the analyses are:

$$\log(\text{Foreclosures}_i^{2007-2012}) = \alpha + \beta \text{Mean Capital}_i^{2003-2006} + \Gamma X_i + \varepsilon_i \tag{4}$$

with observations weighted by county population as of 2000.

Here, X_i is a vector of county-level controls that is designed to address the possibility that, for instance, banks with lower capital might be located in counties that are economically more marginal or otherwise more prone to higher foreclosures. X_i includes standard demographic factors from the 2000 Census such as the log of county population, unemployment rate, population density, poverty and education rates, costs of rental housing, racial makeup, log income per capital, and others. Additionally, X_i includes a measure of the mean FICO score of residents taking out mortgages in the county as of year 2000.[27]

In order to control for the mechanical relation between greater mortgage volume and greater numbers of foreclosures, I include in X_i the log total number of open mortgages at the end of the 2003-2006 period over which I measure bank capital in the counties. I include house prices as of 2000 as a standard demographic control to capture economic characteristics of counties, and then I also include a measure of the percentage change in house prices from 2000 to 2009 in order to focus on an analysis of the riskiness of borrowers who receive loans in these counties, rather than in the house price declines the counties experience, which can also drive foreclosures for many types of borrowers (Palmer, 2014).[28] Finally, X_i includes the percentage of banks in each county (with weighting based on total deposits of each bank in each county) that fall into each of the size groupings of small, medium, and large.

[27] I use credit scores as of year 2000 because this is before many of the more questionable lending practices of the last decade began (see Rajan et al., 2015) and thus these are likely a more accurate measure of the overall credit quality of a county and not as subject to selection effects that inordinately prioritized giving mortgage credit to borrowers with high credit ratings.

[28] Including this measure of house price changes does have the potential to bias the estimate towards showing a less impact of bank capital on loan risk than is actually the case, due to the impact that foreclosures themselves have on house prices. Thus, the estimates from this regression should be seen as a conservative lower bounds.

I also consider two other supplemental analyses. First, I investigate whether there is evidence that bank supervisors respond to changes in the risk of assets that banks take on to their balance sheets, either via formal enforcement actions or via informal means. If this is the case, and bank supervisors are able to effectively monitor and respond to changes in bank risk that might be induced by varying capital levels, then it could potentially mitigate some of the concerns that might be raised by a link between bank capital and risk taking. I discuss the specifics these analyses, and their results, in greater detail in Section A.2.3 of the online appendix.

Second, I investigate whether one mechanism by which banks with lower capital opt for higher-risk strategies is through a reduction in investment in institutional infrastructure to collect on loans that have defaulted. Such a reduction in infrastructure investment would tend to yield higher profits in economic environments with low default rates and higher losses in times of greater defaults. Section A.2.4 of the online appendix discusses the specific form these analyses and their results in greater detail.

5 Results

I start by considering a very simple, unconditional plot of bank capital against loan default rates. To construct this I divide all small business loans into fifty contiguous bins based on the amount of capital of the bank that made each loan, at the time of the loan's origination. Then, for each bin, I calculate the percentage of loans which ultimately default in that bin. Finally, I plot these results for all fifty bins in Figure 5.[29] Several very clear patterns are evident.

First, there is a strong association between higher levels of capital and lower levels of loan risk for all of the regulatory measures of bank capital. Interestingly, the relationship is opposite for the measures that use market capitalization. As I show below, this effect is entirely due to the fact that during periods of strong economic performance (as measured, for instance, based on unemployment levels and stock market returns), banks are willing to take greater risks in loan making, and these periods also tend to boost the stock values of banks, leading to a spurious correlation. Once I add in simple macro-economic controls for unemployment, stock market returns, and interest rates, the relationships reverse to match

[29] Figure A.1 in the online appendix depicts these plots for home mortgage loans, with all key features closely matching those for small business loans discussed here.

those for the regulatory measures of capital. Another interesting pattern is that across the capital measures of banks and their parents, there is a consistent hierarchy in terms of the tightness of fit of the relationships. In particular, the leverage capital ratio, which does not use risk weighted assets, has the weakest relationship, whereas the ratio of Tier 1 plus Tier 2 capital to risk weighted assets has the strongest relationship with loan risk. I examine these relationships and this hierarchy more formally below.

I now proceed to the first of the conditional logistic regression analyses described in Section 4. These model the likelihood of loan default as a function of bank capital at time of loan origination as well as macroeconomic and bank-level predictors. Table 3 presents the results of this initial analysis for the small business loans in my sample.[30] The different panels of Table 3 present results for several of the various metrics of bank capital that I consider, while the columns depict varying specifications of the basic model articulated in Equation 1. Column (1) of Table 3 is the direct analog of Figure 5, using only bank capital to predict default risk. It shows a very strong and statistically significant relationship between bank capital and loan risks. Coefficients in the logistic model are interpreted by exponentiating them, after which they represent the proportional change in the relative probability of default compared to the null of not default. Thus, for example, the coefficient of -0.163 in Panel A, column (1) indicates that an additional percentage point in a bank's Tier 1 capital ratio is associated with a 15% reduction in the ratio of default probability to the probability of no default. Evaluated at the mean default probability over the whole SBA data set, which is 18.5%, this translates to a reduction of approximately 2.3 percentage points in default risk.

Interestingly, panels D and E of Table 3, which consider metrics of capital based on total market capitalization of banks, show strongly positive coefficients, though they are an order of magnitude smaller than coefficients in panels A through C. According to these analyses, higher levels of capital appear to be associated with *greater* loan risk, not less. In order to make sense of this seeming discrepancy in results, column (2) adds a set of macroeconomic controls to the regression, controlling for stock market returns, interest rates, and unemployment, each measured as of the month of loan origination.

[30] I present the results of these same analyses for home mortgage loans in Section A.2.1 of the online appendix.

The results of this additional analysis are clear. During periods of low unemployment and high stock market returns, banks make substantially riskier loans. This could be because better economic times make banks more optimistic about the prospects of riskier business ventures, or because banks feel that they will be less impacted by individual loan defaults if economic times as a whole are prosperous, or it could be for an additional set of reasons. The specifics behind such a phenomenon go well beyond the scope of this paper and are not necessary to fully evaluate for its current purposes. Nevertheless, these prosperous economic times also drive up the stock prices of banks, expanding the market cap based capital measures far more so than the regulatory based measures. Once these economic factors are controlled for, the capital metrics show consistently negative associations with loan risk, and are strongly significant in all cases except for the Tier 1 capital ratio.

Column (3) of Table 3 adds bank fixed effects, while column (4) considers both macro controls and bank fixed effects. In essence, these analysis measure, for a given bank, whether the bank makes riskier loans during periods in which its capital levels drop. They help to control for bank-specific unobservable factors, such as manager risk preferences or abilities to manipulate capital ratings, that are discussed in Section 4 above. Finally, columns (5) and (6) of Table 3 add in year fixed effects, first on their own and then in conjunction with the bank fixed effects and the macro-economic predictors. These control nonparametrically for the factors considered in the macro controls as well as any other factors that consistently influence all banks in a given time period. In essence then, the year fixed effects show that even within a given time period, banks with more capital make significantly less risky loans. Column (6) then combines the bank and year fixed effects - the coefficients remain significantly negative for both of the market based capital measures. The regulatory measures retain effects of comparable size and direction to those of the market based measures in this final specification, but lose their statistical significance.[31] The estimates remain economically significant as well, though obviously they are far smaller than the model without the bank and year fixed effects.

Based solely on these parameter estimates, a very small change in capital ratios might be anticipated to have a limited economic effect. But larger changes, which have been proposed repeatedly in recent

[31] In Section 5.2 and Section A.2.5, I fit the models in Column (6) for subsets of the data based on bank size and amount of default insurance and show that in many of these key specifications, the results are highly statistically significant.

policy discussions, would still have a sizeable impact on overall bank risk taking. For instance, using the coefficient estimate of -0.01 in column (6) for the ratio of market capitalization to total assets, an increase in that ratio by five percentage points would be associated with a reduction in mean loan default probability of 18.5% to 17.76%, or in other words, 0.74 percentage point absolute reduction and a 4% proportional reduction.

Although this might seem modest, it is important to keep in mind that leverage amplifies the effect of losses. A bank funding 90% of its loan assets with borrowed money, something very typical of the banks in this sample, that experiences a 1% loss on its assets will experience a 10% loss on its equity value. With a mean loss severity on defaulted small business loans of 77%, an increase of defaults by 0.74 percent would translate to an increase in total losses by 0.57% which then gets amplified to an increase in total equity losses of 5.7%. Furthermore, as I document below in Section 5.2 the effect of bank capital on the largest banks is far greater than this aggregate measure indicates, as are the effects for those small business loans with the smallest amount of default insurance, as I document in Section A.2.5 of the online appendix.

