The Geographic Diversification of U.S. Bank Lending: Should Banks Focus or Diversify?

Nada Mora[1]

Abstract

This study examines whether banks diversify lending when demographic variation in their home market affords them excess deposit funds. In particular, banks with a branch presence in areas with more seniors (as a fraction of the population) have access to a greater deposit base supplied by seniors, independent of local economic business activity and bank demand. As a result, this gives banks an exogenous motivation to diversify their lending portfolio outside their local area. This research finds support for this view based on annual data reported by U.S. banks since 1996 on their new small business lending. This study also finds that banks that geographically diversified lending pre-crisis (instrumented with seniors) performed better in the period beginning with the 2007 financial crisis and its aftermath.

Keywords: Banks, diversification, risk.

1. Introduction

A longstanding question in banking research is whether banks that diversify their business perform better or worse as a result. Diversification is thought to increase bank returns and lower bank risk because more diversified banks are better able to withstand specific adverse outcomes. Although the caveat applies that even diversified banks remain vulnerable to common shocks affecting all banks as was evident during the 2007-2009

[1] Lecturer, Olayan School of Business, American University of Beirut. Email Address: drnadamora@gmail.com.

Acknowledgments: I would like to thank two anonymous reviewers and participants from the Lebanese University's 2018 International Academic Conference on Economics and Business for their valuable comments.

financial crisis[2]. Many studies have assessed whether diversification improves bank performance. The results of the literature are inconclusive.

A significant number of papers find that diversified banks, in fact, perform worse. One reason is that when banks diversify, they move into areas in which they have less expertise or they grow larger and, therefore, less manageable. Another reason is that the decision to diversify is itself dependent on other factors. A bank may choose to diversify precisely because it is already performing poorly. For example, Campa and Kedia (2002) find evidence of self-selection in the diversification decision. They show that this endogeneity is reflected in a negative correlation between the firm's value and its choice to diversify. Therefore, existing research might not have perfectly isolated the independent effect of diversification.

To get around the endogeneity problem, this research takes advantage of the documented fact that cities with a larger fraction of seniors have plenty of deposit funds that are independent of local economic activity (Becker, 2007). Compared with other age groups, seniors (the population over 65 years in age) have been documented to hold a larger share of their asset portfolio in bank deposits according to various Surveys of Consumer Finances (SCF) conducted by the Federal Reserve (Bricker, Dettling, Henriques, Hsu, Jacobs, Moore, Pack, Sabelhaus, Thompson, & Windle, 2017[3].) For example, Becker (2007) cites the 2001 SCF showing that more than 20% of seniors hold certificates of deposits, more than other age groups. By the 2016 SCF, while all age groups hold fewer certificates of deposit, 12.5% of seniors continue to hold compared with 4.6% of the population under 65 years. Becker (2007) also shows that seniors hold higher absolute levels and this is also observed in the 2016 SCF (e.g., the median holding of certificates of deposit in thousands of 2016 dollars for the age group below 35 was 4, for 65-74 was 25, and for 75 and over was 40).

Becker then shows that the fraction of seniors in a Metropolitan

[2] Interconnectedness between financial institutions that are partly the result of diversification also amplifies systemic risk. As a result of this externality, Ibragimov, Jaffee, and Walden (2011) show that the optimal outcome for society is one with reduced risk-sharing.

[3] For more detail on asset holdings by different age groups, refer to the SCF Chartbook (2016).

Statistical Area (MSA) has a positive effect on overall deposits per capita but does not affect local demand for loans and deposits. In the second stage, he shows that this instrumented deposit supply (and therefore local loan supply) has a positive and significant effect on local economic activity in the MSA. Local area deposit funding is an important driver of lending to local bank-dependent borrowers including small firms and is significantly associated with more business startups in the local area.

The contribution and starting point of this research is to conjecture that banks in metropolitan areas with more seniors supplying steady deposit funds will diversify their lending outside the local area, more so than banks in cities with a limited senior population. After exhausting local businesses and startups, it is expected that banks seek viable lending opportunities outside their local area. Therefore, the contribution of this paper falls into two parts: (1) First, it is tested whether banks in geographic markets characterized by ample deposit funds diversify their loan portfolio. And (2) This demographically driven lending diversification is related to bank performance. To do so, this research takes advantage of the time period surrounding the 2007-2009 financial crisis to relate pre-crisis lending diversification to crisis and post-crisis bank risk measures. One advantage of an exogenously driven lending diversification motive is that this research can shed light on whether diversified banks are more stable or not, which is ambiguous, both in theory and empirically.

The main findings are: (1) Banks significantly diversify their new lending portfolio in response to the greater deposit supply afforded by a savings-rich senior population. And importantly, (2) this geographic diversification outside the local area helps insulate banks. This is consistent with the prediction of standard diversification theory. Post-crisis, bank Return on Assets (ROA) was higher, standard deviation of ROA was lower, and banks experienced less nonperforming loans when they had diversified pre-crisis. This effect is isolated by instrumenting lending diversification with the senior share of the population in the bank's branch markets.

The rest of this research is organized as follows: In Section 2, the existing literature on bank diversification is reviewed and common findings to inform this research framework are extracted. In Section 3, the data used and the research methods applied are outlined. The results are discussed in Section 4 and in Section 5 concluding notes are presented.

2. Literature review

In this section, the potential positive and negative effects of diversification on bank performance and what has been identified in the literature are reviewed. The literature's findings are made difficult by the fact that there are different causes for why banks might choose to either diversify or to focus activity.

In the traditional risk-sharing theory, more diversified banks should perform better. For example, they are less prone to failure due to an idiosyncratic shock to one of their borrowers because their lending portfolio is less lumpy. This argument is general and applies to all types of business diversification by banks – whether lending diversification (number of borrowers, sectoral, or geographic) as well as other types of diversification (asset and income) that are reviewed in the literature below.

