Essays in Financial Intermediation

Essays in Financial Intermediation

A dissertation presented

by

Vitaly M. Bord

to

The Committee for the PhD in Business Economics

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Business Economics

Harvard University

Cambridge, Massachusetts

April 2018

Dissertation Advisor:
Professor David Scharfstein

Author:
Vitaly M. Bord

Essays in Financial Intermediation

Abstract

This dissertation contains three essays on how differences among financial intermediaries affect the provision of financial services. The first essay, "Bank Consolidation and Financial Inclusion," focuses on deposit-taking and identifies the adverse effects of bank consolidation on lower-income depositors through the higher deposit account fees larger banks charge. In the second essay, "Large Banks and Small Firm Lending," Victoria Ivashina, Ryan D. Taliaferro and I examine the large and persistent shift in the composition of lenders to small businesses following the housing market crash in 2007. In the third essay, "Risk, Lending, and Organizational Form," I focus on mortgage lending and investigate how differences in organizational form drive both risk-taking during the real estate boom and subsequent performance after the real estate market crash.

Contents

List of Tables

List of Figures

Acknowledgments

First, I am deeply thankful to my advisors Victoria Ivashina, David Scharfstein, and Adi Sunderam for their advice and guidance throughout my graduate school experience. The essays in this dissertation have also benefited from discussions with members of the Harvard Economics Department and the Finance Unit at Harvard Business School: John Campbell, Yin Chen, David Choi, Anastassia Fedyk, Robin Greenwood, Sam Hanson, Nihar Shah, and Jeremy Stein, as well as the Maxwell Crew.

Most of all, I am grateful to my family and friends who have supported me over the past six years, with good food, great company, and wonderful encouragement. It would have been hard to persevere without their support.

To my parents

Introduction

The banking industry has undergone tremendous change in the past thirty years. As the industry has grown and become more competitive, it has also consolidated. Once prevalent community banks and credit unions have become minor players in the overall financial system, as large banks, diversified both geographically and across business lines, gained market share. As these changes take place, it has become ever more important to understand the the benefits and drawbacks inherent in the changing makeup of the US financial system, as exemplified by trade-offs between profitability and inclusion, diversification and contagion, and efficiency and risk.

This dissertation contains three essays on a common theme: how differences among financial intermediaries affect the provision of financial services. The first essay, "Bank Consolidation and Financial Inclusion," focuses on deposit-taking and identifies the adverse effects of bank consolidation on lower-income depositors through the higher deposit account fees larger banks charge. In the second essay, "Large Banks and Small Firm Lending," Victoria Ivashina, Ryan D. Taliaferro and I examine the large and persistent shift in the composition of lenders to small businesses following the housing market crash. In the third essay, "Risk, Lending, and Organizational Form," I focus on mortgage lending and investigate how differences in organizational form drive both risk-taking during the real estate boom and subsequent performance after the real estate market crash.

In Chapter 1, I examine the effects of bank consolidation on a seldom-considered group: low-income depositors. The existing literature on bank consolidation has focused on lending, and has mainly shown positive or neutral effects on efficiency, commercial loan rates, and small-business lending. The possible effects of consolidation on depositors have remained mostly unexamined, despite the large benefits to low-income depositors of having bank accounts and the potentially high costs of financial exclusion. I begin to fill this gap in the literature by providing evidence that the expansion of large banks has negative consequences for low-income depositors, due to the higher minimum balances and higher fees large banks charge on their deposit accounts, relative to small banks. Using a difference-in-differences methodology, I compare acquisitions of small banks by large banks to similar acquisitions by other small banks.

I find that account fees and minimum required balances increase after acquisitions at branches bought by large banks, and deposit outflow is higher, particularly in low-income areas. Using a measure of alternative financial services that are substitutes for bank accounts, I show that some of the deposit outflow corresponds to depositors who leave the banking system altogether. Moreover, there are long-term consequences to becoming unbanked due to the acquisition of small banks by large banks: households in areas affected by bank consolidation are more likely to experience financial hardship—such as evictions—after subsequent personal financial shocks, consistent with these households facing difficulty in accumulating emergency savings without bank accounts. The findings in this paper bring into focus the distributional effects of bank consolidation, and the contrast between the beneficial effects of mergers on efficiency and lending, and the detrimental consequences to low-income depositors.

Chapter 2 addresses the market structure effects of the contraction of credit in 2007-2008 by large, geographically diversified banks. My co-authors Victoria Ivashina and Ryan D. Taliaferro and I exploit the fact that not all geographical areas, and therefore not all banks, were exposed to the 2006-2007 decline in real estate prices. Focusing on counties that did not experience a significant drop in the real estate market, we compare geographically diversified banks with and without exposure to the real estate shock from their other locations. Large banks affected by the real estate shock contracted their credit to small firms, even in unaffected counties. By contrast, healthy banks—those not exposed to real estate price shocks— captured market share in both loans and deposits, by expanding their operations and even entering new banking markets. The turmoil from the real estate crisis presented an opportunity for healthy banks, and their expansion was comparable to the effects of geographical banking deregulation in the 1990s. Despite this expansion, the net effect of the contraction in credit was negative. Counties with a higher presence of affected banks in 2008 had lower aggregate credit and deposit growth, and lower entrepreneurial activity, through 2015. These findings highlight how the propagation of shocks from one market to another by diversified institutions can lead to large and persistent shifts in the composition of market participants.

Chapter 3 focuses on a third type of difference among financial institutions–differences in organizational form. Specifically I use this setting to test the credit supply explanation, that the cause was excessive risky lending to subprime borrowers, and that credit demand view of the housing crisis, that demand for bigger mortgages by middle and high income families also played a role in the real estate boom and bust.

First, I document that during the recent financial crisis, real estate loans originated by banks under-performed those of mutual institutions such as credit unions. Banks were more likely to fail, and they experienced higher delinquency and charge-off rates in their real estate portfolios. The results are not due to selection on observables, nor due to differences in regulation or mortgage origination and servicing. I find that differences in the types of mortgage lending originated by mutual institutions and banks are consistent with both the credit supply and credit demand explanations of the housing crisis; however, only the credit supply view is consistent with banks' under-performance. Although differences in the riskiness of real estate lending during the financial crisis do not fully explain the worse performance of banks' real estate portfolios, my results suggest that lending to subprime borrowers was a driver of the crisis, but overextension of credit to middle and high-income borrowers was not.

Chapter 1

Bank Consolidation and Financial Inclusion: the Adverse Effects of Bank Mergers on Depositors

1.1 Introduction

The U.S. banking industry has undergone dramatic consolidation over the past twenty-five years. In 1994, community banks with inflation-adjusted assets under $10 billion comprised 57% of deposits and 70% of all bank branches; by 2016, these numbers had fallen to 20% and 44%, respectively.[1] As the role of large banks in the U.S. financial system has grown, it has become more important to understand the impact of bank size on the provision of financial services. While there is an extensive literature on both the positive and negative effects of bank consolidation on efficiency and lending, the impact on depositors and the distributional effects remain less understood.[2]

[1] Sources: Summary of Deposits from the Federal Deposit Insurance Corporation and FFIEC Reports of Condition and Income (Call Reports).

[2] DeLong (2001), Avery and Samolyk (2004), Berger *et al.* (2005), and DeYoung *et al.* (2009), among others, examine the effects of consolidation on efficiency and lending. Prager and Hannan (1998) examine the effect of increased market power after mergers on deposit rates; Park and Pennacchi (2009) present evidence of deposit rate decreases after small bank acquisitions by large banks.

In this essay, I begin to fill this gap in the literature by providing evidence that the expansion of large banks has negative consequences for low-income depositors. Acquisitions of small banks by large banks cause some low-income depositors to exit the banking system, due to the high fees large banks charge on deposit accounts. Existing literature suggests that being "unbanked"—not owning a checking or savings account—has high long-term costs, including decreased ability to save for emergencies.[3] As I show, this lack of emergency savings lowers a household's ability to withstand personal financial shocks.

To explore how bank consolidation impacts low-income depositors, I first document that larger banks charge higher fees and higher minimum required balances on their deposit accounts using an extensive new dataset of account fees.[4] Figure 1.1 illustrates this relationship between bank size and average fee (left panel) and minimum account balance to avoid the fee (right panel) on checking accounts. These higher fees and minimum balances matter because survey evidence suggests that households respond to fees when making decisions on changing their bank providers or exiting the banking system altogether (Kiser, 2002; Federal Deposit Insurance Corporation, 2015). According to the FDIC Survey of Unbanked and Underbanked Households, almost 50% of households who do not currently have a bank account had one in the past; many cite high fees as one of the reasons for leaving the banking system.

To estimate the causal impact of consolidation on depositors, I use a difference-in-differences approach and compare, within the same county and year, the branch-level outcomes of acquisitions of small banks by large banks ("treatment" group) to outcomes of acquisitions by other small banks ("control" group). The main concern about a causal interpretation of this methodology is that whether a large or small bank is the acquirer may be correlated with factors that drive depositor behavior. For example, it is possible that large banks acquire worse-performing small banks or small banks with branches in neighborhoods that experience an increasing trend in the percentage of low-income households. However, based on observable characteristics, I find no evidence to support this type of selection. Zip codes where treatment

[3] See Barr and Blank (2008), Burgess and Pande (2005), and Celerier and Matray (2017).

[4] This finding has also been established by prior studies such as Board of Governors of the Federal Reserve System (2003), Hannan (2006), and Stavins (1999).

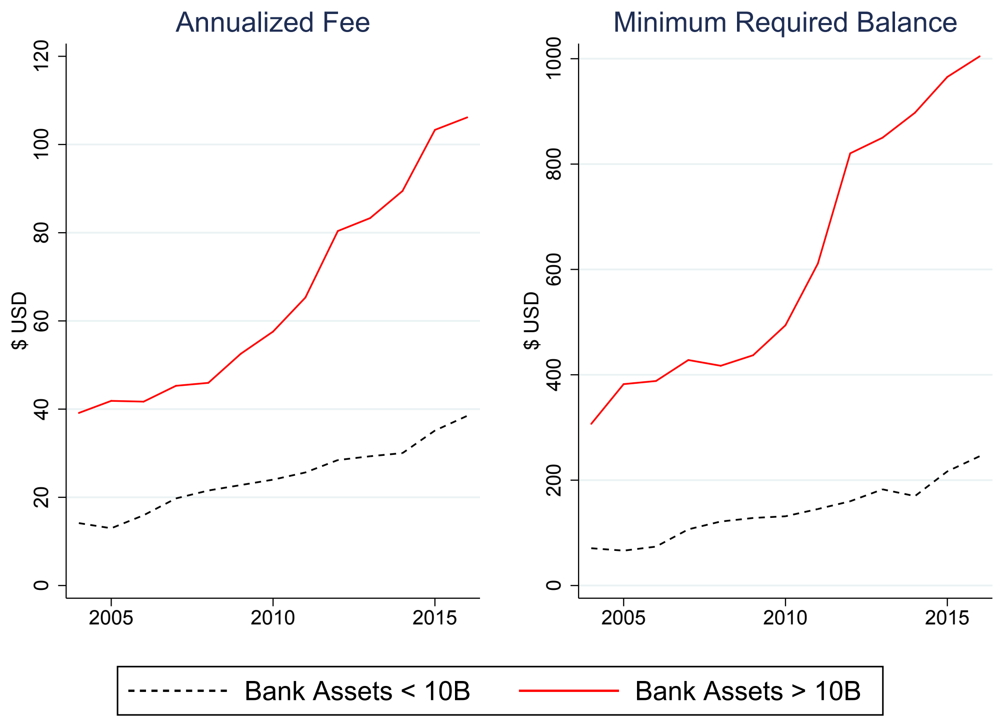


Figure 1.1: *Checking Account Fees and Required Minimum Balances by Bank Size*

This figure shows the average annualized fees (left panel) and average minimum balances held in the account to avoid the fee (right panel) for regular (non-interest bearing) checking accounts. All data are from RateWatch.

and control branches are located are similar in both levels and trends of economic and demographic variables such as income, unemployment rate, and the percentage of low-income households. Similarly, I employ a propensity score matching procedure and show that my results are almost the same after restricting the analysis to a sample of treatment and control mergers matched on observable bank characteristics.

I further address concerns regarding causality in two other ways. First, I create an instrumental variable based on the finding that acquirers are more likely to buy target banks that are close to them geographically (Granja *et al.*, 2017). Specifically, my main instrument is based on a target bank's geographic proximity to other large banks—calculated as the percentage of branches owned by large banks in 1994—in the zip codes where the target bank operates. This instrument plausibly satisfies the exclusion restriction; its effects on subsequent outcomes such as deposit growth come only through its effects on the acquisition decision. Second, my results are robust to limiting the analysis to a plausibly exogenous subset of peripheral branches, branches located in zip codes that contain less than 5% of the acquired bank's deposits. Since these branches are not central to the bank's operations, it is unlikely that any of their characteristics drive the acquisition decision; thus, for these branches, the acquisition is arguably exogenous.

This essay yields three primary sets of results. First, depositors leave small banks acquired by large banks, at least partially due to higher fees and required minimum balances after the acquisition. Branch-level deposit growth is lower at treatment branches than at control branches, corresponding to deposit outflow of about 1.8% per year after the merger. This effect is concentrated in the four years immediately after the merger and cumulatively, the difference in deposit growth between treatment and control branches is 12 percentage points over this period. Fees and minimum balances increase at treatment branches post-merger and, consistent with the hypothesis that depositors respond to the increased fees, deposit outflow is stronger in areas with more low-income households. Deposit outflow is also higher after a plausibly exogenous increase in large bank fees and minimum balances due to a regulatory change in 2011. Although other factors may also drive depositor outflow, neither preferences for small banks nor differences in customer service between small and large banks explain these results. In addition, increases in market power following acquisitions do not drive these findings. Thus, the effects of consolidation due to higher large bank fees are distinct from the effects due to increased market power (Prager and Hannan, 1998; Garmaise and Moskowitz, 2006).

Second, I find evidence consistent with some depositors, particularly those in low-income neighborhoods, exiting the banking system completely. My proxy for the presence of unbanked households is the number of check cashing facilities per capita in each zip code. This is an appropriate proxy because check cashing and formal banking services are substitutes; households who cannot, or choose not to, maintain a deposit account but receive checks turn to check cashing facilities. By five years after the merger, the number of check cashing facilities per capita in zip codes containing treatment branches increases by approximately one check cashing facility per seven zip codes. These results are stronger in zip codes with more than one branch involved in the acquisition, in zip codes with few other small bank branches, and in zip codes with more low-income households. These findings cannot be explained by differential trends in economic or demographic characteristics at treatment and control zip codes. Furthermore, post-merger branch closures do not drive the increase in check cashing facilities.

Third, there are long-term negative real consequences to becoming unbanked due to consolidation. Households in treated zip codes are less likely to withstand unemployment shocks during the Great Recession. Using zip code level evictions data from AIRS, I find that zip codes that had unemployment growth above the median in 2006-2010 experience higher rates of evictions than control zip codes with similar unemployment growth. This finding is consistent with more households in treated zip codes lacking the emergency savings needed to withstand shocks, due to not having bank accounts.

These findings have policy implications. Currently, when regulators decide whether to approve a bank merger or acquisition, they consider several channels through which the merger may impact firms and consumers. First, they examine the overall effect on competition for deposits by considering how measures of concentration might change after the merger. In addition, they often separately consider the possible impact of the merger on small business lending, since small and large banks engage in small business lending differently (Stein, 2002; Berger *et al.*, 2005). The findings in this essay suggest that in addition to small business lending, policy makers should also consider the differential impact of mergers on depositors, especially lower-income ones who may be substantially impacted by a rise in bank fees.

In this essay, I take as exogenous the differences in fees and minimum balances between small and large banks. However, existing literature suggests two explanations for the higher fees that large banks charge. First, deposit accounts at large banks provide some services that accounts at small banks do not, such as more extensive branch and ATM networks. Second, large banks' access to wholesale funding sources reduces their reliance on retail depositors as a source of funding (Park and Pennacchi, 2009). Because of access to wholesale funding, large banks pay lower deposit rates on interest-bearing accounts and charge higher fees on transaction accounts. Both explanations are consistent with intrinsic differences between small and large banks—unrelated to the costs of low-income depositors—driving the difference in fees. There is no evidence that lack of efficiency by large banks or absence of profit maximizing behavior by small banks explain large banks' higher fees (Kovner *et al.*, 2014; DeYoung and Rice, 2004).

The closest paper to mine is Celerier and Matray (2017), which examines the effects of one aspect of the changes in the banking industry: competition. By contrast, I examine the impact of a related but opposing mechanism: consolidation and the emergence of large banks, irrespective of market concentration. Using variation in state branch banking deregulation laws, Celerier and Matray show that increased competition after deregulation led to higher branch density and caused previously unbanked households to open new bank accounts, especially in historically excluded areas. This essay, on the other hand, focuses on consolidation, which also partially resulted from the deregulation laws, and led to the predominance of large banks with their higher fees. While there are forces that pull people into the banking system (such as the branch density examined by Celerier and Matray), my findings suggest that there are also countervailing forces pushing them out.

More generally, this essay contributes to several strands of existing literature. First, there is an extensive literature on the effects of bank consolidation and mergers, which mainly finds positive effects on efficiency (DeLong, 2001; DeYoung *et al.*, 2009; Hannan and Prager, 2006), and commercial loan rates (Erel, 2011), and neutral or positive results on small business lending (Berger *et al.*, 1998; Peek and Rosengren, 1998). A smaller literature documents negative effects as well. For example, increased market power due to mergers increases crime (Garmaise and Moskowitz, 2006). In addition, Nguyen (2017) finds a negative effect from large mergers on mortgage and small business lending due to branch closings. Complementary to this literature on lending, I examine the effect on depositors, and focus on the pricing of retail bank accounts. To my knowledge, this is the first paper that considers the effect of acquisitions on deposit account fees and required minimum balances, and estimates the impact on financial inclusion.[5] Prager and Hannan (1998) and Park and Pennacchi (2009) also present evidence of the negative effects of mergers on depositors but they focus on deposit rates. Finally, related papers examine

[5]Fees and minimums are more relevant than deposit rates for retail depositors, particularly lower-income ones. Amel and Hannan (1999) find that the supply of deposits in checking accounts does not seem to respond to the interest rates paid, while Stavins (1999) shows that deposits in checking accounts are sensitive to some fees.

size-related financing frictions that drive differences in lending and funding between small and large banks (Stein and Kashyap, 2000; Kishan and Opiela, 2000; Williams, 2017).

Second, I contribute to the literature on the determinants and consequences of financial inclusion. A rich literature has found several factors that impact a household's banking status including household characteristics and preferences (Rhine *et al.*, 2006; Barr *et al.*, 2011), as well as bank branch density (Celerier and Matray, 2017). In addition, studies in both the US and developing countries have documented the positive effects of having a bank account on savings rates and asset accumulation (Ashraf *et al.*, 2006; Prina, 2015; Celerier and Matray, 2017). I add to this literature by examining the consequences of an unbanked household's lack of savings after the household contends with a financial shock.

The rest of the essay is organized as follows. Section 1.2 outlines the existing research on the differences in fees between large and small banks and discusses the impact fees may have on depositors. Section 1.3 presents the data and methodology for the analysis of mergers, while Section 1.4 performs this analysis and examines what happens to deposit growth, fees, and the number of unbanked households. Section 1.5 examines the real and financial consequences for households pushed out of the banking system. Section 1.6 concludes.

1.2 Bank Consolidation and Bank Fees

In this section, I address the way through which bank consolidation may negatively impact low-income depositors. First, I document that large banks have higher fees and higher minimum required balances, relative to small banks. Second, I discuss existing survey evidence on the prevalence of financially fragile and unbanked households, who may find it difficult to pay high account fees and minimum required balances. These households' survey responses suggest that some of them respond to high account fees or minimum required balances by closing their deposit accounts and exiting the banking system.

1.2.1 Large Banks and Account Fees

Using an extensive new dataset, I document that large banks charge higher fees on their deposit accounts relative to smaller banks, as has been shown in several prior studies (Board of Governors of the Federal Reserve System, 2003; Hannan, 2006; Stavins, 1999). Although I take this documented difference as exogenous for my analysis of consolidation, I briefly discuss two possible explanations: differences in services provided by the accounts and differences in access to wholesale funding. For both of these explanations, the difference between small and large banks' fees is due to intrinsic differences between banks, and is not driven by large banks discriminating against low-income depositors.

Data

I use a new dataset of bank account and product fees from RateWatch. RateWatch surveys commercial banks, thrifts, and credit unions, and provides fees and rates for a wide variety of deposit accounts, including checking, interest checking, savings, and money market deposit accounts.[6] RateWatch also collects data on the minimum required balances needed to avoid the monthly fee, as well as fees for other types of products and services such as loan applications, ATM usage, and overdraft protection.

The advantage of this dataset is that it contains a panel of posted fees and rates for more than 1000 banks, tracking each branch over time. I avoid problems of prior studies, which either inferred the level of bank fees from bank-level revenue data in the quarterly Reports of Condition and Income (Call Reports) or used small, repeated cross-section samples of several hundred banks. Like with other deposit account survey datasets, the disadvantage is that the dataset includes only posted fees and the minimum balance needed to avoid the fee. I do not observe whether the account fee can be avoided in other ways, such as using direct deposit or debit card transactions.

Throughout the analysis, I follow the Federal Reserve's definitions and characterize as

[6] In many cases, RateWatch collects data on several different accounts for each bank. For each bank, I keep the account with the lowest fee, as this is the account most relevant for lower-income households.

"small" those banks that have less than \$10 billion in assets, in inflation-adjusted 2016 dollars.[7] Similarly, I define as "large" banks that have more than \$10 billion in assets.[8] The Office of the Comptroller of the Currency (OCC) defines community banks as those with less than \$1 billion in assets. My results are robust to using this definition instead.

Differences between Small and Large Banks

Next, I use data from RateWatch to show that transaction accounts at large banks have higher fees and higher minimum required balances needed to avoid the fees. Specifically, I run regressions of the form:

$$f_{b,c,t} = \alpha Large_{b,t} + \beta Large_{b,t} \times After2011_t + \lambda_{c,t} + \epsilon_{b,c,t} \tag{1.1}$$

$f_{b,c,t}$ are deposit account fees and minimums needed to avoid the fee for bank b in county c in year t, $Large_{b,t}$ is an indicator for whether the bank's assets exceed \$10 billion, and $\lambda_{c,t}$ county-year fixed effects. By including county-year fixed effects, I compare a bank's deposit account to those of other banks nearby, thus ruling out that the results are driven by market structure or differences in the economic characteristics of the areas where small and large banks have branches.[9] I include the interaction between $Large_{b,t}$ and $After2011_t$, an indicator for the post-2011 period, because fees and minimum balances on large banks' deposit accounts increase around the passage of the Durbin Amendment (See Figure 1.1). The Durbin Amendment to the Dodd-Frank Wall Street Reform and Consumer Protection Act, passed in 2010 and implemented in 2011, capped the interchange fee that large banks with more than \$10 billion in assets could charge on debit card transactions. Because it decreased the profitability of

[7]The Federal Reserve calls banks with less than \$10 billion in assets "community banks."

[8]Many previous studies of bank size split banks by whether they are single market or multi-market, rather than by size, e.g. Hannan (2006), Park and Pennacchi (2009), among others. My preferred specifications use size since the higher fees large banks charge are due to size-related advantages such as access to wholesale funding. In addition, the set of multi-market banks used by Hannan (2006) is highly correlated with the set of large banks used in this essay and my results are robust to defining large as multi-market.

[9]There is an existing literature on the effect of market structure on bank fees. Hannan (2006) finds that large, multi-market banks have higher fees than single-market banks, and a higher concentration of multi-market banks increases the fees the single-market banks charge. Azar *et al.* (2016) argue that local deposit market concentration, measured taking into account cross-ownership of banks, explains the cross-section of bank fees and interest rate spreads.

deposit accounts for these large banks, as a response, these banks increased fees and minimum balances on their accounts.[10]

Table 1.1 presents the results of equation 1.1: large banks have higher fees and higher required minimum balances relative to small banks, and this difference is not driven by the Durbin Amendment increase. In columns 1 and 2, the dependent variables are checking account fee and minimum balance to avoid the fee; in columns 3 and 4, I repeat the analysis for interest checking accounts. Standard errors are clustered at the county level, but double-clustering at both county and bank levels produces similar standard errors. In addition, large banks' higher market shares also do not drive these findings. In Table A.1 of Appendix A.3, I repeat the analysis restricting only to counties in which the average small bank share is higher than the average large bank share. In these counties, it is unlikely that differences in fees are driven by large banks' market power, and the results are similar. For the purposes of this essay, I focus on transaction account fees, but the results are similar for savings and money market deposit accounts.

There are several possible explanations for why large banks charge higher fees. One reason is product differentiation; large banks' deposit accounts provide extra services and amenities that consumers value and are willing to pay for. For example, large banks have more extensive branch and ATM networks, and are more likely to provide mobile banking and wire transfer services. In Table A.2 of Appendix A.3, I show that disparities in bank amenities explain part of the difference in fees between small and large banks. Using checking account fees as the dependent variable, I repeat the analysis of Table 1 but also include proxies of bank branch network (columns 1-2), a measure of the number of other services the bank provides (column 3), and a measure of customer service (column 4). Controlling for these characteristics explains approximately 70% of the difference in fees.

A second possible reason for large banks' higher fees is that they are willing to pay less for the marginal dollar of retail deposits. Park and Pennacchi (2009) present a model in which large banks are able to access wholesale funding sources, such as large uninsured deposits and the Federal Funds and repo markets, and thus have a funding advantage over small banks,

[10]See Kay *et al.* (2014) for a further discussion of the effects of the Durbin Amendment.

Table 1.1: *Deposit Account Fees*

This table shows the results of a regression of equation 1.1, estimating the difference in deposit account fees and minimum balances between small banks with less than \$10 billion in assets and large banks with more than \$10 billion in assets. The dependent variables are the fee on checking accounts (column 1), the average minimum balance needed to avoid the fee (column 2), fee on interest checking accounts (column 3), and the minimum balance on interest checking accounts (column 4). Each observation corresponds to a bank-county-year triple and I include county-year fixed effects. $Large_{b,t}$ *is an indicator for whether the bank has more than \$10 billion in assets, in inflation-adjusted 2016 dollars.* $Large_{b,t} \times After2011_t$ *is the interaction between this indicator and an indicator for the 2011-2015 period. Standard errors are shown in parentheses and are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.*

Dependent Variable:	Regular Checking		Interest Checking	
	Fee	Min	Fee	Min
	(1)	(2)	(3)	(4)
Large	10.739***	132.852***	34.230***	1368.052***
	(1.375)	(9.980)	(1.535)	(74.440)
Large x After 2010	25.308***	311.121***	20.187***	1897.964***
	(1.205)	(11.324)	(1.485)	(86.201)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	73,203	70,184	75,937	74,736
Within R-squared	0.075	0.116	0.119	0.099

since wholesale funding is cheaper than equity funding. In this model, large banks pay lower deposit rates on their accounts relative to small banks.[11] Insofar as banks use transaction account fees to raise funding, the Park and Pennacchi model suggests that large banks would also charge higher fees on transaction accounts, in addition to lower rates on interest-bearing accounts.

Under this explanation, any impact large banks' higher fees have on lower-income depositors is due to differences between banks, and is not related to depositors' incomes. Both explanations imply that banks value the marginal dollar of deposits from a high-income depositor and from a low-income depositor the same. An alternative explanation is that large banks have higher fees to price discriminate against lower-income depositors, either because they are

[11]Because large banks have lower funding costs, they also charge lower rates on loans. Erel (2011) provides evidence that commercial loan rates decrease after mergers.

more costly to large banks, or because they are more costly generally, relative to high-income households. This explanation is unlikely for two reasons. First, existing literature suggests that large banks are not less efficient than small banks, and have economies of scale in terms of infrastructure, salaries, and other costs (Wheelock and Wilson, 2012; Kovner *et al.*, 2014). This suggests that the cost to a large bank of a lower-income depositor should not be higher than the cost to a small bank. Second, if lower-income depositors are more costly than high-income depositors, then the question arises of why small banks do not also increase fees in order to discriminate against low-income households. Existing literature suggests that small banks do not have systematically lower profits, which would occur if they accepted costlier low-income depositors (DeYoung and Rice, 2004). In addition, as I show in Table A.4 of Appendix A.3, branches of small banks bought by other small banks are not more likely to fail or undergo subsequent mergers, relative to branches of small banks bought by large banks.

Thus, large banks' higher fees are likely driven not by differences in the costs of lower-income depositors, but by differences in account amenities and funding costs. For the purposes of this essay, I do not distinguish between these two explanations. Instead, I take the difference in fees as given and examine its effects on depositors.

1.2.2 Bank Fees and the Unbanked

In this section, I discuss the prevalence of unbanked and financially fragile households in the US. Survey evidence suggests that despite the high costs of not having a bank account, some lower-income depositors already at the margin of staying in the formal banking system decide to leave the banking system altogether due to high fees and high required minimum balances.

According to the FDIC National Survey of Unbanked and Underbanked Households (henceforth FDIC Survey), approximately 7% to 8% of US households are unbanked: they do not have any bank or credit union deposit accounts.[12] Lower-income households are more likely to be unbanked, with approximately 28% of households with an annual income of less than $15,000 and 20% of those with an annual income of less than $30,000 without bank

[12] An additional 5-8% are underbanked, which means they have a bank account but still use some deposit account alternatives such as check cashing, money orders, or prepaid cards.

accounts. Similarly, unbanked rates are higher among households with a single female head of household (18%) and minority households (17%).[13]

Households without access to the formal banking system have to instead utilize alternative financial services, also called fringe banking services. These products, which are essentially bank deposit account substitutes, include check cashing facilities, prepaid cards, money orders, and bill pay outlets. Check cashing facilities are establishments that immediately cash a consumer's checks, for a 3-5% fee. The unbanked use stores with bill pay centers (such as Wal-Mart) to pay their credit card or utility bills and turn to wire transfers and money orders in order to pay individuals or transfer money. Note that these deposit account alternatives are distinct from fringe banking services that are loan alternatives, such as pawn shops and payday lenders. Estimates of the monetary cost of fringe banking services range considerably but most estimates are on the order of $200 to $400 per year (Barr, 2004; Good, 1999).[14]

Although some unbanked households have never had bank accounts, the FDIC survey suggests that many used to be part of the formal banking system. Almost 50% of unbanked households surveyed by the FDIC had a bank account at some point in the past, and 30% of them mentioned high account fees and minimum balances as one of the reasons for leaving the banking system. Another 23% offer the unpredictability of fees as a reason for being unbanked. These percentages are consistent with the finding that a large percentage of the US population is financially fragile, unable to come up with even a relatively small sum of money if it were necessary. For example, using data from the TNG Global Economic Crisis Survey, Lusardi *et al.* (2011) find that 25% of Americans cannot come up with $2,000 within 30 days at all, and another 20% would have to sell some possession or turn to payday lending. Similarly, a Federal Reserve survey conducted in 2014 found that 44% of households would either be unable to produce $400 immediately or would have to borrow the money or pawn some possessions

[13] All calculations reflect data from the FDIC Unbanked/Underbanked Surveys of 2009-2015. These estimates are consistent with prior surveys (Rhine *et al.*, 2006).

[14] Even though the costs of using these services may be higher than the costs of maintaining a deposit account, households may still choose to be unbanked due to: convenience of fringe banking services' hours of operation (Consulting, 2000); unpredictability and high cost of other bank account fees such as overdraft fees (Melzer and Morgan, 2015; Servon, 2013); and decisions based on irrational or incorrect/incomplete information, similar to Agarwal *et al.* (2009) and Bertrand and Morse (2011).

(of Governors fo the Federal Reserve System, 2017). The growing presence of large banks with high fees and minimum balances may mean that these households can no longer afford their bank accounts.

At first glance, the high percentage of unbanked households who used to have bank accounts seems to contradict the general decrease in the fraction of households that are unbanked. According to the Federal Reserve Survey of Consumer Finances, the percentage of households without a transaction account decreased from 15% in 1989 to 7% by 2013. Celerier and Matray (2017) find that following banking deregulation laws, banks expanded their branch networks and more individuals entered the banking system. However, these findings are complementary, not contradictory, and show the counteracting forces that impact the unbanked. As the total number of bank branches increased from 64,000 in 1994 to 84,000 by 2016, the share owned by large banks increased from 30% to 56%. The growth and expansion of the banking industry that Celerier and Matray (2017) examine led to increased competition and reduced the unbanked population. At the same time, consolidation and growth of the largest banks provides a countervailing force that pushes some depositors out of the banking system. An increase in competition without the accompanying consolidation may have reduced the percentage of households without bank accounts even further. In Table A.3 of Appendix A.3, I use the FDIC survey data to show that at the MSA-level, the presence of large banks is positively correlated with the probability of being unbanked. I also confirm the Celerier and Matray (2017) finding that increased branch density leads to a lower probability of being unbanked, and show that this effect is driven mainly by small banks. Higher branch density by large banks increases the probability of being unbanked.

1.3 Empirical Design and Identification

Having discussed the survey evidence, I next turn to a causal estimation of the effects of bank consolidation on depositors. To test whether large banks' high fees and required minimum balances cause depositors to leave the banking system, I examine the effects of mergers in which a large acquirer buys a small target bank. Because banks that are acquired might differ from the general population of banks, I implement a difference-in-differences methodology and compare these acquisitions to similar cases in which the acquirer is another small bank.

An acquisition of a small bank by a large bank, relative to by another small bank, is an exogenous shock to the acquired institution only if whether the acquirer is large or small is randomly assigned. The threat to exogeneity is that large and small banks have different types of acquisition targets, in which case comparing the two types of acquisitions would be invalid. I address this threat to exogeneity in three ways. First, throughout the analysis, I show that there are no pre-trends in the main outcome variables and that it is only after their acquisitions, that small banks acquired by large banks and those acquired by other small banks experience differences in outcomes. Next, in Section 1.3.2, I show that, at the time of the acquisition, the two types of target banks are similar based on the household characteristics of the zip codes where their branches are located. This suggests that acquirers are not targeting certain banks based on the different customers of those banks. Finally, in Section 1.3.3, I discuss my instrument for whether the acquirer is a large bank, and in Section 1.4.1, I show that my results are robust to restricting my analysis to peripheral branches that are arguably unrelated to the merger.

1.3.1 Empirical Methodology and Data

In this section, I lay out my difference-in-differences methodology and discuss the advantages of using a control group of acquisitions by small banks.

The core of my identification is a comparison of small banks that are acquired by large banks (“treatment" group) to those that are acquired by other small banks (“control" group). Figure 1.2 graphically presents the benefits of using small banks acquired by other small banks as a control group in a simplified, univariate context. It is a plot of branch-level *forward-looking* deposit growth: the growth at time 0 is calculated from the year before to the year after the merger. I use the branch-level deposit growth from the FDIC as a proxy for changes in depositor entry into and exit from the acquired bank, since I do not observe individual depositors’ banking decisions.