5.1 Comparing Predictive Performance of Capital Measures

Considering all of these different capital measures makes it possible to evaluate not just the size and significance of the coefficients for bank capital in the various specifications, it also allows for an examination of which measures contribute to the most accurate predictions of loan risk overall, as measured by the AUC values for the different analyses. Doing so reveals a few interesting results. I start by considering the models in column (1) of Table 3, in other words, the models using only the capital metric and no other predictors. Using risk-weighted assets gives a consistent boost of several points of AUC, and using Tier 1 plus Tier 2 capital, rather than just Tier 1, gives an additional couple of points of AUC, such that, for instance, the AUC using the ratio of Tier 1 capital to total assets is approximately 0.6 whereas the AUC using the ratio of Tier 1 plus 2 capital to risk weighted assets is approximately 0.65. These differences in AUCs are highly statistically significant as well, with the null hypothesis of the true AUCs being the same being rejected in each comparison at a level of $p < 0.001$.

To the extent that firms manipulate risk weights in order to simultaneously increase the risk of their portfolio loans while decreasing their net risk weighted assets (as documented by Behn et al. (2016)), capital measures based on risk weighted assets will become less predictive of loan risk, not more. Thus, these results show that even accounting for distortions and manipulations in risk weighted assets that inarguably exist, capital measures using those assets still deliver additional predictive value compared to unweighted assets. Similarly, the finding that the measure of Tier 1 plus Tier 2 capital delivers statistically significant superior performance to Tier 1 alone when predicting these portfolio risk decisions adds nuance to the findings of Haldane (2012) that simpler and less expansive capital metrics have superior performance for predicting short performance in the recent financial crisis. I next consider the predictive performance of the models that use more controls beyond just bank capital. Adding year fixed effects diminishes the differences in AUC between the metrics and adding bank fixed effects all but eliminates them, with each capital metric yielding a nearly identical AUC in the rightmost column that includes the full set of controls.

Another way to compare predictive performance is to consider how large of an effect on loan risk is associated with a shift in each of the capital variables. The coefficients in Table 3 are not directly comparable to each other because each of the capital variables is measured on a somewhat different scale - Tier 1 capital ratio clearly ranges over a different set of values than total capital ratio, for instance. But, Table A.5 in the online appendix considers each bank's capital ratio measured as a quantile compared to all other banks. As such, the meaning of moving up a given quantile in different capital variables is much more comparable.

The results from the quantiles tell a similar picture regarding the different predictive ability of the capital metrics. In the first column of Table A.5, for instance, moving up a percentile in total capital ratio has more than twice the impact (-0.031) than moving up a quantile of the leverage ratio (-0.015). But, when considering the final column with the full set of controls, the coefficients are all much closer in size. Together, these results suggest that if one had to consider a single capital metric on its own, a metric that uses risk weighted assets is preferable to one that uses unweighted assets, and a metric that considers both Tier 1 and Tier 2 capital is preferable to one that considers only Tier 1. But, if one is considering bank fixed effects - in essence, if one is examining deviations from a bank's prior capital levels - then the metrics in general are roughly the same in their predictive ability.

5.2 The Role of Bank Size

The results that I have discussed so far have covered all banks in the sample. Are these results consistent across different types of banks? Are they being driven primarily by small or large banks? In order to investigate these questions, I divide my sample into three groups based on the total assets of the parent company. The smallest group corresponds with bank holding companies with \$17 billion or less in risk weighted assets. The mid-size group corresponds with bank holding companies with risk weighted assets between \$17 billion and \$345 billion, and the large group corresponds with bank holding companies with risk weighted assets between \$345 billion and \$1.7 trillion.

Table 4 displays the results of these analyses by bank size. Two facets of the table are striking. First, the relationship between low capital and high loan risk is consistent and significant both across different measures of capital and across medium and large bank sizes.[32] Yet, the magnitude of the effects vary substantially. The coefficient estimates for the largest category of banks are often twice as large or more as those for mid-sized banks. In general, the coefficients for these largest banks center around -0.1. Evaluated as before at the mean default probability for SBA loan, this suggests that a one percentage point increase in capital ratios is associated with approximately a 1.7 percentage point degrease in default probability, from 18.5% to 16.8%. Capital levels are an important predictor of risk for both mid-size and large banks, but it is the largest banks, with risk weighted assets greater than \$345 billion, which are by far the most sensitive to it. These results suggest, for instance, that the risk-taking incentive effects of low capital levels may be especially large banks due to the supplemental "too big to fail" insurance they enjoy as compared to smaller institutions.

[32]The effect is by far the most attenuated for the smallest of banks, never reaching statistical significance in any of the formulations presented here. Less restrictive formulations with fewer controls do still show some significant results for this smallest bank size group.

5.3 Additional Metrics of Loan Risk

Do banks with lower capital make riskier loans along more dimensions than just probability of default? Here, I summarize results that are presented in greater depth in Section A.2.2 of the online appendix. When considering the losses banks suffer on small business loans, conditional on those loans defaulting, Table A.7 indicates that a move of ten percentile points in the market capitalization based ratios (e.g. a bank moving from the median capitalization, compared to other banks, to the 60th percentile of capitalization) is associated with a reduction of approximately 10% in the average losses suffered given default of a small business loan. Thus, banks with lower capital not only make loans that are more likely to default, when those loans do default, the banks suffer greater losses, adding an additional dimension to the risk they are taking on.

Table A.8 shows that an each extra percentage point in the ratio of the Tier 1 capital to risk weighted asset ratio is associated with banks purchasing an extra 1.6 percentage points of default insurance, on average, even conditional on all available borrower risk characteristics. Banks with low capital could in theory use such default insurance to partially or totally mitigate the impact of the added loan risk that they take on (at least for their small business loans made through the SBA's 7a program). Instead, low capitalized banks double down on their risks, taking out less insurance than banks with higher capital, even conditional on all observable loan risk factors. Finally, as Table A.9 shows, a one percentage point increase in regulatory capital ratios is associated with a decrease in two to six basis points of interest on loans. This result comes with the caveat that only 30% of the small business loans and 16% of the home mortgage loans have information on the loan interest rate. It does still at least suggest that banks are aware of the added risks they are taking on and charging higher interest rates to compensate for those, as would be expected.

5.4 Instrumental Variables Specification

Table 5 presents the results of the IV analyses, designed to address potential identifications problems that may remain after implementing the other empirical strategies discussed in Section 4. Because I implement the IV using an a linear probability model, rather than the conditional logistic model, the coefficients are not directly comparable to those in my main specifications. But, the results in this table nevertheless confirm the consistent relationship between higher capital for banks and lower loan risk. Notably, the coefficients in the IV specifications are consistently greater in magnitude than those for the comparable naive specifications (though in both specifications those coefficients are consistently negative). This indicates that to the extent that biases exist, at least for the time period of the IV, they actually operate to bias coefficients on the capital variables upwards. The most natural explanation for this likely lies in the effects documented by Behn et al. (2016), that banks may simultaneously pair increases in loan risk with manipulation of their capital ratios to, for instance, make risky assets carry lower risk weights than safer assets.

5.5 Impact on Real Economy

Given the large role of banks in allocating credit throughout the economy, do the impacts of capital on risk taking that have been documented above have appreciable effects that can be measured beyond just the balance sheets of banks? Table 6 presents results on inquiries into this question. The analyses use the average capital levels of banks in US counties during the pre-crisis period (2003 to 2006) to predict logged foreclosures during the recent financial crisis (2007 to 2012). While the leverage ratio has no significant effect (mirroring results from throughout this paper in which it is consistently the least predictive measure), both the Tier 1 ratio and total capital ratios are strongly predictive of foreclosures.

A one point increase in average Tier 1 ratio is associated with an approximately 4.4% decrease in logged foreclosures for a county in the crisis period, and a one point increase in average total capital ratio is associated with an approximately 4.0% decrease in logged foreclosures for a county in the crisis period. A one standard deviation move in total capital ratio amongst counties (equal to a change of 1.5 in the Tier 1 capital ratio) would be associated with a 6.6% reduction in foreclosures over the period. With over 2.5 million mortgage foreclosures occurring from 2007 to 2012 nationally, a five point increase, which is on par with increases in capital that have been required post-crisis, could have prevented 550,000 foreclosures, or in other words, reduced the magnitude of the foreclosure crisis by approximately one fifth.

5.6 Supplemental Analyses

I summarize here results of several supplemental analyses that are presented in greater depth, along with full regression tables, in Section 4.4 of the online appendix. Section A.2.1 considers analyses of home mortgage loans equivalent to those presented primarily for small business loans in the body of this paper. Table A.6, for instance, indicates that a one percentage point increase in total capital ratio is associated with a reduction in mortgage default probability from a mean of 2.5% down to 2.4% (a 4% proportional reduction in risk), and that a five percentage point increase in that capital ratio would be associated with a reduction from 2.5% to 2.0% mean default rate. Particularly when amplified by large leverage ratios, changes in loan risk of this magnitude associated with the impact of bank capital are quite significant. These estimates also line up closely with the macro-estimates presented above. In this loan level model, an increase in capital levels by five points is associated with a 20% reduction in defaults, from the macro geographic model above, the same increase is associated with a drop in foreclosures of roughly 15%.