But even the theory is ambiguous. A valuable role that banks provide is monitoring loans. On the one hand, Diamond (1984) shows that banks should be as diversified as possible and best economize on the fixed costs of monitoring. Monitoring the bank is not necessary as long as the bank (the delegated monitor) is adequately diversified. On the other hand, banks might be worse at conducting due diligence, screening, and monitoring a borrower if information problems increase with diversification and expansion. Diamond (1984) abstracts from this and simply models diversification as an increase in the number of entrepreneurs in the financial intermediary's portfolio.

Information problems can be significant when the borrower is not geographically near or is from a different industry than the bank's existing portfolio of borrowers[4]. Distance continues to matter despite the increasing availability and ease of collecting hard information on borrowers through credit reports. This technological change has made soft information less valuable than in the past but still significant – in particular for small

[4] In many theoretical papers such as Winton (1999), the treatment covers both lender diversification across different geographic regions and across different industry sectors. And in other papers such as Rajan, Servaes and Zingales (2000), the information problem stems from agency problems within the divisions of the financial conglomerate as it expands and engages in multiple activities.

business lending. The (continued) importance of distance is shown by Petersen and Rajan (2002) and Brevoort and Hannan (2006). Winton (1999) illustrates how monitoring incentives can be weakened when banks diversify. For example, when the downside risk is large, a diversified bank reduces its standard deviation of returns by reducing the likelihood of left tail events. This might induce it to reduce monitoring because of monitoring costs. This in turn leads to lower returns. Moreover, learning costs and other information problems impede frictionless expansion into new regions and new sectors so it might be better for a bank to be specialized than diversified. Internal frictions between divisions in diversified firms can also cause resources to flow to less efficient divisions, lowering overall firm value (Rajan et al., 2000). The empirical evidence reviewed next is also ambiguous in its conclusions. The empirical review covers different types of diversification.

2.1 The adverse evidence

Many empirical papers find results supportive of the negative view of diversification outlined in the theory. For example, DeLong (2001) shows that bank mergers that concentrate both activity and geography increase stockholder value. In an assessment of the financial crisis, the Federal Deposit Insurance Corporation (FDIC) (2010) documented that banks with above-average out-of-territory lending were also associated with higher nonperforming loans and greater risk overall. Loutskina and Strahan (2011) find that mortgage lenders with concentrated mortgage loans in a few local markets experienced higher profits and equity returns than diversified lenders during 2007-2008. Therefore, these papers conclude that geographic diversification has adverse consequences.

Other papers examine effects of broader asset and income diversification. Laeven and Levine (2007) use Bankscope data across countries and time to find that Tobin's q for a diversified bank is materially lower than it otherwise would be if the same bank were split into separate banks, each specialized in one activity. To measure diversification, Laeven and Levine introduce two distinct measures: (i) "Asset Diversity", defined as [1 – (net loans – other earning assets)/total earning assets]; and (ii) "Income Diversity", defined as [1 – (net interest income – other operating

income)/total operating income]. Higher values indicate higher diversification of the bank's business model.

The latter income-based measure is similar to one introduced by Stiroh (2004), who also finds that diversified banks perform worse. In several research papers, Stiroh (2004) and Stiroh and Rumble (2006) find that a greater share of nontraditional income stemming from noninterest income activities such as fees and trading did not reduce the volatility of bank net operating income. Indeed, increasing reliance on noninterest income (or a Herfindahl index (HHI) measure of income diversification) is associated with lower risk-adjusted returns.

Additional evidence of adverse effects stemming from asset diversity is provided by Goetz, Laeven, and Levine (2013) examining U.S. banks in the pre-crisis period. In addition to controlling for diversity measures applied previously by Laeven and Levine (2007), they examine the geographic diversification of bank holding company assets based on the location and assets of each of the subsidiaries of the bank (a 1 – HHI assets by state measure). To address the endogeneity problem highlighted in the introduction, they rely on the exogenous timing of interstate bank deregulation together with a gravity model borrowed from the international trade literature to instrument a bank's expansion choices and thus the geographic diversification of its assets. Interestingly, even this instrumented measure is associated with lower market values and other problems for the bank such as increased insider loans. An earlier paper by Acharya, Hasan, and Saunders (2006) takes advantage of the finer disaggregation of loans reported by Italian banks for different industries and sectors, compared with only broad loan categories reported in the regulatory reports of U.S. banks such as total commercial and industrial (C&I) lending. They also find that various HHI-type loan diversity measures are associated with lower returns and greater risk, as reflected in higher Non-Performing Loans (NPL) and standard deviation of returns[5].

[5] Albeit they caution that the results for Italy might not fully generalize because of high volatility in the economy during the 1990s as well as the dominance of state banks in the Italian context.

2.2 The positive evidence

So far, the upshot of all the literature reviewed is that banks should proceed with caution when diversifying their activities, whether income-based, loan-based, broad asset-based, or through expanding the geographic headquarters of the different entities within a banking organization.

Nonetheless, there is one smaller strand of literature that finds favorable evidence for bank diversification on bank risk outcomes. What is especially interesting about this literature is that the diversification benefits arrive from the geographic diversification of bank funding sources, and in particular, bank deposit funding. One of the earlier papers is by Deng and Elyasiani (2008) who show that the geographic diversification of a bank's branch deposits at either the state- or MSA-level is associated with increased value and lower risk. They exploit the annual mid-year snapshot of branch deposit distribution reported in the FDIC's Summary of Deposits (SOD).

Several recent papers add to the supportive evidence for geographic diversification of deposits and are also based on U.S. bank SOD data (Goetz, Laeven, & Levine, 2016; Meslier, Morgan, Samolyk, & Tarazi, 2016; Levine, Lin, & Xie, 2016). What is especially interesting about Goetz et al.'s (2016) finding is that it runs counter to their earlier (2013) study even though they use exogenous inter- and intrastate bank deregulation and a gravity model to instrument their diversification measures in both papers (and the analysis in both papers is to the pre-crisis period). While they do not try to reconcile the different implications of their work, it appears that there is something special about the diversification of deposit funding across geographies, and this advantage is not gained when banks try to diversify their assets. Similarly, Meslier et al. (2016) find a positive effect of diversification with some nuance regarding the benefits, according to different asset sizes and whether diversification is measured at the Metropolitan Statistical Area (MSA)-county level or at the state level[6]. While these studies examine the effect of geographic diversification on bank performance and riskiness, Levine et al. (2016) instead examine how

[6] Relatedly, this sensitivity of the results to the definition of the geographic area might be a second reason for the difference between the two studies by Goetz et al. (2013 and 2016). The HHI diversification measure is across MSAs in their more recent paper while it is across wider areas (states) in their earlier paper.

geographic diversification affects bank-funding costs. They find that funding costs are reduced in more diversified banking organizations.