Figure 1.2 illustrates two notable advantages to using acquisitions by other small banks as

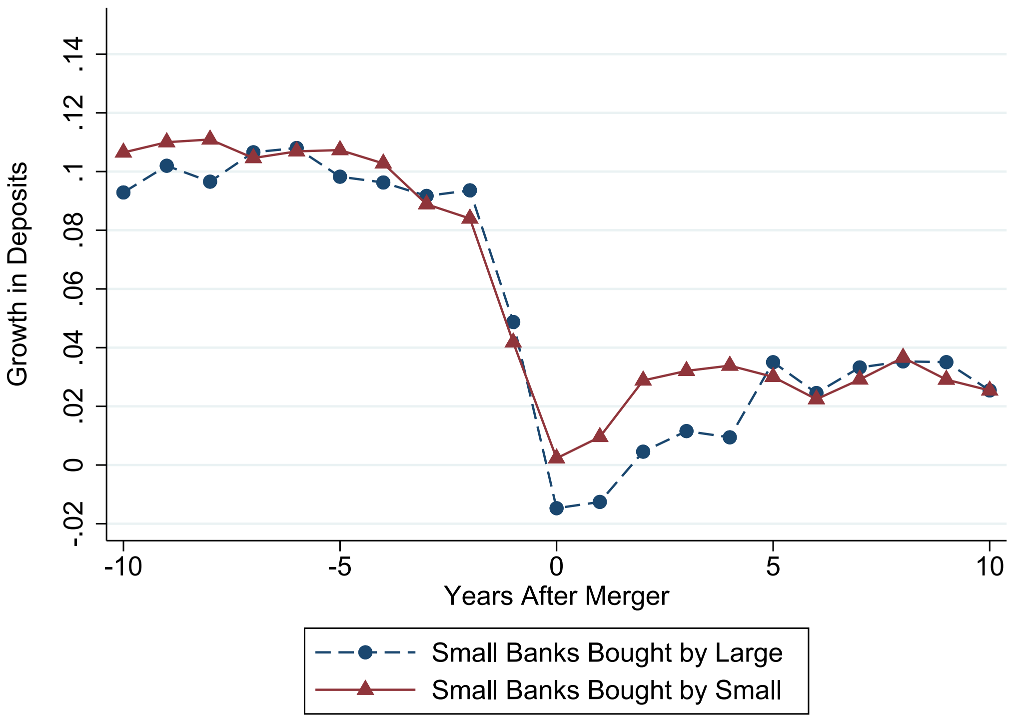


Figure 1.2: *Deposit Growth after Mergers*

This figure plots deposit growth for small banks of less than $10 billion in assets that are acquired either by large banks with more than $10 billion in assets ("treated" group) by other small banks ("control group"). All assets are inflation-adjusted to 2016 dollars. Deposit growth is calculated as the growth from the current period to the next period. Year 0 corresponds to June 30th prior to the merger.

the control group. First, deposit growth begins decreasing two years prior to the acquisition, suggesting that whether a bank is acquired or not is endogenous. Thus, comparing the treatment group to non-acquired banks would give biased results. Second, prior to the acquisition, the treatment and control banks experience fairly parallel trends in deposit growth. This suggests that the acquisitions are unlikely to be endogenously driven by differences in deposit growth. As I discuss later in Section 1.4.1, Figure 1.2 also previews my finding that branches of treated banks, those acquired by large institutions, experience lower growth rates in the 4-5 years following the merger, relative to branches of control banks.

To test how bank consolidation effects depositors, I perform a difference-in-differences analysis comparing, within the same year and county, branches of treated banks (small banks bought by large banks) with branches of control banks (small banks bought by other small banks), before and after the merger. Specifically, I run regressions of the form:

$$Y_{i,b,c,t} = \alpha_{c,t} + \beta_i + \tau_{b,t} + \delta \text{Bought by Large}_b \times \text{Post}_{b,t} + \epsilon_{i,b,c,t} \tag{1.2}$$

$Y_{i,b,c,t}$ is an outcome variable such as deposit growth, calculated as change in log deposits, or account maintenance fees for branch i of bank b in county c at time t. $\alpha_{c,t}$ are county-year fixed effects and β_i are branch fixed effects, which I include to capture any time-invariant branch characteristics. $\tau_{b,t}$ are event-time fixed effects, included to control for any general pre- and post-merger trends. My main coefficient of interest is δ, the coefficient on Bought by Large$_b \times$ Post$_{b,t}$, the interaction between the indicator for a small bank bought by a large bank (the treatment group) and the indicator for the post-merger period. δ captures the difference, within the same county and year, between the treatment and control group, after the merger relative to before.

I obtain bank branch location and deposit information from the FDIC Summary of Deposits, fee and minimum balance data from RateWatch, and bank financial statement data from the FFIEC's call reports. I supplement this with zip code characteristics from the Census, zip code level income data from the IRS's Statistics of Income, and data from the Census's County Business Patterns (CBP) and from Infogroup on the number of check cashing facilities, payday lenders, pawnshops, and total number of establishments in each zip code.

Using the FDIC's Summary of Deposits and the Chicago Federal Reserve's Bank Merger datasets, I create a panel dataset of bank and thrift branches and identify all ownership changes that occurred. Due to some inconsistencies in the Summary of Deposits branch-level identifier, I supplement this dataset with branch-level data from SNL Financial, as well as my own algorithm that matches branches by address. The end result is a panel dataset that tracks characteristics of each branch over time, for the time period 1994-2016. The advantage of this dataset, and of using the FDIC Summary of Deposits data, is that it provides yearly branch-level deposits. There is no public dataset on depositor banking relationships, so I proxy for depositor behavior by the deposit growth rates at the branch-level following the merger. Using this dataset, I am able to track ownership changes of each branch, as well as changes in address. I include branch divestitures in my sample, although limiting my analysis strictly to cases when a whole bank is bought does not change my results. I only consider mergers in which the target was a small bank with inflation-adjusted assets of less than $10 billion, and discard all cases in which the target was a failed bank.[15] I am left with 3,753 mergers, 680 in which the acquirer is a large bank and 3,073 in which the acquirer is a small bank. These mergers correspond to 15,139 branches.

1.3.2 Exogeneity and Summary Statistics

Having described my methodology, I now examine whether there are differences between treatment and control groups either at the bank-level or in the economic or demographic characteristics of the zip codes where the banks operate. Although small banks acquired by large banks differ from those acquired by small banks, these differences are unlikely to be a threat to exogeneity.

Table 1.2 presents summary statistics for the target banks, as of the year prior to their acquisition. Column 1 shows the difference between treated and control branches and banks; column 2 presents the t-statistics of the difference; and column 3 presents the mean for the control sample. Since my analysis includes county-year fixed effects, I include county-fixed effects when calculating the branch-level statistics.[16] Small banks bought by large banks differ from those bought by small banks on several dimensions. First, they tend to be bigger. Branches of the treatment group are bigger in terms of deposits, and these banks have more branches and more assets. In addition, treated banks have a lower ratio of deposits to assets and a lower tier 1 capital ratio. The fact that large banks acquire bigger small banks is consistent with Granja, Matvos, and Seru (2017), who find that acquirers of failed banks buy banks that are similar to themselves in terms of geographic footprint and business lines.

However, Table 1.2 shows little evidence of differences that would be a threat to exogeneity.

[15] I winsorize the branch data at the 1% level to exclude outliers. I also drop observations which have a low quality of the identifier I created to track each branch over time, as well as reorganizations—acquisitions of banks by other banks in the same bank holding company. I limit the analysis to branches that have deposit growth data for the time period from two years before the merger to two years after. Not constraining my sample does not change the results.

[16] Specifically, the summary statistics are calculated as $y_{i,b,c,t} = \alpha + \text{Bought by Large}_{b,t} + \lambda_c + \epsilon_{i,b,c,t}$, where λ_c are county fixed effects and Bought by Large_b is the indicator for treatment. I do not include fixed effects for the bank-level summary statistics.

Table 1.2: *Merger Target Summary Statistics*

This table shows the summary statistics for "treated" and "control" mergers. Column 1 shows the difference between treated and control branches and banks (treated minus control); column 2 presents the t-statistics of the difference; and column 3 presents the mean for the control sample. Treated banks are small banks with less than $10 billion in inflation-adjusted assets that are acquired by large banks with more than $10 Billion in assets. Control banks are small banks with less than $10 billion in inflation-adjusted assets that are acquired by other small banks. The branch-level summary statistics in Panel A are reported after adjusting for county fixed effects. Pct Cons Loans is the percent of the bank's portfolio in consumer loans. Core Deposits are the sum of demand deposits, deposits in NOW and ATS accounts, money market deposit accounts (MMDA), other savings deposits, and time deposits under $100,000 (FDIC (2013)). Pct Pastdue and NonAcc Loans is the total of loans pastdue and nonaccrual loans as a fraction of all loans. Tier1 Ratio is the bank's Tier 1 capital ratio. All data are from the FDIC's Summary of Deposits and the FFIEC call reports.

Panel A: Branch Variables	Difference	T-stat	Control Mean
Deposits in MM	9.347	5.461***	41.558
Checking Acct Fee	1.000	0.601	3.344
Checking Acct Minimum	89.375	0.450	356.523
Treated Branches	5636		
Control Branches	9503		
Panel B: Bank Variables	**Difference**	**T-stat**	**Control Mean**
Infl-adj Assets in MM	732.221	5.708***	712.365
Number of Branches	5.196	6.487***	3.092
Number of Counties	1.051	4.877***	1.634
Loans/Assets	0.013	1.294	0.625
Pct Cons Loans	0.013	1.207	0.101
Pct Real Estate Loans	0.025	1.600	0.664
Deposits/Liabilities	-0.024	4.333***	0.929
Core Deposits/Liabilities	-0.016	1.840*	0.783
Net Income/Assets	0.004	5.372***	0.006
Pct Pastdue and NonAcc Loans	-0.004	1.298	0.026
Net Chargeoffs/Loans	-0.002	3.521***	0.006
Tier 1 Ratio	-1.580	2.478**	15.170
Treated Banks	680		
Control Banks	3,073		

For example, one threat to exogeneity would arise if large acquirers bought banks that perform worse, and the worse performance of these banks drove depositors to leave. If this were the case, any difference in subsequent outcomes between my treatment and control group would be to due to selection rather than the treatment effect of having been bought by a large bank. Table 1.2 suggests that this is not the case. There is no evidence that the treatment group performs worse before the merger; in fact, the treatment group has higher income and lower net charge-offs than the control group.

A related threat to exogeneity is the possibility that these two types of banks have different types of customers or experience differential local macroeconomic shocks that drive both the acquisitions and the subsequent outcomes. Although I cannot rule this out completely since I do not have data on each bank's customers, evidence on observable zip code level characteristics suggests that this is not the case. Table 1.3 presents summary statistics on the zip codes where the branches of the treatment and control banks are located and reveals few differences.[17] First, in Panel A, I examine yearly zip code level measures of income and economic activity and show that there is no difference in these observable characteristics across the locations of the two types of branches. To capture demographic and socio-economic data, in Panel B, I examine differences based on zip code data from the 2000 Census. Small bank branches acquired by large banks tend to be located in more populated urban areas. However, there is no evidence that these areas have more lower-income households, a higher ratio of unemployed households, or that the change from 2000 to 2010 in unemployment, median income, or other characteristics is higher in treated zip codes (Panel C). As I show in Table 1.11 and discuss further in section 1.4.3, measures of local economic activity and economic characteristics of the households experience no trends around mergers. These results suggest that based on observable characteristics, the zip codes where the branches of the two types of banks are located are comparable in both levels and trends.

[17] As above, the summary statistics account for county fixed effects.

Table 1.3: *Branch Location Summary Statistics*

This table shows the summary statistics for the demographic and economic characteristics of zip codes in which "treated" and "control" bank have branches. Column 1 shows the difference between treated and control zip codes (treated minus control); column 2 presents the t-statistics of the difference; and column 3 presents the mean for the control sample. Treated banks are small banks with less than $10 billion in inflation-adjusted assets that were acquired by large banks with more than $10 billion in assets. Control banks are small banks with less than $10 billion in inflation-adjusted assets that were acquired by other small banks. All summary statistics are reported after adjusting for county fixed effects. The characteristics in Panel A are time-varying and are reported as of the year prior to the acquisition. Panel B presents characteristics as of the 2000 Census, and Panel C presents changes in characteristics between 2000 and 2010. Data on number of establishments and number of check-cashing facilities come from Infogroup. Deposit and branch information comes from the FDIC's summary of deposits. Data on the percent with AGI< $25,0000 and percent receiving the Earned Income Tax Credit (EITC) come from the IRS Summary of Statistics. Data for Panels B and C come from the Census.

Panel A: As of Merger Year	Difference	T-stat	Control Mean
Num. Check Cashers/Num. Households	0.041	0.879	0.677
Num. Branches/Num. Households	0.181	0.308	2.858
Pct with AGI < 25K	-0.000	0.085	0.420
Pct Receiving EITC	0.001	0.143	0.157
Panel B: As of 2000 Census	**Difference**	**T-stat**	**Control Mean**
Num. of Households	192.642	1.666*	7,795.345
Pop. Density	0.040	1.067	0.849
Pct Black	0.199	0.885	8.539
Pct Hispanic	0.137	0.506	8.203
Pct under Age 25	-0.013	0.093	33.361
Pct 65+	0.168	1.467	13.990
Pct with Bachelors Degree	0.034	0.217	16.020
Pct of Owner Occupied Housing	-0.008	0.027	69.088
Household Median Income	-31.727	0.099	45,367.934
Pct in Labor Force	-0.081	0.640	63.599
Pct Unemployed	0.008	0.203	3.267
Pct Living below Poverty Level	0.032	0.225	11.126
Panel C: Changes from 2000 to 2010	**Difference**	**T-stat**	**Control Mean**
Pct Unemployment	0.029	0.392	4.749
Pct in Labor Force	0.105	1.052	-4.829
Median Income	-128.890	0.925	11,965.554
Pct Living in Poverty	0.048	0.606	-1.524
Treated Zip Codes	5,991		
Control Zip Codes	11,901		

1.3.3 Instrumental Variables

Although the treatment and control acquisitions are similar based on observable characteristics, the possibility of unobserved selection remains a concern. In this section, I discuss the instrumental variables I use and present the first-stage results.

It is possible that although there are no differences in zip code economic and demographic characteristics of the two types of acquisitions, there may be still be differences in the characteristics of the customers of the specific institutions since banking markets are highly localized (Gilje *et al.*, 2016; Nguyen, 2017). Consider the following hypothetical scenario: customers of small banks acquired by large banks are in areas—neighborhoods within zip codes—that are becoming poorer. For instance, unemployment due to local establishment closures may lead a bank's customers to leave the banking system because they feel that they cannot afford to keep their accounts. This bank would then be bought by a large bank since the acquirer knows that the higher fees it charges will have little impact on the depositor base; the low-income depositors are leaving anyway. By contrast, customers of small banks acquired by small banks are in areas that are well-off financially, and these small banks are not acquired by large banks because the large banks know that depositors may react negatively to the higher fees. In this hypothetical example, the acquisition decision and the difference in depositor outcomes is driven by differences in the customer bases of the two acquired banks; acquisition by a large bank is correlated with, but does not cause, depositor exit.

To rule out endogeneity similar to this example, I turn to instrumental variables based on geographic proximity and similarity of loan portfolios. As Granja *et al.* (2017) show, acquirers of failed banks are similar to the acquired banks based on geography and business strategy. This is also the case for non-failure bank mergers. Almost all acquirers in my sample have branches in the same state as the acquisition: 58% have branches in at least one of the counties the acquired bank is located in; 28% of acquirers have branches in at least one of the same zip codes. Relying on this fact, I use as my instrument the percentage of large banks near the acquired bank. Because contemporaneous proximity to large banks might also be endogenous, I calculate this measure as of 1994. Specifically, for each zip code where the target bank has branches, I first calculate the percentage of branches owned by large banks in 1994. Next, I weigh each zip code by the percentage of acquired bank branches located there. This weighted average is a bank-level measure of the presence of large bank branches in 1994.

Thus, I estimate the effect on deposit growth of mergers with a large acquirer driven by the proximity to large banks. The exclusion restriction is that the percent of nearby branches owned by large banks in 1994 affects deposit growth only through its effects on the acquisition decision. The instrument would fail to address the threat to exogeneity only if areas with more large banks in 1994 are also associated with other demographic or economic changes in the late 1990s and 2000s that drive deposit outflow. This is unlikely, especially for the latter half of my sample, and restricting my analysis to the 2000-2016 period does not change my results. This instrument solves the endogeneity problem of the above example because it only captures the part of the acquisition decision driven by proximity, rather than the customer base.

Columns 1 and 2 of Table 1.4 present the results of the first stage regressions using the geographic proximity instrument. Because the treatment indicator Bought by Large_b is a binary variable, I follow Wooldridge (2010) and first estimate the probability of treatment using a probit (Column 1). I then use the predicted value from the probit as an instrument for treatment using two stage least squares (2SLS). Column 2 presents the first stage, which is strong, with an F statistic greater than 10, so I can reject the possibility of a weak instrument.

I also use an alternative instrumental variable based on potential acquirer loan portfolio characteristics. For each target bank, I calculate the Euclidian distance between its loan portfolio and a weighted portfolio of all large banks with branches in the same county. Similar to Granja *et al.* (2017), the Euclidian distance is calculated over the share of real estate, consumer, and commercial and industrial loans held by the bank as of June prior to the merger. This distance is a measure of similarity between the acquired institution and possible large acquirers. If a potential acquirer has a similar loan portfolio to the target, it is more likely to acquire the target due to potential synergy in lending. The probit and first stage regressions using this instrument are presented in columns 3 and 4 of Table 1.4, and are similar to the results in columns 1 and 2. Because the first instrument is based on geographic proximity, is as of 1994, and uses zip code level variation, it is my preferred specification.[18]

[18] The first-stage F-statistic for the second instrument is 11. Using both instruments results in an F-statistic of 12.7.

Table 1.4: *IV First Stage*

This table runs the first stages of the instrumental variables regressions. Column 1 displays the results of the probit regression of treatment on the geographic proximity instrument: the percentage of branches in each zip code owned by large banks in 1994, weighted by the percentage of the bank's branches in each zip code. Column 2 presents the first stage of 2SLS, instrumenting for Bought by Large$_b \times$ *Post*$_{b,t}$ *by the predicted value of Bought by Large*$_b$ *interacted with Post*$_{b,t}$*, e.g.* $\widehat{BoughtbyLarge}_b \times Post_{b,t}$*. Column 3 and 4 use the Euclidian distance between the target bank's and potential acquirers' loan portfolios. Standard errors are clustered at the county level (Columns 2 and 4). * p<0.10, ** p<0.05, *** p<0.01.*

Instrument:	Large Br Density 94		Euclidian Dist. to Large	
Dependent Variable:	Bought by Large$_b$ Probit (1)	Bought by Large$_b$x Post$_{b,t}$ OLS: First Stage (2)	Bought by Large$_b$ Probit (3)	Bought by Large$_b$x Post$_{b,t}$ OLS: First Stage (4)
Branch Density 1994	0.580***			
	(0.102)			
$\widehat{\text{Bought by Large}_b}$ x Post$_{b,t}$		1.368***		
		(0.127)		
Euclidian Dist. to Large			5.108**	
			(2.179)	
$\widehat{\text{Bought by Large}_b}$ x Post$_{b,t}$				1.367***
				(0.122)
County-Year Fixed Effects		Yes		Yes
Branch Fixed Effects		Yes		Yes
Observations	3,753	186,564	3,753	186,564
Within R-squared	0.007	0.417	0.004	0.406

1.4 Results

In this section, I estimate the causal effect of bank consolidation using my difference-in-differences methodology and the geographic proximity instrument. I first establish that immediately after the acquisition, more deposits flow out of treated branches than out of control branches. Consistent with higher fees and higher minimum balances being a driver of this outflow, fees and minimum required balances increase at treated branches after acquisitions. In addition, deposit outflow is stronger in areas where households are more likely to respond to higher fees and required minimum balances by leaving the bank. Finally, using a proxy for the presence of unbanked households, I present evidence consistent with some of these depositors leaving the banking system altogether.

1.4.1 Deposit Growth

I first examine the impact of bank consolidation on depositors at the acquired branches, using the forward-looking branch-level deposit growth rate as a proxy for changes in depositor entry into and exit from each branch. If some depositors respond to acquisitions of small banks by large banks by leaving the bank—for whatever reason—then relative to deposit growth at control branches, growth at treatment branches should be lower after the merger.

In Table 1.5, I implement the difference-in-differences methodology of equation 1.2, and find that, consistent with Figure 1.2, deposit growth decreases at treated branches after acquisitions, relative to control branches. All regressions include county-year fixed effects, so that the main variables of interest measure the differential decrease in deposit growth for treated banks compared to control banks after, relative to before, the merger in the same county and year. Standard errors are clustered at the county level, but are robust to clustering at both the county and merger level. Column 1 presents the OLS result. In column 2, my instrument is the percentage of large bank branches in each bank's zip codes in 1994. In column 3, I use my

Although in all cases, the F-statistic is greater than the rule of thumb of 10 suggested by Angrist and Pischke (2001), using the second instrument by itself or with the first is more likely to result in problems of weak instruments. Because just-identified instrumental variable analysis is median-unbiased, I present my results using just the first instrument and use the second instrument separately as a robustness check. My results are similar when using both instruments, but the magnitude of the coefficients is slightly larger.

Table 1.5: *The Effect of Consolidation on Deposit Growth*

This table presents the results of the difference-in-differences specification of equation 1.2, estimating the effect of bank consolidation on deposit growth. The dependent variable is branch-level deposit growth rate, calculated as the growth from the current year to the following year. Bought by $Large_b \times Post_{b,t}$ is the interaction between the treatment effect, whether the acquirer is a bank with more than $10 billion in assets, and the post-merger indicator. Column 1 presents the results from the OLS regression. In Column 2, I use as my instrument the percent of nearby branches that were owned by large banks in 1994. In Column 3, I use as my instrument the distance between the target bank's and possible acquirers' loan portfolios. Column 4 limits the sample to peripheral branches, which are in zip codes with less than 5% of the bank's deposits. Column 5 restricts the analysis to a propensity-score matched sample of mergers. County-year fixed effects and branch fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

Dependent Variable:	Branch Deposit Growth				
	OLS	IV		Sub-sample	
		Large Br Density 1994	Euclidian Dist to Large	Peripheral Branches	Matched Sample
	(1)	(2)	(3)	(4)	(5)
Bought by $\text{Large}_b \times \text{Post}_{b,t}$	-0.015***	-0.018**	-0.018**	-0.021***	-0.022***
	(0.004)	(0.008)	(0.008)	(0.007)	(0.008)
County-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes
Branch Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	186,564	186,564	186,564	97,099	102,827
Within R-squared	0.149	0.149	0.149	0.170	0.163

alternate instrument, based on the Euclidian distance between the bank's lending portfolio and a weighted average lending portfolio for large banks. The results are very similar in all cases and show that a treatment merger causes deposit growth rates to be lower by approximately 1.5-1.8 percentage points per year. The fact that the IV results are larger in magnitude than the OLS results is likely due to the OLS result not taking into account that small banks acquired by large tend to be in slightly more urban areas, which are likely to have generally higher deposit growth. Figure 1.3 presents the full set of yearly coefficients and 95% confidence intervals from a fully-saturated regression using my preferred instrumental variable. Consistent with the univariate analysis in Figure 1.2, Figure 1.3 shows no pre-trends and finds that the effect on deposit growth is concentrated in the first few years after the merger. Cumulatively, the first four years after the merger account for a 12 percentage point difference in deposit growth. As expected, the difference in growth rates does not persist long-term; after the initial adjustment period of 4 years, the difference between the two groups disappears. This is consistent with similar long-run depositor entry into, and exit from, acquired banks for both the treated and control groups. Small and large banks are both viable and in equilibrium, depositors choose which bank best suits their needs. The only changes happen around mergers, when some depositors leave the treated banks.

In columns 4 and 5, I show that my results are robust to limiting to subsamples for which the concern of endogeneity is mitigated. First, in column 4, I restrict my sample to peripheral branches, branches located in zip codes in which the bank has less than 5% of its deposits.[19] Even if large banks choose which small banks to acquire based on the consumer profiles of those banks, focusing on peripheral branches, whose consumers would not have an effect on the overall strategy or operations of the bank, should avoid this issue. For these peripheral branches, the merger is plausibly exogenous since the acquirers are not selecting based on the characteristics of these branches. Finally, in column 5, I limit the analysis to a propensity-score matched sample of mergers. Using the bank characteristics from Table 1.2, I estimate a propensity score of being acquired by a large bank, and match each bank bought by a large acquirer to a nearest neighbor, a similar bank undergoing a merger in the same year that

[19] The results are robust to using 2% or 1% of deposits, or 5% of branches.

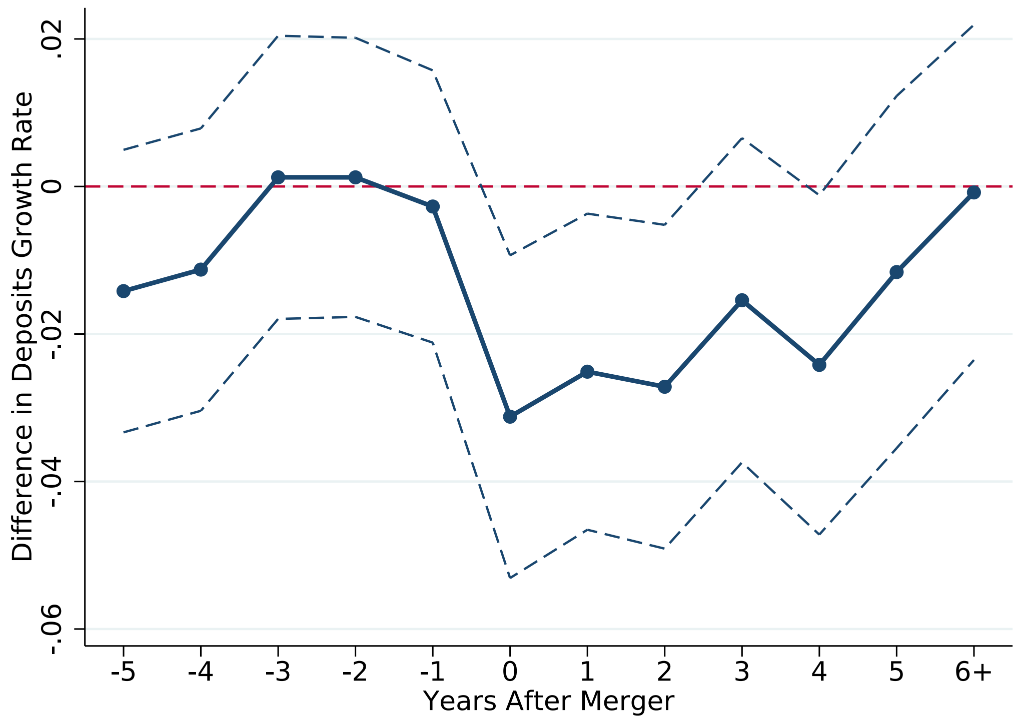


Figure 1.3: *The Effect of Consolidation on Deposit Growth*

This figure plots the yearly coefficients from a difference-in-difference regression estimating the effect of consolidation on deposit growth. The solid line shows the coefficients on the interactions between treatment, whether the bank was acquired by a large bank, and indicators for each years after the merger. Dashed lines capture the 95% confidence intervals. Year 0 corresponds to June 30th prior to the merger.

is bought by a small bank. Following Crump *et al.* (2009), I keep only observations with a propensity score between 0.1 and 0.9. Comparing acquisitions that are similar in characteristics mitigates the concern that selection on pre-merger characteristics by the acquirers drives the results. Although the sample sizes in columns 4 and 5 are smaller, the results are consistent with the OLS and IV analysis.[20]

[20]The larger magnitude of the coefficient on Bought by $\text{Large}_b \times \text{Post}_{b,t}$ is likely due to the fact that large acquirers tend to target better-performing banks, whereas in the matched sample, I compare banks of similar performance prior to the merger.

Robustness and Alternative Explanations

I perform several further robustness tests in Table 1.6 to rule out alternative explanations of my results. First, I show that the results are not driven by differential increases in market power by large banks nor by differences in regulatory approaches to approving mergers (e.g. regulators approving a large bank's purchase of a small bank only in economically dire situations). In column 1, I exclude counties in which the acquirer had a branch prior to the merger, and in column 2, I restrict the sample to mergers for which there was no increase in average concentration across the counties where the branches of the target bank were located. I measure concentration by the Herfindahl-Hirschman Index (HHI) of deposits, calculated as the sum of the squared market shares of each bank in the county. The coefficients are similar to the baseline results of Table 1.5 and highlight that my results capture not the effects of increased market power due to consolidation, but the effects of underlying differences between large and small banks. Next, in column 3, I exclude branches that changed address after the merger to rule out that branch relocations, rather than acquisitions by large banks, drive depositor exit. Finally, I address the concern that my regressions over-estimate the true coefficient because some depositors leave treated branches for control branches, thus inflating my coefficient. In column 4, I restrict the sample so as to remove any zip codes in which both treated and control branches are present. This is a sufficient restriction since I show in Section 1.4.3 that when depositors leave, they tend to go to nearby banks.

1.4.2 Deposit Account Fees and Required Minimum Balances

Having established that branches of treated banks experience lower deposit growth after the merger, I next examine why this happens. Although there are multiple factors that may drive outflow, in this section I focus on higher fees and required minimum balances, as discussed in Section 1.2. Not only do fees and required minimum balances increase after treatment acquisitions, but the deposit outflow is strongest in a) low-income areas, where households are more likely to respond to these higher account prices; and b) for mergers taking place after a plausibly exogenous increase in large bank fees and required minimum balances. In addition,

Table 1.6: *The Effect of Consolidation on Deposit Growth - Robustness*

*This table presents several robustness check for the results in Table 1.5. The dependent variable is branch-level deposit growth rate, calculated as the growth from the current year to the following year. Bought by Large$_b\times$ Post$_{b,t}$ is the interaction between the treatment effect, whether the acquirer is a bank with more than \$10 billion in assets, and the post-merger indicator. Column 1 includes only instances in which the acquirer did not have a branch in the same county as the target branch. Column 2 limits the analysis to mergers that did not result in an increase in the average HHI across the counties in which the target bank has branches. Column 3 excludes observations corresponding to address changes after the merger. Column 4 restricts the analysis to zip codes that had only treatment branches or control branches, but not both. All regressions use my preferred instrument, the percent of nearby branches that were owned by large banks in 1994. County-year fixed effects and branch fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.*

Dependent Variable:	Branch Deposit Growth			
	Out of County Acquirer	No HHI Increase	Excluding Address Chgs	No Other Mergers Near
	(1)	(2)	(3)	(4)
Bought by Large x Post	-0.019*	-0.027*	-0.016**	-0.024**
	(0.011)	(0.016)	(0.008)	(0.010)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Branch Fixed Effects	Yes	Yes	Yes	Yes
Observations	107,526	78,162	163,448	132,442
Within R-squared	0.148	0.148	0.149	0.149

these results hold even for reorganizations within the same bank holding company and for acquisitions that likely did not result in decreased customer service. I cannot rule out, nor do I maintain, that other factors such as depositor preferences do not play a role in the deposit outflow. However, taken together, the evidence I present is consistent with higher fees and higher required minimum balances driving at least part of the outflow.

Table 1.7 repeats the difference-in-differences analysis using checking account fees (column 1), checking minimum balances (column 2), interest checking fees (columns 3), and interest checking minimum balances (column 4) as the dependent variables. The main coefficient of interest, as before, is Bought by Large$_b\times$ Post$_{b,t}$, the interaction between the treatment indicator and the post period indicator. The table presents the results using the geographic proximity

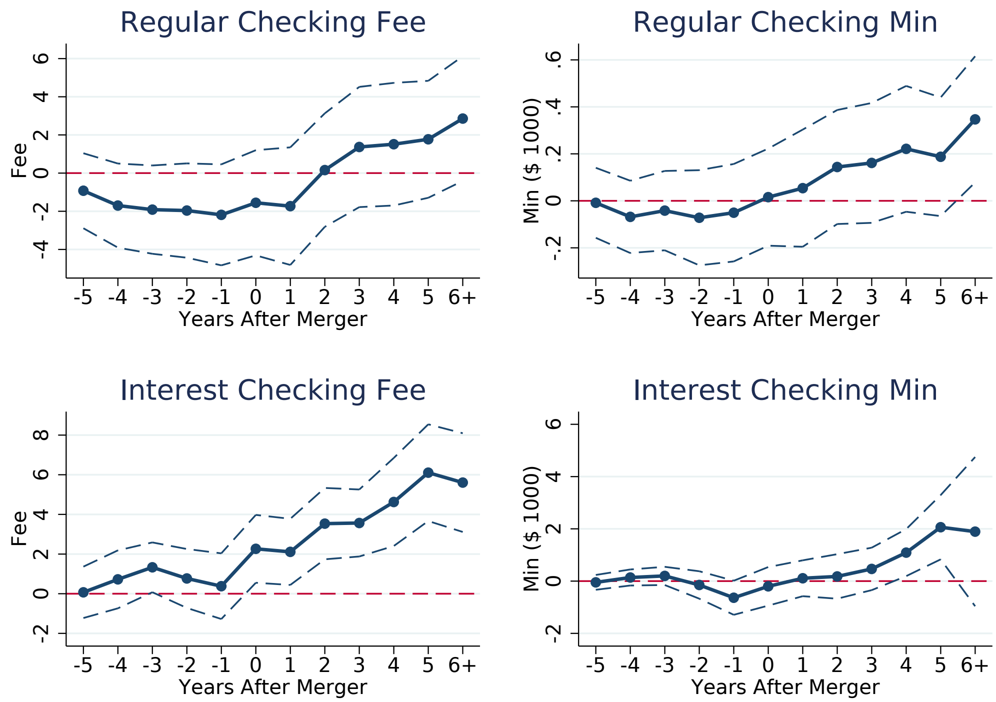


Figure 1.4: *The Effect of Consolidation on Fees and Minimum Balances*

This figure plots the yearly coefficients from a difference-in-difference regression estimating the effect of consolidation on deposit account fees and required minimum balances. The dependent variables are checking account fees (top left), checking account minimum balances (top right), interest checking account fees (bottom left) and interest checking account minimum balances (bottom right). The plot shows the coefficients on the interactions between treatment, whether the bank was acquired by a large bank, and indicators for each year after the merger. Dashed lines capture the 95% confidence intervals. Year 0 corresponds to June 30th prior to the merger.

instrument, and in all cases, the coefficient is positive and significant; small banks bought by large banks experience fee increases after the merger, relative to the control group. Figure 1.4 shows that just as with the deposit growth coefficients in Figure 1.3, there is no evidence of pre-trends prior to the merger. However, unlike the deposits growth coefficients, there is no time variation in the coefficients: fees and minimums increase after the merger and remain increased. The results are also robust to using the alternative instrument and restricting to peripheral branches or the propensity-matched sample.

On average, the regular (interest) checking account fee increases by approximately \$12 (\$34) per year and minimum balances increase by \$200 (\$600). The increase in mimimum balances is relatively similar to the types of financial shocks that many households state they would

Table 1.7: *The Effect of Consolidation on Bank Fees*

*This table presents the results of the difference-in-differences specification of equation 1.2, estimating the effect of bank consolidation on deposit account fees. The dependent variable are: branch-level annualized checking account fee (column 1); checking account minimum balance needed to avoid the fee (column 2); annualized interest checking account fee (column 3); and interest checking account minimum balance needed to avoid the fee (column 4). Bought by Large$_b \times$ Post$_{b,t}$ is the interaction between the treatment effect, whether the acquirer is a bank with more than \$10 billion in assets, and the post-merger indicator. All regressions use my preferred instrument, the percent of nearby branches that were owned by large banks in 1994. County-year fixed effects and branch fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

Dependent Variable:	Regular Checking		Interest Checking	
	Fee	Min	Fee	Min
	(1)	(2)	(3)	(4)
Bought by Large$_b \times$ Post$_{b,t}$	12.130***	228.498***	34.609***	623.131**
	(3.561)	(73.190)	(8.502)	(274.437)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Branch Fixed Effects	Yes	Yes	Yes	Yes
Observations	28,341	26,738	31,598	30,845
Within R-squared	0.002	0.087	0.039	0.051

not be able to overcome without difficulty (of Governors fo the Federal Reserve System, 2017). Although a yearly increase in deposit account fees of just \$15-\$30 a year seems small, this is probably an underestimate, especially for poorer households. Lower income households tend to overuse fee-based bank services such as overdrafts and these services also tend to have higher fees in large banks.