Section A.2.3 looks for evidence of bank supervisors responding to changes in loan risk, either via formal enforcement actions or informal tools they may use to require banks making riskier loans to shift to larger capital buffers. Table A.10 shows no evidence of supervisory response along either dimension - coefficients are statistically and economically insignificant and show no evidence of bank supervisors responding via formal or informal enforcement actions to changes in loan risk.

Section A.2.4 considers whether investments in institutional infrastructure to collect on defaulted loans can explain some of the pattern in which banks with lower capital suffer larger losses given default on small business loans as compared to higher capitalized banks. Table A.7 demonstrates that even after controlling for a large range of observable loan risk characteristics, the coefficient size and significance linking low capital to high losses given default remains unchanged. This suggests that lower capitalized banks do indeed have less capacity to collect effectively on defaulted loans, since if the losses given default were being driven totally by borrower risk characteristics, then conditioning on those risk characteristics should have diminished the role of bank capital.

Finally, Section A.2.5 considers the main default analyses for small business loans but with separate models fit for groups of loans based on the amount of default insurance banks purchased on them. Table A.11 shows that the effects of bank capital on influencing loan risk are strongest for those loans with the smallest amount of default insurance. This result may perhaps be seen as relatively intuitive. In a thought experiment, if default insurance were taken to the limit of 100%, then the risk incentives of banks should become irrelevant as no actual risk would be involved.

6 Robustness Checks

6.1 Stability of Relationship over Time

Are the relationships demonstrated in these analyses between bank capital and loan risk stable over time? Are they driven primarily by a particular time period? Do the bank fixed effects credibly control for unobservable factors such as strategic preferences for risk-taking? To investigate these questions, I consider a number of specifications in which I fit models over smaller sub-periods of my total sample period. Specifically, I consider a set of rolling, five-year long windows of time and fit my main models using bank and year fixed effects (plus the standard macro controls I've used previously) to each of those windows. Thus, for instance, for the window beginning in year 2003, I consider all loans originated between 2003 and 2007. I then plot the coefficient estimate and confidence interval from each model for the key capital variables in question. Figure 6 shows these coefficient plots for the small business loans.[33]

There are several notable features of the plots. First off, the center of mass is clearly and consistently at negative coefficient estimates for the capital variables. With over 50 separate models depicted in the plots, it is perhaps not surprising that a few yield positive coefficients for the capital ratios, yet all of these are small in magnitude and statistically insignificant. By contrast, large and statistically significant negative coefficients characterize a great many of the plots. Despite this consistency in the overall center

[33]Figure A.2 in Section A.2.1 of the online appendix gives the corresponding plot for home mortgage loans.

of mass for the plots, there are notable differences amongst the capital variables. The metrics that use risk-weighted assets, rather than total assets, are more consistently and significantly negative, a result that mirrors some of the findings from the analyses using the entire sample period and presented, for instance, in Table 3. Overall, then, these results using data over five-year subsets of the total sample period help to confirm that there is a robust and persistent relationship between bank capital and loan risk over the banks and loans that I examine.

6.2 Additional Bank Predictor Variables

Throughout the main analyses in this study, the only bank-specific variables that I have considered explicitly in my analyses have been the capital related ones. For instance, I have dealt with bank size by breaking banks into separate groups based on size, rather than by including size as a specific predictor in my analyses. I have done this for a set of specific reasons. First of all, variables like size, as measured by bank assets, will tend to have highly correlated movements with bank capital. A shock that reduces or increases the value of a bank's assets, for instance, will generally have a similar kind of effect on the bank's levels of capital. The near collinearity in movement between these variables can make identification of their separate effects difficult to impossible.

Another reason to omit other types of bank-specific variables is that they are likely to be highly influenced by a bank's risk preferences and strategy. For instance, lower capital levels may lead a bank to adopt a riskier strategy, and that riskier strategy may then lead the bank to shift the maturity balance of its loans or the sources of its debt financing. Since the goal of this investigation is to test whether low levels of capital cause banks to adopt riskier strategies, then by including other bank variables that also reflect risk preferences, I am essentially conditioning on the outcome that I am trying to measure.

Nevertheless, as a robustness exercise, I present variations of my basic analyses that also use a set of additional predictor variables. In particular, I consider the ratio of a bank's total outstanding loans and leases to its total deposits.[34] Additionally, I consider the ratio of a bank's deposits to its total liabilities. This may help to distinguish different types of banks - for instance, smaller and more local banks that

[34]This was the variable that Jorda et al. (2017) found to be far more predictive of financial crises (when measured on a systemic level across all banks) than bank capital levels themselves.

rely more on deposits as opposed to larger banks that adopt more aggressive strategies and fuel greater leverage through capital market debt offerings than they could achieve with deposits alone.

I also consider the ratio of the total value of loans made to officers and directors of a bank compared to a bank's total loan volume. This is designed to capture some of the corporate governance issues considered by Gorton and Rosen (1995). Finally, I explicitly include a bank's log total assets, which might arguably control for bank size in ways more nuanced than via the models fit separately on loans from each of three sizes of banks as I present above.

Section A.1.1 of the online appendix presents the full set of regression tables showing the results of these analyses. Two patterns are evident. First, while adding these additional predictors reduces the size of the capital coefficients somewhat, the reductions are very modest and the general pattern of sign and statistical significance of the bank capital coefficients remains largely unchanged. Second, none of these other variables are consistently significant predictors of loan risk on their own, after controlling for bank capital and the other controls in the specifications. The one predictor that is most often statistically significant in the analyses is the loan to deposit ratio, which has a consistently negative sign. This suggests that banks with lower liquidity may choose to make safer loans. If so, this is an interesting result, though one relatively orthogonal to the focus of this paper.

7 Conclusion

A critical question for financial economics has long been how a firm's investing incentives will be impacted by its funding mix of debt versus equity. Because banks have such a high proportion of debt funding relative to other firms in the economy, and because the investing decisions of banks have such large consequences for the broader economy, understanding ways in which the amount of equity banks have impacts the investments they make is important for many both theoretical and practical reasons. Yet, economic theory has been divided in its predictions on this matter, with some models and sets of assumptions suggesting that low equity relative to debt will result in high levels of asset risk, while other models and assumptions suggest the opposite. Empiricists who have sought to tackle this problem have been constrained by an inability to access detailed information on the specific assets that particular banks hold on their portfolios.

This paper has sought to shed new light on these questions by developing and analyzing a novel data set of over 1.8 million small business and mortgage loans, held on portfolios of banks, each precisely matched to the bank that originated it and the level of equity (capital) of the bank at the time that it originated the loans. The analyses here have shown that across a wide range of metrics of risk such as probability of default or loss given default, across a range of time periods, a range of types of banks, and types of loans, banks with lower capital make systematically riskier loans. These results are significant not just for understanding the ways banks structure their balance sheets, but also for understanding the impact that banks and their investing incentives have on the broader economy. According to the findings in this paper, the riskier loans originated by lower-capitalized banks played a significant role in increasing the number of mortgage foreclosures experienced during the recent financial crisis. Accordingly, the findings suggest that a view of bank capital solely as a loss absorbing buffer is inadequate. Any discussion of capital levels by policymakers or others must take into account the substantial role that it plays in influencing bank investment decisions.

That lower capitalized banks would make such consistently riskier loans presents somewhat of a puzzle from a theoretical perspective. Models that focus on bank incentives such as debt rollover costs predict the opposite relationship, that banks with more capitalization should be better positioned and incentivized to specialize in riskier loans. More heuristically speaking, banks with more capital should be better able to withstand losses on risky loans, and indeed, some of the riskiest investments in the economy, those in tech startups, for instance, are made by venture capital firms that are funded almost exclusively with equity.

Theoretical models that predict an association between low capital and high risk, better fit the results of this study, but also imply that high risk investments by low-capital banks will be less efficient from a private and public perspective. Does this indicate then that the relation between low capital and high risk bank investments demonstrated in this study represents a social inefficiency? The results here are at most suggestive. Therefore, there is much important research to be done in the future, both to examine whether the results presented here can be explained by a credible model that does not also imply inefficient investing by low-capitalized banks, and to look for direct empirical evidence that can either confirm or rebut the possibility that low capital leads banks to make investments that are not just riskier, but less efficient.

References

Acharya, Viral V, Philipp Schnabl, and Gustavo Suarez, 2013, Securitization without risk transfer, *Journal of Financial Economics* 107, 515–536.

Admati, Anat, Peter M DeMarzo, Martin F Hellwig, and Paul Pfleiderer, 2017, The leverage ratchet effect, *Journal of Finance, Forthcoming* .

Admati, Anat, Peter M DeMarzo, Martin F Hellwig, and Paul C Pfleiderer, 2013, Fallacies, irrelevant facts, and myths in the discussion of capital regulation: Why bank equity is not socially expensive, *Working Paper, Stanford University* .