The contribution of this research to the literature is to tie together the two strands: deposit funding on the liability side with lending diversification on the asset side.

3. Data and method

This research makes use of bank-level and census data to assess whether banks in metropolitan areas with more seniors diversify lending. To evaluate this question, data from the 2000 census is combined with data from the Community Reinvestment Act (CRA). Since 1996, the CRA mandates that reporting banks provide a detailed distribution of their (small business) lending in different geographic areas. Small business loans are loans in the amount less than $1 million and banks report both originations and purchases of these loans in a calendar year. Reporting institutions required to provide detailed loan information are the larger banks, which is a downside of the data (e.g., the most recent reporting cutoff in 2017 is $1.226 billion in assets). Nonetheless, the sample of reporting banks is sizeable (there were still 700 banking organizations in 2012 when the individual reporting banks are aggregated to the holding company). And up to 2004, the reporting net was cast wider with an asset cutoff of $250 million that was increased to $1 billion in 2005. The main results hold even on the earlier 1996-2004 sample[7].

Aside from studying the real effects of reporting disclosures, community and low-income development, or whether distance still matters (Brevoort & Hannan, 2006), the CRA data remains relatively unexploited. A related paper by Keeton (2009) does not directly examine diversification (or its

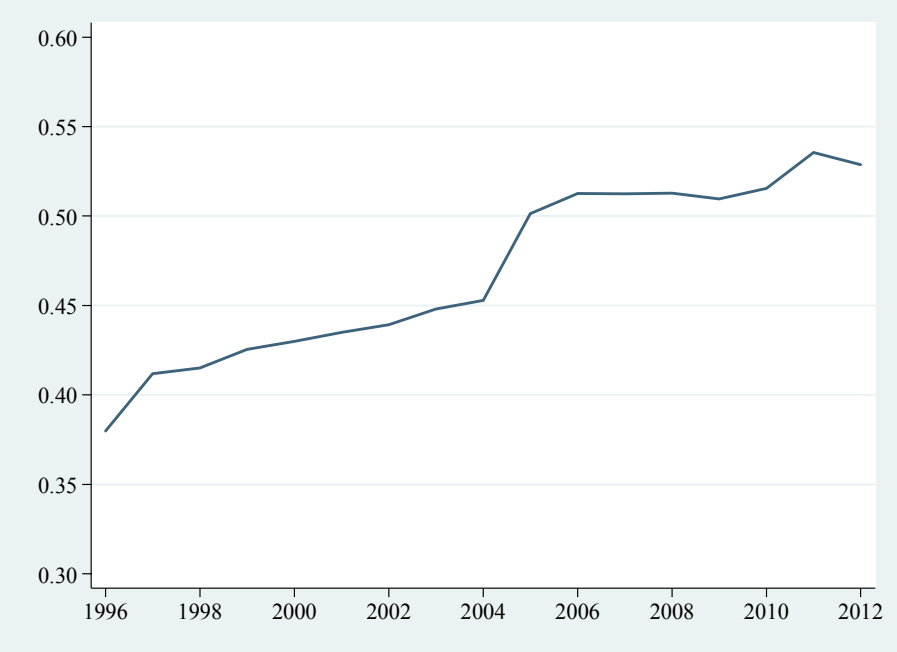
[7] Other studies appear to find real effects of CRA disclosure on banks, in that banks that became exempt from reporting in 2005 experience a decrease in their commercial NPL (Dou & Zou, 2017). They attribute this to public pressure on reporting banks (due to the form requesting not only geographic information but also requesting loans to geographic areas further disaggregated by income groups relative to the median income). If anything, however, this finding would suggest that the sample of reporting banks should experience worse outcomes for bank risk, in particular for those banks that feel pressured to geographically diversify their lending. That effect runs counter to the null hypothesis in this study.

exogenous drivers) but also uses the CRA data to assess whether multi-market banks experiencing slow growth in their local market due to the 2001 recession reduced loan originations at all locations. He finds that banks only reduced loan origination growth significantly at their primary banking office.

From a theory point of view, the costs that banks might incur in learning about a new industry or sector when diversifying are expected to be higher than the costs that banks incur when specializing in the same industry but diversifying the location of the borrowers, they lend to outside their immediate local area. Together with the reduced cost in acquiring information about distant borrowers because of information technology innovation in the last two decades, this implies that geographic diversification of lending should have gone up since 1996. This hypothesis matches the data as shown in Figure 1. The geographic area is the MSA for urban areas and counties for rural areas. Each bank's lending across geographic areas is used to construct a 1 – HHI measure for lending diversification. As with the other measures constructed in the literature, such as asset diversity and income diversity, a higher value reflects higher diversification. The average across banks is plotted in Figure 1. It has gone up from 0.38 in 1996 to 0.53 by 2012 (with the caveat about the reporting cutoff change in 2005 discussed earlier).

In addition to Census and CRA data, this paper relies on the quarterly Call Reports to collect information on bank characteristics such as size, liquidity, and risk outcomes. The regressions also control for a bank's local economic conditions such as house price growth, income per capita, income growth, and population size. These data are sourced from the Federal Housing Finance Agency (FHFA) for house prices and the Bureau of Economic Analysis (BEA) for income data.

Figure 1: Average geographic diversification of lending over time



Source: Community Reinvestment Act (CRA) Data on total small business loan annual surveys, 1996-2012. Total small business loans are the sum of originations and purchases. Data are aggregated to the county level for each bank topholder from which a 1 - HHI for geographic diversification is constructed (range between 0 and 1). A higher value indicates a more diversified loan market. A distinct geographic area is a Metropolitan Statistical Area (MSA) for urban areas and a county for rural areas.