Fees at treatment branches increase after mergers because they converge to the fees of the acquirers. Figure 1.5 illustrates this in a univariate setting, plotting in the left panel checking account fees at treated branches and at branches owned by their acquirers in the same state as the treated branch. The right panel similarly plots checking account fees at control branches and at branches owned by their acquirers in the same state as the control branches. Fees at treated branches are low prior to the acquisition, but increase afterwards and converge to the fees of the acquiring institutions. By contrast, fees at control branches remain low; there is little difference between the fees of control branches and their acquirers before or after the acquisitions.

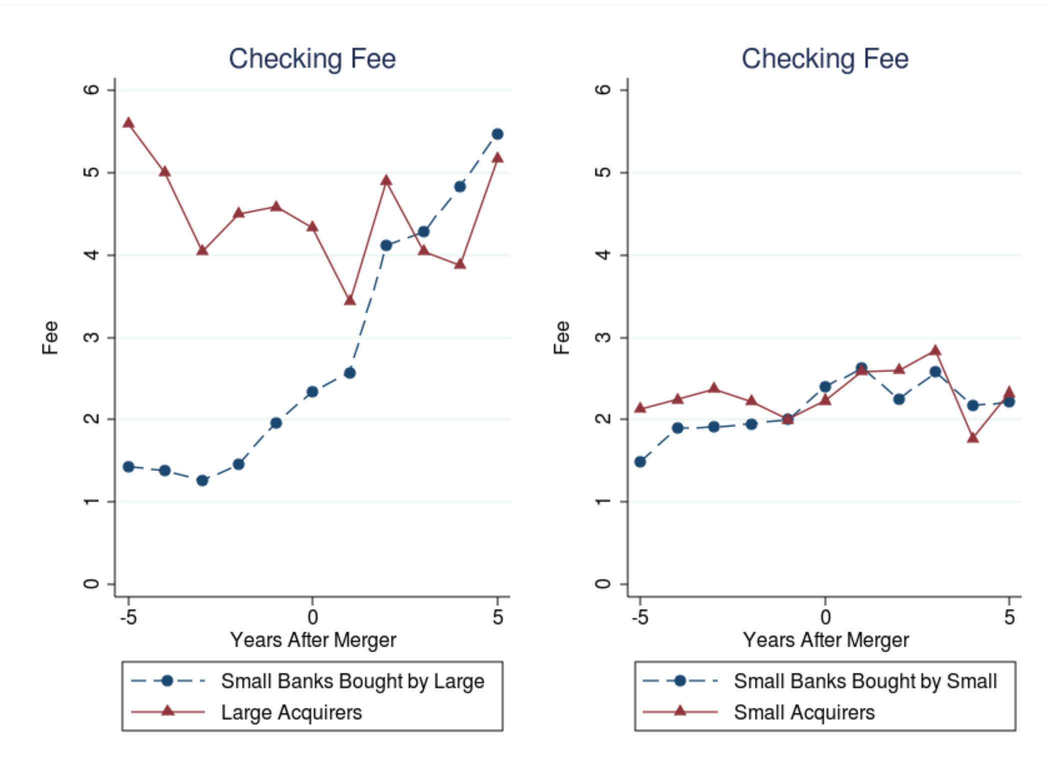


Figure 1.5: *Convergence of Checking Account Fees after Mergers*

The left panel of the figure plots checking account fees for small banks acquired by large banks (treated group) and their acquirers. The right panel plots checking account fees for small banks acquired by other small banks (control group) and their acquirers. Large banks are defined as those with more than $10 billion in inflation-adjusted 2016 dollars; small banks are defined as those with less than $10 billion in inflation-adjusted assets. Year 0 corresponds to June 30th prior to the merger.

To test the hypothesis that depositors leave due to increased fees, I next examine whether the results are stronger in areas with more low-income households, who are less likely to be able to bear the increased cost of holding a deposit account. I run a triple difference regression of the form:

$$Y_{i,b,c,z,t} = \delta \text{Bought by Large}_b \times \text{Post}_{b,t} + \chi \text{LowInc}_z \times \text{Bought by Large}_b \times \text{Post}_{b,t} \qquad (1.3)$$
$$+\phi \text{LowInc}_z \times \text{Post}_{b,t} + \alpha_{c,t} + \beta_i + \tau_t + \epsilon_{i,b,c,z,t}$$

As before, $Y_{i,b,c,z,t}$ is the deposit growth of branch i of bank b in zip code z and county c at time t. LowInc_z is an indicator for whether z is a low-income zip code. If depositors in low-income areas are more likely to leave the acquired branch, then χ should be negative and significant.

Table 1.8 presents the results of this triple-difference regression using different measures of low income zip codes. These measures are indicators for whether the branch is in a zip code that is *above* the median of the distribution of: the percent of households living below the poverty line in 2000 ($\text{I}\{\text{Pct Poverty}\}_z$; column 1); the percent of households with less than \$30,000 in income in the year prior to the merger ($\text{I}\{\text{Pct AGI} < \$25000\}_z$; column 2); the percent of households receiving the Earned Income Tax Credit (EITC), a government subsidy mainly aimed at working single mothers, in the year prior to the merger ($\text{I}\{\text{Pct EITC}\}_z$; column 3). In all cases, the interaction terms are negative and significant—deposit outflow is higher in lower-income neighborhoods, which are more likely to have trouble meeting the increased fees and minimum balances.

Next, I exploit a plausibly exogenous variation in fee increases caused by the implementation of the Durbin Amendment to the Dodd-Frank Act in 2011. The Durbin Amendment limited the debit card interchange fees for banks with more than \$10 billion in assets, and in response, many of these banks increased account fees (Figure 1.1; Kay *et al.* (2014)). In column 4 of Table 1.8, I test whether the post-merger deposit outflow at treated branches is stronger after passage of the Durbin Amendment by implementing a triple-difference with an indicator for the period after 2011, After2011_t. As expected, the coefficient on Bought by $\text{Large}_b \times \text{Post}_{b,t} \times \text{After2011}_t$ is negative and significant.

Alternative Explanations

I address two additional alternative explanations for my findings that large banks' higher fees and required minimum balances cause deposit outflow at acquired branches. First, low-income households may prefer small banks to large banks for reasons unrelated to deposit account fees, perhaps based on the name of the bank. In column 1 of Table 1.9, I present evidence on reorganizations that suggests preference for small banks does not exclusively drive my results. I define reorganizations as those mergers in which both the acquirer and the target are part

Table 1.8: *The Effect of Consolidation on Deposit Growth, by Area Income*

*This table presents the results of the triple differences specification of equation 1.3, estimating how the effect of bank consolidation on deposit growth differs by income and by degree of fee increase. The dependent variable is branch-level deposit growth rate, calculated as the growth from the current year to the following year. Bought by $Large_b \times Post_{b,t}$ is the interaction between the treatment effect, whether the acquirer is a bank with more than $10 billion in assets, and the post-merger indicator. I{Pct Poverty} is an indicator variable for whether the percentage of households in poverty is above the median of the distribution. I{Pct AGI < $25000} is an indicator variable for whether the percentage of households with income of less than $25,000 is above the median of the distribution. I{Pct EITC} is an indicator variable for the percentage of households who receive the Earned Income Tax Credit (EITC), a tax credit mainly aimed at working single female heads of household, is above the median of the distribution. After2011 is an indicator for the period 2011-2016, when large bank fee and required minimum balances increased substantially due to the Durbin Amendment. All regressions use my preferred instrument, the percent of nearby branches that were owned by large banks in 1994. County-year fixed effects and branch fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

Dependent Variable:	Deposit Growth			
	(1)	(2)	(3)	(4)
Bought by $Large_b \times Post_{b,t}$	-0.011**	-0.008	-0.016**	-0.015*
	(0.005)	(0.007)	(0.008)	(0.008)
Bought by $Large_b \times Post_{b,t} \times$ Post $\times$ I{Pct Poverty}	-0.011***			
	(0.004)			
Bought by $Large_b \times Post_{b,t} \times$ I{Pct AGI $<$ $25000}		-0.009*		
		(0.005)		
Bought by $Large_b \times Post_{b,t} \times$ I{Pct EITC}			-0.012**	
			(0.006)	
Bought by $Large_b \times Post_{b,t} \times$ After2011				-0.021*
				(0.011)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Branch Fixed Effects	Yes	Yes	Yes	Yes
Observations	183,432	115,507	183,432	186,564
Within R-squared	0.15	0.15	0.14	0.06

of the same bank holding company and I focus on reorganizations that took place prior to 2001. Following branch banking deregulation, many bank-holding companies reorganized since they no longer had to own different banks in different states but could instead rearrange their banks as one. In 40% of the reorganization mergers in my sample, the reorganization did not even involve a name-change, and it is possible that depositors were not even aware of these mergers.[21] Column 1 reruns the difference-in-difference methodology limiting my sample just to reorganizations. The results are somewhat weaker than in Table 1.5, but are consistent with the prior results. The smaller coefficient suggests that preference for small banks may drive some of the deposit outflow, but it is not the only driver.[22]

Second, differential changes in hours or customer service at acquired and control branches could explain my results. Survey evidence suggests that households sometimes switch banks due to a lack of convenience or poor customer service (Kiser, 2002), and it is possible that large banks have worse customer service and curtailed hours. To rule out this explanation, in column 2 of Table 1.9, I restrict my sample to mergers that likely did not result in changes to customer service. I measure the level of customer service as the number of full-time bank employees divided by the number of branches, and I only include mergers for which the customer service level at the acquirer was higher than at the target bank. The results are again similar to those of Table 1.5, which suggests that changes in customer service probably do not drive my findings.

1.4.3 Where do the Depositors Go?

Having established that some depositors leave treated branches after the acquisitions, and that this deposit outflow is at least partially driven by higher fees and higher required minimum balances, I next show that acquisitions of small banks by large banks cause an increase in the number of check cashing facilities in the zip code. This is consistent with consolidation driving

[21] For example, Suntrust Bank of Atlanta bought Suntrust Bank of South Florida in 2000 and subsequently changed its name to Suntrust Bank.

[22] Because Ratewatch started collecting fees in 2003, I am unable to verify that fees increase following these reorganizations. However, some reorganizations in the 2000s do result in higher fees and minimums. Higher fees at larger banks within the same bank holding company are consistent with more extensive branch and ATM networks as the explanation for large banks' higher fees. In most cases, depositors are not allowed to use services such as branches and ATMs of other banks even within the same holding company.

Table 1.9: *The Effect of Consolidation on Deposit Growth: Hours and Customer Service*

*This table presents a robustness check for the results of Table 1.5. The dependent variable is branch-level deposit growth rate, calculated as the growth from the current year to the following year. Bought by Large$_b \times$ Post$_{b,t}$ is the interaction between the treatment effect, whether the acquirer is a bank with more than \$10 billion in assets, and the post-merger indicator. In column 1, I include only reorganizations that took place between 1994 and 2000. A reorganization is defined as a merger between an acquirer and a target within the same bank holding company. In column 2, I restrict the sample to mergers in which the acquiring bank has a higher percentage of full time employees divided by total number of branches than the target. County-year fixed effects and branch fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.*

Dependent Variable	Branch Deposit Growth	
	Reorganizations (1)	Increased: Employees per Branch (2)
Bought by Large$_b \times$ Post$_{b,t}$	-0.013*	-0.021**
	(0.007)	(0.011)
County-Year Fixed Effects	Yes	Yes
Branch Fixed Effects	Yes	Yes
Observations	65,002	82,421
Within R-squared	0.160	0.156

some depositors out of the banking system. This result is not driven by selection or differential trends in economic characteristics, nor by branch closures.

A novel dataset from Infogroup allows me to proxy for the percentage of unbanked individuals in a zip code by the number of check cashing facilities per capita in the zip code. The disadvantage of this dataset is that, as with all the other data I use, I cannot track individuals' decisions. The advantages of the Infogroup dataset are two-fold. First, it allows me to identify the number of check-cashing facilities, which is a good proxy for the number of unbanked households. Check cashing facilities are substitutes for deposit account alternatives—unbanked households turn to check cashers to cash their employment, government assistance, and other checks. In the FDIC Survey, more than 45% of unbanked households, and more than 50% of unbanked households who used to have a bank account, use check cashing facilities.

Second, as I discuss further in Appendix A.1, the Infogroup dataset allows me to distinguish between check cashing facilities and payday lenders, even though both types of establishments are in the same 6-digit NAICS code. Whereas check cashing outlets are substitutes for bank deposit account services, payday lenders are substitutes for bank personal loans, and a bank account is often necessary to receive a payday loan. In the FDIC Survey, only 8% of unbanked households use payday lending services. If bank consolidation pushes some depositors out of the banking system due to higher deposit account fees, the zip code should experience an increase in demand for check-cashing facilities, but not in demand for payday lenders. I perform this robustness test later in this section.

Using the proxy from Infogroup, I test whether the number of check cashing facilities increases after bank mergers using a zip code level version of equation 1.2. Specifically, I run regressions of the form:

$$CC_{z,c,t} = \alpha_{c,t} + \beta_z + \tau_t + \delta \text{Bought by Large}_z \times \text{Post}_{z,t} + \epsilon_{i,z,t} \tag{1.4}$$

As before, I include county-year fixed effects, $\alpha_{c,t}$, zip code fixed effects β_z and event-time fixed effects $\tau_{z,t}$. Bought by Large$_z$ is an indicator for whether the zip code had a treatment branch or a control branch.[23] The dependent variable is the number of check cashing facilities per 10,000 residents. Columns 1 of Table 1.10 presents the OLS results and column 2 presents the IV. The magnitude of the coefficient on Bought by Large$_z$ $\times$ $Post_{z,t}$ is small, but in absolute terms, treated zip codes increase their ratio of check cashing facilities per 10,000 residents by approximately 0.045 more than control zip codes. Figure 1.6 presents the full set of yearly coefficients and 95% confidence intervals from a fully-saturated regression that checks for pre-trends. Since there are costs to opening a new check cashing facility, new entry takes time.[24] By 5 years after the acquisition, the difference between treated and control zip codes is responsible for approximately a 0.075 increase in the number of check cashing facilities per 10,000 residents, representing an increase of one check cashing facility per 7 zip codes, on

[23] There are few zip codes with both types of branches and they are excluded from my analysis.

[24] The increase in year 0 coefficient relative to the year -1 coefficient is likely due to the way the number of check cashing facilities is measured. Unlike deposits, which are as of June 30, the number of check cashing facilities is as of December 31st of each year.

Table 1.10: *The Effect of Consolidation on Check-Cashing Facilities*

*This table presents the results of the zip code-level difference-in-differences specification 1.4, estimating the effect of bank consolidation on the number of check cashing facilities. The dependent variable is the number of check-cashing facilities per 10,000 residents. Bought by Large$_z \times$ Post$_{z,t}$ is the interaction between the treatment effect, whether the acquirer is a bank with more than $10 billion in assets, and the post-merger indicator. Big Merger$_z$ is an indicator for whether the number of branches involved in the merger was above the median (which is 1). Few Small Br$_z$ is an indicator for whether the percent of other small bank branches in the zip code at the time of the merger was below the median. See Table 1.8 for definitions of I{Pct Poverty}$_z$ and I{Pct EITC}$_z$. All regressions use my preferred instrument, the percent of nearby branches that were owned by large banks in 1994. County-year fixed effects and zip code effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

Dependent Variable:	Num Check-Cashing Facilities / Population					
	OLS	IV				
	(1)	(2)	(3)	(4)	(5)	(6)
Bought by Large$_z \times$ Post$_{z,t}$	0.042**	0.045**	0.036	0.012	0.011	-0.001
	(0.017)	(0.023)	(0.023)	(0.026)	(0.024)	(0.035)
Bought by Large$_z \times$ Post$_{z,t} \times$ Big Merger$_z$			0.087***			
			(0.030)			
Bought by Large$_z \times$ Post$_{z,t} \times$ Few Small Br$_z$				0.061**		
				(0.024)		
Bought by Large$_z \times$ Post$_{z,t} \times$ I{Pct Poverty}$_z$					0.065***	
					(0.023)	
Bought by Large$_z \times$ Post$_{z,t} \times$ I{Pct EITC}$_z$						0.079**
						(0.033)
County-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Zip Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	123,118	123,118	123,118	123,118	123,079	67,125
Within R-Squared	0.002	0.002	0.002	0.002	0.002	0.003

average.

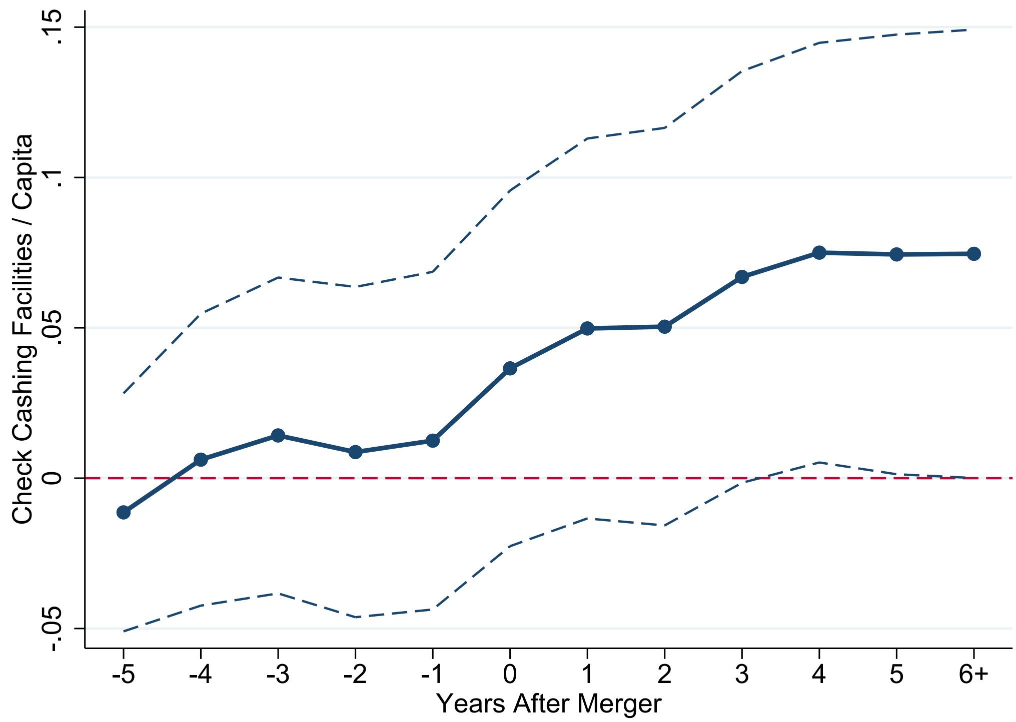


Figure 1.6: *The Effect of Consolidation on Check Cashing Facilities*

This figure plots the yearly coefficients from a difference-in-difference regression estimating the effect of consolidation on check cashing facilities. The dependent variable is the number of check cashing facilities per 10,000 residents. The plot shows the coefficients on the interactions between treatment, whether the bank was acquired by a large bank, and indicators for each year after the merger.Dashed lines capture the 95% confidence intervals. All variables are measured as of December 31st, and Year 0 corresponds to the year prior to the merger.

Next, I test whether the increase in the number of cash-checking facilities is larger in areas where more households are affected by the merger and where there are more lower-income households. In column 3, I run a triple difference, interacting my main variables with Big Merger, an indicator for whether the number of branches involved in the merger is greater than 1, the median. In column 4, I interact with Few Small Br, an indicator for whether the percent of other small bank branches, those uninvolved in any mergers, is lower than the median. Finally, I follow Table 1.8 and interact with indicators for whether the zip code is above the median in the percentage of households living in poverty (I{Pct Poverty}$_z$; column 5) and percentage of households receiving the EITC (I{Pct EITC}$_z$; column 6).[25] In all cases, the interaction term is positive and significant, and the coefficient on Bought by Large$_z$ $\times$ $Post_{z,t}$ is generally not significant. Thus, the increase in check cashers is concentrated in areas where more depositors were affected by the merger, where there are few other small bank branches for depositors to go to, and where there are more low-income households who find it more difficult to pay the increased fees and minimums.

Robustness and Alternative Explanations

In this section, I examine two possible alternate explanations for the increase in unbanked households following treatment acquisitions. The first potential concern, as before, is selection; namely, it is possible that both consolidation and the number of check cashing facilities are driven by trends in economic characteristics. If zip codes that experience higher growth rates of low-income households also experience higher rates of treatment mergers, this could explain the results I find above. Table 1.3 shows that based on cross-sectional observable characteristics, this alternative explanation does not seem to hold, and the instrumental variables analysis also helps address this concern. However, to further resolve this issue, in Table 1.11, I run the difference-in-differences methodology on several economic and demographic variables to show that they reveal no trends around the time of the mergers. In column 1, I use as my dependent variable the number of payday lending stores and pawnshops per 10,000 residents. If the increase in check cashing facilities is driven by higher percentages of low-income households, then I should also observe an increase in payday stores and pawnshops. However, this is not the case. In column 2, the dependent variable is the number of other establishments—excluding check cashers, payday lenders, and pawnshops—per 10,000 residents. In column 3, the dependent variable is log amount of mortgages originated. In column 4-6, I use as the dependent variable the average zip code average adjusted gross income (AGI), the percentage of filers with income of less than $25,000, and the percentage of filers that receive the EITC, respectively. In all cases, the coefficient on Bought by Large$_b$$\times$ Post$_{b,t}$ is not significant.

[25] The results are similar when using indicators for a zip code above the median in percent of households with adjusted gross income (AGI) less than $25,000 or zip codes with below median AGI.

Table 1.11: *Trends in Zip Code Characteristics around Consolidation*

*This table presents the results of the zip code-level difference-in-differences specification 1.4, estimating the effect of bank consolidation on other demographic and economic trends. The dependent variable in column 1 is the number of payday lenders per 10,000 residents. In column 2, the dependent variable is the total number of other establishments, excluding check cashing facilities and payday lenders per 10,000 residents. In column 3, the dependent variable is log mortgages originated, and in column 4, it is the average adjusted gross income (AGI) from the IRS's Statistics of Income. In column 5, the dependent variable is the percentage of households with AGI less than $25,000. In column 6, the dependent variable is the percentage of households receiving unemployment benefits. Bought by Large*$_z \times$ *Post*$_{z,t}$ *is the interaction between the treatment effect, whether the acquirer is a bank with more than $10 billion in assets, and the post-merger indicator. All regressions use my preferred instrument, the percent of nearby branches that were owned by large banks in 1994. County-year fixed effects and zip code fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

	Other Zip Code Characteristics					
Dependent Variable:	Payday Lenders	All Other Establishments	Log Mtg Orig	Ave AGI	Pct AGI < $25,000	Pct Unemployed
	(1)	(2)	(3)	(4)	(5)	(6)
Bought by Large$_z \times$ Post$_{z,t}$	0.011	0.064	-0.011	-0.064	0.001	-0.012
	(0.013)	(0.071)	(0.013)	(0.279)	(0.001)	(0.038)
County-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Zip Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	123,118	123,103	123,118	104,149	104,149	53,087
Within R-squared	0.001	0.002	0.004	0.004	0.004	0.001

A second alternative explanation for my results is that demand for check cashing facilities increases because of branch closures, rather than higher fees and required minimum balances (Nguyen, 2017). This explanation is unlikely because, although large acquirers do tend to close more branches than small acquirers, all the target banks in my sample are small and few branches are closed in these cases. In the first year after a merger, treatment zip codes experienced an average of 0.09 branch closures, as opposed to 0.08 for control zip codes. By three years after the merger, these numbers rise to 0.25 and 0.20 branches, and by five years after the merger to 0.35 and 0.28, respectively. This is equivalent to 0.3 branch closures *per merger* in the first year after the merger, 0.7 branch closures by 3 years after the merger, and 1 branch closure by 5 years after the merger. By contrast, the average merger between two large banks with more than $10 billion in assets each results in 13 branch closures in the first year. In Table 1.12, I rerun the difference-in-differences regressions limiting to zip codes that did not experience a branch closure 3 years after the merger (column 1) and 5 years after the merger (column 2). Next, I test a related explanation—that individuals leave the banking system due to recession-related job-losses. In column 3, I drop recession years and the years immediately following (2001-2002, 2008-2010) from my analysis. In all three cases, the results are very similar to those in column 2 of Table 1.10.

Depositor Switching Behavior

Next, I further examine the finding that the increase in check-cashing facilities is lower in areas with more branches of unacquired small banks. If depositors leave acquired banks due to the higher fees and minimum balances, one would expect that some of those who leave go to other nearby small banks. In Table A.5 of Appendix A.1, I test this hypothesis by repeating the difference-in-differences analysis of equation 1.4 using as the dependent variable deposit growth at branches of banks that do not undergo mergers or acquisitions. In each zip code, I calculate, separately, the zip code level deposit growth of small and large banks that do not experience a merger. As before, I define treated zip codes as zip codes that contain at least one branch of a small bank acquired by a large bank, whereas control zip codes are those that contain at least one branch of a small bank acquired by another small bank and conduct my

Table 1.12: *The Effect of Consolidation on Check-Cashing Facilities - Closures and Recessions*

*This table presents the results of the difference-in-differences specification 1.4, estimating the effect of bank consolidation on the number of check cashing facilities. The dependent variable is the number of check-cashing facilities per 10,000 residents. Bought by Large$_z \times$ Post$_{z,t}$ is the interaction between the treatment effect, whether the acquirer is a bank with more than \$10 billion in assets, and the post-merger indicator. Column 1 excludes all zip codes in which a branch of the merged institution was closed within 3 years after the merger. Column 2 excludes all zip codes in which a branch of the merged institution was closed within 5 years after the merger. Column 3 excludes recession years and the years immediately following a recession (2001-2002, 2008-2010). All regressions use my preferred instrument, the percent of nearby branches that were owned by large banks in 1994. County-year fixed effects and zip code fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

Dependent Variable:	Num Check-Cashing Facilities / Population		
	No Branches Closed after:		Excluding Recessions
	3 Years	5 Years	
	(1)	(2)	(3)
Bought by Large$_b \times$ Post$_{b,t}$	0.060**	0.057**	0.048*
	(0.027)	(0.027)	(0.026)
County-Year Fixed Effects	Yes	Yes	Yes
Zip Fixed Effects	Yes	Yes	Yes
Observations	95,859	89,882	80,148
Within R-squared	0.002	0.002	0.002

analysis at the zip code level.

After the acquisition, deposit growth at branches of other small banks increases slightly in treated zip codes, relative to control zip codes. By contrast, there is no effect on branches of large banks or branches of banks further away from the merger. First, in column 1, I use as the dependent variable the deposit growth at other small banks located in the same zip code as an acquisition. In column 2, the dependent variable is deposit growth at branches of large banks. In columns 3 and 4, the the dependent variable is zip code level deposit growth at branches of small and large banks, respectively, in zip codes adjacent to ones in which merger take place. The results are generally consistent with some depositors leaving small banks after they are acquired by large banks, and going to branches of unacquired small banks. Figure A.1 presents the set of yearly coefficients and 95% confidence intervals from a fully-saturated regression that checks for pre-trends. The coefficients are noisy, but the first panel—correponding to branches of other small banks— shows a general increase after the merger, whereas there is no effect on other sets of branches.

Because the average zip code has 3 branches of other small banks and 1.8 branches of acquired banks, the results suggest that the deposit growth at other small bank branches in the same zip code corresponds to approximately 50% of the deposit outflow estimated in Table 1.5.[26] In addition to going to other small, depositors may also leave acquired institutions for credit unions. I do not have branch-level data on credit union deposits, but on average, credit union branches are one-third the size of commercial bank branches, and the average zip code has 2-3 credit union branches. Thus, accounting for depositors who switch to credit unions is approximately equivalent to accounting for an extra small bank branch in the zip code. If this assumption is correct, even accounting for credit unions, depositor switching behavior only corresponds to 60% of the deposit outflow from acquired branches.

1.5 Real Consequences of Becoming Unbanked

Having documented the effects of mergers on depositors, I examine the consequences of becoming unbanked due to bank consolidation. There is evidence from developing countries that having a bank account improves a household's ability to save (Ashraf *et al.*, 2006; Burgess and Pande, 2005; Prina, 2015). Thus, those without bank accounts are less likely to be able to save and to smooth temporary shocks to earnings (Barr and Blank, 2008).

In this section, I test the mechanism of unbanked status causing a decreased ability to smooth consumption and withstand a financial shock. My hypothesis is that a personal financial shock, such as unemployment, has a larger effect on unbanked households, who have an impaired ability to save due to their unbanked status, and hence have less savings to rely on. Using a novel dataset of evictions from AIRS, I run a difference-in-differences analysis

[26] Acquired branches also tend to be larger than branches of unacquired banks.

testing whether households in treated zip codes that experience an unemployment-related zip code level shock during the Great Recession are more likely to undergo financial hardship than similar households in control zip codes, who also experience the shock. I find evidence supporting this hypothesis: households in treated zip codes that experience the shock are more likely to become evicted during the Great Recession.

Data and Methodology

My data on households' real and financial consequences come from American Information Research Services, Inc, and contains the number of evictions by zip code for each year from 2009 to 2016. I use as my measure of financial hardship the per capita number of evictions from 2009 to 2012.

Using the evictions dataset, I test whether treated zip codes—zip codes in which a large bank buys a small bank—had higher rates of evictions after the unemployment-related shock than households in control zip codes that also experience a the shock. I run regressions of the form:

$$f_z = \beta Shock_z + \gamma Treated_z + \delta Shock_z \times \text{Bought by Large}_z + X_z + \lambda_c + \epsilon_z \tag{1.5}$$

f_z is the percent of households evicted between 2009 and 2012, calculated as the number of evictions from AIRS divided by the total number of households in the zip code, as of 2000. Z_z are zip code level controls and λ_c are county fixed effects. Zip code controls include log number of households, population density, median income, whether the zip code is urban or rural, and percentages of households that are: black, Hispanic, aged 25-34, living in owner-occupied housing, in the labor force, unemployed, with earnings, and living in poverty. All zip code controls are as of the 2000 Census.

The ideal measure of a personal financial shock would be an individual's exogenous and unexpected unemployment. Because I do not observe individual-level employment, I instead use zip code level (or county level) shocks related to the Great Recession. Specifically, $Shock_z$ is one of the following measures of unemployment : 1) indicator for whether the county level increase in unemployment was above the median between 2006-2010 or 2) indicator for whether the zip code level increase in unemployment was above the median between 2000 and 2010.

Bought by Large_z is the zip code level treatment indicator from previous sections. My main coefficient of interest is δ, the coefficient on the interaction term between treatment and the personal financial shock.

The advantage of using measures of unemployment related to the Great Recession is that this was a very powerful shock, and business cycles have a stronger negative effect on lower-income and unskilled workers (Krusell *et al.*, 2009; Mukoyama and Sahin, 2006). Unskilled workers have a higher risk of becoming unemployed in recessions than skilled workers (Mincer, 1991). Thus, the employment shock is more likely to have a stronger effect on the same households who are at risk of leaving the banking system after acquisitions by large banks. This reduces concern that the set of households that drive the results below are not the same set of households who leave the banking system in Table 1.10.

Since the financial crisis began in earnest in 2008, I limit my analysis to mergers that take place between 2002 and 2007, and consider outcomes in subsequent years. To focus on areas where more households are likely to have become unbanked, I limit the sample to zip codes with more than 1 branch involved in the merger (the median number of acquired branches per zip code is 1).

Results

Table 1.13 presents the results, showing that treatment mergers coupled with personal financial shocks lead to higher evictions. The dependent variable in columns 1 and 2 is zip code level evictions from 2009-2012 divided by the total number of households. Using either measure of financial shock, households in treatment zip codes are more likely to be evicted during the Great Recession, by 0.004-0.005 evictions per household.[27] This corresponds to approximately 9,000 evictions in my sample. As expected, the coefficient on the zip code measure of financial shock is positive and significant. The high magnitude of the coefficient on the interaction of the shock and treatment, relative to magnitude on the coefficient on Shock_z, is due to the fact that I limit the analysis to zip codes where more than 1 branch undergoes a merger. In zip codes

[27]Relative to an estimate of approximately 6% of households who faced eviction during the time period (Desmond and Shollenberger, 2015).

Table 1.13: *The Effect of Consolidation and Financial Shocks: Evictions*

*This table presents the results of the difference-in-differences specification 1.5, estimating the effect of the interaction between bank consolidation and financial shocks on households. In columns 1 and 2, the dependent variable is the percent of households evicted during the 2009-2012 period. In column 3 and 4, the dependent variable is average zip code rent price during the 2009-2012 period. Bought by Large$_z$ is the treatment effect, whether the acquirer is a bank with more than \$10 billion in assets. County Unempl Shock$_c$ is an indicator for whether the county unemployment increase from 2006 to 2010 was above the median. Zip Unempl Shock$_z$ is an indicator for whether the zip code unemployment increase from 2000 to 2010 was above the median. Bought by Large$_z \times$ County Unempl Shock$_c$ and Bought by Large$_z \times$ Zip Unempl Shock$_z$ are the interaction terms between the treatment indicator and the financial shock. All regressions use my preferred instrument, the percent of nearby branches that were owned by large banks in 1994. County fixed effects and zip code controls are included in each regression. Zip code controls include log number of households, population density, median income, whether the zip code is urban or rural, and percentages of households that are: black, Hispanic, aged 25-34, living in owner-occupied housing, in the labor force, unemployed, with earnings, and living in poverty. All zip code controls are as of the 2000 Census. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

Dependent Variable:	Percent Households Evicted		Rent Prices	
	(1)	(2)	(3)	(4)
Bought by Large$_z$	-0.002	-0.001	-118.741	-101.956
	(0.002)	(0.002)	(76.551)	(73.222)
Bought by Large$_z \times$ County Unempl Shock$_c$	0.005*		32.452	
	(0.003)		(87.349)	
Zip Unempl Shock$_z$		0.002**		-69.761
		(0.0002)		(48.612)
Bought by Large$_z \times$ Zip Unempl Shock$_z$		0.004*		15.237
		(0.002)		(67.959)
County Fixed Effects	Yes	Yes	Yes	Yes
Zip Controls	Yes	Yes	Yes	Yes
Observations	941	941	824	824
Within R-squared	0.263	0.263	0.446	0.446

where only 1 branch is acquired, the effect of the shock remains strong, but the effect of the interaction term is smaller in magnitude and not statistically significant.

In column 3 and 4, I rule out that the increase in evictions is driven by higher rent prices. To do so, I use as my dependent variable average rent prices in the zip code during the 2009-2012 period. The interaction between Bought by Large_z and the measures of financial shock is not statistically significant, which shows that the combination of shock and treatment does not affect rent prices.

1.6 Conclusion

I estimate the effects of bank consolidation on depositors and find that the higher fees and minimum balances that large banks charge on their accounts cause lower-income depositors to leave the banking system. I find that relative to acquisitions of small banks by small banks, acquisitions by large banks lead to increased fees and minimum balances and deposit runoff at target institutions. Increases in the number of check cashing facilities several years after the merger suggest that some depositors leave the banking system altogether and instead turn to fringe banking services, such as check cashers. There are economically significant real and financial consequences to this phenomenon; households in areas affected by bank consolidation are more likely to be evicted when faced with an unemployment-related personal shock.

My results deal with a portion of the population that has only recently been studied in the finance literature. Low income households, and especially unbanked households, do not participate much in the traditional financial system and so have little effect on it. Yet there are almost 9 million unbanked households in the US, and the fringe banking industry is popular and growing. In 2016, consumers spent $5.6 billion in fees on check cashing, prepaid cards, billpay and money orders, which suggests that low-income and unbanked households constitute important markets.[28] Understanding how these groups respond to changes in the financial system may help explain why they do not play a larger role in it.

---

[28] See Schmall and Wolkowitz (2016).

Chapter 2

Large Banks and Small Firm Lending[1]

2.1 Introduction

This essay documents a significant and persistent withdrawal of the largest U.S. banks from small business lending following the Great Financial Crisis. However, the full effect on the credit supply to small borrowers is partially mitigated, as other banks opportunistically expand their market share.[2] Importantly, we show that significant changes in composition of credit providers to small firms persist in nearly ten years following the initial shock.