Admati, Anat, and Martin Hellwig, 2013, *The Bankers' New Clothes: What's Wrong with Banking and What to Do about It* (Princeton University Press).

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru, 2017, Policy intervention in debt renegotiation: Evidence from the home affordable modification program, *Journal of Political Economy* 125, 654–712.

Becker, Bo, and Victoria Ivashina, 2015, Reaching for yield in the bond market, *The Journal of Finance* 70, 1863–1902.

Behn, Markus, Rainer FH Haselmann, and Vikrant Vig, 2016, The limits of model-based regulation, *Working Paper, European Central Bank* .

Beltratti, Andrea, and René M Stulz, 2012, The credit crisis around the globe: Why did some banks perform better?, *Journal of Financial Economics* 105, 1–17.

Bernile, Gennaro, Vineet Bhagwat, and P Raghavendra Rau, 2017, What doesn't kill you will only make you more risk-loving: Early-life disasters and ceo behavior, *The Journal of Finance* 72, 167–206.

Bidder, Rhys, John Krainer, and Adam Shapiro, 2017, Drilling into bank balance sheets: Examining portfolio responses to an oil shock, *Working Paper 2017-03, Federal Reserve Bank of San Francisco* .

Buser, Stephen A, Andrew H Chen, and Edward J Kane, 1981, Federal deposit insurance, regulatory policy, and optimal bank capital, *The Journal of Finance* 36, 51–60.

Chen, Nan, Paul Glasserman, Behzad Nouri, and Markus Pelger, 2017, Contingent capital, tail risk, and debt-induced collapse, *Review of Financial Studies, Forthcoming* .

Colomer, Nora, 2012, The return of small business loan securitization, *Asset Securitization* 12, 16–19.

Craig, Ben, William Jackson, and James Thomson, 2006, Small firm credit markets, SBA-guaranteed lending, and economic performance in low-income areas, *Federal Reserve Bank of Cleveland Working Paper* .

de Andrade, Flavio, and Deborah Lucas, 2009, Why do guaranteed sba loans cost borrowers so much?, *Working Paper, Northwestern University* .

de Rugy, Veronique, 2007, Banking on the SBA - big banks, not small businesses, benefit the most from SBA loans programs, *Mercatus on Policy* 2.

Demirguc-Kunt, Asli, Enrica Detragiache, and Ouarda Merrouche, 2013, Bank capital: Lessons from the financial crisis, *Journal of Money, Credit and Banking* 45, 1147–1164.

Diamond, Rebecca, and Timothy McQuade, 2016, Who wants affordable housing in their backyard? an equilibrium analysis of low income property development, *NBER Working Paper w22204* .

Ferreira, Fernando, and Joseph Gyourko, 2015, A new look at the us foreclosure crisis: Panel data evidence of prime and subprime borrowers from 1997 to 2012, Technical report, National Bureau of Economic Research.

Flannery, Mark J, 1989, Capital regulation and insured banks choice of individual loan default risks, *Journal of Monetary Economics* 24, 235–258.

Gan, Jie, 2004, Banking market structure and financial stability: Evidence from the Texas real estate crisis in the 1980s, *Journal of Financial Economics* 73, 567–601.

Gandhi, Priyank, and Hanno Lustig, 2015, Size anomalies in us bank stock returns, *The Journal of Finance* 70, 733–768.

Gennotte, Gerard, and David Pyle, 1991, Capital controls and bank risk, *Journal of Banking & Finance* 15, 805–824.

Gorton, Gary, and Richard Rosen, 1995, Corporate control, portfolio choice, and the decline of banking, *Journal of Finance* 1377–1420.

Granja, Joao, Gregor Matvos, and Amit Seru, 2017, Selling failed banks, *Journal of Finance* 72, 1723–1784.

Haldane, A, 2012, The dog and the frisbee, *Speech given at the Federal Reserve Bank of Kansas City's 36th Economic Policy Symposium "The Changing Policy Landscape"* .

Hellmann, Thomas F, Kevin C Murdock, and Joseph E Stiglitz, 2000, Liberalization, moral hazard in banking, and prudential regulation: Are capital requirements enough?, *American economic review* 90, 147–165.

Hill, Julie Andersen, 2012, Bank capital regulation by enforcement: An empirical study, *University of Indiana Law Journal* 87, 645.

Jensen, Michael C, and William H Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of financial economics* 3, 305–360.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil, 2014, Securitization and loan performance: A contrast of ex ante and ex post relations in the mortgage market, *Review of Financial Studies* 27, 454–483.

Jorda, Oscar, Bjorn Richter, Moritz Schularick, and Alan M Taylor, 2017, Bank capital redux: Solvency, liquidity, and crisis, *Working Paper, National Bureau of Economic Research* .

Keeley, Michael C, 1990, Deposit insurance, risk, and market power in banking, *The American Economic Review* 80, 1183–1200.

Kelly, Bryan T, Hanno Lustig, and Stijn Van Nieuwerburgh, 2016, Too-systemic-to-fail: What option markets imply about sector-wide government guarantees, *American Economic Review* 1106, 1278–1319.

Keys, Benjamin J, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig, 2010, Did securitization lead to lax screening? evidence from subprime loans, *The Quarterly Journal of Economics* 125, 307–362.

Kim, Daesik, and Anthony M Santomero, 1988, Risk in banking and capital regulation, *The Journal of Finance* 43, 1219–1233.

Marcus, Alan J, 1984, Deregulation and bank financial policy, *Journal of Banking & Finance* 8, 557–565.

Merton, Robert C, 1977, An analytic derivation of the cost of deposit insurance and loan guarantees an application of modern option pricing theory, *Journal of Banking & Finance* 1, 3–11.

Mian, Atif, and Amir Sufi, 2015, Foreclosures, house prices, and the real economy, *Journal of Finance* 70, 2587–2634.

Myers, Stewart C, 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147–175.

O'hara, Maureen, and Wayne Shaw, 1990, Deposit insurance and wealth effects: the value of being "too big to fail", *The Journal of Finance* 45, 1587–1600.

Palmer, Christopher, 2014, Why did so many subprime borrowers default during the crisis: Loose credit or plummeting prices?, *Working Paper, UC Berkeley Haas School of Business* .

Posner, Eric A, 2015, How do bank regulators determine capital-adequacy requirements?, *The University of Chicago Law Review* 1853–1895.

Rajan, Uday, Amit Seru, and Vikrant Vig, 2015, The failure of models that predict failure: Distance, incentives and defaults, *Journal of Financial Economics* 111, 237–260.

VanHoose, David, 2007, Theories of bank behavior under capital regulation, *Journal of Banking & Finance* 31, 3680–3697.

Williams, Michael, 2015, HMDA coverage of the mortgage market, master's thesis, *University of Minnesota, Hubert Humphrey School of Public Affairs* May 18, 2015.