The main methodology is described as follows: The first set of regressions test whether seniors have a positive effect on geographic lending diversification. Since the lending diversification dependent variable is only available annually from the CRA, this regression is on a bank-year panel from 1996 to 2012. The dependent variable, LDIV_GEO (=1-HHI), is regressed on the fraction of seniors in the bank's deposit market, controlling for other measures of bank diversification (such as asset diversity) and bank characteristics (such as size and capital) based on year-end Call Report data

– all denoted by X, as well as local economic activity in the bank's market (such as income per capita) denoted by M[8]. The regression includes bank fixed effects to control for bank heterogeneity, year dummies to control for time heterogeneity, and the standard errors are clustered at the bank level.

$$LDIV\_GEO_{it} = \alpha_i + \beta Seniors_{it} + \gamma X_{it} + \phi M_{it} + \varepsilon_{it} \quad (1)$$

Hypothesis 1: $\beta > 0$. A related test confirms that a higher senior share has a positive effect on the total deposits per capita in the bank's market, which in turn (instrumented) is associated with higher lending diversification. The results are presented in Table 3 discussed in the next section.

The second stage tests how lending diversification affects bank risk outcomes. These regressions instrument LDIV_GEO with seniors due to the endogeneity of diversification choice with bank risk outcomes. To also ensure predetermined diversification, these regressions relate a bank's pre-crisis diversification (PRE_LDIV_GEO) to its post-crisis risk outcome (Y), where a higher value of Y denotes better performance and lower risk. Pre-crisis variables are averaged over the 5-year quarterly data 2001:Q4-2006:Q4 and post-crisis variables are averaged over the 5-year quarterly data from 2007:Q4-2012:Q4. This second stage regression is a cross-sectional regression with each observation corresponding to one bank. The standard errors are robust to heteroscedasticity.

$$Y_i = \alpha + \delta PRE\_LDIV\_GEO_i^{IV} + \gamma X_i + \phi M_i + \varepsilon_i \quad (2)$$

Hypothesis 2: $\delta > 0$. Common measures of outcomes tested in Table 4 include Non-Performing Loans (NPL), profitability (annualized Return on Assets (ROA)), the standard deviation of ROA, and the Z-score. The Z-score measures the number of standard deviations that a bank's rate of return on assets must fall to cause the insolvency of the bank. It is defined as the sum of average ROA and average capital-to-asset ratio divided by the

[8] To determine a bank's market, information on the location of the bank's (deposit) branches is taken from the Summary of Deposits. Each branch is assigned to a particular geographic area: an MSA if it is an urban market and a state-county if it is a rural market. Economic activity from the BEA, Census, and FHFA is assigned to each of the geographic areas. Then a branch-deposit-weighted measure of economic activity is computed for the overall banking organization at time t.

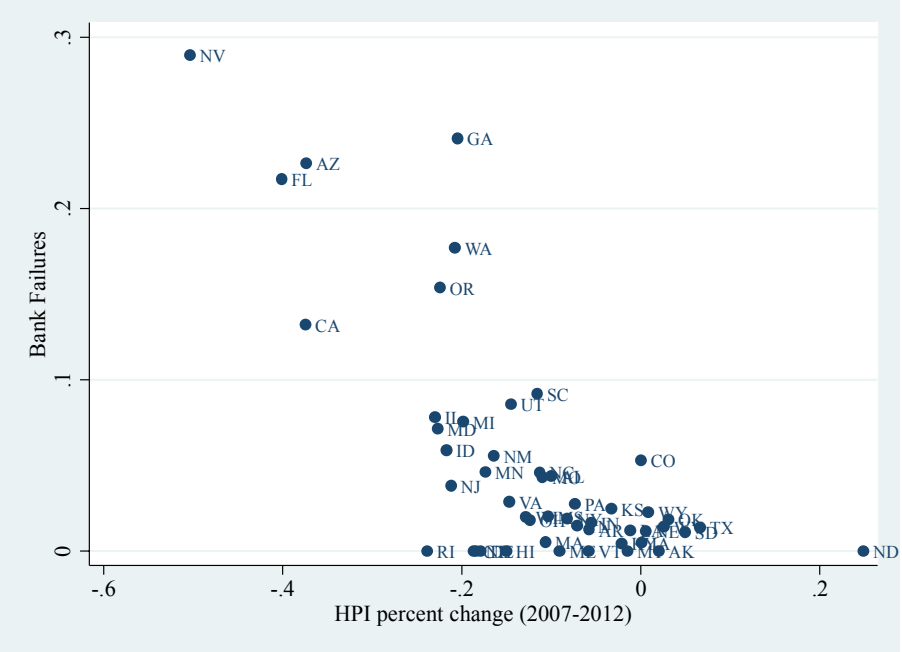
standard deviation of ROA computed over a period of rolling windows of 8 quarters.

Before turning to the formal regression analysis in the next section, Figure 2 motivates the hypothesis for a favorable effect from geographic lending diversification. It shows that bank failures (summing to the state level) in the crisis and post-crisis period are highly related to the house price growth in the state where the bank is located. For example, states such as Nevada, Arizona, Georgia, and Florida experienced greater declines in house prices and were also associated with the highest share of bank failures. This figure supports Aubuchon and Wheelock (2010) who examine the correlation between bank failures at the state-level and house prices, as well as the correlation of failures with other local economic activity measures. Their review can be distilled into the view that despite deregulation since the 1980s, banks continue to operate in a small number of markets. As a result, banks are especially exposed to economic shocks locally.

Approaching this question from a different perspective, Morgan, Rime and Strahan (2004) show that bank integration across states has reduced economic volatility in a state's economic growth. Figure 3 presents stylized evidence relating the fraction of seniors in a state's population to the state's bank failures. This figure shows a negative correlation between the two series, in line with the hypothesis that seniors supply relatively more local deposits that might be channeled into diversified out-of-area lending. Florida and Alaska are outliers for other reasons but the overall evidence is consistent with the Becker hypothesis and its extension to a diversification study in the next section[9].