To isolate the credit supply effect, we exploit the fact that not all geographical areas were exposed to the 2006-2007 decline in real estate prices. Larger banks tend to be more geographically diversified, and leading up to the crisis, many of the largest banks operated in areas that had different real estate price evolution. Exposure to losses in one geographical area constitutes a shock to a bank's aggregate activities (Peek and Rosengren, 1997; Chava and Purnanandam, 2011; Schnabl, 2012, among others). Thus, as our main result, we consider geographically diversified ("large") banks with exposure to the real estate shock at the aggregate level, and examine their small business lending activity in counties that did not experience a significant drop in real estate prices. We benchmark these results against other large banks that

[1]Co-authored with Victoria Ivashina and Ryan D. Taliaferro

[2]In normal times, competition for borrowers is hindered by adverse selection, but when banks experience a negative shock, the information gap is reduced, facilitating lender substitution (Dell'Ariccia and Marquez, 2004; Darmouni, 2016).

were not exposed to real estate shocks in their portfolio; the latter, though also geographically dispersed, tend to be banks that are significantly smaller in size.

The core identifying assumption in the analysis is that small business lending is comparable across different banks. This assumption is similar to the one underlying within-borrower analysis that is typically used to identify the effects of credit supply shocks in countries with a centralized credit registry. In Khawaja and Mian (2008) and in much of the literature that builds on this work, there is no comparison of bank characteristics, and aggregate statistics indicate substantial heterogeneity in the bank sample. However, in these studies, the banks are assumed to be comparable in their decisions to lend to the same firm. Similarly, we assume that banks are comparable in making small business loans in a given county. To strengthen the cross-bank comparison, the main control sample of banks includes other geographically diversified banks that did not have exposure to real estate bust. In addition, we show that the findings are generalizable to comparing the largest banks to small, local banks in unaffected areas.

Our data on small business lending come from the Federal Reserve's Community Reinvestment Act (CRA) dataset and cover the period between 2005 and 2015. The advantage of these data is that they provide information about loan origination to U.S. small businesses (loans smaller than $1 million in size) at the county level for all deposit-gathering institutions except small community banks. Notably, small firms tend to borrow from banks and do so locally (Petersen and Rajan, 2002; Liberti and Mian, 2009; Agarwal and Hauswald, 2010). So, while the CRA data do not identify individual borrowers, it is feasible to assess the overall economic impact on small creditors using county-level information.

The economic magnitudes are substantial: in unaffected counties – those that do not fall into the top quartile of real estate depreciation between June 2006 and December 2007 – large banks exposed to the real estate decline elsewhere cut their small business lending by 25.2 percentage points more, from 2006 to 2008, that similar healthy banks. Over this period, healthy large banks increased their lending to small firms, even as exposed banks slashed their lending. It was not until 2009 (the trough of the widespread economic recession) that healthy banks' lending to small firms started to drop. After 2009, healthy banks' lending stayed the same and even increased, whereas the lending of exposed banks continued to decline through 2010. The results on the extensive margin are very similar: following the collapse in real estate prices, exposed banks were more likely than healthy banks to stop operations and close their branches in counties unaffected by the real estate shock. Healthy banks, on the other hand, expanded their operations and entered new markets, particularly ones that were highly competitive or historically difficult to enter.

Reallocation of lending toward healthy banks is also mirrored in the deposit market. We find that healthy banks substantially expanded their deposit taking activities, mostly by entering new counties. Whereas historically, deposits are sticky, the market share gain of healthy banks following the financial crisis was a standard deviation above the long-run historical average growth in market share. Overall, the deposit reallocation impact among the banks observed during the financial crisis was similar in magnitude to the disruptive effect of geographical deregulation of 1980s. This deposit reallocation is consistent with Mester *et al.* (2007) who show that information collection through deposit-taking and borrower monitoring are interlinked. However, the opportunity to capture deposits is not necessarily directly connected to lender substitution; it is plausible these are just independent responses to a negative shock to competing lenders.[3]

Finally, we examine the persistence of the change in the composition of small business lending and deposit-taking activity, and whether this change is associated with subsequent firm employment and new business creation outcomes. We show that the offsetting expansion and growth by healthy banks does not fully make up for the cutbacks by exposed banks, neither during nor after the financial crisis. Counties with a larger presence of exposed banks experience slower overall growth in both deposits and loans, and these effects persist through 2015.[4] In addition, these counties also experience higher unemployment, a higher decrease in the number of small business, and a lower rate of new business creation during the

[3] As with loan issuance, in comparing deposit-taking across banks with different characteristics, we assume that all banks offer the same deposit product to a given client (that is, there is no matching between clients and banks). Indeed, in our sample, deposits are insured by the Federal Deposit Insurance Corporation (FDIC), which makes deposits from different banks perfect substitutes.

[4] These results contrast with those in Greenstone *et al.* (2015), and we elaborate on this difference more in Section 2.5.

financial crisis. The decrease in the number of small business is concentrated in the smallest firms—those with fewer than 20 employees—and persists through 2014, as does the lower rate of new business creation.

The primary insight of this essay is to show the large and persistent change in the composition of small business lenders: we find that large banks retreat from this segment, leaving a long-lasting economic impact, despite the effect being partially offset by other creditors. In addition, studying credit availability to small firms is important in and of itself. Small business represents approximately 98% of all business establishments, 46% of GDP, and 40% of employment in the US.[5] Small firms are bank-dependent, and estimates of the bank supply contraction for large firms (e.g., Ivashina and Scharfstein (2010)) likely understate the corresponding effect on small firms.

More broadly, the findings in this essay contribute to several strands of literature. The evidence for the 2007-2008 period is consistent with the overall severe contraction in credit supply that followed the collapse of real estate prices in the US. Our focus, however, is on small business credit, and specifically, on the impact of the largest banks, as well as the competitive forces that counteract this effect. Recent work by Chen, Hanson, and Stein (2017) reiterate some of our results, emphasizing that the withdrawal of large banks from small lending was particularly pronounced among the top-4 banks.[6] We also see our work as contributing to the extensive and growing research on the role of geographically diversified banks in transmitting shocks across different markets. This includes Peek and Rosengren (1997) and Peek and Rosengren (2000), as well as more recent papers by Chava and Purnanandam (2011), Schnabl (2012), and Cetorelli and Goldberg (2012).

The outline of the essay is as follows. Section 2.2 describes the data sources. Section 2.3 details the identification strategy. Section 2.4 reports the key empirical results, which demonstrate the propagation of distant shocks through the balance sheets of large banks. Section 2.5 explores aggregate and long-term effects, and section 2.6 concludes.

---

[5] "Small Business GDP: Update 2002-2010," Small Business Administration, January 2012. "Small Firms, Employment and Federal Policy," Congressional Budget Office, March 2012. We are using 100 employees as a threshold.

[6] See Brian Chen and Stein (2017) for further discussion on the complementarity of their findings.

2.2 Data Sources

2.2.1 The CRA Data

The data on small business loan originations used in our study are collected by the Federal Financial Institutions Examination Council (FFIEC) under the auspices of the Community Reinvestment Act (CRA) enacted by U.S. Congress in 1977. Unlike data collected by the FFIEC under the Home Mortgage Disclosure Act (HMDA), which contain loans and applications for home-related loans, the CRA focuses on small business lending. Small business loans are defined as loans not exceeding $1 million. Specifically, the purpose of the CRA is to "encourage insured depository institutions to help meet the credit needs of the communities where they are chartered."[7] To that end, commercial banks and thrifts regulated by the Office of the Comptroller of the Currency (OCC), the Federal Deposit Insurance Corporation (FDIC), the Federal Reserve System, or, formerly, the Office of Thrift Supervision (OTS), must report annual data on the origination and purchases of small loans.

The CRA data contain a complete and unbiased view of small business loans at filing institutions. As part of their compliance with the CRA, institutions report three types of data. First, institutions must report the aggregate number and amount of loans designated as "community development" lending. Importantly, all filing institutions must also report the aggregate number and amount of all small business loans and small farm loans they originated or acquired during the reporting year. The data distinguish between purchased and originated loans. For the purpose of this study, we only include loan originations. Small business loans—the focus of our study—are loans whose original amounts are $1 million or less and that are either commercial or industrial loans or loans secured by non-farm, non-residential real estate. (Small farm loans are loans with original amounts of $500,000 or less that are either secured by farmland or used to finance agricultural production.)[8]

---

[7] Information about the CRA's purpose and data submissions can be found in the "Guide to CRA Data Collection and Reporting": http://www.ffiec.gov/cra/guide.htm. Additional information about the details of how institutions file can be found in the "Interagency Q&A": http://www.ffiec.gov/cra/pdf/2010-4903.pdf."

[8] Small farm loans and loan purchases are small in magnitude relative to small business loan originations, and including them does not qualitatively change our results. On average, small business loan purchases are 4% of small business loan originations and small farm loans are 8% of small business loans. Because the number of banks

The data are disaggregated by size and geographical location. For both small business and small farm loans, institutions are further required to break out the number of loans and total amount originated into the following categories: (i) loans smaller than \$100,000, (ii) loans between \$100,000 and \$250,000, (iii) loans between \$250,000 and \$1 million, and (iv) loans issued to very small enterprises with less than \$1 million in revenues. Institutions report loans by geographical location called "assessment area." Assessment areas are chosen by the banks but must be larger than a census tract and "consist generally of one or more metropolitan statistical division (MSA/MD) or one or more contiguous political subdivisions, such as counties, cities, or towns." Filing institutions must report CRA lending for geographies where its branches and ATMs are located, as well as "surrounding geographies in which it has originated or purchased a substantial portion of its loans." [9] For ease of comparison, we aggregate the data up to the county level.

Only banks above a certain threshold have to file CRA reports, and the assets threshold for reporting institutions increases each year. Notably, prior to 2005, all banks with more than \$250 million in assets, or those with less assets but associated with a bank holding company with more than \$1 billion in assets, had to file information under the CRA. In 2005 (the beginning of our sample), the size requirement was raised to \$1 billion. Since then, it has been adjusted slightly each year: in 2007, only institutions with more than \$1.033 billion in assets had to file, whereas by 2015, the size threshold rose to \$1.221 billion.

We exclude thrift institutions from our sample so that we may have consistent and consolidated balance sheet and income statement information at the holding company (BHC) level. Our final sample includes 648 banks; these banks account for 72% of all bank deposits in 2006. The fact that only large banks have to file might appear restrictive insofar as small firms are more likely to borrow from smaller banks (e.g., Stein, 2002; Berger et al., 2005). However, as of 2006, CRA filers—banks with more than \$1 billion in assets—had on their balance sheets 63% of all small business lending as reported by all banks in their Call Reports. In sum, the CRA

that engage in small farm loans or business loan purchases is small, we are not able to confirm that our results hold for just small farm loans or business loan purchases. Our results remain unchanged when adding small farm loans or business loan purchases to business loan originations.

[9] http://www.ffiec.gov/cra/guide.htm.

data cover a large fraction of U.S. banks and a large fraction of bank lending to small firms.

CRA data complement the small business and farm lending data reported by institutions in Schedule RC-C of the Consolidated Reports of Condition and Income ("Call Report"). The Call Report data are a stock: they report the total number and amount of all small business and small farm loans outstanding. By contrast, the CRA data are a flow: as of the end of each calendar year, the data give insight on the total number and amount of small business loans originated in that year. In contrast to HMDA data or loan data collected through a credit registry in foreign countries, the limitation of CRA data is that they are aggregated by geographical location and by size, and do not identify individual borrowers. As mentioned earlier, the data cover only loans smaller than $1 million in origination size issued by relatively large banks.

Data collected under the CRA is used to assign CRA ratings which the FFIEC takes into account when an institution applies to engage in merger and acquisition (M&A) activity or to open a new branch. In the context of our study, one could be concerned that an increase in small lending by relatively healthy banks could be a response to an increase in regulatory and public scrutiny or a desire to pursue M&A more broadly: those banks that can afford it increase lending in counties that were lagging on CRA compliance. The use of a difference in difference approach at the county level gets around these issues.

2.2.2 Other Data Sources

We use four other main data sources:

Banks' balance sheet data: We obtain quarterly bank and bank-holding accounting data from the Federal Reserve. We use data at the bank holding company (BHC) level (FRB form Y9C), or, if a holding company is not available, we use data at the bank level (Call Report level). Throughout the essay, we use the term "bank" to refer to the consolidated entity.

Deposits: We use Summary of Deposits (SOD) data from the FDIC, measured annually as of June 30th, to ascertain the deposits of each bank across its branches in each county. (Commercial banks that the CRA are a subset of institutions that file FDIC deposit data.) We aggregate deposits by county to the corresponding bank holding company if the bank holding company exists, and to the corresponding commercial bank if it does not. We also obtain branch-level deposit rate data from RateWatch.

The real estate shock: We obtain county-level real estate price index data from FISERV. FISERV publishes Case-Shiller house price indices using same-house repeated-sales data. Although the data are available at the zip-code level, we use county-level information in correspondence with our data on small business lending.

Local economy: Finally, we use annual county-level demographics data as of 2006 from the Census Bureau and employment and establishment data from the County Business Patterns and Census Business Dynamics Statistics datasets.

2.3 Empirical Design

To examine the change in the composition of lenders following a shock, we first establish a causal link between bank-level shocks and effect of their propagation on lending more broadly. To do so, we compare the lending behavior of banks that were exposed to geographies that suffered severe drops in real estate prices with the behavior of relatively unaffected banks. To identify a supply channel, we examine lending behavior by these two types of banks within areas that were not affected by real estate price shocks. We subsequently examine banks' deposit-taking behavior in an analogous way.

To illustrate our methodology, consider the representative cases of the following two banks in our sample: TCF Financial Corp and Amcore Financial Corp. TCF is a medium-sized regional bank with assets of \$14.3B as of June 2006, and primarily operating in the Midwest—Illinois, Minnesota, Michigan, and Wisconsin. TCF had branches in several areas that were severely affected by the real estate crisis, including 12% of its deposits in Oakland and Washtenaw Counties, MI which experienced real estate declines of 15% and 13%, respectively, from June 2006 to December 2007. Because of the size of its presence in these markets, TCF is characterized as "exposed" (to the real estate shock) by our algorithm.

Amcore Financial is also a medium-sized regional bank, with assets of \$5.4B as of June 30, 2006. It also serves the Midwest and all of its branches are in Illinois and Wisconsin. However, unlike TCF, Amcore did not have branches in counties that experienced a big drop in real estate prices. The worst performing county where Amcore had branches was Vermillion County, IL, which experienced a 5% decline in real estate prices, and Amcore had less than 5% of its deposits located there. Because of this, Amcore is characterized as "healthy" by our algorithm. Since both banks serve the Midwest, the overlap in several counties, including two where the real estate prices did not fall—Cook County, IL and Milwaukee County, WI.

The crux of our identification comes from comparing the deposit-taking and lending activities of exposed banks, such as TCF, with healthy banks, such as Amcore, in counties in which they both operate and in which real estate prices did not plummet. This empirical design requires us to focus on banks that are geographically dispersed. To measure geographical presence, for each bank, we compute the Herfindahl-Hirschman Index (HHI) of deposits across the counties the bank operates in, as of June 30, 2006. That is, for each bank, we calculate the percent of total deposits that are in each county, and the HHI is the sum of the squared percentages for each bank. Our main sample is constrained to banks in the lowest quartile of the HHI distribution.[10] The banks in this sample represent 65% of all U.S. bank deposits as of 2006.

We define counties affected by the real estate shock as counties in the bottom quartile of the distribution of the change in the Case-Shiller index from the end of the second quarter of 2006, the quarter in which the real estate market reached its peak, to the fourth quarter of 2007, the beginning of the economic recession. Thus, the *affected counties* are counties with a 2006:Q2–2007:Q4 decline in real estate prices in excess of 2.5% (an average decline of 10%). By comparison, the national Case-Shiller index fell by 6.1% during this period. Since banks in our sample have a broad geographical presence, we calculate each bank's deposit-weighted exposure to the affected counties, and then scale this exposure by the bank's total deposits as of June 30, 2006. By this measure, banks in the bottom quartile are classified as exposed to the

[10] This approach is comparable to Cortes (2013), who defines banks as local if more than two-thirds of their deposits are located in the main MSA or county market where they operate. Our definitions characterize as local all banks that are defined as local by Cortes, and approximately 80% of our local banks are characterized as local by Cortes's definition. Our results are also robust to using a filter the distribution of the number of counties each bank has branches in.

real estate shock.[11] We classify banks as *unaffected* or "healthy" otherwise. Exposed banks have an average deposit-weighted exposure of -5.7%, as opposed to an exposure of only -0.2% for healthy banks.[12]

Deposits are an indirect measure of banks' county-level exposure to real estate shocks. In Figure 2.1, we validate this measure by looking at aggregate metrics that show that banks we classify as exposed indeed experience substantial distress. First, as compared to unaffected banks, banks classified as exposed sustain a substantially larger increase in real estate loans past due (as a fraction of total loans) starting in 2007. Second, these exposed banks also show a rise in net charge-offs on real estate loans. Both of these patterns are consistent with high exposure to the real estate shock.

The two lower panels in Figure 2.1 show evidence for the two mechanisms through which the initial real estate shock led to the contraction in credit supply. The first is capital constraints: the average Tier 1 capital ratio declined for healthy as well as exposed banks, but it sank more, and remained much lower, for exposed banks. Although on paper, even exposed banks had Tier 1 capital ratios higher than the minimum 4% requirement, there was widespread concern at the time that Tier 1 ratios were not representative of banks' true financial health.[13] News outlets reported that this anxiety over banks' financial health led regulators to push banks to raise more capital, making sure that their Tier 1 ratios were much higher than the minimum requirements.[14] As illustrated in the bottom left panel, the sudden rise in banks' Tier 1 ratios in the last quarter of 2008 and in 2009 is consistent with capital injections.

---

[11] The results are robust to other definitions of healthy and exposed based on terciles or quintiles of the distribution. Defining healthy banks as those in the top two quartiles also does not change our results. Measuring the real estate shock from 2006:Q2 to any period from 2007:Q2 to 2008:Q2 does not change our results. Finally, as discussed further in the robustness section, defining healthy and exposed banks based on the number of affected counties the bank has branches in, or the percent of deposits in affected counties, does not change the results.

[12] Because counties that experienced a larger real estate shock were more likely to be urban and large, by our definitions, the average exposed bank had approximately 47% of its deposits in affected counties, while the average healthy bank only had 13% of its deposits in affected counties.

[13] For example, the *Wall Street Journal* pointed out in April 2008 that Citibank and Merrill Lynch had avoided letting certain write-downs impact their income statements and Tier 1 ratios by classifying them as "comprehensive other income." See "A Way Charges Stay off Bottom Line," *Wall Street Journal*, April 21, 2008.

[14] "Banks Told: Lend More, Save More; Can They Do Both? Regulators Want to See More Capital, Regardless," Wall Street Journal, December 26, 2008.

The second mechanism is increased risk of short-term lending. Due to exposed banks' higher charge-offs and falling capital ratios, the risk of short-term lending to these banks increased in both the Federal Funds and repo markets (e.g., Gorton and Metrick, 2012; Afonso, Kovner, and Schoar, 2011). The bottom right panel of Figure 2.1 shows that both healthy and exposed banks had trouble rolling over their short term federal funds and repo debt as those markets became stressed in late 2007, but the effect was greater for exposed banks.

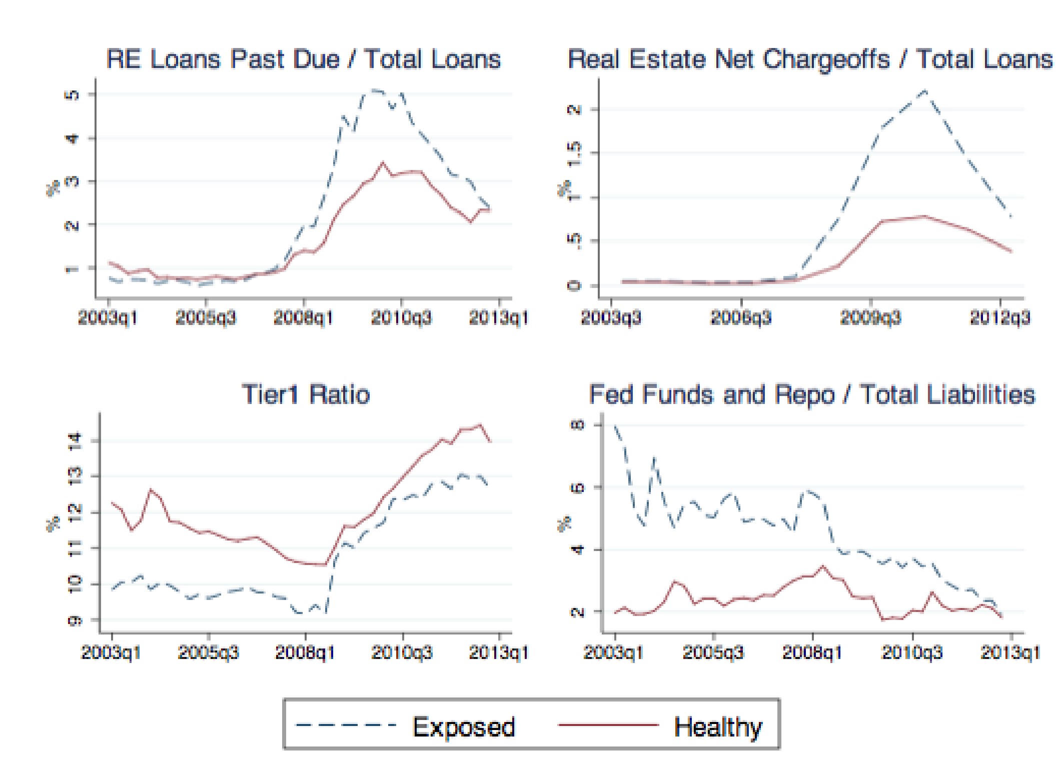


Figure 2.1: *The Effect of the Real Estate Shock on Healthy and Exposed Banks*

This figure shows the effect of the real estate shock on various characteristics of healthy and exposed banks over time.

To quantify the credit supply shock due to the real estate price decline the exposed banks experienced, we examine the subsequent lending behavior of healthy and exposed banks. Importantly, we need the clientèles of the banks in our sample to be comparable within the dimensions that we consider: lending to small firms and deposit-taking. This assumption parallels the borrower fixed effect approach typically used to identify effects of credit supply in countries with a centralized credit registry. The papers using this approach, such as Khawaja and Mian (2008), similarly assume banks to be comparable because they lend to the same borrower, and not because their characteristics look the same. The underlying assumption in such an approach is that different banks lend to the same borrower for the same purpose and/or with the same collateral. In our setting, small business loans originated by different institutions are comparable because they are reported under the auspices of the CRA, enacted "to encourage depository institutions to help meet the credit needs of the communities in which they operate." Thus, from a policy perspective, these loans are seen as comparable.[15] This means that a firm's employment is local, but it does not necessarily mean that the demand for the firm's products is local. The concern is that TCF and Amcore Financial, from our earlier example, both lend to small firms in Cook and Milwaukee counties, but TCF lends to exporting firms (i.e., firms with out-of-county demand) and these firms, in turn, might be affected by the out-of-county shock through a drop in demand. To address this possibility, we look at the trends in lending growth up to the real estate shock.

If exposed banks are indeed lending to exporting firms, we would expect not only a differential collapse in lending following the real estate shock, but a differential increase in lending during the boom in real estate prices. In other words, the positive correlation with real estate prices in counties that experience the real estate boom and bust should arise throughout the cycle and not just in the downturn. Figure 2.2 suggests that this is not the case: the lending patterns from 1996 to 2006 are very similar for both groups, especially in the $250,000 to $1 million category, which comprises the majority of the value of small business lending done by the banks. We test for trends (displayed beneath the figure) in the pre-crisis period of the figure by regressing the amount of each bank's small business lending (scaled by the 2006 level) on our indicator for healthy banks, a linear time trend, and the interaction of the two, using county fixed effects. The results show that there is no differential time trend between the two groups (the interaction is not statistically significant) and the indicator for a healthy bank is not significant in the period prior to the crisis. [16]

---

[15]http://www.ffiec.gov/cra/history.htm.

[16]Alternatively, we test for trends by including the full 1996-2015 sample, year fixed effects and the interactions between each year indicator and our healthy bank indicator. Prior to 2006, none of the interactions are significant, whereas all the interactions from 2007 to 2014 are positive and significant. We also find no difference in pre-trends

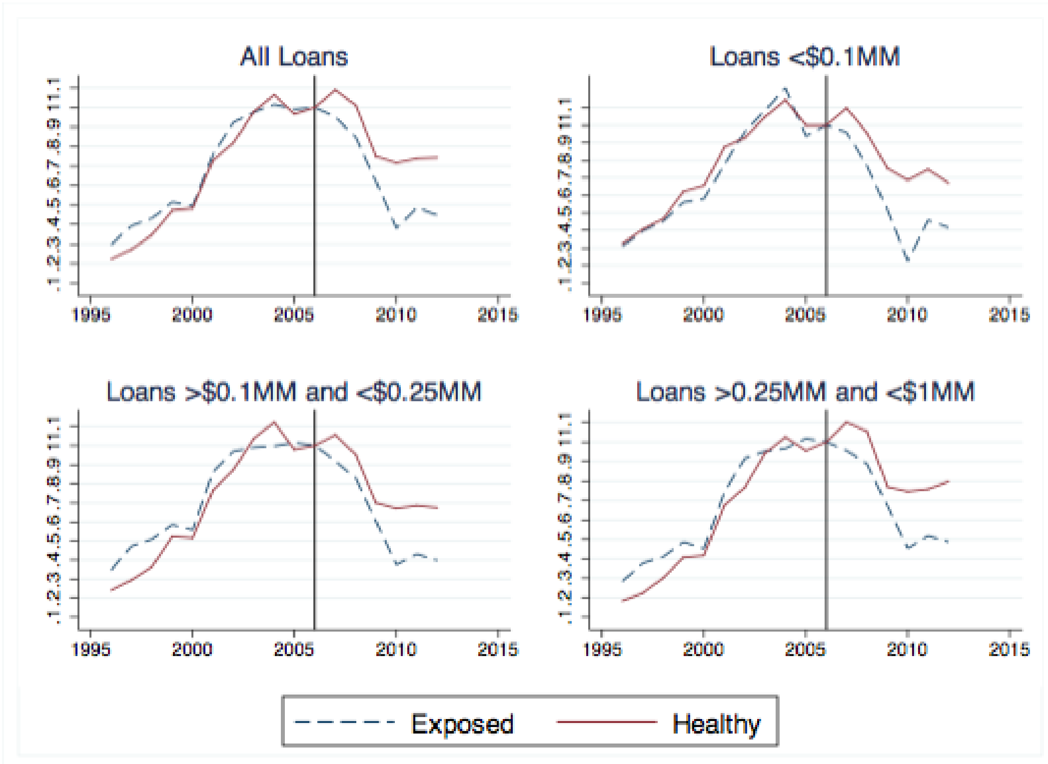


Figure 2.2: *Loan Volume, 1996-2016*

This figure shows the evolution of lending to small firms by loan size. The data are from CRA. Loan volume is indexed to 2006 levels. The figure reports an equal weighted average across counties that did not experience a drop in real estate prices, focusing on the difference between exposed and healthy dispersed banks.

Test of trends in lending by exposed and healthy banks prior to 2006 (top left panel):

$$L_{ilt} = \alpha + \beta G_i + \gamma t + \zeta t \times G_i + \delta_l + \epsilon_{ilt}$$

Healthy bank (G_i)	-0.296
	(0.207)
Linear time trend (t)	0.073***
	(0.022)
Interaction $(t \times G_i)$	0.027
	(0.022)
County Fixed Effects	Yes
Observations	37,559
R-squared	0.156

There are other alternative explanations, which we discuss in the next section.

when using the yearly percent change, the dependent variable we consider in the rest of the analysis.

Table 2.1: *Bank Characteristics*

*This table compares banks exposed to the real estate shock and banks that were unaffected by the real estate shock as of June 30, 2006. The first 3 columns correspond to the full sample. The fourth and fifth columns exclude banks in the top 10 by size. There are no healthy banks in the top 10 banks by size. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

	Exposed banks (Obs.= 79)	Healthy banks (Obs. =195)	Diff	Exposed excluding top10 (Obs.= 69)	Diff
Assets ($billion)	100.77	3.85	96.92***	19.420	15.567***
Number of branches	458.23	67.52	390.71***	183.61	116.08***
Number of counties	64.77	16.01	48.77***	34.855	18.850***
Deposits/Total assets	0.673	0.780	-0.107***	0.714	-0.076***
Loans/Total assets	0.647	0.682	-0.034**	0.661	-0.020
Real estate loans/Loans	0.714	0.747	-0.032	0.733	-0.013
C&I loan /Loans	0.181	0.150	0.031**	0.178	0.028**
Past due/Loans	0.015	0.013	0.002	0.014	0.001
Net charge-offs/Loans	0.001	0.001	0.000*	0.001	0.000
MBS/Assets	0.098	0.083	0.016	0.098	0.015
Contingent claims/Assets	0.001	0.000	0.001	0.001	0.001
Tier 1 capital ratio	0.107	0.119	-0.012**	0.110	-0.008

The evidence discussed above gives credence to our assumption that the small business lending behavior of exposed and healthy banks is similar, despite the differences these two types of banks may have in their overall characteristics. Table 2.1 presents summary statistics for our sample. Because the size distribution of US banks is highly skewed, and because the largest banks are more likely to operate in many different (urban) areas and thus are more likely to have had exposure to strong real estate price decreases, exposed banks are on average significantly larger than healthy banks. As of the end of the second quarter of 2007, the average exposed bank has 458 branches in 65 counties and assets of $100.8 billion, whereas the average healthy bank has 68 branches in 16 counties and assets of $3.9 billion. Exposed banks also have other characteristics that are associated with larger US banks such as lower deposits to assets and tier 1 capital ratios and a higher ratio of commercial and industrial loans to assets. This is also consistent with Figure 2.1.

Despite the differences in overall characteristics, the two types of banks are very similar in aggregate loan characteristics and quality. The percentage of real estate lending was similar for the two types of as of 2006, which highlights that our identification strategy comes where the banks operated and extended real estate loans, rather than have from any differences in lending behavior during the 2002-2006 boom. In addition, some of the large disparities seen in Table 2.1 are due to averages skewed by outliers. Excluding the largest 10 banks drops the average assets of exposed banks to $19 billion and narrows the deposits gap from 11 to 7 percentage points. As shown in the robustness tests, our results remain both economically and statistically significant if we exclude the largest banks. However, due to the distribution of bank size, trying to closely match banks by size substantially reduces the economic relevance of the sample.

2.4 Results

We begin the analysis by examining whether exposure to the real estate shock affects the intensive margin of banks' lending and deposit-taking in counties unaffected by the shock. Table 2.2 presents the basic, univariate results and illustrates that there are marked differences

Table 2.2: *Change in Lending and Deposits: Exposed Banks vs. Healthy Banks*

*This table compares the evolution of deposits and loans from 2006 to 2008 for exposed and healthy banks in our sample. The analysis is constrained to counties that did not experience a collapse in real estate prices. Each observation corresponds to a bank-county pair. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

	Counties unaffected by the direct real estate shock				
	Exposed banks		Healthy banks		
Variable:	Mean	Obs.	Mean	Obs.	Diff.
Number of loans to small enterprise, 2006 (1,000s)	0.215	2,027	0.141	1,587	0.074***
Amount of loans to small enterprise, 2006 ($million)	23.461	2,027	18.757	1,587	4.704**
Percent change between 2006 and 2008:					
Number of small business loans	-0.152	2,027	0.016	1,587	-0.168***
Number of loans that are <100K	-0.198	2,027	-0.008	1,587	-0.190***
Number of loans that are >100K and <250K	-0.111	2,027	0.027	1,587	-0.139***
Number of loans that are >250K and <1M	-0.052	2,027	0.111	1,587	-0.163***
Number of loans to firms with <$1M revenue	-0.232	2,027	-0.101	1,587	-0.131***
Number of loans to firms with >$1M revenue	-0.104	2,027	0.219	1,587	-0.323***
Amount of small business loans	-0.090	2,027	0.077	1,587	-0.167***
Amount of loans that are <100K	-0.146	2,027	0.006	1,587	-0.152***
Amount of loans that are >100K and <250K	-0.074	2,027	0.026	1,587	-0.100***
Amount of loans that are >250K and <1M	-0.038	2,027	0.105	1,587	-0.143***
Amount of loans to firms with <$1M revenue	-0.136	2,027	-0.040	1,587	-0.096***
Amount of loans to firms with >$1M revenue	-0.040	2,027	0.181	1,587	-0.221***
Number of branches	0.019	1,951	0.065	1,557	-0.046***
Deposits/Assets	0.002	1,951	0.008	1,557	-0.006***
Deposits	-0.095	1,951	0.101	1,557	-0.196***

between the two types of banks. In counties that did not experience a decline in real estate prices, healthy banks tended to increase their lending from 2006 to 2008 while exposed banks cut theirs. For example, on average, exposed banks extended 15% fewer loans per county in 2008 than in 2006, whereas unaffected banks increased the number of loans they extended by 1.6%. Similarly, exposed banks cut their total lending by 9%, whereas unaffected banks increased their lending by approximately 8%. This pattern is consistent across all loan size categories. Interestingly, the only category in which both types of banks cut their loan originations was lending to firms with revenues of less than $1 million. Healthy banks cut the number of loans to these firms by 10%, and the loan amount by 4%; exposed firms cut even more, slashing their number of loans to these businesses by 23% and the loan amount by 14%.

The results on banks' deposit-taking activities parallel those on small business lending. Despite deposit insurance, county deposits held by exposed banks shrank on average by 10% between 2006 and 2008, whereas they grew by 10% for healthy banks. The distribution is non-normal, but the medians also suggest a similar story: the median percent change in deposits was 3.8% for exposed banks and 7.7% for healthy banks. Both healthy and exposed banks seem to have expanded the number of their branches during the 2006 to 2008 period, but healthy banks grew more, opening on average 6.5% new branches in a county, whereas exposed banks only opened 2% more branches in each county.

In the follow sections, we first further examine the intensive margin of banks' lending to confirm the preliminary findings that exposed banks cut lending more than healthy banks. Next, we examine the extensive margin and show that exposed banks are more likely to stop all operations in a county during this time period, whereas healthy banks are more likely to enter new counties. Finally, we confirm that similar results hold for the intensive margin of banks' deposit-taking activities.

2.4.1 Intensive Margin

Lending activity 2006–2008

In Table 2.3, we look more formally at lending within counties that are unaffected by the direct real estate shock. We estimate regressions of the form:

$$\triangle L_{it} = \alpha + \beta G_i + \gamma X_i + \delta_l + \epsilon_{il} \tag{2.1}$$

$\triangle L_{it}$ is the change in the logarithm of the amount of small business loans extended by bank i in county l between 2006 and 2008, in millions.[17] G_i, our main variable of interest, is an indicator variable that is equal to 1 for banks classified as healthy and 0 for banks exposed to the real estate shock. β, the coefficient of interest, can be interpreted as the difference in the percent change in lending from 2006 to 2008 between exposed and healthy banks.[18] X_i is the set of bank-level control variables, and δ_l are county fixed effects. By including county fixed effects, we make sure that we are identifying the impact of being a healthy or an exposed bank on lending within each county.

In specification (1), we find that, controlling for the lending volume, healthy banks increased their lending more than affected banks. Specification (2) adds county fixed effects (δ_l) and specification (3) also controls for the log of assets to account for the size of the bank. Further, to control for differences in bank strategy, in specification (4) we control for deposits as a fraction of assets, insured deposits as a fraction of total deposits, loans as a fraction of assets, and real estate loans as a fraction of assets. Specification (5) adds controls for the amount of loans that are past due as a fraction of total loans, the amount of net charge-offs (charge-offs minus recoveries) as a fraction of total loans, the Tier 1 ratio, and the amount of asset-backed securities as a fraction of total assets. All bank variables are measured as of June 30, 2006. Specifications (4) and (5) show that our results are not driven by differences in strategy or differences in

[17] We constrain the analysis to lending in counties in which the bank had a branch since yearly lending to other counties is very volatile. More than 80% of all small business lending is extended to counties in which the bank has a branch. The loans and deposits data are winsorized at the 0.5% level.