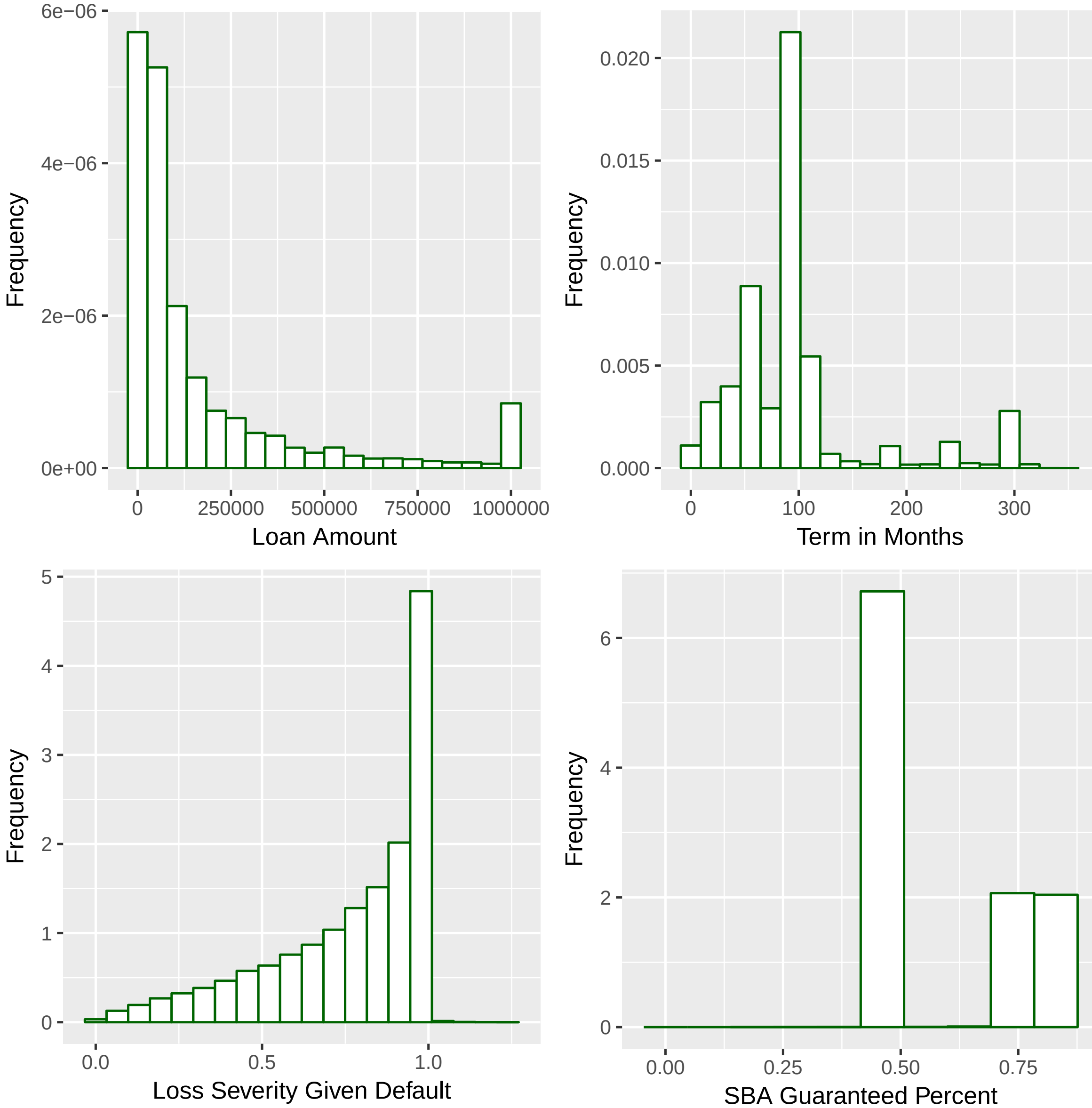

Figure 1. Small business loans - descriptive plots. For interpretability of the plot, loans above $1 million are represented as $1 million in this plot.

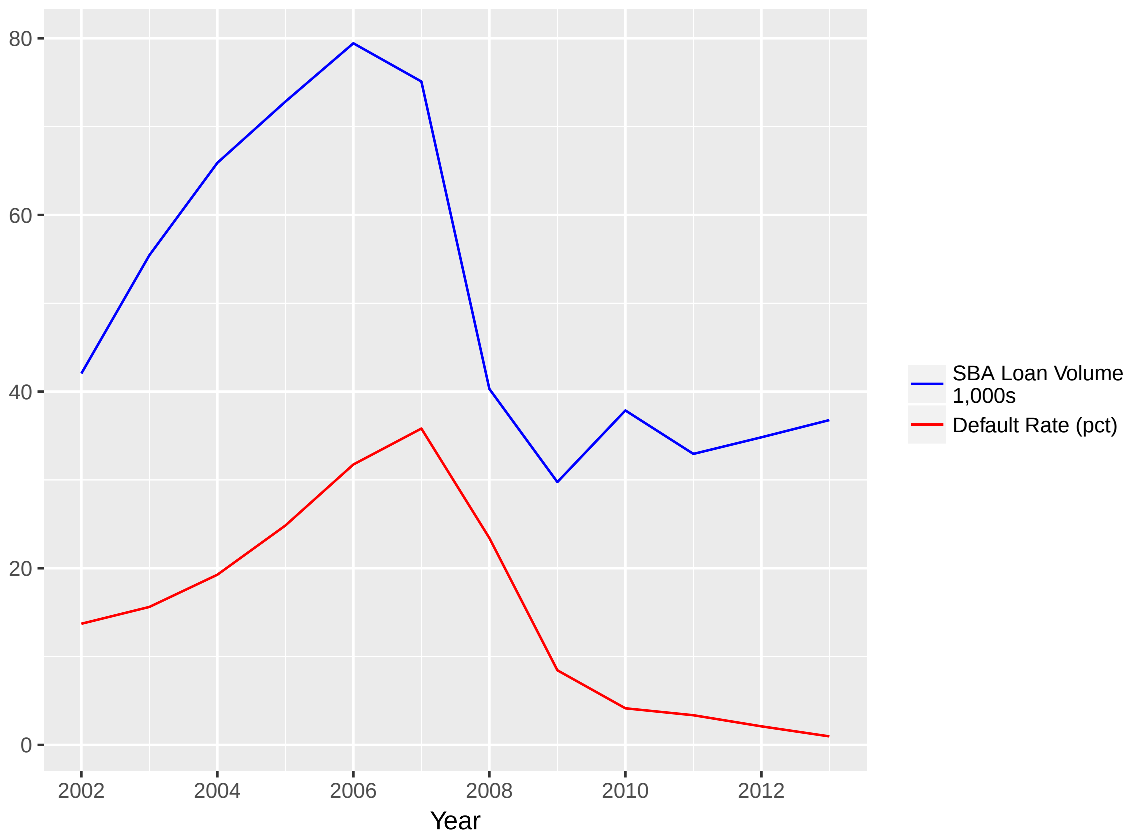

Figure 2. Small business loans - origination and default rates by time.

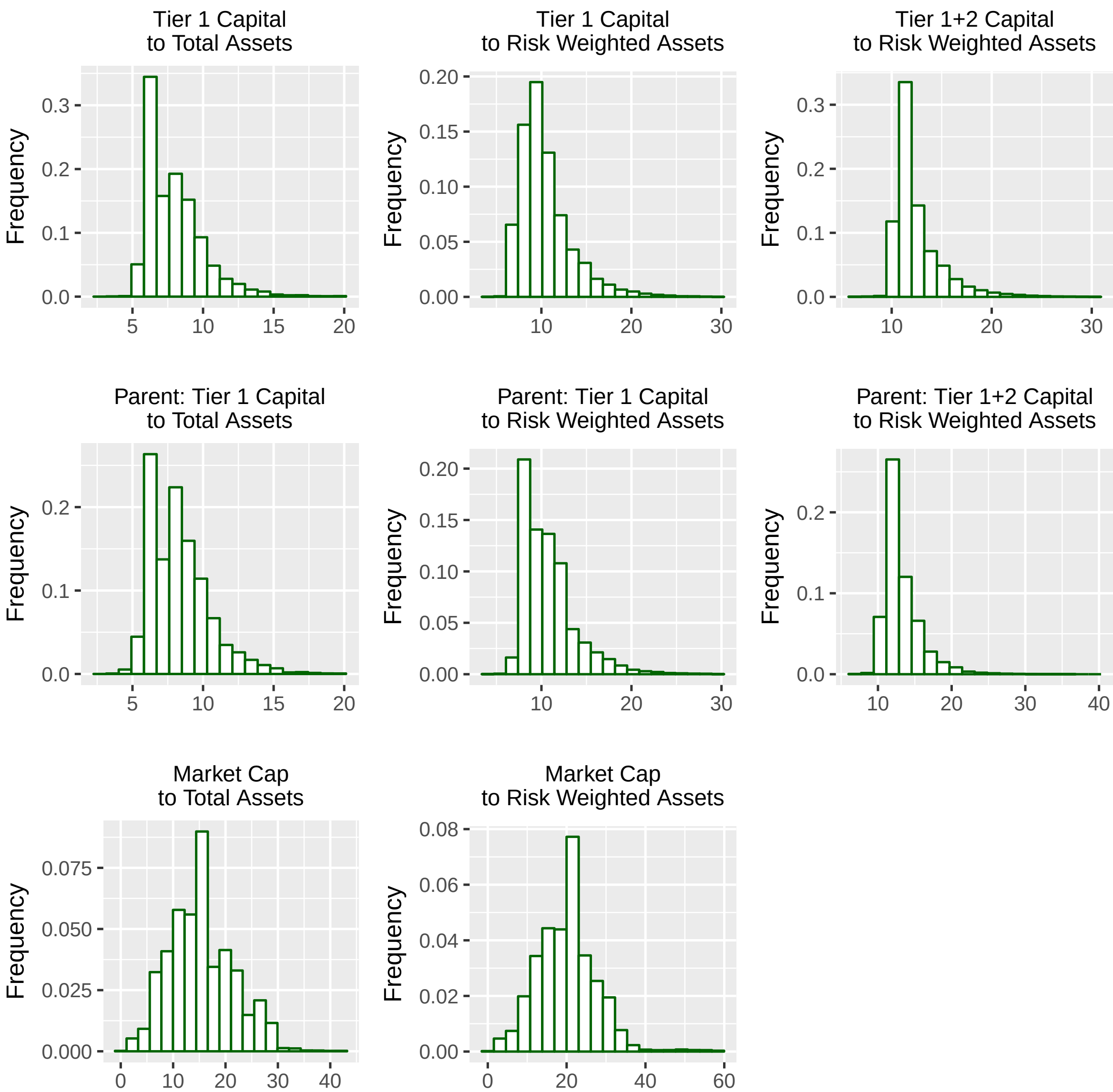

Figure 3. For each small business loan, this plot considers the capital level of the bank that originated the loan at the time of the origination and plots the empirical distribution of those capital levels.