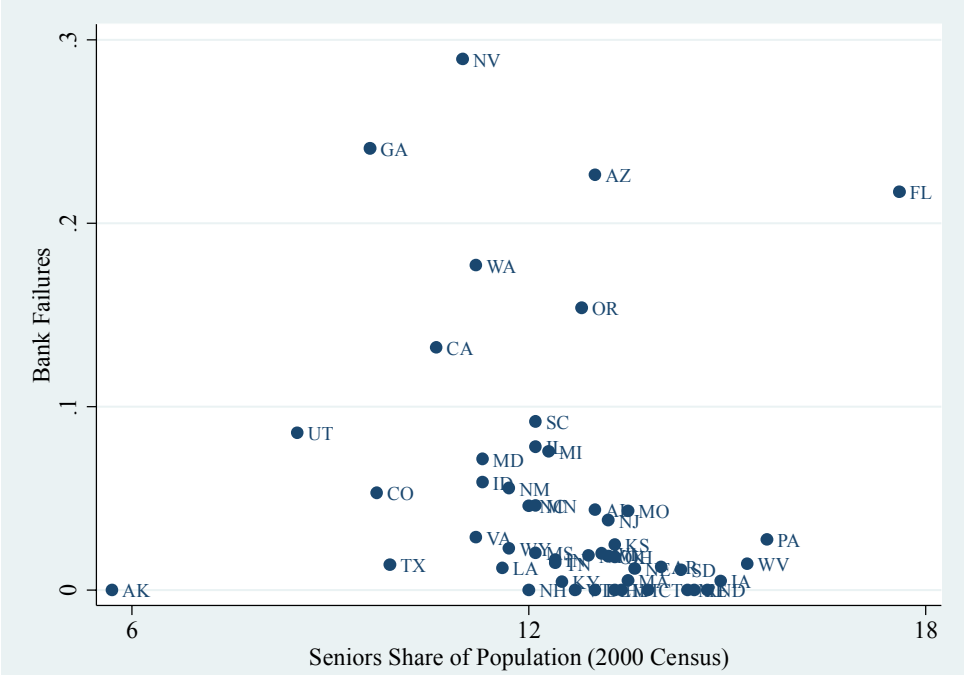
[9] Clearly this is meant to be illustrative only but it is worth noting that the significant negative correlation between seniors and bank failures holds even after conditioning on house price growth.

Figure 2: Bank failures and house price growth



Failures are failed banks and thrifts, 2007-2012 (From the FDIC failed bank list), normalized by the number of banks and thrifts in 2006 (from the Summary of Deposits). House prices are from the FHFA house price index quarterly data (for seasonally adjusted purchase-only data). House price growth is calculated from the peak in 2007:Q1 to 2012:Q4. The data are aggregated to the state level.

Figure 3: Bank failures and the fraction of seniors



Failures are failed banks and thrifts, 2007-2012 (From the FDIC failed bank list), normalized by the number of banks and thrifts in 2006 (from the Summary of Deposits). The data are aggregated to the state level.

4. Results

4.1 The drivers of geographic lending diversification

First, the first-stage regression in Becker's study is replicated. He relates the log of local deposits per capita to the fraction of seniors in an MSA, controlling for house prices, population size and income per capita. In the original study, there are different Census year slices (or pooled Census data). Becker notes however that the fraction of seniors is very stable over time. In this study, the 2000 Census is therefore used because it is closest to

the CRA sample period. The results in Table 1 are similar to those of Becker. Columns (1) and (2) are for MSAs and the 2.45 to 3.04 coefficient estimate on seniors is similar to the 2.57 Becker reports for a pooled 1990 and 2000 Census sample. This coefficient implies that a 0.025 increase in the fraction of seniors in a local area increases local deposits per capita by 0.06. To get a sense of the economic magnitude, see the summary statistics in Table 2. The mean senior share is 12.9% and a 0.025 increase is a 2.5 percentage point increase, which is equivalent to one standard deviation of the senior share. Relative to the standard deviation of the log of local deposits per capita (0.49), the 0.06 effect is economically significant. The economic effect is even greater when examining all geographic areas (AGA) that include rural areas where it makes sense that the local supply of deposits is even more sensitive to the fraction of seniors (columns (3) and (4)).

Table 1: Endorsing the Becker (2007) finding: Local deposits and the prevalence of seniors as a fraction of the population

The dependent variable is the log of deposits per capita

	(1)	(2)	(3)	(4)
	MSA	MSA	AGA	AGA
Senior share	3.038	2.448	4.813	3.807
	(4.56)***	(4.08)***	(17.93)***	(14.37)***
Log population		-0.077		-0.040
		(2.46)**		(4.78)***
Log House price index		-0.510		-0.282
		(1.71)*		(2.20)**
Log income per capita		1.293		1.119
		(5.71)***		(22.20)***
Constant	1.951	-7.687	1.653	-7.613
	(21.00)***	(4.40)***	(37.43)***	(10.74)***
Observations	318	309	2345	2308
Adjusted R-squared	0.04	0.19	0.16	0.35

This table shows the effect of seniors (as a share of population in the different geographic markets) on the log of total deposits per capita. The senior share is the share of the population that is 65 years or older (from the 2000 Census). The regressions control for the log of population, income per capita, and house prices in columns (2) and (4). Each observation is an MSA in columns (1) and (2) and a geographic area in columns (3) and (4), where a geographic area is an MSA for urban areas and a county for rural areas. The t-statistics are reported in parentheses. Standard errors are robust to heteroscedasticity and are clustered at the geographic level (MSA in columns (1) and (2)). ***, **, and * indicate 1%, 5%, and 10% significance respectively.