[18] Interpreting a difference in logarithms as a percent change relies on x being small in the approximation log(1+x)$\approx$ x. If x is large, as it is in some of our observations, the approximation no longer holds. However, using the directly-calculated percent change does not change our results.

Table 2.3: *Change in Lending, 2008 vs. 2006*

*The analysis is constrained to counties that did not experience a collapse in real estate prices. The variable of interest is Healthy bank, equal to 1 if a bank was not exposed to the collapse in real estate prices across the counties in which it operates. Specifications (1)–(5) correspond to the main result (with increasing number of controls). Specification (6) uses as the dependent variable the change in the market share of loan originations. All control variables are as of June 30, 2006. Standard errors (reported in brackets) are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Dependent Variable:	%△ Loan amount					△ Loan mkt shr
	(1)	(2)	(3)	(4)	(5)	(6)
Healthy bank G_i	0.149***	0.162***	0.218**	0.237***	0.252***	0.064***
	(0.049)	(0.046)	(0.087)	(0.079)	(0.076)	(0.019)
Log(SBL loans in county)	-0.063***	-0.091***	-0.092***	-0.100***	-0.103***	
	(0.013)	(0.021)	(0.021)	(0.019)	(0.018)	
Log(Assets)			0.016	0.061***	0.081***	0.017***
			(0.015)	(0.023)	(0.024)	(0.005)
Deposits/Assets				0.749***	0.926***	0.099
				(0.263)	(0.273)	(0.067)
Loans/Assets				0.099	0.227	0.008
				(0.254)	(0.246)	(0.047)
Real estate loans/Total loans				0.126	0.278	0.024
				(0.180)	(0.208)	(0.025)
Net charge-offs/Total loans					17.385	-1.874
					(16.865)	(4.330)
Past due loans/Total loans					-0.083	0.925*
					(2.952)	(0.544)
Tier 1 ratio					1.816**	0.312**
					(0.908)	(0.139)
ABS/Assets					2.586	0.603*
					(2.029)	(0.323)
County Fixed Effects		Yes	Yes	Yes	Yes	Yes
Observations	3,499	3,499	3,499	3,207	3,207	4,264
R-squared	0.062	0.267	0.270	0.300	0.310	0.203

exposure to real estate or to the securitization market. The control variables generally have the expected signs. The log of loan originations in 2006, which is a measure of the bank's activity in the county, is negative and significant, suggesting that banks with more market power cut lending more. The log of assets becomes significant once we control for other bank balance sheet variables.[19] Banks with more deposits over assets cut lending less, probably because the healthy dispersed banks are on average smaller and so have higher deposits as a fraction of assets. Standard errors in all specifications are clustered at the bank level. Clustering at both the bank and county levels does not change the standard errors or the significance of the coefficients. In unreported results, we show that, as Table 2.2 suggests, these effects persist across all loan sizes, but are concentrated among loans to firms with revenues of more than $1 million.

The central takeaway is that the difference in lending between healthy and exposed banks is economically and statistically significant, and robust across specifications. On average, the difference in lending between healthy and exposed banks is 25.2 percentage points per county, when including all of our controls. This difference corresponds to an 8.4-percentage-point difference in the weighted average real estate price decline for affected versus unaffected banks. These results are similar to those of Huang and Stephens (2014), who find that a bank's exposure to a 1% decrease in real estate prices reduces new small business lending by approximately 3–4%. The large economic magnitude of the decline in lending is also consistent with the results for the same period in Ivashina and Scharfstein (2010).

In specification (6), we use as the dependent variable the change in the market share of loan originations between 2006 and 2008. The market share variable we use is a holistic measure that includes banks that enter and exit the county.[20] The growth in lending by healthy banks,

[19]This is potentially due to non-linearities in the relationship between bank performance and assets. As mentioned earlier, exposed banks, which performed worse, tend to be larger. But some of the largest exposed banks actually performed better than smaller banks, a result which is potentially explained by government policies such as TARP, which were primarily targeted at large banks. As we discuss below, our results are unchanged when we run the analysis on a constrained sample that removes the largest banks. When we do so, assets impact the change in the amount of loans in a statistically significant way.

[20]We do not control for the log of loans in 2006 since that would necessitate excluding banks that extend loans in the county only after 2006 and not before. However, limiting the analysis in this way gives virtually identical results.

relative to exposed banks, translates to a market share gain of 6.4 percentage points. By contrast, the average 2-year growth in market share from 1996 to 2006 is 0.08%, with a standard deviation of 4.4%. Thus, the gain in loan origination market share by healthy banks is almost 1.5 times the standard deviation of the previous ten years.

Intensive Margin: Lending Activity Robustness

Next, we perform a wide range of robustness tests. First, in specification (1) of Table 2.4, we address concerns that our measure of exposure does not take into account subprime real estate prices. During the period we consider, it was mostly the decline in subprime real estate prices that drove banks' losses, whereas our measure of exposure is based on the overall price drop. To take this into account, in specification (1), we rerun our algorithm for classifying banks as exposed, only using counties that had more subprime borrowers. We measure the presence of subprime borrowers by the percentage of 2004-2006 mortgage originations in the Home Mortgage Disclosure Act (HMDA) dataset from 2004 which were classified as "high-priced." From 2004 to 2008, the Federal Reserve's Regulation C required lenders to collect and report the spread between the annual percentage rate (APR) on a loan and the yield on Treasury securities of comparable maturity if the spread was "equal to or greater than 3.0 percentage points for a first-lien loan (or 5.0 percentage points for a subordinate-lien loan)."[21] These high-priced loans were more likely to be loans to subprime borrowers (Mayer and Pence, 2008). We focus on counties in the top three quartiles of the distribution of the presence of high-priced loans.[22] The results are very similar in magnitude and suggest that whether the measure of exposure is calculated just using subprime counties or not does not drive our findings.

In addition, we test whether exposure to within-county variation in the presence of subprime

---

[21] See "Rules and Regulations," Board of Governors of the Federal Reserve System, Federal Register, October 24, 2008.

[22] Limiting to counties in the top two quartiles produces qualitatively similar results. The results are also very similar if we define subprime zip codes as those in the top quartile of the presence of high-priced mortgages (similar to the way Atif Mian and Sufi (2013) define subprime zip codes using percent of people with FICO score less than 660). We then keep the top three quartiles of counties that have the highest percentage of the population living in subprime zip codes. The results are also robust to using the county level measure of debt to income in 2006, made available by Amir Sufi on his website, as this measure is also highly correlated with the presence of subprime borrowers (Atif Mian and Sufi, 2013).

Table 2.4: *Change in Lending, 2008 vs. 2006: Robustness*

*In specification (1), we limit the analysis only to counties in the top three quartiles of the percent of high-priced HMDA mortgages originated from 2004-2006. The sample used in specification (2) excludes the ten biggest banks. In specification (3), we discard observations where a given bank had entered after 2002 (i.e., new or non-traditional markets). In specification (4), we remove observations corresponding to mergers and acquisitions and bank failures. In specification (5), we change the definition of the variables; specifically, "large" banks are defined as those in the top quartile of the distribution of the number of counties each bank operates in and "exposed" as the top quartile of the distribution of the number of weak counties each bank operates in. In specification (6), we change our control group from healthy geographically dispersed banks to local banks. Because the analysis is constrained to counties that did not experienced contraction in real estate prices, local banks were also not exposed to the real estate shock. Standard errors (reported in brackets) are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Dependent Variable:	%△ Loan amount					
	Uses measure of subprime	Excludes Top 10 banks	Excludes counties entered post-2002	Excludes mergers	Alt Def'ns of dispersed, healthy	Control group is local banks
	(1)	(2)	(3)	(4)	(5)	(6)
Healthy bank G_i	0.301***	0.248***	0.154**	0.141***	0.161***	0.174*
	(0.086)	(0.089)	(0.067)	(0.055)	(0.060)	(0.103)
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Bank Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,335	2,099	2,188	3,041	3,285	2,426
R-squared	0.330	0.403	0.371	0.302	0.286	0.335

borrowers drives our results. Although we constrain our analysis to counties where real estate prices did not crash, counties can cover large geographic areas and it is possible that healthy and exposed banks differ in the types of within-county areas where they have branches. To address this concern, we create a measure of within-county exposure to subprime areas by calculating, for each bank, the percent of deposits that are in branches located in subprime zip codes. We classify a zip code as subprime if it falls into the top quartile of the distribution of the percent of high-priced HMDA mortgages originated in that zip code between 2004 and 2006. Both exposed and healthy banks have, on average, 15% of their county deposits in zip codes we classify as subprime, so it is unlikely that differences in the types of neighborhoods where branches are located drive our results.

Next, we examine the importance of the largest banks for our findings. In Specification (2), we exclude the ten largest banks from the analysis; despite a lower number of observations, the results are almost identical. This robustness check mitigates concern that differences in the characteristics of healthy and exposed banks—in particular, the much larger size of exposed banks—drives our findings. As the last 2 columns of Table 2.1 showed, the characteristics of healthy and exposed banks are much more similar once the ten largest banks are removed. In addition, in unreported results, we perform our analysis on a propensity-score-matched sample of banks. After matching on the bank-level observables we control for in our analysis, and only keeping matches that lie on the support of the propensity score distribution, we obtain a sample of 57 exposed and 40 healthy banks. The results of this section, and most of the subsequent sections, are robust to using this subsample. We do not focus on this subsample due to the small number of observations and subsequent lack of power and economic relevance.

Another concern is that exposed and healthy banks may have different expansionary policies. For example, it could be the case that exposed banks only entered many of the counties we examine in the early to mid-2000s, during the real estate boom and expansionary monetary policy of the period. If these banks over-expanded and decided to scale back, then it would be natural that from 2006 to 2008, they decreased lending in many of the counties that they had just recently entered. In other words, these might be non-core counties for the bank's business, and as such, it might make sense to cut credit in these counties even if there were no changes in demand. Although this mechanism still represents a contraction in credit that propagates into otherwise healthy geographical areas by large dispersed banks, it is a different channel, and it might have different implications for borrowers. To alleviate this concern, in specification (3), we re-estimate our main results using only counties where a bank had branches before 2002. In addition we use only these counties when creating the dispersed and exposed variables. The effect remains statistically significant, though lower in magnitude at approximately 15%.

In specification (4), we show that the results are robust to excluding all mergers and acquisitions (M&A), including bank failures. The coefficient is statistically significant but smaller in magnitude if we exclude these observations. Thus, withdrawal from non-core counties and M&A-related differences do not fully explain the difference in lending.

Next, we perform robustness checks to determine that our results are not driven by the precise definition of dispersed and exposed banks. In specification (5), we consider at the sample of banks in the top quartile of the distribution of the number counties the bank has branches in, and define "exposed" as the top quartile of this distribution. In specification (6), we compare exposed banks with local banks. For methodological reasons, our main control sample is constrained to large, geographically-dispersed banks. However, banks with a large geographical presence may respond to shocks elsewhere (shocks that are different from, but contemporaneous to, the decline in real estate markets), not just in the counties we analyze. In that sense, looking at the small, local banks in unaffected counties as a control group provides insight since these local banks' lending reflects only local conditions. Again, although smaller in magnitude, the coefficients are fairly similar in statistical and economic significance.

Similar to the parallel trends test for evolution of credit reported in Figure 2.2, we run placebo tests (unreported) in which we use the difference in lending during the pre-crisis period (e.g., the percent change in lending from 2003 to 2005) as the dependent variable. In these regressions, the healthy bank indicator G_i is not statistically significant, confirming that the difference in lending between healthy and exposed banks arises solely in the pre-crisis period. We also run our results using as the dependent variable the difference between the 2006-2008 and 2004-2006 growth rates, as an alternative way to control for pre-trends.

Finally, it is possible that other events that arose during the early stages of the financial crisis—rather than the real estate shock—drive our results. One alternative explanation is that the larger banks that were exposed to the real estate shock had greater commitments to off-balance-sheet asset-backed commercial paper (ABCP) vehicles. When the ABCP market froze in late 2007, banks that had existing commitments to these vehicles had to provide liquidity and/or credit support to them. As explained in the methodology section, this would have implications for the specific type of channel at work, although we would still be identifying a supply channel. In addition, controlling for the amount of ABCP liquidity and credit commitments as a fraction of assets does not change our results.[23]

2.4.2 Extensive Margin: County Exit and Entry, 2006–2008

Having established that exposure to the real estate shock affected the intensive margin of banks' lending, we next investigate the extensive margin: does exposure to the shock affect whether a bank maintains a lending presence in an area. Although banks engage in small business lending in counties where they do not have branches, on average, more than 85% of all small business loans (by number and volume) are extended in counties with a branch of the lending bank. Because of this, we use as our measure of whether a bank maintains a lending presence whether the bank has a branch in the county.

First, we examine whether exposed banks are more likely to exit a county than unaffected banks. The dependent variable is an indicator equal to 1 if a bank that had branches in the country as of June 2006 no longer has branches in that county as of June 2008, and 0 otherwise. As before, the central explanatory variable is *Healthy bank*. Specification (1) of Table 2.5, Panel A, estimates the effect of being a healthy bank on exit from a county, using OLS with county fixed effects. Specifications (2) and (3) use probit, and because probit produces inconsistent estimates when using fixed effects, we instead include county-level controls. These covariates include the change in real estate prices from June 2002 to June 2006 and the change in real

[23]Our results also remain unchanged when controlling for the difference in ABCP liquidity and credit exposure from 2006 to 2008, which is a measure of how much liquidity and credit support banks had to provide during that period.

Table 2.5: *Extensive Margin*

*In Panel A, the dependent variable is an indicator equal to 1 if a bank that had branches in the country as of June 2006 but no longer has branches in that county as of June 2008, and 0 otherwise. Specification (3) excludes observations corresponding to exit due to bank failure and M&A activity. In Panel B, the dependent variable is equal to 1 if the bank entered the county between June 2006 and June 2008, and a probit is estimated. In all specifications, the coefficients reported are marginal effects. Standard errors (reported in brackets) are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Panel A: Exit

Dependent Variable:	Exit from county 2006-2008		
	OLS	Probit	Probit (excl M&A)
	(1)	(2)	(3)
Healthy bank G_i	-0.208***	-0.036***	-0.003*
	(0.073)	(0.011)	(0.001)
County Fixed Effects	Yes		
County Controls		Yes	Yes
Bank Controls	Yes	Yes	Yes
Observations	3,897	3,828	3,494
Within R-squared	0.290	0.165	0.218

Panel B: Entry

Dependent Variable:	Entry into county 2006-2008			
	(1)	(2)	(3)	(4)
Healthy bank G_i	0.047***	0.100***	0.022	0.026*
	(0.014)	(0.034)	(0.019)	(0.029)
-Log (Num new banks 1996-2005)		-0.044***		
		(0.014)		
Healthy bank		0.023*		
× -Log (Num new banks 1996-2005)		(0.012)		
Deposits HHI (2006)			-0.013*	
			(0.007)	
Healthy bank × Deposits HHI			0.013*	
			(0.008)	
Rice and Strahan (2010) index				-0.015**
				(0.006)
Healthy bank × Rice-Strahan index				0.012*
				(0.006)
County Controls	Yes	Yes	Yes	Yes
Bank Controls	Yes	Yes	Yes	Yes
Observations	6,484	6,484	6,484	1,421
Within R-squared	0.082	0.091	0.086	0.187

estate prices from June 2006 to December 2007; the debt-to-income ratio, the total population, total number of households, household median income, housing density, percent of households below the poverty line, the unemployment rate, and the percentages of households working in finance, construction and real estate, all as of 2006. All coefficients are reported as marginal effects at the mean of the distribution except for the coefficient of interest, the healthy bank indicator, which is a binary variable. Reported standard errors are clustered at the bank level; clustering at both the bank and county level or just at the county level produces similar results.

Either estimation approach suggests that healthy, unaffected banks are less likely to exit counties that did not experience a real estate decline.[24] For example, specification (2) of Panel A suggests that healthy banks are 4 percentage points less likely to exit a county relative to an exposed bank in the same county (compared to an unconditional mean of 10%). Specification (3) of Panel A drops observations that correspond to exit due to bank failures and mergers and acquisitions (M&A). The coefficient on Gi is much smaller in magnitude because M&A activity accounts for a large portion of exits. That said, the coefficient is still statistically significant, which implies that exposed banks are more likely to exit a county even if they do not fail and are not acquired.

In Table 2.5, Panel B, we examine whether healthy banks are also more likely to enter counties where they did not have branches before 2006. For this analysis, each observation corresponds to a bank-county pair where the bank did not have any branches in the county in 2006, but did have branches in an adjacent county. The dependent variable is an indicator equal to 1 if the bank entered the county in 2007 or 2008, and 0 otherwise. We use probit regression and control for county covariates. As before, the coefficients presented are marginal effects. The coefficient on G_i is positive and significant, suggesting that healthy banks were 4.3 percentage points more likely to expand into new counties than exposed banks.

Specifications (2)–(4) further test whether healthy banks are relatively more likely than exposed banks to expand into counties that are traditionally difficult to enter. In specification (2), we measure the difficulty of entry into a county by the number of banks that had entered that county in the previous 10 years. The explanatory variable of interest is the negative of the

[24]The results are also robust to using a fixed-effects logit.

log of the number of banks that entered the county from 1996 to 2005; a larger value implies that the county is harder to enter. The coefficient on this variable is negative and significant, while the coefficient on its interaction with G_i is positive and significant, implying that exposed banks are less likely to expand into hard to enter counties, and healthy banks are relatively more likely to.[25]

Another measure for the difficulty of expanding into a county is concentration, as measured by the HHI of deposits. More concentrated markets may be harder to enter because a few banks control most of the market share and consumers may have longer relationships with one of these banks. Specification (3) supports this hypothesis. The coefficient on the Deposits HHI variable is negative and significant, whereas the coefficient on the interaction between the HHI and G_i is positive and significant. Again, exposed banks are less likely to enter into concentrated markets, whereas healthy banks are relatively more likely to.

A final measure of difficulty of entry is an index compiled by Rice and Strahan (2010). This is a state-level index that measures the barriers to cross-state entry that a state imposes on its banking markets. The index uses the values 0 to 4, which correspond to how many of the following restrictions a state imposes: a minimum age of 3 for institutions of out-of-state acquirers; a ban on de novo branching; a ban on acquisition of individual branches by out-of-state institutions; and a deposit cap of 30% for each institution. In specification (4), we use this variable as a proxy for difficulty of entry and restrict the observations to the set of out-of-state counties that each bank can expand into. As expected, the coefficient on this variable is negative and significant, but the interaction with G_i is positive and significant. Exposed banks are less likely to enter counties in states with restrictions, but healthy banks are relatively more like to do so. Note that, in specifications (3) and (4), the coefficient on Gi is positive, but no longer significant. This implies that healthy banks do not expand more everywhere, but instead focus on counties which they otherwise would have trouble entering.

Table 2.6: *Deposits and Branches*

*This table examines whether healthy banks are more likely to expand their number of branches and enter new counties. Specifications (1)–(2) focus on the change in the log of deposits and in number of branches in counties where the bank already has branches, respectively. In specification (3), the dependent variable is equal to 1 if the number of branches increased, 0 if it stayed the same, and -1 if it decreased, and an ordered probit is estimated. Specification (5) includes observations corresponding to both exit from and entry into a county. Standard errors (reported in brackets) are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Dependent Variable:	%△ Deposits (1)	%△ Branches (2)	%△ Branches (Ordinal) (3)	%△ Deposits per branch (4)	%△ Deposits market share (5)
Healthy bank G_i	0.076*	0.059**	0.223*	-0.061*	0.025***
	(0.042)	(0.027)	(0.122)	(0.082)	(0.009)
County Fixed Effects	Yes	Yes		Yes	Yes
Bank Controls	Yes	Yes	Yes	Yes	Yes
Observations	3,207	3,207	3,184	3,207	5,353
Within R-squared	0.335	0.270	0.058	0.313	0.232

2.4.3 Intensive Margin: Branches and Deposits, 2006–2008

Having established that healthy banks cut lending less, and expand their lending presence more, than exposed banks in unaffected counties, we next examine whether banks behave similarly with respect to deposit-taking. To do so, we repeat the analysis of specification (5) of Table 2.3 using the change in the log of deposits as the dependent variable. Specification (1) of Table 2.6 shows that the healthy bank indicator is again positive and significant, implying that healthy banks increase their deposits by 7 percentage points more than exposed banks. In comparison to the effect on lending, this might seem small. However, deposits—especially retail deposits—tend to be very sticky, and capturing new deposits may be harder than capturing new borrowers. Note that since this specification limits the sample to counties in which the bank had a branch both in 2006 and 2008, the inter-county expansion established in Table 2.5 does not explain these results. However, just Table 2.5 shows that healthy banks expand into new counties, it is possible that they also expand their branch networks in existing counties.

Next, we examine whether, in parallel with increasing their deposits, healthy banks also increase their number of branches more than exposed banks. In specification (2) of Table 2.6, the dependent variable is the percent change from 2006 to 2008 in the number of branches. The coefficient on the Healthy Bank indicator G_i is positive and significant, and suggests that healthy banks increased their number of branches by 6 percentage points more than exposed banks. However, the distribution of the percent change of branches is highly skewed; the median change in the number of branches is 0 for both types of banks.

To ensure that a few outliers do not drive our results, in specification (2) we replace the dependent variable with an indicator variable that is equal to 1 if the number of branches increased, -1 if it decreased, and 0 if it stayed the same. We run this regression using an ordered probit model, and controls for county covariates rather than county fixed effects, because probit produces inconsistent coefficients when using fixed effects. The reported coefficients are the marginal effects of each independent variable on the probability of expansion, evaluated at the mean of the variable's distribution. As before, standard errors are clustered at the bank level,

[25] All coefficients, including the interactions, are reported as marginal effects at the means of the relevant variables using the Stata command *margins*.

but clustering at both the bank and county levels produces similar results. Our results remain strong and statistically significant: healthy banks are more likely to expand their number of branches in the counties they are already in. These results are robust to excluding observations corresponding to merger and acquisition activity.

In specification (4), we investigate whether this increase in branches can fully explain the increase in branch deposits. Using deposits per branch as the dependent variable, we find no difference between healthy and exposed banks, suggesting that healthy bank grow their deposit bases mainly through new branches.

Finally, we examine the economic significance of the increase in deposits by large banks. In specification (5), we consider the change in the deposit market share as the dependent variable.[26] Our definition of market share, as in Table 2.3, is holistic and captures both entries and exits. The healthy bank indicator G_i is positive and significant at the 1% level, which suggests that, relative to exposed banks in the same county, healthy banks increase their market share of deposits in the county. The estimate of a difference of 2.5 percentage points is economically significant within the historical context. Using deposit data from 1994 to 2006, we estimate that during this time period, the average market share change over any two year period is a decrease of 0.06 percentage points (after de-meaning by year and by county, consistent with our regressions).[27] The standard deviation is approximately 2 percentage points. Thus, the difference between healthy and exposed banks of 2.5 percentage points is larger than a one-standard-deviation difference in deposit market share. Similarly, Strahan (2003) finds that post-geographical deregulation, small banks, in aggregate, lost 2% of their deposits share as a result of increased competition. This implies that the increase in the market share for the average healthy bank, relative to an exposed bank, is comparable to the impact of geographical deregulation.

Why did depositors flock to the new branches that healthy banks opened? One possible

---

[26] For ease of interpretation, we use the change in market shares rather than the percent change. We also do not control for loans or deposits as of 2006 because that would limit our sample to those bank-county pairs in which the bank had branches in 2006.

[27] We consider a change over two years so as to be comparable with our main analysis for the two-year period of 2006–2008.

explanation is that these banks were able to offer lower deposit rates, relative to exposed banks. Alternatively, depositors responded not to the deposit rates, but to the apparent safety of the healthy banks, as discussed in Section 3. We discern between these two explanations by examining changes in rates and deposit growth between the two types of institutions.

First, using branch-level deposit rate data from RateWatch, we re-estimate specification (3) of Table 2.6 using the change in deposit rates from 2006 to 2008 as the dependent variable.[28] The results are presented in panels A and B of Table 2.7. In panel A, we use as the dependent variable the deposit rates corresponding to an account with a balance of $10,000.[29] These are the accounts with the most data in RateWatch and have been used in prior literature as a proxy for the deposit rates a bank is able to offer (Itamar Drechsler and Schnabl, 2017, e.g.). Thus, Panel A estimates differences in the rates between healthy and exposed banks on different types of insured accounts. Column (1) uses the change in the money market rate as the dependent variable; column (2) uses the change in the rate on 3-month Certificates of Deposit (CDs), and columns (3) and (4) use the change in the rate on 12-month and 5-year CD, respectively. In all cases, there is no statistically significant difference between the rates that healthy and exposed banks offer.

However, depositors are unlikely to respond to rates on a balance of $10,000 because during our time-period, deposits are fully insured up to a total of $100,000 per depositor (the deposit insurance limit was raised to $250,000 in October 2008). Deposits most likely to run in a crisis are short-term, uninsured accounts, and so in Panel B, we repeat the analysis of Panel A but using changes in rates on accounts with a balance of $100,000. The table show that for short-term accounts, such as money market accounts and 3-month CDs, healthy banks decrease their rates more than exposed banks from 2006 to 2008. For longer-term accounts such as 12-month and 5-year CDs, there is no difference between the two types of banks. By offering relatively higher rates on large short-term accounts, exposed banks sought to acquire

[28]For each bank, we obtain the mean change in deposit rates at the county level (across all branches in that county) and use this as the dependent variable. Using the median as the dependent variable does not change our results.

[29]Deposit rates schedules are tiered and tiers vary by banks. RateWatch provides the rates corresponding to accounts with a specific balance such as $10,000, $100,000, and so on.

Table 2.7: *Change in Deposit Rates, 2008 vs. 2006*

*This table compares the change in deposit rates from 2006 to 2008 for healthy banks and for exposed banks. In Panel A, the dependent variable is the change in the deposit rate for a $10,000 balance in a money market account (specification (1)), 3 Month CD (Specification (2)), 12 Month CD (Specification (3)), and 5 Year CD (Specification (4)). In panel B, the dependent variable is deposit rate for a $100,000 balance in each of those accounts. Both panels include Standard errors (reported in brackets) are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Panel A: Account with $10,000 balance

Dependent Variable:	Change in the deposit rate:			
	Money market (1)	3-month CD (2)	12-month CD (3)	5-year CD (4)
Healthy bank G_i	-0.185	0.040	0.204	-0.251
	(0.197)	(0.221)	(0.262)	(0.236)
County Fixed Effects	Yes	Yes	Yes	Yes
Bank Controls	Yes	Yes	Yes	Yes
Observations	1,423	1,415	1,427	1,370
R-squared	0.400	0.503	0.433	0.482

Panel B: Account with $100,000 balance

Dependent Variable:	Change in the deposit rate:			
	Money market (1)	3-month CD (2)	12-month CD (3)	5-year CD (4)
Healthy bank G_i	-0.404*	0.412*	-0.084	-0.176
	(0.238)	(0.221)	(0.200)	(0.278)
County Fixed Effects	Yes	Yes	Yes	Yes
Bank Controls	Yes	Yes	Yes	Yes
Observations	1,037	885	942	898
R-squared	0.514	0.686	0.659	0.748

more deposit funding at a time when overnight repo and federal funds were becoming more stressed (Afonso *et al.*, 2011). However, they were mainly unsuccessful in this attempt.

Table B.1 of Appendix B shows at a cross-sectional level that despite the difference in rates, deposit growth in large, uninsured deposits was higher at healthy banks. In this table, each observation corresponds to a bank, not a bank-county pair. In specification (1), we regress the growth in bank-wide deposits from 2006 to 2008 on the bank control variables from before and our healthy bank indicator G_i. Consistent with the results of Table 2.6, healthy banks' deposits grew faster than exposed banks'. In specifications (2) and (3), we further separate this deposit growth into growth in insured and in uninsured deposits. As the results show, whereas the growth in insured deposits is about the same, healthy banks experienced higher growth in uninsured deposits.

Overall, our findings indicate that despite the lower deposit rates they received, uninsured depositors sought the safety of healthy banks. Although these gains did not come from offering higher rates on retail deposits, there is also some weak evidence that healthy banks used advertising to attract depositors. Between June 2006 and June 2008, healthy banks spent more on advertising expenses as a percentage of total assets than exposed banks, but the difference is not statistically significant.

2.5 Persistence and Economic Impact of the Effects

We have shown that healthy banks outpace exposed banks in the intensive margin of increasing loans and deposits, as well as the extensive margin of entry and exit. In this section, we show that the changes in market structure we document persisted in the long-term and did not subside after the period of turmoil from 2006 to 2008. Although isolating the long-term impact of the real estate shock is hard due to general upheaval in the financial markets as well as policy interventions, this fact should make it more difficult for us to find persistent changes in the market structure.

Table 2.8 examines persistence of the differences between exposed and healthy banks: the dependent variable is the change from 2006 to 2015 in the variable of interest. In specifications

Table 2.8: *Long-Term Effects*

*This table tests whether the differences in gains in loans and deposits between healthy and exposed banks persist in the long term. In all specifications, the dependent variable is defined as the change from 2006 to 2015. Standard errors (reported in brackets) are clustered at the bank level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

	2006-2015			
Dependent Variable:	%△ Loans	△ Market share (loans)	%△ Deposits	△ Market share (deposits)
	(1)	(2)	(3)	(4)
Healthy bank G_i	0.325***	0.109***	0.135*	0.042***
	(0.086)	(0.028)	(0.079)	(0.011)
County Fixed Effects	Yes	Yes	Yes	Yes
Bank Controls	Yes	Yes	Yes	Yes
Observations	2,534	3,587	2,545	3,587
R-squared	0.387	0.276	0.480	0.324

(1) and (2), the dependent variables are the percent change in small business originations and the change in the market share of small business loan originations, respectively. The coefficient on G_i is positive and significant in both regressions, suggesting that healthy banks extend more loans and capture more of the loan origination market for exposed banks. In specifications (3) and (4), the dependent variables are for deposits rather than small business loans. In both cases, healthy banks increase their market shares more than exposed banks, and the difference is both statistically and economically significant. The change in market shares from 2006 to 2015 is 10.9 percentage points higher for healthy banks in lending, and 4.2 percentage points higher in deposits.

In the previous tables, we measured the differences between healthy and exposed banks within the same county. Next, we examine the overall effect on a county's local economic conditions. The dependent variables in specifications (1)–(4) of Table 2.9, Panel A are calculated as aggregates across all healthy banks in a county for the 2006 to 2008 period. The main independent variable of interest is *Exposed county*, the indicator variable for whether a county is exposed to the real estate shock through the presence of branches of affected banks.[30] Consistent with our conclusions in the preceding sections, the county-wide percent changes in both loans and deposits of healthy banks increase more in counties with a larger presence of exposed banks. The overall market share of healthy banks also increases more in these counties. In specifications (5)–(6), the dependent variables are total loan and deposit growth across all banks in the county. The results indicate that although healthy banks increase lending and deposits more in areas that have been further exposed to the shock, they do not fully make up for the impact of the shock. Counties with a higher presence of exposed banks had loan growth of 9 percentage points less, and deposit growth of 2.5 percentage points less, than similar counties that were not exposed to the shock. Specification (7) tests whether a county's exposure to the real estate shock affects its concentration, as measured by the deposit HHI. Although the effect is negative, it is not significant. Since we have shown that healthy banks are more likely to enter new and concentrated markets, one might expect the deposit HHI to decrease more in the long run as new entrants gain market share.

Next, we examine the long-run aggregate effects. Table 2.9, Panel B replicates Panel A, but uses the change from 2006 to 2015 for all dependent variables. The results are very similar to those of Panel A. Notably, it appears that in the long run, the deposit concentration in exposed counties decreases, as healthy banks enter new, concentrated markets and begin to grow their market shares.

As with the results in previous tables, although the coefficients in this table appear to be small in magnitude, they are economically significant in the historical context. For example, specification (4) of Panel B shows that between 2006 and 2015, healthy banks in exposed counties grew their aggregate market share by 9.8 percentage points more than in other counties. By comparison, Jayaratne and Strahan (1997) find that high-profit banks increase their market shares by 6.7 percentage points in the 6 years following geographic deregulation, relative to a comparable 6 years pre-deregulation.[31] Similarly, Strahan (2003) shows that HHI

---

[30]Consistent with prior definitions, *Exposed county* equals 1 in counties in top quartile of exposure, calculated as the deposit-weighted exposure measure across banks in the county.

[31]The percentage change in deposits from 2006-2012, a comparable 6 years, is 7.4 percentage points.

Table 2.9: *Aggregate Effect*

*This table examines the aggregate effect at the county level. County exposure is measured as the deposit-weighted average, across banks, of the real estate exposure of banks in each county. The main variable of interest, Exposed County, is equal to 1 if the county lies in the top quartile of this distribution. In Panel A, all dependent variables are changes from 2006 to 2008. In Panel B, the dependent variables are changes from 2006 to 2015. All specifications include county controls. The analysis is constrained to counties that did not experience a real estate drop. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Panel A: 2006 to 2008

Dependent Variable:	%$\triangle$ Healthy Loans	$\triangle$ Healthy Market Share (Loans)	%$\triangle$ Healthy Deposits	$\triangle$ Healthy Market Share (Deposits)	%$\triangle$ Loans	%$\triangle$ Deposits	$\triangle$ Deposit HHI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposed County	0.10*	0.068***	0.031***	0.035***	-0.090*	-0.025*	-68.787
	(0.059)	(0.017)	(0.012)	(0.009)	(0.054)	(0.014)	(47.740)
County Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	704	704	704	704	716	716	716
R-squared	0.030	0.054	0.142	0.059	0.027	0.168	0.128

Panel B: 2006 to 2015

Dependent Variable:	%$\triangle$ Healthy Loans	$\triangle$ Healthy Market Share (Loans)	%$\triangle$ Healthy Deposits	$\triangle$ Healthy Market Share (Deposits)	%$\triangle$ Loans	%$\triangle$ Deposits	$\triangle$ Deposit HHI
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exposed County	0.156*	0.095***	0.058***	0.098***	-0.414***	-0.076***	-138.688**
	(0.090)	(0.029)	(0.019)	(0.017)	(0.103)	(0.026)	(68.913)
County Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	672	672	672	672	690	690	690
R-squared	0.109	0.045	0.256	0.093	0.130	0.238	0.124

decreases by 76 points in the post-deregulation period, which is generally consistent with the results of specifications (7) of Panels A and B.

Finally, we examine whether these differences in lending and deposit-taking have an effect on the real economy. Using the Census County Business Patterns (CBP) dataset, we examine how employment and of the number of operating establishments differ between counties with and without a large presence of exposed banks.[32] In essence, we rerun the analysis of Table 2.10 using changes in employment and number of firms as the dependent variables. Because the recession did not start until the end of 2007 and changes in firm employment lagged the changes in small business lending, we use the period from 2007–2009 instead of 2006–2008 in our analysis.

In specification (1) of Panel A of table 10, we use the percent change in employment as the dependent variable. Specification (2) uses the percent change in the total number of establishments. We then disaggregate the change in all establishments into the percent change in establishments with 1–19 workers (specification (3)), 20–49 workers (specification (4)), and more than 50 workers (specification (5)). The first specification shows that there is a statistically significant difference in the change in county employment between exposed and unexposed counties. On average, the decrease in employment is 1.5 percentage points higher in magnitude for exposed counties than unexposed counties. Considering that the total drop in employment from 2007 to 2009 was 7%, this is an economically significant difference. The second specification shows a similar difference in the number of establishments: exposed counties explain a 1 percentage-point higher drop in number of establishments during the time period. The next three columns of the table show that this decline in the number of establishments mainly comes from smaller firms with fewer than 50 employees, as there is no difference between exposed counties and non-exposed counties in the number of establishments with more than 50 employees. This is intuitive since our exposed county variable captures the presence of banks that decreased small business lending and this type of funding is probably less important for larger firms. The last column shows that this change in the number of

[32]Unfortunately, the CBP has data for county employment, and number of establishments split by firm size, but not employment split by firm size.