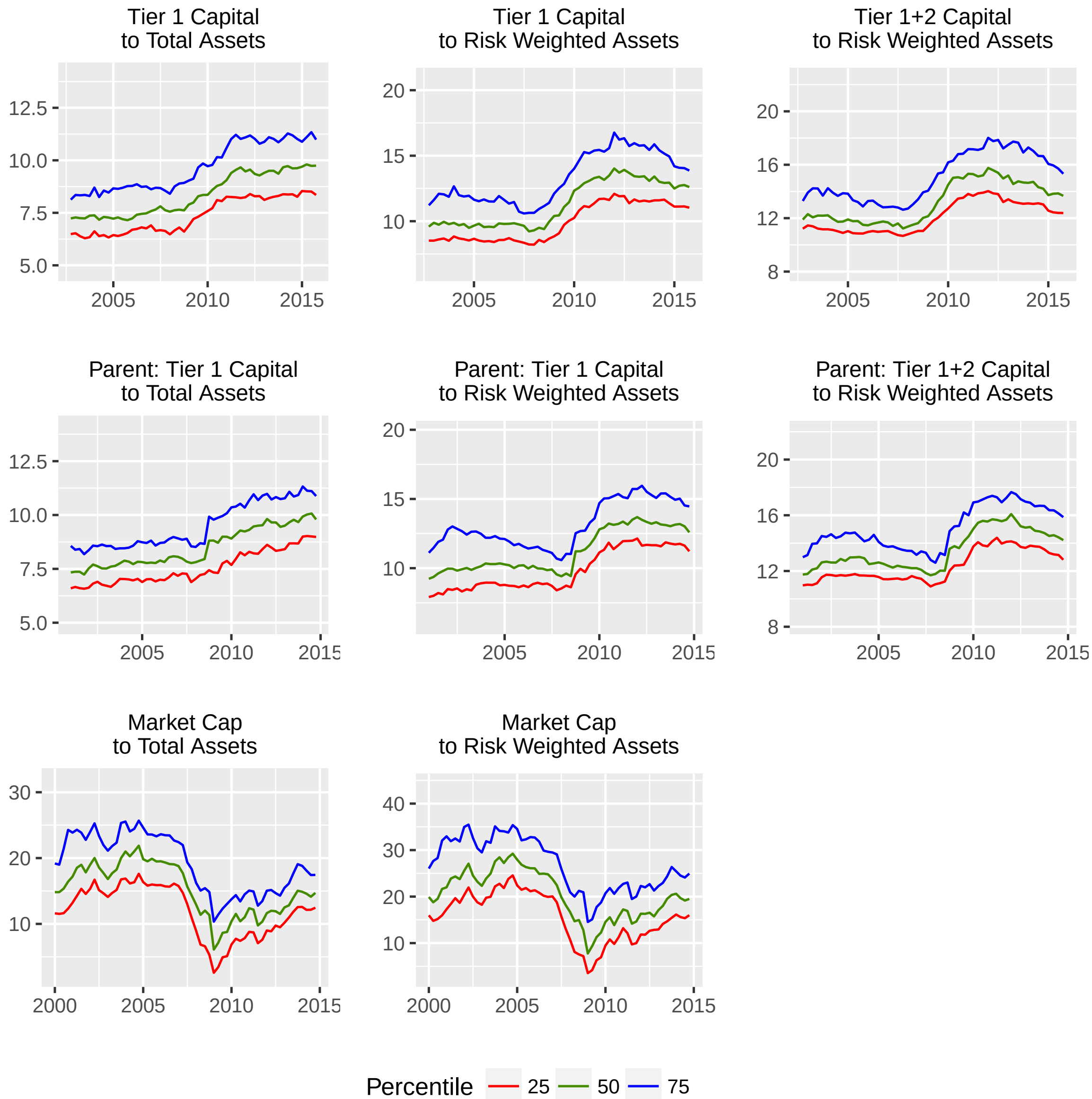

Figure 4. Capital Ratio Quantiles. Figure considers the 100 largest banks that report call data and depicts the 25th, 50th, and 75th percentiles of the capital ratios across these entities.

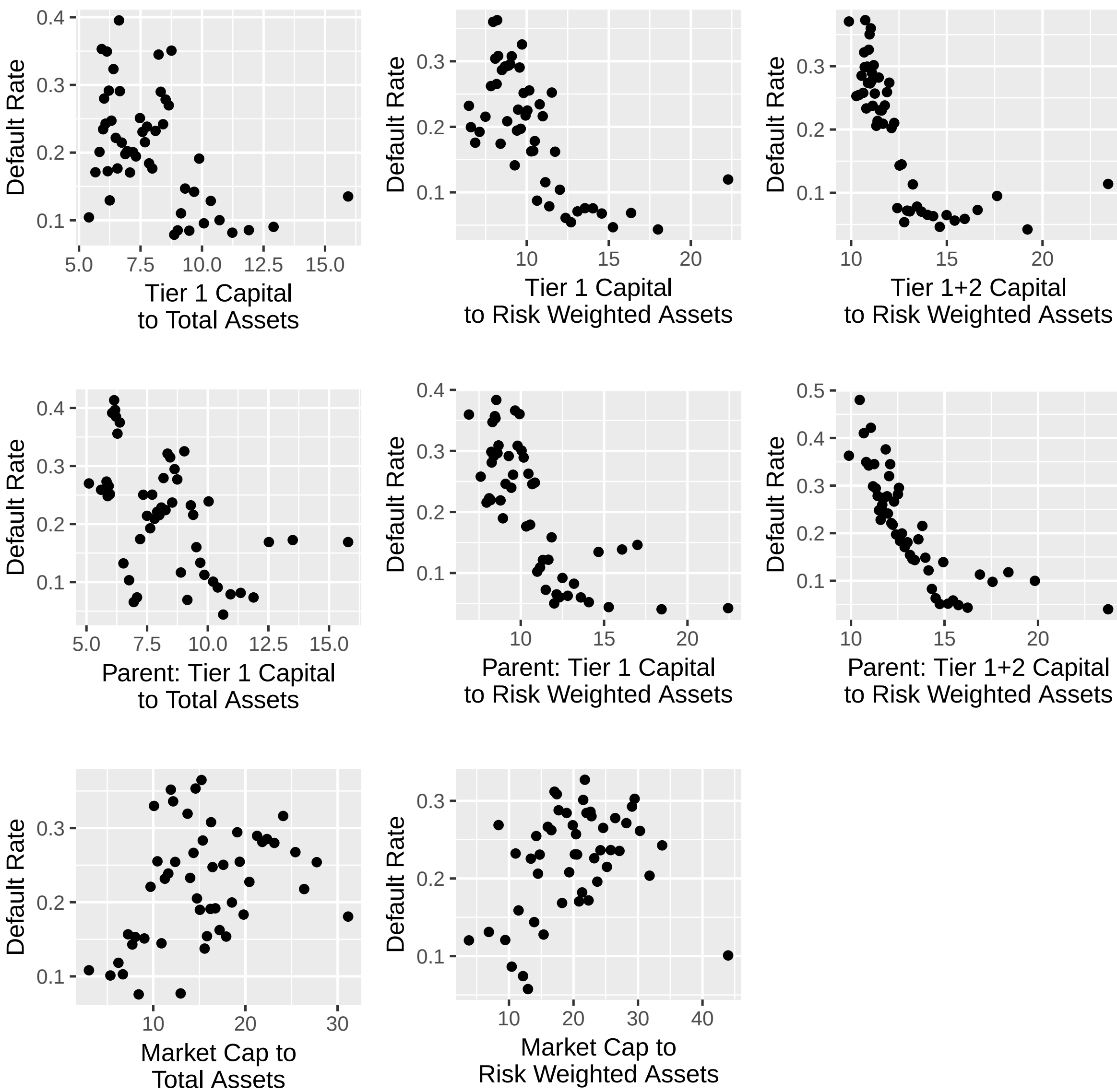

Figure 5. Default risk against bank capital - small business loans. To create these plots, I discretize each of the continuous capital metrics into 50 separate bins, chosen such that each bin will contain roughly the same number of total loans. I then calculate the percentage of all loans in each bin that default and plot this against the middle of each bin in each of the respective graphs.