Table 2: Summary statistics and correlations

	Mean	Std. Dev.	25 %	Median	75 %	N
Summary Statistics for Regression Bank-Year Panel, 1996-2012						
Geographic loan diversification	0.462	0.314	0.153	0.480	0.750	13855
Senior share in the bank's market (branches)	0.129	0.025	0.112	0.126	0.140	13855
Asset diversification	0.563	0.218	0.406	0.556	0.724	13855
Income diversification	0.400	0.201	0.261	0.381	0.519	13846
Loan 5-sector diversification	0.389	0.180	0.258	0.406	0.532	13853
Capital ratio (book capital to assets)	0.097	0.031	0.080	0.091	0.106	13855
Size (log assets)	14.036	1.371	13.066	13.669	14.570	13855
Liquidity ratio (liquid assets excl. MBS/ABS)	0.201	0.119	0.117	0.179	0.261	13855
Wholesale funding ratio	0.235	0.120	0.150	0.217	0.295	13855
Log of income per capita in the bank's market	10.408	0.258	10.223	10.411	10.585	13753
Growth of income per capita in the bank's market	0.038	0.027	0.025	0.040	0.053	13753
Growth of house prices in the bank's market	0.040	0.061	0.005	0.040	0.065	13753
Log of total population in the bank's market	13.839	1.693	12.608	13.793	15.299	13753
Log of total deposits per capita in the bank's market	2.897	0.489	2.574	2.853	3.171	13855
Geographic branch deposit diversification	0.355	0.330	0.000	0.309	0.658	13855
Summary Statistics for Cross-Sectional Bank Data: Post-Crisis Bank Risk on Pre-Crisis Diversificatio						
Nonperforming loans to loans	0.032	0.033	0.013	0.023	0.040	1154
ROA (Net income before extraordinary items)	0.0013	0.017	-0.001	0.005	0.009	1154
Std. deviation of ROA	0.010	0.013	0.002	0.005	0.012	1154
Z-score	41.228	30.948	18.021	34.859	56.535	1154
Log Z-score	3.089	1.050	2.371	3.289	3.905	1154
Geographic loan diversification (pre-crisis)	0.453	0.297	0.159	0.473	0.716	1154

Pairwise correlation coefficients for the main diversification measures and the instrument

	(1)	(2)	(3)	(4)	(5)	(6)
1 Geographic loan diversification	1.000					
2 Senior share in the bank's market	0.112	1.000				
3 Asset diversification	-0.125	0.037	1.000			
4 Income diversification	0.262	-0.065	0.019	1.000		
5 Loan 5-sector diversification	0.312	-0.046	0.112	0.346	1.000	
6 Geographic branch deposit diversification	0.837	0.155	-0.104	0.254	0.321	1.000

This table presents summary statistics for the main regression variables. Each bank refers to a U.S. banking organization (topholder). The summary statistics are computed over the sample for which lending diversification is available for the CRA reporting banks (annual 1996-2012). The bank characteristics in Table 3 are year-end measures based on the fourth quarter. Post-crisis bank risk outcomes are computed using the quarterly Call Report data from 2007:Q4 to 2012:Q4 (5-year average) and pre-crisis measures are computed based on the period 2001:Q4 to 2006:Q4 (5-year average). The correlations reported are statistically significant at the 1% level. Data Sources: Call Reports, Community Reinvestment Act (CRA) submissions, the FDIC's Summary of Deposits, FHFA, BEA, and U.S. Census.

The first set of main results testing Hypothesis 1 is presented in Table 3. In the first three columns, the geographic diversification of bank lending is directly regressed on the senior share instrument. Column (1) does not control for local economic activity and population, while column (2) controls for these variables. The effect of senior share is statistically significant at the 1% level in this bank-year panel for CRA reporters. The economic effect is also significant. For example, the 2.26 effect estimated in column (2) implies that a one standard deviation in the senior share of 0.025

increases geographic lending diversification by 0.057[10]. This is economically significant when compared to mean (standard deviation) of lending diversification of 0.462 (0.314). Column (3) shows that the effect of senior share on geographic lending diversification is somewhat lower in economic magnitude and is only statistically significant at the 10% level for the smaller sample of bank-years during 1996-2004 before the CRA changed the disclosure reporting criteria from $250 million to $1 billion in asset size.

Similar effects are demonstrated for instrumental variable (IV) regressions in the last two columns of Table 3. These regressions use the senior share to instrument either the log of total local deposits per capita or the geographic branch deposit diversification. The results are somewhat imprecise for log total deposits but that is attributable to the less precise first stage results in this conditional CRA bank-year sample compared with all geographic areas included in Table 1. The first stage effect of senior share on log of total deposits per capita is 2.56. But the economic effects of a 0.025 increase in the senior fraction leads to a very similar increase in geographic diversification (0.056), as expected when compared with regressing directly on the instrument in the first three columns. There is a similar effect in column (5) when instrumenting the geographic diversification of branch deposits with seniors[11].

[10] More intuitively, this comparison is also similar to moving from a 25 percentile senior share (0.112) to a 75 percentile senior market share (0.140) (Table 2).

[11] As shown in Table 2, the correlation between geographic lending diversification is highest with the geographic branch deposit diversification (the correlation coefficient is 0.84). This makes sense to the extent that there is a natural synergy between deposit-taking and lending. It is also worth noting that the correlation of lending geographic diversification is second most correlated with loan sector diversification, which also makes sense that banks that diversify lending geographically are also more likely to diversify lending across sectors and industries. Note that the five sectors are limited to the coarse groupings in the Call Reports (C&I, real estate, individual, agriculture, and other).