Table 2.10: *Impact on the Real Economy*

*This table examines the effect at the county level on the real economy. County exposure is measured as the deposit-weighted average, across banks, of the real estate exposure of banks in each county. The main variable of interest, Exposed County, is equal to 1 if the county lies in the top quartile of this distribution. In Panel A, all dependent variables are changes from 2007 to 2009. In Panel B, the dependent variables are changes from 2007 to 2014, apart from the last column which is 2007 to 2013. All specifications include county controls. The analysis is constrained to counties that did not experience a real estate drop. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Panel A: 2007 to 2009

Dependent Variable:	%△ Employment	%△ Num of Establishments	%△ Num of Establishments, 1-19 employees	%△ Num of Establishments, 20-49 employees	%△ Num of Establishments 50+ employees	% Firm Births
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed County	-0.015**	-0.009**	-0.008**	-0.021*	-0.003	-0.010***
	(0.007)	(0.004)	(0.004)	(0.011)	(0.014)	(0.004)
County Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	716	716	716	716	716	715
R-squared	0.136	0.149	0.119	0.117	0.033	0.383

Panel B: 2007 to 2014

Dependent Variable:	%△ Employment	%△ Num of Establishments	%△ Num of Establishments, 1-19 employees	%△ Num of Establishments, 20-49 employees	%△ Num of Establishments 50+ employees	% Firm Births
	(1)	(2)	(3)	(4)	(5)	(6)
Exposed County	0.003	-0.015**	-0.018***	0.008	0.017	-0.029***
	(0.011)	(0.007)	(0.007)	(0.015)	(0.019)	(0.010)
County Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	715	715	715	715	715	715
R-squared	0.175	0.292	0.283	0.081	0.107	0.466

small businesses is driven by less business creation in exposed counties.[33] Exposed counties experience less new firm creation, as a percentage of total firms, relative to counties with a smaller presence of exposed banks.

Panel B of Table 2.10 repeats the analysis for the 2007 to 2014 time period.[34] The results show that due to the presence of exposed banks, exposed counties experience a higher drop in the number of small businesses, which lasts through 2014. This effect seems to only be persistent for the smallest firms, those with fewer than 20 employees. Firms with more than 20 employees recover much faster after 2010. Again, the last column shows that new firm creation is lower in exposed counties.

Our results contrast with those of Greenstone *et al.* (2015), who find that while the decrease in small business lending has a statistically significant impact on employment, the effect is not economically meaningful. One reason for this difference comes from the fact that Greenstone *et al.* (2015) use a subsample of small business lending data from the CRA: lending to firms with less than $1 million in revenues. This subsample comprises 45% of all small business lending originated in 2005 (and 48% of the number of loans originated). In terms of employment, to match this data to the U.S. Census Longitudinal Business Database (LBD), Greenstone *et al.* (2015) consider firms with less than $1 million in revenues to be equivalent to firms with fewer than 20 employees.[35] Firms with fewer than 20 employees represent approximately 18% of all employment. The Small Business Administration (SBA) generally classifies as "small" firms with fewer than 500 employees (the average small firm in the SBA sample has annual revenues of $14 million), and firms with fewer than 500 employees represent about 55% of all employment. Thus, although the result for the subsample of very small firms is interesting, the SBA definition makes an argument for looking at the full CRA sample, especially when considering broader economic implications.

---

[33]Data on firm births come from the Census Business Dynamics Statistics.

[34]The latest CBP data is as of the end of 2014 so we cannot examine the period ending in 2015 as we do in our other tables. The latest data on Business Dynamics is as of 2013.

[35]It is possible that some firms with fewer than 20 employees receive small business loans but have more than $1 million in revenues. If that is the case, the effect of the decline in lending supply on their employment would not be captured when only considering loan supply for loans to firms of less than $1 million in revenues.

2.6 Conclusion

Between 2008 and 2010, U.S. unemployment rose to the highest levels in thirty years, and GDP per capita fell by 3% in a single year. While these adverse outcomes were widely felt across the economy, their causes were more localized. This essay studies the propagation of these local shocks into the broader economy and the offsetting forces that led to major changes in the composition of financial institutions.

We find that banks affected by a real estate shock in their portfolio substantially reduce their lending in local markets unaffected by the shock, relative to less exposed banks in the same markets. Further, we find that exposed banks were more likely to exit an unaffected market by closing all branches there. However, in parallel, healthy banks used their stronger balance sheets to enter new markets and to expand their activities in both lending and deposit-taking in existing markets. These gains in market shares remain in the long run and are comparable in magnitude to changes resulting from the geographic deregulation of the US banking sector.

Chapter 3

Risk, Lending, and Organizational Form: Lessons from Small Lenders

3.1 Introduction

Although more than ten years have passed since the onset of the housing crisis of 2007-2008, the origins of the real estate boom and bust are still being debated. Two prevalent explanations are the credit supply-based view of Mian and Sufi (2009, 2016) and the credit demand-based view of Adelino *et al.* (2016).[1] The credit supply view argues that increased credit to households that were traditionally denied credit, such as subprime consumers, led to higher real estate prices and higher subsequent defaults. By contrast, the credit demand view argues that increased supply to previously denied households does not fully explain what happened during the real estate boom and bust, and that middle and high-income households also played a role. Adelino *et al.* (2016) argue that larger mortgage sizes were associated with households with higher income growth, and that it was demand for credit, perhaps fueled by irrational expectations of house price growth, that further bid up house prices and led to higher subsequent defaults.

In this essay, I propose to test these two channels using a novel context: the differences in lending and performance between stock financial institutional, such as commercial banks, and

[1] See also Favilukis *et al.* (2016), Foote *et al.* (2012), among others.

mutual institutions, such as credit unions. The main difference between banks and mutual institutions is that whereas banks are profit-maximizing firms seeking to create value for their shareholders, mutual institutions have no outside shareholders. They are owned by their depositors and thus their incentives are different. For these reasons, banks are more likely to engage in risk-shifting, whereas mutual institutions have an incentive to avoid even appropriate levels of lending risk and "live a quiet life." These differences between banks and mutual institutions in their approach to risk-taking allow me to both examine the extent to which banks engaged in lending consistent with the credit supply or credit demand view of the housing crisis, and to determine which type of risk-taking –and which explanation–determines subsequent losses by financial institutions.

Although differences in ex-ante risk-taking are difficult to identify, the differences between banks and mutuals in their ex-post performance are striking. Figure 3.1 presents the failure rates of banks and mutual institutions from 2007 through 2014. From 2003 to 2006, no banks or credit unions in the US failed. However, between 2007 and 2014, hundreds of institutions were shut down by their regulators. Although both institutions experienced an increase in failures relative to the pre-crisis period, Figure 3.1 presents a striking contrast between the two types of institutions. Many more banks failed, even though the total number of banks is lower than the total number of mutual institutions during this time period. At the peak of the failures in 2010, 140 banks and 20 mutual institution failed.

I begin by documenting the significant under-performance of bank real estate lending portfolios, relative to those of similar credit unions. During the financial crisis period of 2007-2012, banks experienced higher delinquency and charge-off rates on their loan portfolios, relative to mutuals. This difference is driven mainly by residential real estate lending. Although for my sample of small institutions, delinquency and charge-off rates are generally low, for some institutions they are high enough to lead to failure. Consistent with Figure 3.1, I find that banks fail at a much higher rate, relative to mutual institutions, during the period. I confirm that these results are not driven by differences in regulators, in exposure to MBS, or in percentage of mortgages bought or sold. I argue that there is no evidence to suggest that these differences in performance are driven by differences in mortgage rates or by differences

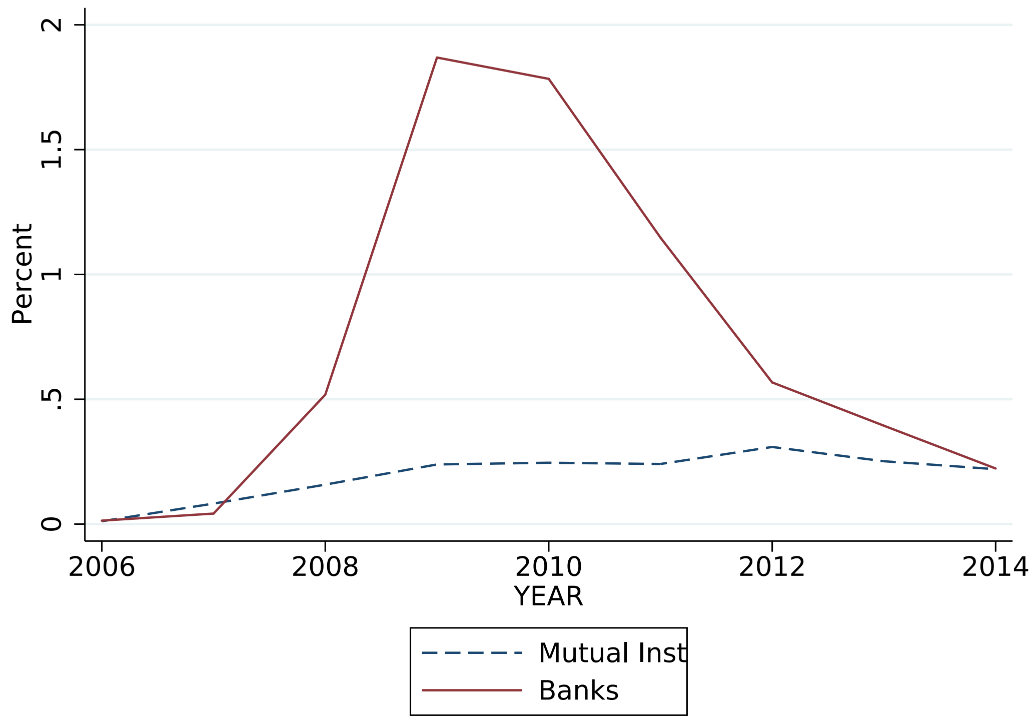


Figure 3.1: *Bank and Mutual Institution Failures*

This figure shows the percent of banks and mutuals that fail each year from 2007 to 2013. Banks are defined as commercial banks and stock state savings banks. Mutual institutions are defined as credit unions and mutual state savings banks.

in non-profit status.

Next, I examine differences in the mortgage lending of the two types of institutions during the real estate boom of 2002-2006. Consistent with the credit supply view, I find that mutual institutions engaged in lower rates of subprime lending relative to banks. In addition, whereas the growth in banks' subprime lending was higher in areas zip codes with more latent demand and a higher ratio of lower income households, mutual institutions did not increase subprime lending–or mortgage lending generally–in these areas. Furthermore, banks' decreased lending standards more than mutual institutions, decreasing denial rates in zip codes experiencing higher growth in the number of applications, and in zip codes with higher latent (subprime) demand.

In addition, consistent with the credit demand view, I find that whereas for banks, there was a strong correlation between the growth in zip code Adjusted Gross Income (AGI) and average mortgage amount during the boom period, there was no such correlation for mutual institutions. Although banks and mutual institutions originated larger mortgages to higher income applicants throughout the whole 1998-2006 sample, there was no difference between the boom and pre-boom period for mutual institutions. By contrast, for banks, a 1% increase in applicant income resulted in a 1.15% increase in average mortgage size, relative to the pre-boom period.

Finally, I test whether the credit supply or credit demand view best explains the under-performance of banks' real estate lending portfolios. I find that overextension of credit to zip codes with high growth in AGI does not explain the higher delinquencies and charge-offs that banks experienced, suggesting that the credit demand view does not explain subsequent bank performance. By contrast, the percentage of subprime lending in 2004-2006 is positively correlated with subsequent bank performance, although its explanatory power is minimal. Overall, my results suggest that the credit supply view helps explain more of bank performance than the credit demand view, it is not the sole driver of the differences in performance.

Throughout my analysis, I restrict my sample of mutual institutions and banks to those that are comparable. Specifically, I exclude internet banks and large banks with more than $10 billion in assets, as well as very small credit unions with less than $25 million in assets. I also exclude credit unions that are likely not open to the community–those that have "Firemen", "Teachers," "Employees," and other professions in the title. Similarly, I exclude credit unions with potential membership less than half of the population of the county in which they operate. Finally, for the main analysis, I restrict all institutions to have branches in only 1 county. By utilizing these restrictions and implementing county-year fixed effects, I compare the performance of institutions that have arguably similar borrowers, within the same county.

Although banks and mutual institutions differ on dimensions other than organizational form, I show that for the small banks and mutuals in my sample, these other differences are unlikely to matter. Specifically, banks and mutuals in my sample undergo similar growth in assets, deposits, and loans during the sample period, which suggests that small banks' ability to issue equity does not strongly impact their lending choices. Similarly, credit unions' non-profit status, and ability to offer lower lending rates, does not affect the comparability of my sample and does not drive my results.

This essay adds to several strands of literature. First, it examines the relationship between organizational form, lending, and risk-taking. The existing literature on whether mutual institutions are less efficient than banks is mixed.[2] However, both the theoretical and empirical literature supports the view that mutual institutions take on less risk. Rasmusen (1988) builds a model of uninformed depositors who do not observe the mutual institutions' assets, and shows that in this context, mutual institutions take (inefficiently) less risk than stock institutions. Esty (1997) uses data on savings and loans associations to show that thrifts that convert from mutual to stock ownership engage in riskier projects and increase profit volatility.

Several papers have addressed the performance of credit unions during the financial crisis. Cororaton (2017) finds that credit unions do not cut lending as much as banks during the 2007-2012 period. Smith (2012) documents that credit unions outperformed banks during and after the financial crisis, especially in terms of commercial and industrial lending. He finds that unlike banks, credit unions increased their commercial lending post 2008, and that their delinquencies and charge-offs on commercial loans were lower than for banks. My results confirm these findings. As I discuss later, making comparisons in terms of commercial lending is difficult because banks are much more involved in commercial lending than credit unions are. Because of this, I examine all types of lending, and overall failure rates. Finally, Chatterji *et al.* (2013) document that following the financial crisis, as public opinion began to perceive banks as destroying communities through the consequences of their risky lending behavior, deposits began to flow into credit unions. They argue that that the reputation of credit unions as community-oriented institutions helped them gain deposits and grow in the aftermath of the financial crisis. My essay complements Chatterji *et al.* (2013) in so far as I provide a mechanism for why credit unions', and other mutuals', reputation is actually valid and helps them to outperform banks.

This essay also adds to the literature examining how mortgage lending contributed to

[2] For example, Fried *et al.* (1993) find substantial variation and inefficiencies among credit unions, whereas Altunbas *et al.* (2001) find that stock banks are not more efficient than mutual banks using German data.

the real estate boom and bust. This literature can roughly be divided into two views. The supply-based view, which argues that supply of credit to subprime borrowers drove real estate prices and subsequent delinquencies (Mian and Sufi, 2009, 2016; Favilukis *et al.*, 2016). The demand-based view, on the other hand, argues that subprime borrowing does not fully explain the housing boom and bust, and that demand for credit from middle and high-income households also contributed (Adelino *et al.*, 2016; Foote *et al.*, 2012). This essay adds to the literature by attempting to distinguish between these two views by examining which view is consistent with the differences in lending and performance between banks and mutual institutions.

The rest of the essay is organized as follows. Section 3.2 presents some background on mutual and stock institutions, as well as my methodology and the data used. Section 3.3 details my findings that banks underperformed mutual institutions during the 2007-2012 period, presents robustness checks, and rules out alternative explanations. Section 3.4 examines whether the supply or demand-based views of the real estate crisis best explain the differences in banks' and mutual institutions' lending and performance. Section 3.5 concludes.

3.2 Background, Data, and Methodology

3.2.1 Financial Institutions and Organizational Form

In this section, I discuss the main differences between the major types of banking institutions in the US. In the United States, there are four main types of banking institutions: commercial banks, credit unions, thrifts, and state savings banks.

Commercial banks are private stock institutions regulated by the Federal Reserve, the Federal Deposit Insurance Corporation (FDIC), and the Office of the Comptroller of the Currency (OCC), depending on whether they are members of the Fed and whether they have national or state charters. Commercial banks comprise the majority of banking institutions (by volume) in the US, with approximately 5000 banks and $16 trillion in assets as of the end of 2017.

Credit unions are cooperative non-profit mutual organizations, regulated by the National

Credit Union Association (NCUA). Because they are mutual institutions, the residual claim on a credit union is held by the depositors, who each get one vote. Credit unions are founded around a common affiliation and a restricted membership. For example, some credit unions were formed to serve the municipal employees of a certain geographic area. However, for many credit unions, membership is limited to residents of a certain county or state, and over time, more credit unions have either limited their membership based on geography or based on organizations that anyone can join for a small fee. This means that often, households can choose between using a local bank or the credit union for their banking services. Essentially, despite their closed membership, many credit unions compete with other depository institutions for customers. Importantly, credit unions are non-profit institutions and pay no income taxes. I will discuss the importance of this distinction in section 3.3. According to the National Credit Union Association (NCUA), as of the end of the 2017, the 5600 credit unions had 111 million members, roughly one-third of the population of the United States. They also had over $1.3 trillion of assets.

Thrift institutions, also called savings and loan associations, were originally founded to encourage "thrift" by primarily taking deposits from individuals and lending money via residential mortgages. In the 1800s and early 1900s, commercial banks were not heavily engaged in this market, especially in the years after the Great Depression. Thrifts were regulated by the Office of Thrift Supervision before 2011, and are now regulated by the FDIC and OCC. State savings banks are similar to thrifts in that the state saving bank charter was created by a few states, mostly on the East Coast, in the 1800s to engage in consumer banking. Both thrifts and state savings banks began as mutual institutions, but many converted to stock charters. In particular, after the Savings and Loan Crisis in the early 1980s, many thrifts went bankrupt and many others converted from mutual organization into stock organizations.

Although there several different types of depository institutions and several different regulating agencies, they can be roughly divided into mutual institutions (credit unions, mutual thrifts, and mutual savings banks) and stock institutions (commercial banks, stock thrifts, and stock savings banks).[3] The central question of this essay is whether mutuals take on

[3]There are some differences among the various types of institutions. For example, depositors of state savings

less risk, as theory suggests, and whether this can in fact be a mitigating effect during booms when banks underwrite too many risky loans.

As Rasmusen (1988) and Cororaton (2017) points out, mutual institutions do not allow for any mechanism through which shareholders can discipline managers. Each depositor gets one vote, and ownership is non-transferrable, and usually not tied to the amount of deposits a shareholder has. Thus, even when shareholders can vote for management, such as in mutual thrifts and credit unions, a single shareholder cannot amass enough votes to institute his own policies, as he could do in a stock institution. Since there is no mechanism to mitigate the agency problem between management and shareholders, the managers of mutual institutions are more likely to take perks such as fringe benefits and low managerial effort, and are less likely to try to increase profits (Jensen and H.Meckling, 1976). The management of mutuals has incentives to not engage in risky projects because of their perk-taking and relative disinterest in profits. By contrast, managers of stock institutions will engage in risky projects and in fact can take on too much risk since the fact that depositors do not observe assets means that depositors cannot withdraw their money if the institution takes on more risk than they want. A different mechanism explains the lack of risk-taking by mutuals is that they internalize the negative effects on the borrower if the borrower defaults (Cororaton, 2017). In either case, existing theory suggests that banks should engage in more risky lending relative to mutual institutions. For the purposes of this essay, I do not take a stand as to whether this relatively higher risk-taking is efficient or not.

3.2.2 Data and Methodology

Throughout the essay, I refer to commercial banks and stock savings banks as “banks” and to credit unions and mutual savings banks as mutual institutions or “mutuals” Due to data limitations, I constrain my analysis to consider only commercial banks, credit unions and savings banks. I do not have any stock or mutual thrifts in my sample.

Data for this analysis come from the Statements of Financial Conditions (Call Reports)

banks do not have voting rights.

compiled by the Federal Reserve for commercial banks and by the National Credit Union Association for credit unions. I consider the years 2003-2015 so as to use data both on the pre-crisis boom and the financial crisis and its aftermath. I start my analysis in 2003 to avoid the effects of the 2001 recession. I supplement the Call Report data with information on commercial bank branch locations from the FDIC's Summary of Deposits files and the county-level demographics data from the Census. I obtain quarterly real estate data from Zillow.

The main regressions that I run are difference-in-differences of the form:

$$Y_{ict} = \alpha + \beta Bank_i \times Post_t + \delta X_{it} + \lambda_i + \chi_{ct} + \epsilon_{ict} \tag{3.1}$$

Here, Y_{ict} are the outcomes I consider, which include whether institution i in county c failed at time t, and fraction of loans the institution charged off at time t and fraction of delinquent loans the institution had on its balance sheet. I consider several different types of lending, including real estate, consumer, and business loans. Business loans include both commercial and industrial (C&I) lending, as well small business loans. Note that the type of business lending the institutions do may not be comparable: commercial banks originate a great deal of C&I loans, whereas mutual institutions engage in very little C&I lending. X_{it} are control variables, which include the following variables lagged one year: log of assets, deposits as a fraction of assets, loans as a fraction of assets, real estate loans as a fraction of total loans, committed credit lines as a fraction of loans, and the equity ratio. λ_i are institution fixed effects, which control for any time-invariant differences between banks and mutual institutions.

The main variable of interest is $Bank_i \times Post_t$, which is the interaction between an indicator for whether institution i is a bank and an in indicator for the post-real estate market crash period of 2007-2012. The coefficient on this variable shows whether mutuals, on average, had better financial performance than banks during the crisis period, relative to the pre-crisis period. Better financial performance, especially in terms of delinquencies and charge-offs, implies that the loans mutual institutions made were of higher quality than loans made by similar banks.

3.2.3 Identification and Comparability

In this section, I discuss the two main concerns regarding the comparability of my analysis: comparability of the institutions' borrowers and comparability of the institutions' behaviors.

One threat to comparability is the possibility that banks and mutuals cater to different types of borrowers through their choices of where to locate. For example, if mutual institutions are more likely to be in high-income counties, their better performance may be due to differences in the incomes of the households they lend to . In this case, differences in the institutions' lending portfolio risk would be mechanical. To mitigate this issue, I estimate all regressions using county-year fixed effects χ_{ct} and limit my analysis only to institutions with branches in a single county to control for differences in borrower risk across geographies.[4] Thus, I compare, within each county-year pair, the performance of small banks and mutuals, who likely only lend to borrowers in the county where the branches are located.

A related threat to comparability is the issue that all credit unions have a closed membership. To rule out differences in membership as the driver of my findings, I limit my analysis just to "open" credit unions, for which membership is likely based on either geography or association with large organizations. Although I do not observe the membership criteria directly, I exclude from my sample unions whose membership is based on employee organizations or other limited groups by excluding all credit unions with the words "Employee", "Teachers", "Police", "Firefighters" and other professions in the name. I also exclude all credit unions for which the reported potential membership is less than half of the population of the county the credit union is domiciled in. Finally, I exclude all internet banks, large banks with more than $10 billion in assets, and all tiny credit unions with assets less than $25 million.

Table 3.1 presents summary statistics on my main sample as of June 30, 2002. My sample consists of 4722 banks and 683 mutuals. Unsurprisingly, the two types of depository institutions are somewhat different. Mutuals have a slightly higher fraction of deposits to assets–85% to only 83% for banks–and a higher ratio of loans to assets. They are also more likely to engage in real estate and consumer lending, as opposed to business lending. Importantly, the two types

[4] Branch locations are not available for credit unions prior to 2010 so I limit my sample to credit unions with 2 branches or fewer. The median bank with branches in only 1 county has 2 branches.

Table 3.1: *Summary Statistics*

This table compares banks and mutuals June 30, 2002. Assets = assets (in millions); Deposits/Assets = deposits as a fraction of assets; Loans/Assets = loans as a fraction of assets; Equity/Assets = equity ratio;Unused Credit Lines/Loans = commitments, including undrawn lines of credit, scaled by total loans, Res RE Loans/Assets = residential real estate loans scaled by assets; Cons Loans/Assets = consumer loans scaled by assets; Bus Loans/ Loans = commercial and industrial, as well as small business, loans, scaled by total loans.

		Full Sample	
Variable	Difference	Stock Inst	Mutual Inst
Assets	317.818	131,848.282	131,530.463
Deposits / Assets	-0.023	0.830	0.854
Loans / Assets	-0.063	0.597	0.660
Equity / Assets	0.003	0.117	0.114
Unused Credit Lines / Loans	0.005	0.106	0.100
Res RE Loans / Loans	-0.244	0.278	0.521
Cons Loans / Loans	-0.229	0.122	0.351
Bus Loans / Loans	0.398	0.518	0.120
N		4,722	683

of institutions in my sample are, on average, of the same size and have the same equity ratio. Thus, differences in size and how well-known the institution is, as well as in capital adequacy, are unlikely to drive my results. Because the institutions are different based on observable characteristics, in unreported results, I confirm that my findings hold when constrained to an even more comparable, propensity-matched sample of banks and mutuals.

By including county-year fixed effects and limiting to mutuals and banks whose membership is relatively open, I make sure that my findings are not driven by any mechanical differences in the borrowers of the two types of institutions. However, it is possible that differences in their lending may be driven by other fundamental differences between the institutions, unrelated to organizational form and risk-taking. One major difference between mutual and stock institution is the ability to grow and issue equity. Whereas a bank can issue equity to take advantage of a new investment opportunity, a mutual institution cannot, and this constraint may drive the type of lending the institutions engage in.

Although this constraint may play a role in the differences between large banks and credit unions, for the small institutions in my sample, it is not an issue. After excluding observations that correspond to mergers and acquisitions, the average yearly growth in assets is 5.3% for banks and 5.2% for mutuals, and the difference is not statistically significant. Similarly, the average growth in deposits and in loans is 4.5% for banks and 4.3% for mutuals, and again, the difference is not statistically significant in either case.

This is consistent with the view that that issuing equity is costly, either due to asymmetric information and moral hazard issues, or due to the costs of placing the equity, either directly to private shareholders or through financial intermediaries. Whereas larger banks have access to whole-sale funding, small banks do not and due to the costliness of equity, they tend to fund themselves mainly through deposits (Park and Pennacchi, 2009). Thus, even though mutual institutions cannot issue equity, this does not disadvantage them, and does not affect their lending decisions relative to those of small banks.

Note that this comparability only holds for my sample of small institutions. When one includes larger banks, banks do tend to grow faster than mutuals. Similarly, banks undergo mergers and acquisitions much more often than mutual institutions, and so the comparability of growth rates holds only for non-acquiring institutions.

Because of this, generalizing my findings to a general comparison of mutuals and banks may not be appropriate. However, this does not affect the main finding that organizational form plays a role in risk-taking. I do not estimate the size of this effect relative to other factors, and it is quite possible that relative to the ability to issue equity, the effect of organizational form is minor. However, it is not the case that organizational form does not matter at all.

3.3 Performance during the Financial Crisis

In this section, I address the question of whether the performance of banks differed from that of mutual institutions during the 2007-2012 period, as would be consistent with banks engaging in more risky lending. I find that banks did underperform mutual institutions: banks were more likely to fail and had higher charge-offs and delinquencies, particularly on their real estate loan portfolios.

3.3.1 Failure, Charge-offs, and Delinquencies

Table 3.2: *Probability of Failure*

*This table presents the results of the difference-in-differences specification of equation 3.1 of Section 3.2, estimating the effect of organizational growth on failure rates. The dependent variable is an indicator for whether the institution failed in that year. Bank $\times$ Post Period is the interaction between an indicator for whether the institution is a bank and the 2007-2012 crisis period. County-year fixed effects and institution fixed effects are included in each regression. Control variables include the following variables lagged one year: log of assets, deposits as a fraction of assets, loans as a fraction of assets, real estate loans as a fraction of total loans, committed credit lines as a fraction of loans, and the equity ratio. All variables are measured as of December 31 of each year. Standard errors are shown in parentheses and are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.*

Dependent Variable:		Failure	
	(1)	(2)	(3)
Bank x Post Period	0.011***	0.010***	0.010***
	(0.002)	(0.002)	(0.003)
County-Year Fixed Effects	Yes	Yes	Yes
Institution Fixed Effects		Yes	Yes
Controls			Yes
Observations	44,616	37,540	36,821
Within R-squared	0.001	0.005	0.002

I begin by examining the effect of organizational form on failure rates. Table 3.2 presents the results. The dependent variable is an indicator for whether the institution failed in that year. The main variable of interest is $Bank_i \times Post_t$, the interaction between an indicator for whether the institution is a bank multiplied by an indicator for the post housing crash 2007-2012 period. Column 1 includes county-year fixed effects, whereas column 2 adds institution fixed effects and column 3 adds institution-year controls. The controls are the following variables lagged one year: log of assets, deposits as a fraction of assets, loans as a fraction of assets, real estate loans as a fraction of total loans, committed credit lines as a fraction of loans, and the equity ratio. In all columns, the coefficient on $Bank_i \times Post_t$ is approximately the same and equal to 0.01, suggesting that the probability of failure is 1% higher each year for a bank than for a mutual institution.

I next examine *why* banks were more likely to fail during the 2007-2012 period, by comparing their financial performance to that of mutuals in Tables 3.3 and 3.4. First, in column 1 of Table

Table 3.3: *Effect of Organizational Form on Charge-offs*

*This table presents the results of the difference-in-differences specification of equation 3.1 of Section 3.2, estimating the effect of organizational growth on charge-offs as a fraction of loans. The dependent variable is total charge-offs as a fraction of loans in Column 1. In columns 2 through 4, the dependent variables are charge-offs on real estate loans as a fraction of real estate loans, charge-offs on consumer loans as a fraction of consumer loans, and charge-offs on business loans a fraction of business loans. Bank $\times$ Post Period is the interaction between an indicator for whether the institution is a bank and the 2007-2012 crisis period. County-year fixed effects and institution fixed effects are included in each regression. Control variables include the following variables lagged one year: log of assets, deposits as a fraction of assets, loans as a fraction of assets, real estate loans as a fraction of total loans, committed credit lines as a fraction of loans, and the equity ratio. All variables are measured as of December 31 of each year. Standard errors are shown in parentheses and are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.*

Dependent variable:	Chargeoffs/Loans			
	Total (1)	RRE (2)	Consumer (3)	Business (4)
Bank x Post Period	0.002***	0.001***	0.001	0.001
	(0.000)	(0.000)	(0.001)	(0.001)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Institution Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	36,812	40,228	40,962	34,938
Within R-squared	0.020	0.004	0.004	0.001

3.3, I repeat the analysis of column 3 of Table 3.2, using charge-offs as a fraction of loans as the dependent variable. The coefficient on $Bank_i \times Post_t$ is positive and significant, which suggests that banks did have higher charge-offs during the financial crisis. Next, I examine 3 types of lending: real estate lending in column 2, consumer lending in column 3, and business lending in column 4. Note that throughout my analysis, I focus solely on residential real estae lending since credit unions do not engage in commercial real estate lending. The results suggest that the higher charge-offs are driven solely by the real estate portfolio; banks and mutual institutions do not seem to differ in the performance of their consumer or business loans. The coefficient on $Bank_i \times Post_t$ is 0.001 and the baseline rate of charge-offs as a fraction of loans for banks during the 2007-2012 period is 0.004, so the difference between banks and mutuals during the crisis explains approximately 25% of banks' baseline rate of charge-offs. Note that the rate of 0.004 during the 2007-2012 period is very low, and this is due to the fact that the sample is composed of small banks. Larger banks had higher charge-offs during the financial crisis.

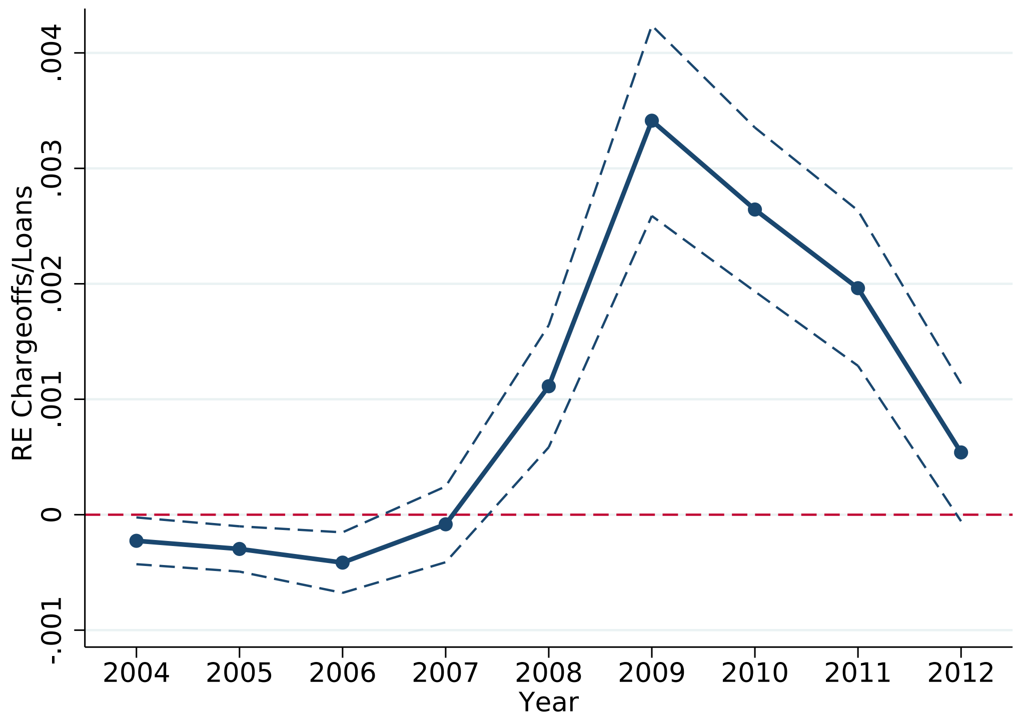


Figure 3.2: *Yearly Coefficients: Charge-offs*

This figure plots the yearly coefficients from a difference-in-difference regression estimating the effect of organizational form on real estate charge-offs as a fraction of real estate lending. The solid line shows the coefficients on the interactions between whether an institution is a bank, and indicators for each years after the merger. Dashed lines capture the 95% confidence intervals. Banks are defined as commercial banks and stock state savings banks. Mutual institutions are defined as credit unions and mutual state savings banks.

Figure 3.2 shows the difference between banks and mutual institutions each year. Consistent with Figure 3.1, charge-offs as a fraction of loans increase for banks relative to mutual institutions through 2009, before slowly decreasing again to the pre-crisis baseline by 2013.