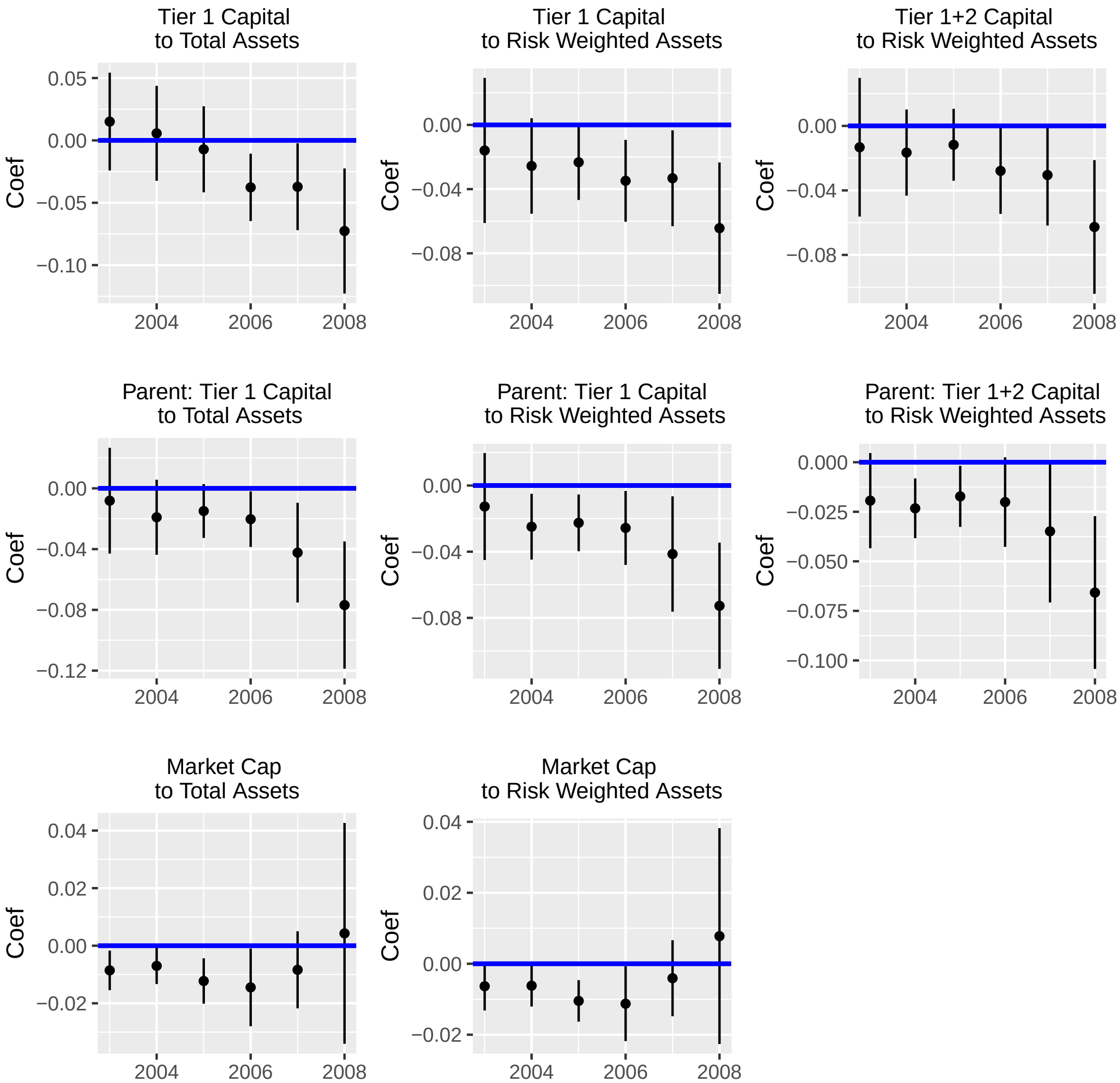

Figure 6. Five-year subsets of data - small business loans. These figures consider models fit over rolling, five-year long windows of time. For each time window, I calculate the coefficient estimate and confidence interval (95%) for the key capital variable. For instance, the coefficient and confidence interval associated with the year 2003 on these plots represents a model fit using data from 2003 to 2007. All specifications include macroeconomic controls plus bank and year fixed effects with robust errors clustered at the bank level.

Table 1
Top Lenders in Small Business Data. Names given are of the ultimate parent holding company of the banks. Some lenders changed names and/or ultimate holding companies during the sample period. In these cases, the most recent parent name is given. "Loans" gives the total number of loans by each lender that appear in the data set. "% Total Loans" represents the percentage of all loans in the data that were originated by each lender. "Cumulative %" simply keeps a running sum of all percentages so that one may see, for instance, what percentage of all total loans in the data were originated by the top *n* lenders.

Bank Name	Loans	% Total Loans	Cumulative %
Bank of America Corporation	65,539	11.21	11.21
JP Morgan Chase & Co.	51,148	8.75	19.96
U.S. Bancorp	31,915	5.46	25.42
Royal Bank of Scotland	31,795	5.44	30.86
Center Financial Corporation	25,155	4.3	35.17
PNC Financial Services Group	21,575	3.69	38.86
Huntington Bancshares Inc.	16,725	2.86	41.72
Hibernia Corporation	16,174	2.77	44.48
Zions Bancorporation	15,922	2.72	47.21
Popular, Inc.	12,560	2.15	49.36
Capital One Financial	12,003	2.05	51.41
Toronto Dominion Bank	9,380	1.6	53.01
Allied Irish Banks	8,320	1.42	54.44
BB&T Corporation	7,476	1.28	55.72
Umpqua Holdings Corporation	7,327	1.25	56.97
Keycorp	6,513	1.11	58.08
Compass Bank	5,729	0.98	59.07
Citigroup Inc.	5,245	0.9	59.96
Comerica Inc.	4,577	0.78	60.75
Associated Banc-Corp	4,065	0.7	61.44

Table 2
Top Lenders in Home Mortgage Data. Names given are of the ultimate parent holding company of the banks. Some lenders changed names and/or ultimate holding companies during the sample period. In these cases, the most recent parent name is given. "Loans" gives the total number of loans by each lender that appear in the data set. "% Total Loans" represents the percentage of all loans in the data that were originated by each lender. "Cumulative %" simply keeps a running sum of all percentages so that one may see, for instance, what percentage of all total loans in the data were originated by the top *n* lenders.

Bank Name	Loans	% Total Loans	Cumulative %
Royal Bank of Scotland	91,623	7.63	7.63
Bank of America Corporation	73,608	6.13	13.75
BB&T Corporation	71,919	5.99	19.74
Wells Fargo & Company	44,754	3.73	23.46
U.S. Bancorp	31,603	2.63	26.09
Suntrust Banks, Inc.	24,572	2.05	28.14
BNP Paribas	21,419	1.78	29.92
PNC Financial Services Group	21,047	1.75	31.67
Fifth Third Bancorp	20,593	1.71	33.39
Regions Financial Corporation	18,928	1.58	34.96
Toronto Dominion Bank	16,649	1.39	36.35
National Bank of the Great Lakes	15,433	1.28	37.63
Amsouth Bancorporation	12,447	1.04	38.67
Fulton Financial Corporation	11,537	0.96	39.63
Washington Mutual Bank	10,788	0.9	40.53
JP Morgan Chase & Co.	10,677	0.89	41.42
Charter One Financial, Inc.	9,321	0.78	42.19
Northwest Bank	9,169	0.76	42.96
Arvest Bank Group, Inc.	8,733	0.73	43.68
Southtrust Corporation	8,605	0.72	44.4

Table 3

Small business loans - Basic Regulatory Capital. This table presents the results from a conditional logistic regression. The dependent variable is whether an individual loan defaults or not. Capital ratios are represented in percentage units. Thus, if a bank's tier 1 capital divided by total assets yields 0.12, then the value of the corresponding dependent variable would be 12. AUC measures the area under the receiver operator characteristic (ROC) as a measure of goodness of fit. Errors are clustered at the bank level.