Table 3: The drivers of lending diversification

	(1)	(2)	(3)	(4) IV	(5) IV
Instrument					
Senior share in the bank's market	1.937	2.260	1.930		
	(2.96)***	(3.21)***	(1.80)*		
Log of total deposits per capita				0.883	
in the bank's market				(1.84)*	
Geographic branch deposit					0.394
diversification					(4.44)***
Controls					
Asset diversification	-0.030	-0.029	-0.022	-0.026	-0.024
	(1.76)*	(1.73)*	(1.15)	(0.84)	(1.53)
Income diversification	-0.009	-0.010	-0.017	0.106	-0.005
	(0.70)	(0.75)	(1.05)	(1.52)	(0.42)
Loan 5-sector diversification	0.016	0.017	0.018	0.046	0.031
	(0.43)	(0.45)	(0.41)	(0.61)	(0.88)
Capital ratio	-0.150	-0.199	-0.287	-0.170	-0.227
	(1.11)	(1.46)	(1.72)*	(0.60)	(1.93)*
Size (log assets)	0.083	0.082	0.082	-0.040	0.051
	(9.16)***	(8.84)***	(6.35)***	(0.54)	(4.83)***
Liquidity ratio	-0.117	-0.108	-0.156	-0.126	-0.089
	(3.23)***	(3.05)***	(2.99)***	(1.68)*	(2.66)***
Wholesale funding ratio	0.008	0.015	-0.015	0.183	0.043
	(0.23)	(0.47)	(0.42)	(1.40)	(1.40)
Log of income per capita		-0.168	-0.165	-0.705	-0.078

in the bank's market		(2.78)***	(1.93)*	(2.28)**	(1.49)
Growth of income per capita		0.101	0.177	0.660	0.030
in the bank's market		(1.88)*	(2.38)**	(2.05)**	(0.58)
Growth of house prices		0.040	-0.179	-0.030	0.037
in the bank's market		(1.09)	(2.60)***	(0.33)	(1.11)
Log of total population		0.019	0.026	0.055	-0.009
in the bank's market		(1.88)*	(1.49)	(1.88)*	(1.08)
Observations	13844	13739	8576	13739	13739
R2	0.19	0.20	0.17	.	0.29
First-stage *F*-test				6.08**	42.03***

This table shows the effect of seniors (as a share of population in a bank's geographic deposit funding market) on the bank's decision to geographically diversify its lending. The sample period is 1996 to 2012, with the exception of column (3), which is limited to the period from 1996 to 2004 before the CRA changed the reporting criteria asset size from $250 million to $1 billion. The regressions in columns (4) and (5) are the second stage of 2SLS that use the senior share to instrument deposit funding measures. The t-statistics are in parentheses. Standard errors are robust to heteroscedasticity and are clustered at the bank (topholder) level. All specifications are panel regressions with fixed effects for banks and year dummies. The reported R2 is within-R2. ***, **, and * indicate 1%, 5%, and 10% significance respectively.

4.2 The relationship between pre-crisis diversification and post-crisis bank risk

The second set of main results testing Hypothesis 2 is presented in Table 4 and discussed in this section. Table 4 relates the demographically driven lending diversification identified in Table 3 to common bank risk outcomes examined in the diversification literature (Acharya et al., 2006, Goetz et al., 2016). To test whether banks can escape the stylized fact that bank failures continue to reflect local conditions as documented in Aubuchon and

Wheelock (2010), post-crisis bank risk outcomes (2007-2012 quarterly average) are regressed on instrumented pre-crisis geographic lending diversification (2001-2006 quarterly average). The regressions also control for other bank characteristics as well as the respective pre-crisis risk measure for each of NPL in column (1), ROA in column (2), standard deviation of ROA in column (3), the Z-score in column (4), and the log of Z-score in column (5). Post-crisis economic activity measures in a bank's deposit market are also controlled for.

The results support the second part of the diversification benefit hypothesis: banks that exogenously diversified their lending geographically in the years prior to the financial crisis later performed significantly better, both statistically and economically. For example, a one-standard deviation to instrumented geographic lending diversification (0.297) is associated with a 0.004 decrease in the volatility of ROA. This is a significant effect when evaluated relative to mean post-crisis volatility of ROA equal to 0.010 (the summary statistics in the middle panel of Table 2). Similarly, a one-standard deviation increase in instrumented geographic lending diversification improves the Z-score by 13.74, which is also economically material relative to mean post-crisis Z-score of 41.23[12].

[12] Note that the summary statistics for the Z-score and its components, annualized ROA and the standard deviation of ROA, are similar to Stiroh and Rumble (2006) and other studies. Many studies also take the log of Z-score because it is more normally distributed than the Z-score, which tends to be skewed. This is confirmed to be the case in this sample too. The effect on log Z-score is 0.29 compared with a mean (standard deviation) of this variable of 3.09 (1.05).

Table 4: The relationship between pre-crisis diversification and post-crisis bank risk

	(1)	(2)	(3)	(4)	(5)
	NPL	ROA	Std. Dev. ROA	Z-score	Log Z-score
Pre-crisis measures					
Geographic loan	-0.029	0.016	-0.014	46.278	0.985
diversification	(1.60)	(2.07)**	(2.52)**	(2.86)***	(2.11)**
Asset diversification	-0.024	0.012	-0.010	29.277	0.760
	(3.38)***	(4.09)***	(3.86)***	(5.05)***	(4.25)***
Income diversification	0.002	-0.002	0.002	-19.412	-0.084
	(0.20)	(0.61)	(0.80)	(2.99)***	(0.43)
Loan 5-sector diversification	-0.003	0.001	0.004	-30.280	-0.293
	(0.32)	(0.21)	(1.27)	(3.28)***	(1.10)
Capital ratio	0.056	-0.001	0.044	19.196	0.658
	(1.10)	(0.02)	(2.27)**	(0.71)	(0.76)
Size (log assets)	0.002	-0.001	0.001	-0.989	-0.062
	(1.19)	(1.86)*	(2.11)**	(0.69)	(1.49)
Liquidity ratio	-0.013	0.014	-0.008	29.256	0.931
	(0.88)	(2.51)**	(1.41)	(2.53)**	(2.61)***
Wholesale funding	0.037	-0.019	0.016	-63.983	-1.550
ratio	(2.91)***	(2.81)***	(3.47)***	(6.55)***	(4.91)***
NPL	1.165				
	(6.72)***				
ROA		0.512			
		(3.49)***			