In Table 3.4, I repeat the analysis of 3.3 using delinquent loans as a fraction of loans as the dependent variable. Because of discrepancies between how banks and credit unions report delinquencies, the measure I use is very broad. I define delinquencies to be any non-accruing loan as well as loans 30 or more days past due. As with charge-offs, banks have higher rates of delinquent loans than mutual institutions, and this is mainly driven by the real estate lending

Table 3.4: *Effect of Organizational Form on Delinquencies*

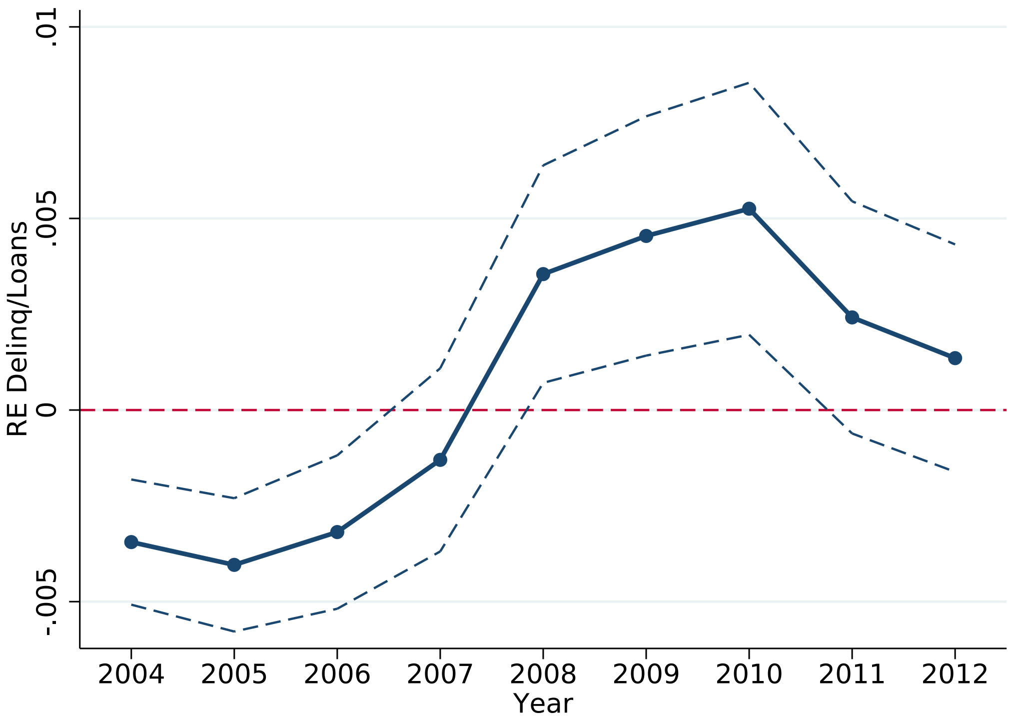
*This table presents the results of the difference-in-differences specification of equation 3.1 of Section 3.2, estimating the effect of organizational growth on delinquencies as a fraction of loans. The dependent variable is total delinquencies as a fraction of loans in Column 1. In columns 2 through 4, the dependent variables are delinquent real estate loans as a fraction of real estate loans, delinquent consumer loans as a fraction of consumer loans, and delinquent business loans a fraction of business loans. Delinquent loans are defined as non-accrual loans and loans more than 30 days past due. Bank $\times$ Post Period is the interaction between an indicator for whether the institution is a bank and the 2007-2012 crisis period. County-year fixed effects and institution fixed effects are included in each regression. Control variables include the following variables lagged one year: log of assets, deposits as a fraction of assets, loans as a fraction of assets, real estate loans as a fraction of total loans, committed credit lines as a fraction of loans, and the equity ratio. All variables are measured as of December 31 of each year. Standard errors are shown in parentheses and are clustered at the county level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

Dependent variable:	Delinquency/Loans			
	Total	RRE	Consumer	Business
	(1)	(2)	(3)	(4)
Bank x Post Period	0.002	0.004**	-0.003	0.000
	(0.002)	(0.002)	(0.003)	(0.006)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Institution Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	36,819	40,228	40,960	34,938
Within R-squared	0.004	0.003	0.000	0.001

portfolio. The difference between banks and mutual institutions explains approximately 10% of banks' delinquent loans. Figure 3.3 plots the difference between banks and mutual institutions each year. Again, the difference increases sharply until about 2010 before decreasing back to the pre-crisis normal.

Both Figures 3.2 and 3.3 show that the difference between banks and mutual institutions was negative prior to 2007. It is possible that banks were engaged in less risky lending than mutual institutions before the crisis, but it is more likely that this difference is driven by rates. As I discuss further in the next subsection, it is possible that banks were originating more adjustable-rate mortgages (ARMs) that had very low introductory "teaser" rates that drove down banks' delinquencies and charge-offs prior to the real estate crash.

Figure 3.3: *Yearly Coefficients: Delinquency*



This figure plots the yearly coefficients from a difference-in-difference regression estimating the effect of organizational form on real estate delinquencies as a fraction of real estate lending. The solid line shows the coefficients on the interactions between whether an institution is a bank, and indicators for each years after the merger. Dashed lines capture the 95% confidence intervals. Banks are defined as commercial banks and stock state savings banks. Mutual institutions are defined as credit unions and mutual state savings banks.

3.3.2 Alternative Explanations

In the above section, I showed that banks were more likely to fail and had lower delinquencies and charge-offs than comparable mutual institutions, consistent with organizational form affecting risk-taking. In this section, I rule out several alternative explanations for my results.

First, one possible explanation for these results is that banks and mutuals sold or securitized loans and mortgages at different rates, and it is differences in securitization or mortgage servicing that drive the results. To rule this out, in columns 1 and 2 of table 3.5, I replicate columns 2 of Table 3.3 and 3.4 only keeping institutions that did not sell, service or securitize mortgages. I also exclude institutions that held mortgage-backed securities (MBS). The results are very similar. Generally, banks securitized and sold mortgages at higher rates than credit

Table 3.5: *Alternative Explanations*

*Columns 1 and 2 exclude from the analysis any institution that has a non-zero income from mortgage servicing. Columns 3 and 4 constrain the analysis just to state savings banks. In columns 1 and 3, the dependent variable is charge-offs on real estate loans as a fraction of real estate loans. In columns 2 and 4, the dependent variable is delinquent real estate loans as a fraction of real estate loans. Bank $\times$ Post Period is the interaction between an indicator for whether the institution is a bank and the 2007-2012 crisis period. County-year fixed effects and institution fixed effects are included in each regression. Control variables include the following variables lagged one year: log of assets, deposits as a fraction of assets, loans as a fraction of assets, real estate loans as a fraction of total loans, committed credit lines as a fraction of loans, and the equity ratio. All variables are measured as of December 31 of each year. Standard errors are shown in parentheses and are clustered at the county level. * p<0.10, ** p<0.05, *** p<0.01.*

	Excl Mtg Serv		State Savings Banks	
Dependent Variable:	Chargeoffs (1)	Delinq (2)	Chargeoffs (3)	Delinq (4)
Bank x Post Period	0.002*** (0.000)	0.006*** (0.002)	0.001** (0.000)	0.004* (0.002)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Institution Fixed Effects	Yes	Yes		
Controls	Yes	Yes	Yes	Yes
Observations	35,751	35,406	3,042	3,042
Within R-squared	0.005	0.003	0.014	0.005

unions, but both types of institutions in my sample are small and did not engage much in securitization.

Second, since credit unions make up the majority of my mutual institutions sample, and credit unions differ from banks in several aspects of their operations, it is possible that differences other than organizational form drive these results. Specifically, credit unions are regulated by the NCUA, whereas most banks in my sample are regulated by the FDIC. Although definitive rules exist for when a loan becomes delinquent, institutions, in practice, have more leeway in determining charge-offs. If the NCUA is more lax in its enforcement of these rules, then credit unions may be able to artificially delay charging off a loan, thus lowering their charge-off rates. Because both delinquencies are also lower for credit unions, it is unlikely that differences in regulators drive my results.

However, to rule out more directly that my findings are due to differences in regulators, in columns 3 and 4 of table 3.5, I rerun my analysis focusing solely on state savings banks. State savings banks can be mutual or stock institutions, and both types are regulated by the FDIC. As the results show, mutual state savings banks perform better than stock state savings banks, during the financial crisis.[5]

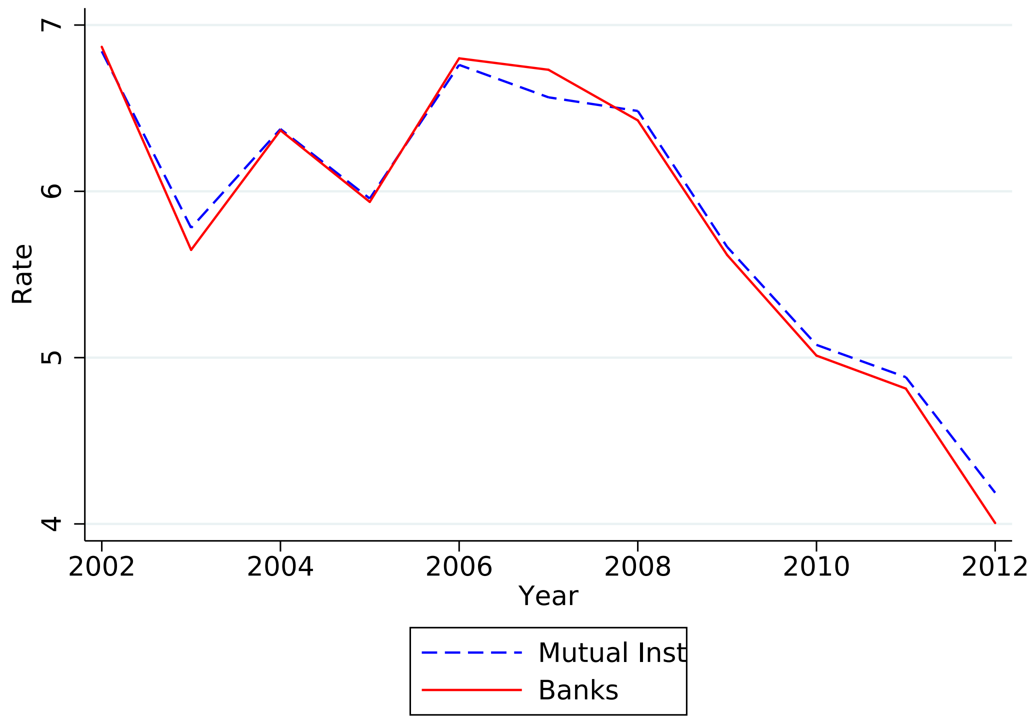


Figure 3.4: *30-Year Fixed Rate Mortgage Rates*

This figure plots the average 30-year fixed rate mortgage rates at origination for banks and mutual institutions. Banks are defined as commercial banks and stock state savings banks. Mutual institutions are defined as credit unions and mutual state savings banks.

A related concern is that because credit unions do not pay income taxes, their sources of funding are essentially cheaper than the sources of funding for banks. Thus, it is possible that credit unions are able to consistently offer lower loan and mortgage rates. If that is the case, even if the set of borrowers is the same, credit unions would consistently have lower rates of delinquencies and charge-offs. I rule this out in two ways. First, columns 3 and 4 suggest that my results are robust to excluding credit unions, so differences in loan rates cannot be the sole

[5] I exclude institution fixed effects because of the low number of observations.

explanation for the results. Second, although credit unions do sometimes offer lower mortgage rates, the difference with rates offered by the small banks in my sample is minute. Figure 3.4 plots the 30-year fixed-rate mortgage rate for banks and mutuals using data from Ratewatch. Although the rates for mutuals are lower during the 2006-2008 period, they are higher from 2009 to 2012. In unreported results, I show that comparing institutions in the same county, the rates for mutual institutions are lower and the difference is statistically significant, but it is not economically significant. On average, the mortgage rates offered by mutual institutions are lower by approximately 5 basis points.[6] Thus, although differences in rates exist, they do not seem large enough to explain the differences in delinquency and charge-off rates.

It is important to note that since I cannot directly observe much information about the borrowers of mutual and stock institutions, I do not distinguish between banks originating loans to riskier borrowers, or banks originating more risky types of loans. Although Figure 3.4 shows that rates on 30-year fixed rate mortgages are comparable, it is possible that banks originated more adjustable rate mortgages (ARMs). If the rates on ARMs were lower during the 2002-2006 boom and then reset much higher after the real estate market crash, this would explain the difference in delinquencies and charge-offs, even if the set of borrowers is exactly the same and banks do not lend more to riskier borrowers. This does not contradict my hypothesis that banks take on more risk–originating ARMs instead of fixed-rate mortgages is inherently more risky.

However, there is some evidence in the literature that suggests that it is not a difference in the types of lending that drives subsequent performance. Foote *et al.* (2012) argue that the changes in rates when ARMs reset were small, and could not be the sole driver of subsequent borrower delinquency and default. This suggests that intrinsic borrower riskiness played a role. I examine more directly the characteristics of mortgages and mortgage borrowers of the two types of institutions in the next section.

---

[6]Figure 3.4 compares the full mutual institution sample with the sample of banks, but the results are almost identical when limiting to just credit unions and banks.

3.4 Supply and Demand-Based Theories of the Real Estate Crisis

In this section, I discuss the supply and demand-based theories of the real estate crisis and examine the extent to which the lending and performance of banks and mutual institutions are consistent with these theories. Having established that bank under-performed mutuals during the 2007-2012, I take this difference in performance as given and examine the differences in risk-taking that led to it.

To do so, I use data from the Home Mortgage Disclosure Act (HMDA) mortgage database. Under HMDA, virtually all depository institutions have to disclose each mortgage originated in a metropolitan statistical area (MSA) where a branch of office of the institution is located. For each mortgage, the institution reports the year the loan is originated, the census tract where it is originated, the size of the loan, the applicant's income, race and other demographic information, whether the loan is sold within a year, and if so whether it is sold to one of the government-sponsored enterprises, to an affiliate institution, or to private securitization. Beginning in 2004, the Federal Reserve's Regulation C required lenders to collect and report the spread between the annual percentage rate (APR) on a loan and the yield on Treasury securities of comparable maturity if the spread was "equal to or greater than 3.0 percentage points for a first-lien loan (or 5.0 percentage points for a subordinate-lien loan)."[7] These high-priced loans were more likely to be loans to subprime borrowers (Mayer and Pence, 2008).

3.4.1 Supply-based View

The current literature presents two main views of risk-taking, mortgage lending, and real estate prices. The first is the supply-based view of Mian and Sufi (2009, 2016). The authors show that during the 2002-2005 period, mortgage lending expanded most in zip codes with high latent demand for credit. These are zip codes that had high denial rates for mortgages in 1996; essentially, these are low-income and subprime zip codes that were traditionally seen as high-risk and thus denied credit. Due to low interest rates and the initial housing boom in the

[7] See "Rules and Regulations," Board of Governors of the Federal Reserve System, Federal Register, October 24, 2008.

early 2000s, banks expanded credit to these households. Mian and Sufi (2009) show that the increase in lending was not associated with an increase in household income; mortgage growth was strongest in zip codes that experienced low or even negative growth in adjusted gross income (AGI). The inflow of credit pushed up real estate prices, and subsequently led to the crash in prices, as well as borrower defaults and foreclosures. Overall, the supply-based view suggests that a loosening of credit standards via increased lending to subprime households was the main cause of the real estate bubble and crisis.

Since mutual institutions outperformed banks from 2007 to 2012, the supply-based view implies that mutual institutions did not engage in as much subprime lending and did not loosen their credit standards on mortgages during the boom period of 2002-2006. I next empirically test this hypothesis.

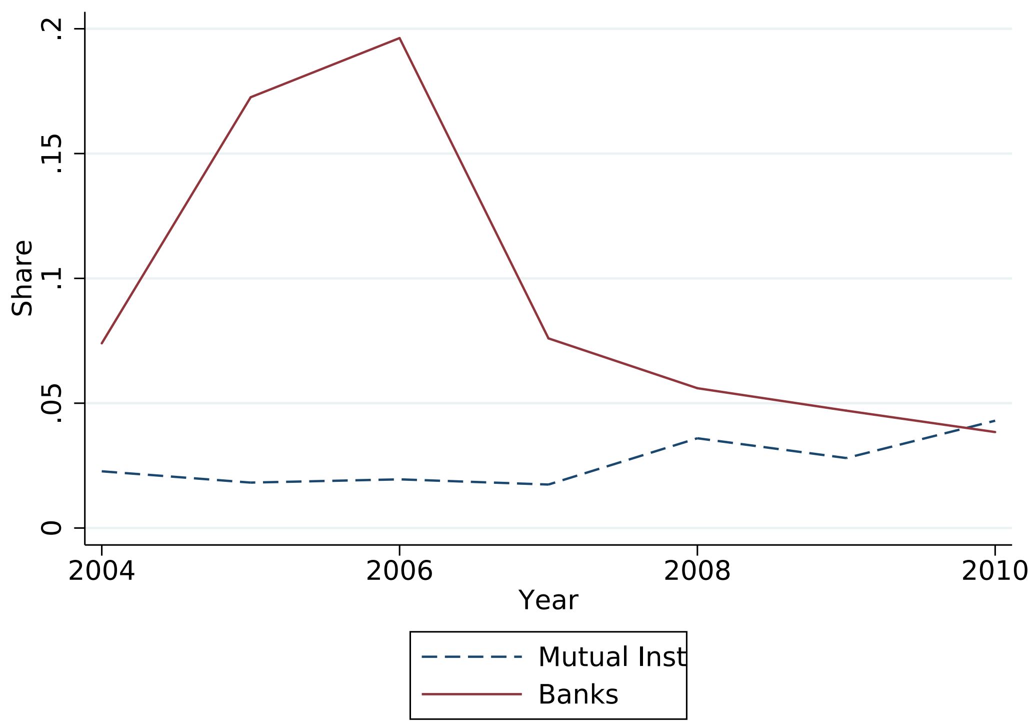


Figure 3.5: *Share of High-priced Lending*

This figure shows the average percent of high-priced lending originated by each type of institution between 2004 and 2010. A loan is defined as high-priced if it is marked as high-priced in the HMDA dataset. Banks are defined as commercial banks and stock state savings banks. Mutual institutions are defined as credit unions and mutual state savings banks.

Figure 3.5 shows the average fraction of high-priced lending originated by banks and mutual institutions. Although I do not observe whether a borrower is subprime or not, one proxy for subprime lending is lending identified as high-priced in HMDA. As the figure shows, banks have a much higher share of high-priced lending. Not only did banks engage in more high-priced lending in 2004, but they grew their shares from 2004 to 2006, whereas mutual institutions did not.

Panels A and B of 3.6 examine more explicitly whether differences in the lending of banks and mutual institutions are consistent with the supply-based view. In Panel A, the dependent variable is the change from 2004 to 2006 in the percent of high-priced lending. The dependent variable in Panel B is the change from 2002 to 2006 in the denial rate, the percentage of applications denied by the institution. Each observation corresponds to a single zip code, and all lending is aggregated by institution type. I include county fixed effects in all regressions. The first two columns examine banks and columns 3-4 examine credit unions.

The results show that risk-taking by banks was consistent with the supply-based view of the real estate crisis. Banks increased their high-priced lending in areas with high latent demand (as measured by the denial rate in 1996) and lower-income areas (as measured by the fraction of households with AGI of less than $50,000 in 2002). By contrast, mutual institutions did not increase their subprime lending in these areas; as can be seen from 3.5, the high-priced lending of mutual institutions was low throughout this period. Similarly, banks lowered their credit standards during the 2002-2006 period, whereas mutual institutions did not. Banks decreased their denial rates in the zip codes where they experienced higher growth in the number of applications, and in zip codes with higher latent demand. Similar to the results of Panel A, there was no relationship between high latent demand or the growth in the number of applications and denial rates for mutual institutions.

3.4.2 Demand-based View

The second view of the housing crisis is the demand-based view of Adelino *et al.* (2016). Adelino *et al.* (2016) argue that the supply-based view does not fully explain the real estate boom and bust, and that middle and high-income households played a role in what happened,

Table 3.6: *High-Priced Lending and Loose Lending Standards*

*This table examines whether the lending of banks and mutual institutions is consistent with the supply-based view of the real estate crisis. In Panel A, the dependent variable is the change from 2004 to 2006 in the percent of high-priced lending the institution originates. In Panel B, the dependent variable is the change from 2002 to 2006 in the percentage of mortgage applications that are denied. In columns 1 and 2, the sample is banks and in columns 3 and 4, it is mutual institutions. All specifications include county fixed effects. Standard errors (reported in brackets) are clustered at the county level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Panel A: High-Priced Lending

Dependent Variable:	Chg in Percent High-Priced Lending '04-'06			
	Banks		Mutuals	
	(1)	(2)	(3)	(4)
Denial Rate 1996	0.034***		-0.007	
	(0.004)		(0.009)	
Frac HH w/Inc < $50K		0.054***		-0.009
		(0.004)		(0.010)
Gr in Ave Agi '02-'06	-0.286***	-0.187***	-0.088*	-0.094**
	(0.020)	(0.016)	(0.049)	(0.048)
Gr in RE Prices '02-'04	0.025***	0.015**	0.045**	0.039**
	(0.008)	(0.007)	(0.018)	(0.017)
County Fixed Effects	Yes	Yes	Yes	Yes
Observations	8,316	8,825	3,045	3,245
Within R-squared	0.111	0.136	0.005	0.004

Panel B: Loose Lending Standards

Dependent Variable:	Chg in Denial Rate '02-'06			
	Banks		Mutuals	
	(1)	(2)	(3)	(4)
Gr. Num Apps '02-'06	-0.010*	-0.013***	0.004	0.004
	(0.005)	(0.005)	(0.005)	(0.005)
Denial Rate 1996		-0.057***		-0.002
		(0.006)		(0.009)
County Fixed Effects	Yes	Yes	Yes	Yes
Observations	8,314	7,837	7,748	7,293
Within R-squared	0.001	0.035	0.000	0.000

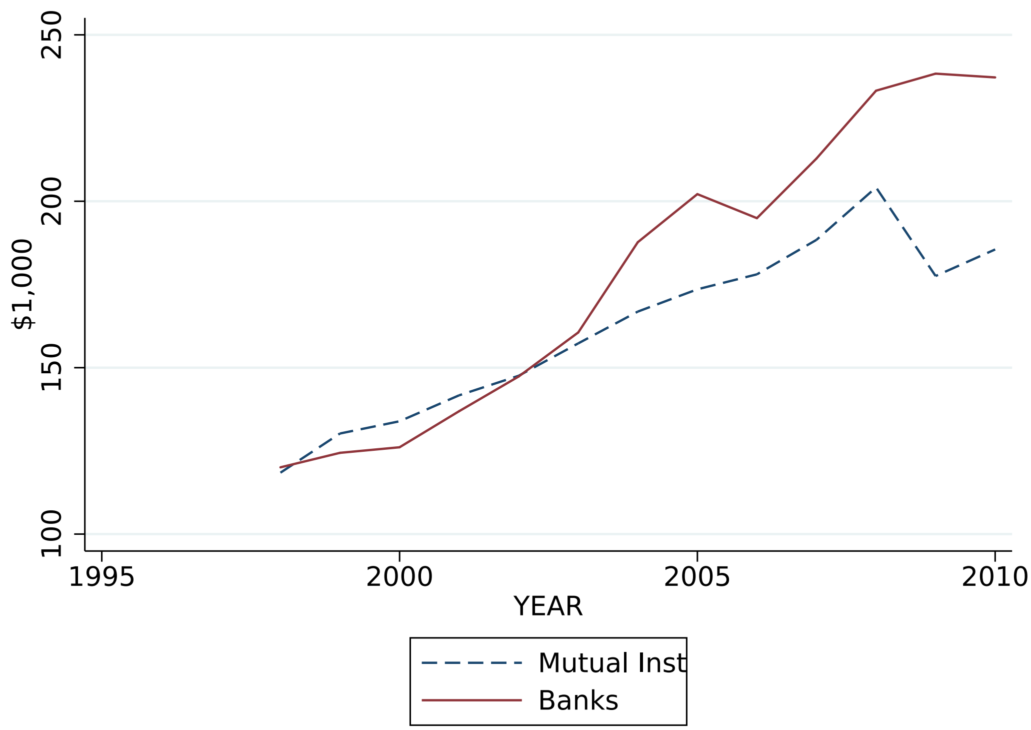


Figure 3.6: *Average Mortgage Size*

This figure shows the average size of the mortgages originated by each type of institution between 1998 and 2010. Banks are defined as commercial banks and stock state savings banks. Mutual institutions are defined as credit unions and mutual state savings banks.

not just subprime households. They show that although mortgage lending increased more in subprime areas, the average mortgage size grew faster in high-income areas. SImilarly, whereas total lending was negatively correlated with AGI growth, growth in the average mortgage amount was positively correlated with AGI growth and with growth in applicant income. Thus, the demand-based view argues that demand for credit from middle and high-income households led to higher average mortgage sizes and, therefore, increased real estate prices. Adelino *et al.* (2016) also show that this subsequently led to defaults, with the share of total defaults increasing faster in non-subprime areas.

Overall, the demand-based view suggests that banks underperformed mutuals due to relative over-extension of mortgage credit to middle and high-income households. Figure 3.6 shows that, consistent with this hypothesis, banks did grow their average mortgage size faster than mutuals, beginning in 2004. In Panel A of Table 3.7, I replicate the results of Adelino *et al.* (2016) for banks and mutual institutions. In columns 1 and 3, the dependent variable is the growth in total mortgage lending from 2002 to 2006, and in columns 2 and 4, it is the growth in the average mortgage size during the same period. As before, each observation corresponds to a zip code and I include county fixed effects. Consistent with the findings of Adelino *et al.* (2016), the growth in total mortgage lending was negatively correlated with the growth in zip code AGI from 2002 to 2006 for both banks and mutuals. However, whereas the growth in average mortgage size was positively correlated with AGI growth for banks, there was no relationship for mutual institutions.

In Panel B, I investigate further whether banks extended bigger mortgages to middle and high-income households, whereas mutual institutions did not. To do so, I examine whether the relationship between the average mortgage amount and income is stronger during the 2003-2006 boom period than the prior 1998-2002 period. I run a pooled regression for the full 1998-2006 period. The dependent variable in Panel B is the log of the average mortgage amount, and again, columns 1 and 2 examine banks and columns 3 and 4 examine mutual institutions. In columns 1 and 3, the main variables of interest are the log of AGI and the interaction between log of AGI and an indicator for the 2003-2006 boom period. In columns 2 and 4, I use log of the applicant income instead of AGI. It is important to consider both since there is evidence of overstatement of applicant income during the 2002-2006 period (Blackburn and Vermilyea, 2009; Mian and Sufi, 2017).

The results are consistent with those of Panel A. Generally, higher incomes lead to higher mortgage sizes for both banks and mutuals. However, for banks, relative to the pre-boom period, an increase in AGI of 1% during the 2003-2006 period leads to an increase in the average mortgage size of 1.08%. Similarly, an increase of 1% in the applicant income leads to an increase in the average mortgage size of 1.15%. By contrast, mutual institutions did not differentially increase average mortgage size more during the boom period.

Figure 3.7 shows this explicitly by presenting how the difference between the growth rates in the average mortgage size and applicant income vary by the income growth. The x-axis displays quintiles of AGI growth from 2002 to 2006. The y-axis displays the difference between

Table 3.7: *High-Priced Lending and Loose Lending Standards*

*This table examines whether the lending of banks and mutual institutions is consistent with the demand-based view of the real estate crisis. In Panel A, the dependent variable in columns 1 and 3 is growth in total mortgage origination from 2002 to 2006. In columns 2 and 4, it is the growth in the average mortgage size. In Panel B, the dependent variable is log of the average mortgage amount. In columns 1 and 2, the sample is banks and in columns 3 and 4, it is mutual institutions. Standard errors are clustered at the county level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Panel A: High-Priced Lending

	Banks		Mutuals	
	Gr Total ’02-’06	Gr Ave ’02-’06	Gr Total ’02-’06	Gr Ave’02-’06
	(1)	(2)	(3)	(4)
Gr AGI ’02-’06	-0.682***	0.159***	-0.341***	0.019
	(0.080)	(0.041)	(0.128)	(0.062)
County Fixed Effects	Yes	Yes	Yes	Yes
Observations	8,313	8,310	7,412	7,812
Within R-squared	0.018	0.004	0.001	0.000

Panel B: Loose Lending Standards

Dependent Variable:	Log Ave Mortgage Amount			
	Banks		Mutuals	
	(1)	(2)	(3)	(4)
Log AGI x ’03-’06 Period	0.079***		-0.020	
	(0.017)		(0.015)	
Log AGI	0.356***		0.389***	
	(0.067)		(0.071)	
Log Appl Inc x ’03-’06 Period		0.148***		-0.002
		(0.020)		(0.013)
Log Appl Inc		0.330***		0.504***
		(0.008)		(0.009)
Zip Fixed Effects	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	52,612	79,849	51,734	78,264
Within R-squared	0.033	0.221	0.005	0.214

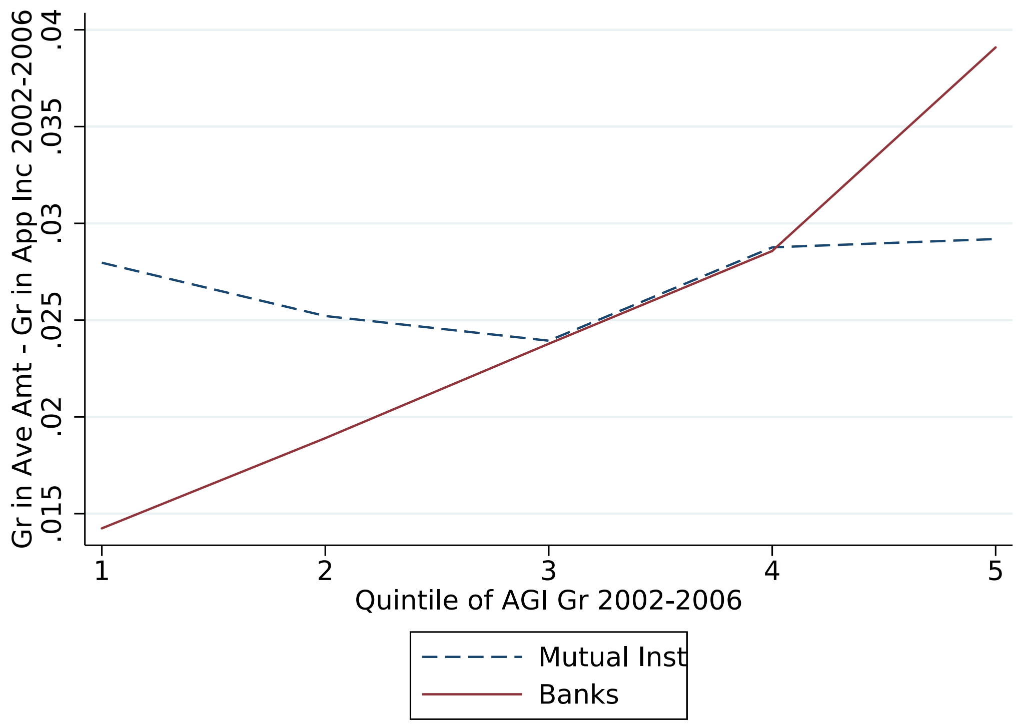


Figure 3.7: *Difference in Mortgage Size and Applicant Income Growth Rates*

This figure shows the difference between the growth in the average mortgage size and the growth in applicant income from 2002 to 2006, by each type of institution and by quintile of adjusted gross income (AGI) at the zip-code level. Banks are defined as commercial banks and stock state savings banks. Mutual institutions are defined as credit unions and mutual state savings banks.

the growth rate in the average mortgage size and the growth rate in applicant income. For mutual institutions, the difference is approximately stable: the average mortgage size does not increase much faster than income in areas where income is growing fast. However, for banks, the difference in growth rates has an upward slope. Average mortgage size grows slightly faster than applicant income in areas where AGI grows slowly, but the difference is very small. By contrast, in areas where AGI grew quickly from 2002 to 2006, the average mortgage size grew much faster than applicant income. This is consistent with banks over-extending credit to middle and high-income households, households in areas where average income grew quickly.

Overall, differences in lending between banks and mutual institutions are consistent with both the supply and demand-based views. Relative to mutual institutions, banks engaged in more high-priced lending and increased their high-priced lending more from 2004 to 2006. At the same time, they extended larger mortgages to middle and high-income borrowers. Next, I examine explicitly which type of risk-taking can explain banks' worse performance.

3.4.3 Risk-taking and Performance

The previous sections show that banks' lending during the real estate boom was consistent with both the supply and demand-based explanations. In this section, I differentiate between the two explanations by investigating which type of risk-taking best explains banks' worse performance during the 2007-2012 period.

To tie risk-taking directly to performance, I create two proxies, one for each explanation of the real estate crisis. My proxy for the supply-based explanation is Pct High-priced, the average fraction of 2004-2006 lending that is defined as high-priced in HMDA. Institutions that had a high percentage of high-priced lending engaged in more risky, subprime lending. My proxy for the demand-based explanation is Gr Ave Amt - Gr App Inc, the difference between the growth in the average mortgage amount and the growth in the applicant income. Institutions that grew their average mortgage size much faster than their applicant incomes overextended credit to middle and high-income households, consistent with the demand-based view of the real estate crisis.

Table 3.8 presents the results. In columns 1 and 2, the dependent variable is charge-offs on real estate loans, as a fraction of real estate loans, and in columns 3 and 4, the dependent variable is real estate delinquent loans, as a fraction of real estate loans. In columns 1 and 3, I use as my main variable of interest the interaction between Pct High-priced and the 2007-2012 Post Period. In columns 2 and 4, I use the interaction between Gr Ave Amt - Gr App Inc and the 2007-2012 Post Period. The table shows that whereas institutions that engaged in more high-priced lending had higher charge-offs and delinquencies during the crisis period, the difference between the growth rate in the average mortgage size and applicant income does not explain subsequent performance. Thus, my results are consistent with the supply and not the demand-based view of the crisis.

However, it is important to note that the explanatory power of Pct High-priced is very low. Adding the interaction between Pct High-priced and the 2007-2012 Post Period does not have a

Table 3.8: *What Drives Banks' Worse Performance?*

*This table examines whether proxies for the supply-based or demand-based view of the real estate crisis best explain banks' worse performance from 2007 to 2012. In columns 1 and 2, the dependent variable is charge-offs on real estate loans as a fraction of real estate loans loans. In columns 3 and 4, the dependent variable is delinquent real estate loans as a fraction of real estate loans. Bank × Post Period is the interaction between an indicator for whether the institution is a bank and the 2007-2012 crisis period. Pct High-priced × Post Period is the interaction between the institution's average percent of high-priced lending from 2004-2006 and the 2007-2012 crisis period. Gr Ave Amt - Gr App Inc × Post Period is the interaction of the difference between the growth rate of the average mortgage size and the growth rate of applicant income from 2002 to 2006, and the 2007-2012 crisis period. Control variables include the following variables lagged one year: log of assets, deposits as a fraction of assets, loans as a fraction of assets, real estate loans as a fraction of total loans, committed credit lines as a fraction of loans, and the equity ratio. Standard errors are clustered at the county level. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Dependent Variable:	RRE Chargeoffs/Loans		RRE Delinq/Loans	
	(1)	(2)	(3)	(4)
Bank x Post Period	0.001**	0.001*	0.008***	0.005*
	(0.001)	(0.001)	(0.003)	(0.003)
Post Period × Pct High-priced	0.002*		0.010*	
	(0.001)		(0.006)	
Post Period		0.001		-0.003
× Gr Ave Amt - Gr App Inc		(0.000)		(0.003)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Institution Fixed Effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Observations	20,908	20,908	20,908	20,908
Within R-squared	0.011	0.011	0.019	0.018

large effect on the coefficient on Bank × Post Period. In addition, the coefficient on Bank × Post Period remains positive and statistically significant in all regressions. Thus, although the supply-based explanation has some predictive power for subsequent performance, it is not the sole explanation for the difference in performance between banks and mutual institutions. Other factors, such as mutual institutions' stronger relationships with their depositors and borrowers, likely play a role.

3.5 Conclusion

In this essay, I attempt to distinguish between the supply and demand-based explanations for the real estate crisis using a novel setting: the differences in risk-taking and performance between mutual institutions and banks. I find that banks underperformed mutual institutions during the 2007-2012 crisis period. Banks were more likely to fail, and had higher charge-offs and delinquencies on their real estate portfolios. This under-performance is likely linked to banks engaging in more risky lending during the real estate boom period. I examine the types of risky lending that banks engaged in during the 2002-2006 boom and find evidence consistent with both the demand and supply-based views. Relative to mutual institutions, banks engaged in more subprime lending and also over-extended credit to middle and high-income borrowers. However, only subprime lending helps explain banks' subsequent under-performance.

Although the evidence for the supply-based view is not conclusive, and does not fully explain the difference in performance between banks and mutual institutions, these findings are one of the first attempts to explicitly distinguish between the supply and demand-based views of the housing crisis. These findings are important both for academic researchers seeking to understand mutual institutions and the link between organizational form and risk, and for policy-makers seeking to understand the risk-taking that led to the worst financial crisis since the Great Depression. Understanding exactly what kind of risk-taking led to such dire effects can help policy-makers and regulators remain vigilant for the future.