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A						
Tier 1 Capital to Total Assets	-0.163***	-0.056*	-0.235***	-0.087***	-0.043	-0.009
	(0.0419)	(0.0314)	(0.0423)	(0.0201)	(0.0322)	(0.0122)
S&P 500 Five-Year Return		0.202**		0.256***		0.161***
		(0.0791)		(0.0568)		(0.0394)
5 Year Treasury Rate		0.235***		0.215***		0.038**
		(0.0388)		(0.0309)		(0.0179)
Unemployment Rate		-0.286***		-0.21***		-0.221***
		(0.0393)		(0.0348)		(0.0234)
Observations	571685	571685	571685	571685	571685	571685
AUC	0.598	0.717	0.725	0.77	0.723	0.775
Panel B						
Tier 1 Capital to Risk Weighted Assets	-0.15***	-0.051***	-0.251***	-0.089***	-0.037**	-0.014
	(0.023)	(0.015)	(0.0437)	(0.012)	(0.0154)	(0.0153)
S&P 500 Five-Year Return		0.166**		0.253***		0.156***
		(0.0816)		(0.0583)		(0.0396)
5 Year Treasury Rate		0.233***		0.217***		0.039**
		(0.0378)		(0.0284)		(0.0179)
Unemployment Rate		-0.277***		-0.186***		-0.22***
		(0.0364)		(0.032)		(0.0236)
Observations	571685	571685	571685	571685	571685	571685
AUC	0.624	0.718	0.736	0.77	0.722	0.775
Panel C						
Tier 1+2 Capital to Risk Weighted Assets	-0.245***	-0.093***	-0.249***	-0.074***	-0.071***	-0.014
	(0.0307)	(0.0229)	(0.039)	(0.0109)	(0.0196)	(0.0146)
S&P 500 Five-Year Return		0.2**		0.258***		0.158***
		(0.0878)		(0.0667)		(0.0397)
5 Year Treasury Rate		0.233***		0.217***		0.038**
		(0.0375)		(0.0287)		(0.0176)
Unemployment Rate		-0.257***		-0.191***		-0.22***
		(0.0322)		(0.0321)		(0.0237)
Observations	571685	571685	571685	571685	571685	571685
AUC	0.65	0.722	0.735	0.77	0.726	0.775
Panel D						
Market Cap to Total Assets	0.02**	-0.019***	0.036*	-0.03***	-0.014**	-0.01***
	(0.0084)	(0.0065)	(0.0213)	(0.0031)	(0.0071)	(0.0022)
S&P 500 Five-Year Return		0.134*		0.083		0.141***
		(0.0746)		(0.1078)		(0.0466)
5 Year Treasury Rate		0.204***		0.216***		0.048**
		(0.045)		(0.0316)		(0.0186)
Unemployment Rate		-0.349***		-0.297***		-0.241***
		(0.0463)		(0.0368)		(0.0302)
Observations	382225	382225	382225	382225	382225	382225
AUC	0.548	0.707	0.698	0.752	0.711	0.756
Panel E						
Market Cap to Risk Weighted Assets	0.012	-0.019***	0.02	-0.024***	-0.014***	-0.006***
	(0.0074)	(0.0053)	(0.014)	(0.0025)	(0.0052)	(0.0023)
S&P 500 Five-Year Return		0.084		0.072		0.147***
		(0.0775)		(0.1112)		(0.05)
5 Year Treasury Rate		0.208***		0.215***		0.046**
		(0.0424)		(0.031)		(0.0184)
Unemployment Rate		-0.353***		-0.288***		-0.235***
		(0.0458)		(0.0384)		(0.0303)
Observations	382225	382225	382225	382225	382225	382225
AUC	0.545	0.707	0.695	0.752	0.711	0.756
Banks FEs	no	no	yes	yes	no	yes
Time FEs	no	no	no	no	yes	yes
Macro Controls	no	yes	no	yes	no	yes

Cluster robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 4
Bank Size Groups - Small business loans. This table presents results of conditional logit regressions in which the dependent variable is a binary indicator of whether or not an individual loan has defaulted. Each analysis is conducted separately for each size group of banks. The smallest group corresponds with bank holding companies with $17 billion or less in risk weighted assets. The mid-size group corresponds with bank holding companies with risk weighted assets between $17 billion and $345 billion, and the large group corresponds with bank holding companies with risk weighted assets between $345 billion and $1.7 trillion. Errors are clustered at the bank level.

Bank Size:	(Small)	(Medium)	(Large)
Panel A			
Tier 1 Capital to Total Assets	0.0	-0.014	-0.12 ***
	(0.0118)	(0.0183)	(0.0352)
Observations	285619	170213	124494
AUC	0.8	0.749	0.712
Panel B			
Tier 1 Capital to Risk Weighted Assets	-0.009	-0.039 *	-0.084 ***
	(0.0161)	(0.0203)	(0.0137)
Observations	285619	170213	124494
AUC	0.8	0.749	0.712
Panel C			
Tier 1+2 Capital to Risk Weighted Assets	-0.008	-0.03	-0.081 ***
	(0.0166)	(0.0208)	(0.0219)
Observations	285619	170213	124494
AUC	0.8	0.749	0.711
Panel D			
Parent: Tier 1 Capital to Total Assets	0.009	0.005	-0.122 **
	(0.0132)	(0.0137)	(0.0502)
Observations	285619	170213	124494
AUC	0.8	0.749	0.712
Panel E			
Parent: Tier 1 Capital to Risk Weighted Assets	-0.001	-0.012	-0.095 *
	(0.0165)	(0.0103)	(0.0556)
Observations	285619	170213	124494
AUC	0.8	0.749	0.712
Panel F			
Parent: Tier 1+2 Capital to Risk Weighted Assets	-0.001	-0.016 **	-0.063 *
	(0.0171)	(0.0079)	(0.0352)
Observations	285619	170213	124494
AUC	0.8	0.749	0.712
Panel G			
Market Cap to Total Assets	-0.003	-0.008 **	-0.002
	(0.0049)	(0.0033)	(0.0149)
Observations	121904	142781	124489
AUC	0.796	0.744	0.712
Panel H			
Market Cap to Risk Weighted Assets	-0.0	-0.011 *	0.001
	(0.0047)	(0.0056)	(0.0113)
Observations	121904	142781	124489
AUC	0.796	0.745	0.711
Banks FEs	yes	yes	yes
Time FEs	yes	yes	yes
Macro Controls	yes	yes	yes

Cluster robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 5

Small business loans - IV specification (year 2006 geographic weights). This table presents the results of an instrumental variables specification for a linear probability model. The response variable is a binary classifier of whether or not a loan defaults during its term. The capital variables are instrumented for using the weighted change in house price index in the areas in which banks operate. Bank market shares for the purposes of this instrument are assessed as of 2006. The macro controls used in these specifications include the national unemployment rate, the five-year returns on the S&P 500 index and the 5-year US treasury rate. These are the same macro variables used in all other regression specifications in this paper. Errors are clustered at the bank level.

	(Naive)	(IV)
Panel A		
Tier 1 Capital to Total Assets	-0.006	-0.04 **
	(0.0054)	(0.0201)
Observations	64194	64194
Adjusted R^2	0.09	0.079
Panel B		
Tier 1 Capital to Risk Weighted Assets	-0.004	-0.039 **
	(0.0033)	(0.0163)
Observations	64194	64194
Adjusted R^2	0.09	0.072
Panel C		
Tier 1+2 Capital to Risk Weighted Assets	-0.002	-0.067 *
	(0.0031)	(0.0365)
Observations	64194	64194
Adjusted R^2	0.09	0.029
Banks FEs	yes	yes
Time FEs	yes	yes
Macro Controls	yes	yes

Cluster robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 6

Geographic Foreclosure Predictions. These models predict the log total number of foreclosures in each US county from 2007 to 2012. The key dependent variable is the average capital level for banks in each county over the 2003-2006 period. Averages are calculated weighting based on the total deposits each bank has in each county. Mean FICO score is measured across all new mortgages originated in year 2000 in each county. House prices are measured in thousands of dollars. '% Small Banks' represents the percentage of deposits in a county in banks with $17 billion or less in assets. '% Medium Banks' represents the percentage of deposits in a county in banks with assets between $17 and $345 billion. Additional controls used in these regressions and not depicted in the table are: log of county population, population density per square mile, per capita income, median year houses were built, % population employed as of year 2000, % population below poverty, % population that is Black, White, or Hispanic, % population that is married, % population that has less than a high school education, a high school education, some college study, or greater than a bachelors degree, and the % of rental units in each county that is ≤ 30% median income, between 30 and 49% median income, and ≥ 50% median income. Data on foreclosures, open mortgages, and FICO scores from CoreLogic. Data on House prices from Zillow. Data on banks in counties from FDIC statement of deposits, all other data from 2000 Census. All calculations here robust to measuring foreclosures per population, rather than log foreclosures, as the dependent variable. Only 869 counties are depicted due to coverage of Zillow house price data. All calculations are robust to a larger set of counties 3168 counties that excludes house prices from the model. Counties are weighted in this regression according to population. All predictors except bank capital and bank size percentages are scaled to have unit variance.

	(Leverage Ratio)	(Tier 1 Ratio)	(Total Ratio)
Tier 1 Capital to Total Assets	-0.0		
	(0.0148)		
Tier 1 Capital to Risk Weighted Assets		-0.044 ***	
		(0.0106)	
Tier 1+2 Capital to Risk Weighted Assets			-0.04 ***
			(0.011)
log(Open Mortgages, 2006)	1.403 ***	1.38 ***	1.388 ***
	(0.0982)	(0.0939)	(0.0955)
Mean FICO, 2000	0.026 ***	0.026 ***	0.026 ***
	(0.0095)	(0.0094)	(0.0094)
House Price, 2000	0.002 **	0.001 **	0.001 **
	(0.0007)	(0.0007)	(0.0007)
% Change House Price, 2000-2009	-0.939 ***	-0.924 ***	-0.936 ***
	(0.0854)	(0.0854)	(0.0867)
Median Household Income, 2000	-0.072	-0.039	-0.047
	(0.0746)	(0.0744)	(0.0748)
% Small Banks	-0.002	-0.001	-0.002
	(0.0017)	(0.0017)	(0.0017)
% Medium Banks	-0.003 **	-0.003 *	-0.003 *
	(0.0015)	(0.0015)	(0.0015)
Observations	869	869	869
Adjusted R^2	0.966	0.967	0.966

Robust standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$