Std. Dev. ROA			0.378		
			(3.17)***		
Z-score				0.004	
				(0.62)	
Log Z-score					0.424
					(9.27)***
Post-crisis controls					
Growth of income per capita	-0.249	-0.102	0.015	-14.441	4.830
in the bank's market	(1.86)*	(1.44)	(0.31)	(0.14)	(1.38)
Growth of house prices	-0.219	0.172	-0.084	162.483	6.714
in the bank's market	(5.76)***	(6.46)***	(4.27)***	(5.74)***	(6.45)***
Constant	0.011	0.010	-0.006	49.155	1.730
	(0.61)	(1.18)	(0.98)	(3.13)***	(3.47)***
Observations	1154	1154	1154	1154	1154
Adjusted R2	0.20	0.18	0.05	-0.00	0.20
First-stage *F*-test	52.42***	64.82***	52.97***	50.51***	50.83***

This table shows the effect of pre-crisis bank diversification measures on bank risk outcomes during and after the financial crisis (the results of 2SLS). The main lending diversification measure is instrumented with the senior share. The regression also controls for other bank diversification measures and characteristics as well as post-crisis economic performance where the bank branches are located. The t-statistics are in parentheses. Standard errors are robust to heteroscedasticity and are clustered at the bank (topholder) level. All specifications are cross-sectional bank regressions, where the pre-crisis values are averaged over 2001:Q4 to 2006:Q4 (5-year) and post-crisis values are averaged over 2007:Q4 to 2012:Q4 (5-year). ***, **, and * indicate 1%, 5%, and 10% significance respectively.

5. Conclusion

There are additional avenues to explore with this research, in particular, the relationship between branch deposit diversification and lending diversification. The methodology of recent papers exploring exogenous determinants of branch deposit diversification through time-staggered bank deregulation and a gravity model of branching out can be combined with the methods of this research to provide a richer framework.

Nonetheless, the preliminary findings of this research are supportive of the view that banks that diversify their loan book perform better. Such banks were less susceptible to the last financial crisis and its immediate aftermath. This paper's findings highlight the importance of addressing the self-selection problem when evaluating the effects of diversification.

Reference List

Acharya, V. V., Hasan, I., & Saunders, A. (2006). Should banks be diversified? Evidence from individual bank loan portfolios. *The Journal of Business*, *79*(3), 1355-1412.

Aubuchon, C. P., & Wheelock, D. C. (2010). The geographic distribution and characteristics of US bank failures, 2007-2010: do bank failures still reflect local economic conditions? *Federal Reserve Bank of St. Louis Review*, *92*(5), 395-415.

Becker, B. (2007). Geographical segmentation of US capital markets. *Journal of Financial Economics*, *85*(1), 151-178.

Brevoort, K. P., & Hannan, T. H. (2006). Commercial lending and distance: Evidence from Community Reinvestment Act data. *Journal of Money, Credit, and Banking*, *38*(8), 1991-2012.

Bricker, J., Dettling, L. J., Henriques, A., Hsu, J. W., Jacobs, L., Moore, K. B., Pack, S., Sabelhaus, J., Thompson, J. & Windle, R. A. (2017). Changes in US family finances from 2013 to 2016: Evidence from the survey of consumer finances. *Federal Reserve Bulletin, 103*(3), 1-42.

Campa, J. M., & Kedia, S. (2002). Explaining the diversification discount. *The Journal of Finance*, *57*(4), 1731-1762.

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

Deng, S. E., & Elyasiani, E. (2008). Geographic diversification, bank holding company value, and risk. *Journal of Money, Credit and Banking*, *40*(6), 1217-1238.

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *The Review of Economic Studies*, *51*(3), 393-414.

Dou, Y., & Zou, Y. (2017). *The real effects of geographic lending disclosures on banks*. Working Paper, NYU Stern School of Business.

Federal Deposit Insurance Corporation (FDIC) (2010). Out-of-Territory Lending. *Superivsory Insights*, 20-21.

Ibragimov, R., Jaffee, D., & Walden, J. (2011). Diversification disasters. *Journal of Financial Economics*, *99*(2), 333-348.

Goetz, M. R., Laeven, L., & Levine, R. (2013). Identifying the valuation effects and agency costs of corporate diversification: Evidence from the geographic diversification of US banks. *The Review of Financial Studies*, *26*(7), 1787-1823.

Goetz, M. R., Laeven, L., & Levine, R. (2016). Does the geographic expansion of banks reduce risk? *Journal of Financial Economics*, *120*(2), 346-362.

Keeton, W. R. (2009). Has multi-market banking changed the response of small business lending to local economic shocks? *Economic Review-Federal Reserve Bank of Kansas City*, *94*(1), 1-36.

Laeven, L., & Levine, R. (2007). Is there a diversification discount in financial conglomerates? *Journal of Financial Economics*, *85*(2), 331-367.

Levine, R., Lin, C., & Xie, W. (2016). *Geographic diversification and banks' funding costs* Working Paper No. w22544. National Bureau of Economic Research.

Loutskina, E., & Strahan, P. E. (2011). Informed and uninformed investment in housing: The downside of diversification. *The Review of Financial Studies*, *24*(5), 1447-1480.

Meslier, C., Morgan, D. P., Samolyk, K., & Tarazi, A. (2016). The benefits and costs of geographic diversification in banking. *Journal of International Money and Finance*, *69*, 287-317.

Morgan, D. P., Rime, B., & Strahan, P. E. (2004). Bank integration and state business cycles. *The Quarterly Journal of Economics*, *119*(4), 1555-1584.

Petersen, M. A., & Rajan, R. G. (2002). Does distance still matter? The information revolution in small business lending. *The Journal of Finance*, *57*(6), 2533-2570.

Rajan, R., Servaes, H., & Zingales, L. (2000). The cost of diversity: The diversification discount and inefficient investment. *The Journal of Finance*, *55*(1), 35-80.

SCF Chartbook (2016). https://www.federalreserve.gov/econres/scfindex.htm.

Stiroh, K. J. (2004). Diversification in banking: Is noninterest income the answer? *Journal of Money, Credit, and Banking*, *36*(5), 853-882.

Stiroh, K. J., & Rumble, A. (2006). The dark side of diversification: The case of US financial holding companies. *Journal of Banking & Finance*, *30*(8), 2131-2161.

Winton, A. (1999). Don't put all your eggs in one basket? Diversification and specialization in lending. Working Paper, University of Minnesota.