References

Adelino, M., Schoar, A. and Severino, F. (2016). Loan originations and defaults in the mortgage crisis: The role of the middle class. *Review of FInancial Studies*, **29** (7), 1635–1670.

Afonso, G., Kovner, A. and Schoar, A. (2011). Stressed, not frozen: The federal funds market in the financial crisis. *Journal of Finance*, **66** (4), 1109–1139.

Agarwal, S. and Hauswald, R. (2010). Distance and private information in lending. *Review of Financial Studies*, **23** (7), 2757–2788.

—, Skiba, P. M. and Tobacman, J. (2009). Payday loans and credit cards: New liquidity and credit scoring puzzles? *American Economic Review*, **99** (2).

Altunbas, Y., Evan, L. and Molyneux, P. (2001). Bank ownership and efficiency. *Journal of Money, Credit, and Banking*, **33** (4), 926–954.

Amel, D. F. and Hannan, T. (1999). Establishing banking market definitions through estimation of residual deposit supply equations. *Journal of Banking and Finance*, **23** (11), 1667–1690.

Angrist, J. and Pischke, J.-S. (2001). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, 2nd edn.

Ashraf, N., Karlan, D. and Yin, W. (2006). Tying odysseus to the mast: Evidence from a commitment savings product in the philippines. *The Quarterly Journal of Economics*, **121** (2), 635–672.

Atif Mian, K. R. and Sufi, A. (2013). Household balancesheets, consumption, and the economic slump. *The Quarterly Journal of Economics*, **128** (4), 1687–1726.

Avery, R. B. and Samolyk, K. A. (2004). Bank consolidation and small business lending: the role of community banks. *Journal of Financial Services Research*, **25** (2/3), 291–325.

Azar, J., Raina, S. and Schmalz, M. C. (2016). Ultimate ownership and bank competition, working Paper.

Barr, M. S. (2004). Banking the poor, brookings Institution Research Brief.

— and Blank, R. M. (2008). Access to financial services, savings, and assets among the poor. *National Poverty Center Policy Brief*, **13**.

—, Dokko, J. K. and Feit, E. M. (2011). *Preferences for banking and payment services among low- and moderate-income households*. Finance and Economics Discussion Series 2011-13, Board of Governors of the Federal Reserve System (U.S.).

Berger, A., Saunders, A., Scalise, J. M. and Udell, G. (1998). The effects of bank mergers and acquisitions on small business lending. *Journal of Financial Economics*, **50** (2), 187–229.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G. and Stein, J. C. (2005). Does function follow organizational form? evidence from the lending practices of large and small banks. *Journal of Financial Economics*, **76** (2), 237–269.

Bertrand, M. and Morse, A. (2011). Information disclosure, cognitive biases, and payday borrowing. *Journal of Finance*, **66** (6), 1865–1893.

Bhutta, N. (2014). Payday loans and consumer financial health. *Journal of Banking and Finance*, **47** (1), 230–241.

Blackburn, M. L. and Vermilyea, T. (2009). The prevalence and impact of misstated incomes on mortgage loan applications, working Paper.

Board of Governors of the Federal Reserve System (2003). *Annual Report to the Congress on Retail Fees and Services of Depository Institutions*. Tech. rep., Federal Reserve System.

Bord, V. M., Ivashina, V. and Taliaferro, R. (2017). Large banks and small firm lending, working Paper.

Brian Chen, S. H. and Stein, J. (2017). The decline of big-bank lending to small business: Dynamic impacts on local credit and labor markets, nBER Working Paper.

Burgess, R. and Pande, R. (2005). Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review*, **95** (3), 780–795.

Celerier, C. and Matray, A. (2017). Bank branch supply and the unbanked phenomenon, working Paper.

Cetorelli, N. and Goldberg, L. (2012). Banking globalization and monetary transmission. *Journal of Finance*, **67** (5), 1811–1843.

Chatterji, A., Luo, J. and Seamans, R. (2013). Competition between organizational forms: Banks vs. credit unions after the financial crisis, working Paper.

Chava, S. and Purnanandam, A. (2011). The effect of banking crisis on bank-dependent borrowers. *Journal of Financial Economics*, **99** (1), 116 – 135.

Consulting, D. (2000). *Survey of Non-bank financial institutions*. Tech. rep., U.S. Department of the Treasury.

Cororaton, A. (2017). The impact of firm objectives on lending: Banks and credit unions in the great recession, working Paper.

Crump, R., Hotz, J., Imbens, G. and Mitnik, O. (2009). Dealing with limited overlap in estimation of average treatment effects. *Biometrika*, **96** (1), 187–199.

Darmouni, O. (2016). Estimating informational frictions in sticky relationships.

Dell'Ariccia, G. and Marquez, R. (2004). Information and bank credit allocation. *Journal of Finance*, **61** (5), 2511–2546.

DeLong, G. L. (2001). Stockholder gains from focusing versus diversifying bank mergers. *Journal of Financial Economics*, **59** (2), 221–252.

Desmond, M. and Shollenberger, T. (2015). Forced displacement from rental housing: Prevalence and neighborhood consequences. *Demography*, **52**, 1751–1772.

DeYoung, R., Evanoff, D. D. and Molyneux, P. (2009). Mergers and acquisitions of financial institutions: A review of the post-2000 literature. *Journal of Financial Services Research*, **36** (2), 87–110.

— and Rice, T. (2004). How do banks make money? a variety of different business strategies. *Economic Perspectives*, **52**, 52–68.

Erel, I. (2011). The effect of bank mergers on loan prices: Evidence from the united states. *Review of Financial Studies*, **24** (4), 1068–1101.

Esty, B. (1997). Organizational form and risk taking in the savings and loan industry. *Journal of Financial Economics*, **44** (1), 25–55.

Favilukis, J., Ludvigson, S. C. and Nieuwerburgh, S. V. (2016). The macroeconomic effects of housing wealth, housing finance, and limited risk sharing in general equilibrium. *Journal of Political Economy*, **125** (1), 140–223.

Federal Deposit Insurance Corporation (2015). *FDIC National Survey of Unbanked and Underbanked Households*. Tech. rep., Federal Deposit Insurance Corporation.

Foote, C. L., Gerardi, K. S. and Willen, P. S. (2012). Why did so many people make so many ex post bad decisions? the causes of the foreclosure crisis, federal Reserve Bank of Bosston Public Policy Discussion Paper.

Fried, H. O., Lovell, C. and Eeckaut, P. V. (1993). Evaluating the performance of us credit unions. *Journal of Political Economy*, **17** (2-3), 251–265.

Garmaise, M. J. and Moskowitz, T. J. (2006). Bank mergers and crime: The real and social effects of credit market competition. *Journal of Finance*, **61** (2), 495–538.

Gilje, E. P., Loutskina, E. and Strahan, P. E. (2016). Exporting liquidity: Branch banking and financial integration. *Journal of Finance*, **71** (3), 1159–1184.

Good, B. A. (1999). Bringing the unbanked onboard. *Federal Reserve Bank of Cleveland Economic Commentary*.

Granja, J., Matvos, G. and Seru, A. (2017). Selling failed banks. *Journal of Finance*, **72** (4), 1723–1784.

Greenstone, M., Mas, A. and Nguyen, H.-L. (2015). Do credit market shocks affect the real economy? quasi-experimental evidence from the great recession and 'normal' economic times, working Paper.

Hannan, T. (2006). Retail deposit fees and multimarket banking. *Journal of Banking and Finance*, **30** (9), 2561–2578.

— and Prager, R. A. (2006). Multimarket bank pricing: An empirical investigation of deposit interest rates. *Journal of Economics and Business*, **58** (3), 256–272.

Itamar Drechsler, A. S. and Schnabl, P. (2017). The deposit channel of monetary policy. *Quarterly Journal of Economics*, **132** (4), 1819–1876.

Ivashina, V. and Scharfstein, D. (2010). Bank lending during the financial crisis of 2008. *Journal of Financial Economics*, **97** (3), 319–338.

Jayaratne, J. and Strahan, P. (1997). The benefits of bank branch deregulation, new York Federal Reserve Economic Policy Review.

Jensen, M. C. and H.Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, **3** (4), 305–360.

Kay, B. S., Manuszak, M. D. and Vojtech, C. M. (2014). Bank profitability and debit card interchange regulation: Bank responses to the durbin amendment, fEDS Working Paper.

Khawaja, A. I. and Mian, A. (2008). Tracing the impact of bank liquidity shocks: Evidence from an emerging market. *American Economic Review*, **98** (4), 1413–42.

Kiser, E. (2002). Predicting household switching behavior and switching costs at depository institutions. *Review of Industrial Organization*, **20** (4), 349–365.

Kishan, R. P. and Opiela, T. (2000). Bank size, bank capital, and the bank lending channel. *Journal of Money, Credit and Banking*, **32** (1), 121–41.

Kovner, A., Vickery, J. and Zhou, L. (2014). Do big banks have lower operating costs? *Economic Policy Review*, (Dec), 1–27.

Krusell, P., Mukoyama, T., Sahin, A. and Anthony A. Smith, J. (2009). Revisiting the Welfare Effects of Eliminating Business Cycles. *Review of Economic Dynamics*, **12** (3), 393–402.

Liberti, J. M. and Mian, A. R. (2009). Estimating the effect of hierarchies on information use. *Review of Financial Studies*, **22** (10), 4057–4090.

Lusardi, A., Schneider, D. and Tufano, P. (2011). Financially fragile households: Evidence and implications. *Brookings Papers on Economic Activity,*, **42** (1), 83–150.

Mayer, C. and Pence, K. (2008). Subprime mortgages: What, where, and to whom?, nBER Working Paper.

Melzer, B. (2011). The real costs of credit access: Evidence from the payday lending market. *Quarterly Journal of Economics*, **126** (1), 517–555.

— and Morgan, D. P. (2015). Competition in a consumer loan market: Payday loans and overdraft credit. *Journal of Financial Intermediation*, **24** (1), 25–44.

Mester, L., Nakamura, L. and Renault, M. (2007). Transactions accounts and loan monitoring. *Review of Financial Studies*, **20** (3), 529–556.

Mian, A. and Sufi, A. (2009). The consequences of mortgage credit expansion: Evidence from the u.s. mortgage default crisis. *Quarterly Journal of Economics*, **124** (4), 1449–1496.

— and — (2016). Household debt and defaults from 2000 to 2010:the credit supply view, working Paper.

— and — (2017). Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. *Review of Financial Studies*, **30** (6), 1832–1864.

Mincer, J. (1991). *Education and Unemployment*. Nber working papers.

Mukoyama, T. and Sahin, A. (2006). Costs of business cycles for unskilled workers. *Journal of Monetary Economics*, **53** (8), 2179 – 2193.

Nguyen, H.-L. (2017). Do bank branches still matter? the effect of closings on local economic outcomes, working Paper.

of Governors fo the Federal Reserve System, B. (2017). *Report on the Economic Well-Being of U.S. Households in 2016*. Tech. rep., Board of Governors of the Federal Reserve System.

Park, K. and Pennacchi, G. (2009). Harming depositors and helping borrowers: The disparate impact of bank consolidation. *Review of Financial Studies*, **22** (1), 1–40.

Peek, J. and Rosengren, E. (1997). The international transmission of financial shocks: The case of japan. *American Economic Review*, **87** (4), 495–505.

— and — (1998). Bank consolidation and small business lending: It's not just bank size that matters. *Journal of Banking and Finance*, **22** (6-8), 799–819.

— and — (2000). Collateral damage: Effects of the japanese bank crisis on real activity in the united states. *American Economic Review*, **90** (1), 30–45.

Petersen, M. and Rajan, R. (2002). Does distance still matter? the information revolution in small business lending. *American Economic Review*, **57** (6), 2533–2570.

Prager, R. A. and Hannan, T. (1998). Do substantial horizontal mergers generate significant price effects? evidence from the banking industry. *Journal of Industrial Economics*, **46** (4), 433–52.

Prina, S. (2015). Banking the poor via savings accounts: Evidence from a field experiment. *Journal of Development Economics*, **115** (16), 31.

Rasmusen, E. (1988). Mutual banks and stock banks. *Journal of Law and Economics*, **31** (2), 395–421.

Rhine, S. L., Greene, W. H. and Toussaint-Comeau, M. (2006). The importance of check-cashing businesses to the unbanked: Racial/ethnic differences. *Review of Economics and Statistics*, **88** (1), 146–157.

Rice, T. and Strahan, P. (2010). Does credit competition affect small-firm finance? *Journal of Finance*, **65** (3), 861–889.

Schmall, T. and Wolkowitz, E. (2016). 2016 financially underserved market size study. Center for FInancial Services Innovation.

Schnabl, P. (2012). The international transmission of bank liquidity shocks: Evidence from an emerging market. *Journal of Finance*, **67** (3), 897–932.

Servon, L. J. (2013). The high cost, for the poor, of using a bank. *The New Yorker*.

Smith, D. (2012). Commercial lending during the crisis: Credit unions vs banks, filene Research Interest Working Paper.

Stavins, J. (1999). Checking accounts: What do banks offer and what do consumers value. *New England Economic Review*, **March/April**, 3–13.

Stein, J. and Kashyap, A. (2000). What do a million observations on banks say about the transmission of monetary policy? *American Economic Review*, **90** (3), 407–428.

Stein, J. C. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *Journal of Finance*, **57** (5), 1891–1921.

Strahan, P. (2003). The real effects of us banking deregulation, federal Reserve Bank of St. Louis Review.

Wheelock, D. C. and Wilson, P. W. (2012). Do large banks have lower costs? new estimates of returns to scale for u.s. banks. *Journal of Money, Credit, and Banking*, **44** (1), 171–199.

Williams, E. (2017). Monetary policy and the funding structure of banks, working Paper.

Wooldridge, J. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 1st edn.

Appendix A

Appendix to Chapter 1

A.1 Datasets and Data Creation

Deposits: Creating a Branch-level Dataset Consistent over Time

For the purposes of my analysis, it is important to have a panel dataset of branches and follow each branch over time. Although the FDIC summary of Deposits (SOD) dataset provides the variable uninumbr, which is meant to create a consistent time-series for each branch, this variable has several drawbacks. First, it is not defined for most branches of thrifts, prior to 2011. Second, there are instances of 2 branches within the same bank "swapping" uninumbrs. For example, from 2003-2008, uninumbr 13448 corresponds to a branch located at 901 North Boll Weevil Circle in Enterprise, Alabama, with deposits of $120-$140 million. Uninumbr 249140 corresponds to a branch at 3680 West Main Street in Dothan, Alabama, with deposits of approximately $50 million. However, in 2009, the branches swap places and from 2009-2016, uninumbr 13448 corresponds to a branch at 3680 West Main Street in Dothan, Alabama, with deposits of approximately $50 million and uninumbr 249140 corresponds to a branch in Enterprise, Alabama with deposits of $120-$140 million.

Third, the uninumbr seems to correspond to a bank's office rather than a fixed geographic location. Offices may move around and switch locations with other offices, especially immediately after mergers. There are many instances of an office (uninumbr) of the target bank closing and an office of the acquiring bank moving to that same location. For the purposes of my paper, however, and for the local population who have accounts at this branch, the closing is irrelevant, since the branch remains open and accessible to depositors.

Most papers either aggregate the SOD data to the county level (e.g. (Bord *et al.*, 2017)) or use the SOD data to identify branches at a single point in time (e.g. (Nguyen, 2017)). For these papers, most inconsistencies regarding the time series of each branch are irrelevant. However, because I use branch-level deposit and fee data and track each branch over time, these inconsistencies may introduce much more noise into the analysis.

To adjust for these differences between what uninumbr tracks and what is required for my paper, I use the SOD data, and augment it with both an algorithm that matches branches based on location and data from SNL Financial, which also has data on bank branches and has its own internal identifier. The algorithm matches branches 1) first based on address within the same identifier (first Federal Reserve RSSD, then FDIC cert), and 2) then based on address accounting for bank mergers. I use fuzzy string matching to account for typos and changes in the address, as well as differences in zip code definitions over time. I check this algorithm and the original uninumbr against the identifier snlbranchkey, the internal SNL identifier for bank branches. Since there is always the possibility of both type I and type II error, for the analysis in this paper, I only keep branch matches which have a very high confidence of being correct.

The final outcome is a dataset that is better lined up than either the SOD or the SNL branch datasets. First, whereas the SOD dataset contains approximately 131,600 unique uninumbrs, they correspond to 128,600 unique branches in my dataset. Second, yearly deposit growth is less noisy in my dataset, with a standard deviation of 0.42 rather than 0.44 and an interquartile range of [-0.30, 0.128], rather than [-0.032, 0.131] for the SOD dataset.

Check-Cashing Outlets: County Business Patterns and Infogroup

The County Business Patterns (CBP) dataset contains information on the number of establishments by zip code and year for each six digit NAICS category. The NAICS category 522390 comprises "establishments primarily engaged in facilitating credit intermediation (except mortgage and loan brokerage; and financial transactions processing, reserve, and clearinghouse activities)." Prior papers have used this NAICS category to identify payday lenders (Bhutta, 2014; Melzer, 2011), the category also includes check cashing facilities. Many check-cashing facilities engage in payday lending, and vice versa, but the two types of activities are distinct and serve as substitutes for two different types of bank services. Payday lending is a substitute for bank consumer lending, whereas check cashing is a substitute for bank deposit account services. For this paper, it is important to disentangle the two types of establishments since I focus on depositors and their demand for deposit account alternatives, not credit alternatives. In addition, having a bank account is a pre-requisite for most types of payday lending.[1]

To disentangle these two types of establishments, I turn to data from Infogroup. Infogroup collects and verifies establishment location data from thousands of yellow and white pages books around the country. For each establishment, it reports the address and name, as well as the detailed NAICS and SIC codes the establishment falls under. I identify check-cashing outlets as those in SIC code 609903, “Check Cashing Services," as well as those that have both “Check" and “Cash" in their names. I identify payday lenders as those in SIC code 614113, “Payday Loans", as well as establishments that have the word “Cash" in their name, but are neither pawnshops, gold stores, nor check-cashing facilities. The zip code level number of payday stores from CBP and from Infogroup have a correlation of 0.8.

[1]Most banks do not engage in small value consumer lending, instead satisfying demand for these loans by extending credit cards. Since large banks are more likely to issue credit cards, demand for payday lending should remain constant or decrease after mergers involving large banks. It is unlikely the number of payday lenders would decrease.

A.2 Cross-Sectional Results from the FDIC Survey

In this section, I use the FDIC Survey to examine the correlation between the presence of large banks and the prevalence of unbanked households. Because the FDIC survey contains geographical data only at the MSA level, I am not able to address issues of causality using the survey data. Instead, I show cross-sectional correlations that are suggestive evidence of a link between the presence of large banks and the prevalence of unbanked households.

FDIC Survey Data

The FDIC's National Survey of Unbanked and Underbanked Households (FDIC Survey) dataset is a nationally representative survey that contains information on households' banking status and use of altenative financial services. The survey is conducted as part of the Consumer Population Survey (CPS) by the census, and was conducted 4 times—in January 2009, June 2011, June 2013, and June 2015. Approximately 47,000 individuals filled out the survey each time, answering questions about whether they have a bank account, whether they use alternative financial services (AFS) and what types, and various socio-economic questions. The detailed, individual-level information on banking status and AFS use is the main advantage of the data. The main limitation of these data is that the geographic identifier for the household's location is available only at the MSA level in the public data. In addition, there are only 4 years of repeated cross-sectional data.

Methodology

Using data from the FDIC survey and FDIC's Summary of deposits, I regress an individual's banking status on the presence of large banks in the MSA and controls. Specifically, I run a regression of the form:

$$Y_{i,m,r,t} = \alpha + \beta \text{Large Bank Presence}_{m,r,t} + \gamma I_{i,m,r,t} + \delta M_{m,r,t} + \lambda_{r,t} + \epsilon_{i,m,r,t} \tag{A.1}$$

The dependent variable, $Y_{i,m,r,t}$, is either an indicator for whether the household i in MSA m in region r surveyed at time t has a bank account or an indicator for whether the household has ever used different types of alternative financial services (AFS), such as check-cashing facilities, money orders, or prepaid cards. The main variable of interest is Large Bank Presence$_{m,r,t}$, which is a measure of the presence of large banks in MSA m in region r at time t. $I_{i,m,r,t}$ are individual-level controls for individual i, and $M_{m,r,t}$ are MSA-level controls. I discuss the specific dependent and independent variables I use below. $\lambda_{r,t}$ are region-year fixed effects. The regions in the survey are Northeast, Midwest, South, and West and time periods are 2009, 2011, 2013 and 2015. Each observation is weighted by its household survey sampling weight, provided by the CPS, to account for the sampling methodology of the survey. Throughout this section, standard errors are clustered at the MSA level.

From the FDIC Survey, I include as household-level controls the following indicator variables corresponding to whether the household: lives in an urban part of the MSA, is black or Hispanic, foreign born, aged 65 or older, unemployed, a homeowner, married, single female head of household. In addition, since the Unbanked Survey and CPS include a limited number of household characteristics and because whether a household is unbanked may also influenced by other MSA-level factors (such as availability of AFS, the banking status and AFS use of the household's network), I also include MSA-level controls. From the 2000 census, I include the housing density (number of households per square mile), the log of total number of households, average family size, the log of median MSA income, and percentages of households that are: living in the urban part of MSA, black, Hispanic, living below the poverty line, aged 65 or older, unemployed, income less than \$10 thousand, and with income between \$10 and \$35 thousand.[2] I also include the yearly MSA deposit HHI and the MSA debt to income ratio as of 2006.[3] Most of these controls have the expected signs. As previous literature has found, households located in urban areas, minority households, and unemployed households, and households with a single female head of household are more likely to be unbanked. Older individuals and those that own their house are less likely to be unbanked.

[2] I use 2000 census data since my first of data is 2009. Using the 2010 data for the variables available from the 2010 Census does not change the results

[3] I compile the MSA-level debt to income data from the county data available on Amir Sufi's website.

Results

Table A.3 presents the results of the regression described above, showing the correlation between the presence of large banks and the prevalence of unbanked households. In column 1, the dependent variable is an indicator for whether the household has a bank account. My main variable of interest is the presence of large banks in the MSA, which is calculated as the ratio of the branches in the MSA that belong to large banks.[4] In column 2, I use as the dependent variable an indicator for whether the household has ever used any deposit alternative financial services; namely: check-cashing facilities, prepaid cards or money orders. Since the bank accounts and deposit AFS use are substitutes, households in areas with a higher presence of large banks are both more likely to be unbanked and more likely to use AFS. These results are both statistically and economically significant. Individuals in an MSA with a Large Bank Presence one standard deviation higher than the mean are approximately 0.5% more likely to be unbanked and 1.4% more likely to use AFS.

Comparison to Celerier and Matray (2017)

At first glance, these results are inconsistent with those of Celerier and Matray (2017) who argue that interstate branching deregulation has decreased, not increased, the percent of unbanked households. Yet the changes in the banking industry have resulted in two counteracting forces that impact the unbanked. On the one hand, as Celerier and Matray (2017) show, the increase in the number of bank branches has decreased the number of unbanked households. On the other hand, the consolidation that followed increased the proportion of large banks, which I argue increase the percent of unbanked households. To clarify this distinction, I follow Celerier and Matray and in column 3, I repeat the regression of column 1 using as the independent variable branch density, calculated as the number of branches per household in the MSA. Consistent with the results of Celerier and Matray, the coefficient is negative and significant, suggesting

[4] Although the existing literature often measure bank presence using share of deposits, share of branches is more applicable in this analysis since it is the existence of a branch, and not its size, that is relevant for a lower-income depositor's decision to open an account. Using a share of deposits would overestimate the presence of large banks because large banks' deposits are driven to a large extent by the large deposit accounts of firms and wealthy individual rather than retail deposits. Using the share of deposits produces qualitatively similar but statistically less significant results.

that in MSAs with more branches per household, households are less likely to be unbanked. In column 4, I include both branch density and the presence of large banks. Both variables maintain their signs from the previous columns. This suggests that the positive relationship between the existence of bank branches and the banked status that Celerier and Matray (2017) find is driven mostly by small banks. A higher presence of large banks, on the other hand, increases the percentage of unbanked households. Since large banks tend to have higher fees, I check in unreported results that MSAs with a higher presence of large banks also tend to have higher checking and savings fees and that individuals in MSAs with higher average fees are more likely to be unbanked.

A.3 Supplementary Tables and Figures

Table A.1: *Robustness: Checking Account Fees*

This table shows the results of a regression of equation 1.1 from Section 1.2, estimating the difference in deposit account fees and minimum balances between small banks with fewer than \$10 billion in assets and large banks with more than \$10 billion in assets. The sample is limited to counties in which the average large bank market share is less than the average small bank market share. The dependent variables in this panel are the annualized fee on checking accounts (column 1), the average minimum balance needed to avoid the fee (column 2), the annualized fee on interest checking accounts (column 3), and the minimum balance on interest checking accounts (column 4). Each observation corresponds to a bank-county-year triple and I include county-year fixed effects. $Large_{b,t}$ *is an indicator for whether the bank has more than \$10 billion in assets, in inflation-adjusted 2016 dollars.* $Large_{b,t} \times After2011_t$ *is the interaction between this indicator and an indicator for the 2011-2015 period. Standard errors are shown in parentheses and are clustered at the county level.*

Dependent Variable:	Regular Checking		Interest Checking	
	Fee	Min	Fee	Min
	(1)	(2)	(3)	(4)
$\text{Large}_{b,t}$	8.483***	113.813***	38.770***	1630.049***
	(2.787)	(22.355)	(2.898)	(154.077)
$\text{Large}_{b,t}$ x After2011_t	25.896***	335.886***	18.748***	1846.370***
	(2.673)	(26.292)	(3.152)	(190.690)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	17963	17099	18563	18316
Within R-squared	0.087	0.138	0.161	0.125

Table A.2: *What Drives Bank Account Fees?*

*This table runs a hedonic regression of checking account fees on account characteristics. The dependent variable is the annualized checking account fee. Large$_{b,t}$ is an indicator for whether the bank has more than \$10 billion in assets, in inflation-adjusted 2016 dollars. Large$_{b,t}\times$ After2011$_t$ is the interaction between this indicator and an indicator for the 2011-2015 period. Branch dispersion is calculated as 1 minus the within-county branch HHI of the bank, which proxies for the dispersion of the bank's branch network within the county. I{Branches in other Counties} is an indicator for whether the bank has branches in another county. I{Branches in other States} is an indicator for whether the bank has branches in another state. Number of Services is the total number of other services the bank provides out of the following: billpay, person to person payments, overdraft line of credit, overdraft protection, mobile banking, domestic wire transfers, and international wire transfers. Num Employees / Num Branches is a proxy for convenience and customer service. It is calculated as the number of full-time employees divided by the number of the bank's branches. Standard errors are shown in parentheses and are clustered at the county-year level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.*

Dependent Variable:	Checking Account Fee			
	(1)	(2)	(3)	(4)
Large$_{b,t}$	10.668***	4.943***	4.270***	3.811**
	(1.374)	(1.524)	(1.538)	(1.652)
Large$_{b,t}$ x After2011$_t$	25.292***	25.698***	24.981***	25.279***
	(1.206)	(1.199)	(1.212)	(1.290)
Branch Dispersion	2.971	2.509	2.584	3.063
	(2.282)	(2.254)	(2.252)	(2.244)
I{Branches in other Counties}		-3.678**	-4.222**	-3.718**
		(1.661)	(1.664)	(1.736)
I{Branches in other States}		8.781***	8.693***	9.101***
		(1.103)	(1.105)	(1.157)
Number of Services			1.027***	0.975***
			(0.199)	(0.202)
Num Employees/Num Branches				0.013
				(0.010)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Observations	73203	73203	73203	69396
Within R-squared	0.075	0.079	0.080	0.081

Table A.3: *Large Banks and the Unbanked - FDIC Survey*

*This table tests the relationship between a household being unbanked and the presence of large banks in the MSA. The dependent variable in columns 1, 3 and 4 is whether the household was unbanked. The dependent variable in column 2 is whether the household used deposit alternative financial services (AFS), namely check-cashing facilities, money orders, or prepaid cards. Large Bank Presence is the share of branches that are owned by large banks with more than $10 billion in inflation-adjusted assets. Branch Density is the number of bank branches divided by the total number of households. The number of households is as of the 2000 Census. I include region-year fixed effects and household and MSA controls. Household controls include indicators for whether the household is: in an urban MSA, black or Hispanic, foreign born, aged 65 or older, unemployed, a homeowner, married, and a single female head of household. MSA-level controls include housing density (number of households per square mile), log of total number of households, average family size, log of median income, and percentage of households that are: urban, black or hispanic, living below the poverty rate, unemployed, households aged 65 or older, with income less than $10 thousand, and with income between $10 and $35 thousand. Standard errors are shown in parentheses and are clustered at the MSA level. * p<0.10, ** p<0.05, *** p<0.01.*

Dependent Variable:	Unbanked	AFS Use	Unbanked	
	(1)	(2)	(3)	(4)
Large Bank Presence	0.021*	0.064**		0.015**
	(0.012)	(0.025)		(0.007)
Branch Density			-0.014***	-0.010**
			(0.004)	(0.005)
MSA Controls	Yes	Yes	Yes	Yes
Year-Region Fixed Effects	Yes	Yes	Yes	Yes
Observations	158459	155922	158459	158459
R-Squared	0.132	0.193	0.132	0.132

Table A.4: *Treated Banks Are Not More Likely to Be Sold*

This table tests whether branches of small banks are more likely to fail, be sold or be closed than branches of large banks. The sample is treated (branches of small banks acquired by a large bank) and control (branches of small banks acquired by other small banks) branches. Small banks are defined as those with less than $10 billion in assets and large banks are defined as those with more than $10 billion in assets. The coefficients are from a regression of subsequent branch events on county-year fixed effects and $Large_{b,t}$*, an indicator of whether the acquirer is a bank with more than $10 billion in assets. In column 1, the dependent variable is an indicator for whether the branch is closed before the end of the sample period. In column 2, it is an indicator for whether the branch is still open 5 years after the merger. In column 3, it is an indicator for whether the branch is sold as part of a bank merger. In column 4, the dependent variable is an indicator for whether the branch subsequently fails as part of a bank failure. In column 5, the dependent variable is an indicator for whether the branch is divested. In column 6, the dependent variable is an indicator for whether the branch subsequently moves locations. I restrict the sample to the first time a branch is involved in a merger during my sample period. Standard errors are shown in parentheses and are clustered at the county level. ** $p<0.10$*, *** $p<0.05$*, **** $p<0.01$*.*

Dependent Variable:	Indicator for Branch Event:					
	Closure	Existence in 5 Yrs	Sale	Failure	Divestiture	Move out of Zip Code
	(1)	(2)	(3)	(4)	(5)	(6)
Bought by $Large_b$	-0.013	0.004	-0.028	-0.001	0.016	-0.001
	(0.028)	(0.018)	(0.037)	(0.001)	(0.016)	(0.007)
County-Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10219	10219	10219	10219	10219	10219
R-squared	0.320	0.727	0.774	0.361	0.603	0.222

Table A.5: *Bank-Switching Behavior*

*This table tests whether deposit growth increases at unacquired branches in close geographic proximity to treated branches, consistent with some depositors going to other small banks, after the acquisition. The dependent variable is the zip code level deposit growth for different samples of branches. Column 1 uses as the dependent variable the average deposit growth at branches of other small banks in the same zip code as an acquisition. Column 2 uses deposit growth at branches of large banks in the same zip code as an acquisition. Columns 3 and 4 use branches of small and large banks, respectively, in zip codes that do not experience an acquisition, but are adjacent to ones that do. County-year fixed effects and zip code fixed effects are included in each regression. Standard errors are shown in parentheses and are clustered at the zip code level. * p<0.10, ** p<0.05, *** p<0.01.*

	Same Zip Code		Adjacent Zip Codes	
	Small	Large	Small	Large
Bought by Large x Post	0.008*	-0.006	-0.002	-0.007
	(0.005)	(0.005)	(0.004)	(0.006)
County-Year Fixed Effects	Yes	Yes	Yes	Yes
Zip Fixed Effects	Yes	Yes	Yes	Yes
Observations	88573	68350	440580	206460
Within R-squared	0.001	0.000	0.000	0.000

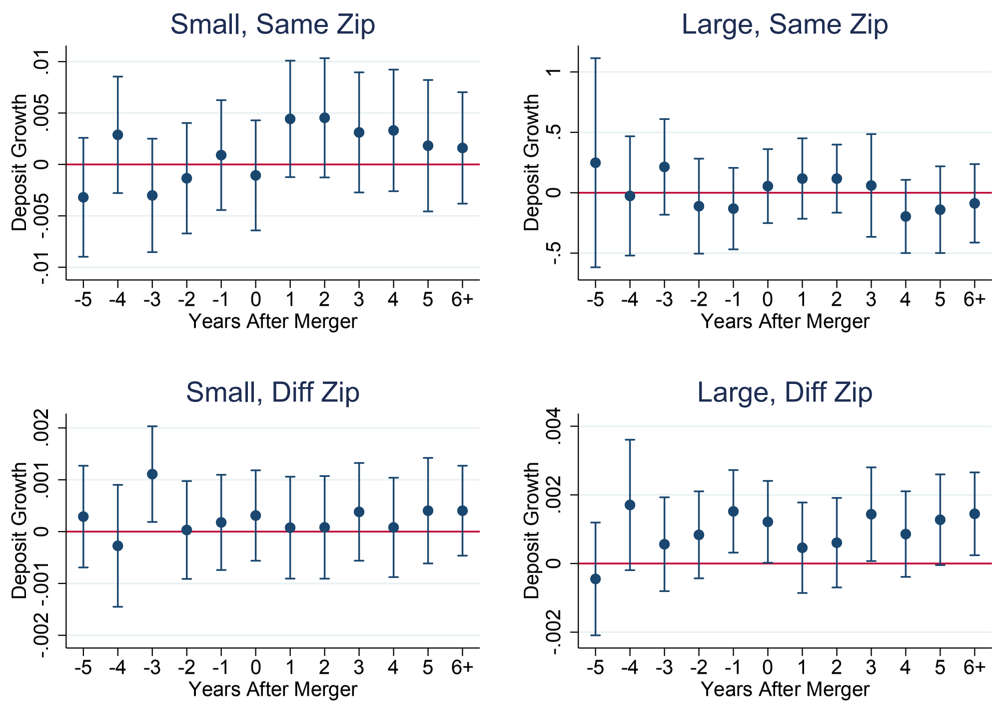


Figure A.1: *Deposit Growth at Other Branches*

This figure plots the yearly coefficients from a difference-in-difference regression estimating the effect of consolidation on deposit growth at unacquired branches. In the top row, the dependent variable is zip code-level deposit growth at branches of small banks in the same zip code (top left) and large banks in the same zip code (top right) as an acquisition. In the bottom row, the dependent variable is zip code-level deposit growth in adjacent zip codes to those that experienced an acquisition. The bottom left panel considers deposit growth at small banks and the bottom right deposit growth at large banks.The plot shows the coefficients on the interactions between treatment, whether the bank was acquired by a large bank, and indicators for each year after the merger. Dashed lines capture the 95% confidence intervals. Year 0 corresponds to June 30th prior to the merger.

Appendix B

Appendix to Chapter 2

B.1 Supplementary Tables and Figures

Table B.1: *Change in Deposit Rates, 2008 vs. 2006*

*This table compares the change in deposit rates from 2006 to 2008 for healthy banks and for exposed banks. Each observation corresponds to a bank and the dependent variables are the change in overall deposit growth from 2006-2008, the change in insured deposits and the change in uninsured deposits, respectively. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.*

Bank-level Growth in Deposits

Dependent Variable:	Change in:		
	Deposits	Insured Deposits	Uninsured Deposits
	(1)	(2)	(3)
Healthy bank G_i	0.049**	-0.002	0.022**
	(0.020)	(0.007)	(0.010)
Bank Controls	Yes	Yes	Yes
Observations	279	279	279
R-squared	0.108	0.352	0.117