The Geographic Flow of Bank Funding and Access to Credit: Branch Networks and Local-Market Competition

Victor Aguirregabiria*
University of Toronto and CEPR

Robert Clark*
Queen's University

Hui Wang*
Peking University

September, 2017

Abstract

The integration of deposit and loan markets may be constrained by the geographic dispersion of depositors, borrowers, and banks. Asymmetric information between geographic locations, monitoring costs, transaction costs, and imperfections in interbank wholesale markets can all serve as frictions to the flow of funds across markets, leaving some with limited access to credit. Banks' branch networks can reduce some of these frictions and increase the flow of funding to geographic locations where credit is in greater demand. However, local market power and economies of scope between deposits and loans at the local level may have a negative impact on the geographic flow of credit. This paper studies empirically the contribution of branch networks, local market power, and economies of scope to this flow. Our results are based on the estimation of a structural model of bank oligopoly competition for deposits and loans in multiple geographic markets using data at the bank-county-year level from the US banking industry for the period 1998-2010. The identification of the model exploits the independence of transitory local shocks between geographic locations which are distant enough from each other. The estimated model shows that a bank's total deposits has a very significant effect on the bank's market shares in loan markets. We also find evidence that is consistent with significant economies of scope between deposits and loans at the local level. Counterfactual experiments show that these economies of scope generate a substantial home-bias in the utilization of funds. Local market power has also a significant negative effect on the geographic flow of credit.

Keywords: Geographic flow of bank funds; Access to credit; Bank oligopoly competition; Branch networks; Economies of scope between deposits and loans.

JEL codes: L13, L51, G21

Victor Aguirregabiria. Address: 150 St. George Street. Toronto, ON, M5S 3G7, Canada. Phone: (416) 978-4358. E-mail: victor.aguirregabiria@utoronto.ca

Robert Clark. Address: 94 University Avenue. Kingston, ON, K7L 3N6, Canada. E-mail: clarkr@econ.queensu.ca

Hui Wang. Address: Guanghua School of Management. Peking University. Beijing, 100871, China. E-mail: jackie.wang@gsm.pku.edu.cn

*Preliminary and incomplete. We want to thank Jason Allen, Heski Bar-Isaac, Lu Han, Ali Hortaçsu, Jean-François Houde, Jakub Kastl, Matt Osborne, Carlo Regiani, Howard Smith, Xavier Vives, and participants in seminars and conferences at the Carey Business School, the University of Toronto, the Vancouver School of Economics, *Industrial Organization of the Financial Sector conference* at the Becker-Friedman Institute, *Jornadas de Economia Industrial* and *CEPR/JIE School and Conference on Applied Industrial Organization* for helpful comments. We are grateful to Paul Lim for excellent research assistance.

1 Introduction

There is evidence that heterogeneity exists in the ability of individuals to access credit. Since access to financing has been linked to entrepreneurship levels, employment, wages, and economic growth (see for instance Gine and Townsend (2004)), this heterogeneity can lead to socio-economic inequality. Moreover, there is mounting concern among policy makers that differences in the ability to access loans is at least partly geographic, with individuals in some regions able to more easily obtain financing than individuals in other regions.

An important determinant of credit provision is the availability of deposits: greater deposits allow banks to make more loans. Unfortunately, in any given region, the demand for loans may not always coincide with the availability of deposits. This would not be a problem in an economy without geographic frictions, as funds would flow from one area to another such that, in equilibrium, the expected rate of return and the risk of the marginal loan would be the same across geographic markets, and the funding of an investment project would not depend on its geographic location. In actual economies, geographic distance between borrowers and lenders can increase asymmetric information, monitoring costs, and transaction costs of liquidity within banks. All these can serve as frictions to the flow of funds across markets and can generate substantial geographic imbalances in the provision of, and access to, credit (*credit deserts*).[1]

Wholesale liquidity markets and bank branch networks can help to alleviate the effects of these frictions. Banks can buy and sell liquidity (deposits) in the interbank wholesale market. However, there are transaction costs involved in using these wholesale markets due to bank precautionary motives and liquidity hoarding (Ashcraft el al., 2011, Acharya and Merrouche, 2012). Banks can also use branching as an instrument to reduce geographic frictions in the flow of credit. By opening branches in multiple locations, a bank can reduce its geographic distance with borrowers, and therefore it can reduce frictions which are related to the geographic distance between lender and borrowers. If transaction costs between branches of the same bank are smaller than the costs of using interbank markets (Coase, 1937), then banks' branch networks may increase the flow of funding to geographic locations where credit is in greater demand.

Two counterbalancing forces can affect negatively the willingness of a bank to transfer funds between its branches: (i) economies of scope between deposits and loans at the branch level, and (ii) local market power. Clients may prefer to have their deposit account and their mortgage in the same bank. For the bank, the cost of managing a deposit account and a loan may be smaller if they

[1] Brevoort and Wolken (2009) and Nguyen (2015) show that the geographic distance between borrowers and lenders have a negative impact on the amount of credit.

belong to the same client. These economies of scope between deposits and loans create incentives to concentrate lending activity in those branches with high levels of deposits, and therefore to limit the geographic flow of liquidity to markets with more need of credit.[2] Local market power implies that a change in the marginal cost of loans (e.g., a reduction in the interbank interest rate) is only partially passed-through to borrowers. As a result, smaller local markets with highly concentrated market structures may not benefit from aggregate shocks in the supply of credit as much as more competitive markets. Local market power can have a negative impact on the geographic flow of credit.[3]

The purpose of this paper is to provide systematic evidence on the extent to which deposits and loans are geographically imbalanced in the US commercial banking industry, and to investigate empirically the contribution of branch networks, economies of scope, and local market power to this imbalance. We focus on three empirical questions: (i) How important is the 'home bias' generated by economies of scope between deposits and loans?; (ii) What is the contribution of local market power to the geographic distribution of bank credit?; and (iii) How did the deregulation that allowed banks to expand geographically in the 1990s and 2000s affect the geographic flow of bank funds?

Our results are based on the estimation of a structural model of bank oligopoly competition for deposits and loans in multiple geographic markets. The equilibrium of the model allows for rich interconnections across geographic locations and between deposit and loan markets such that local shocks in demand for deposits or loans can affect endogenously the volume of loans and deposits in every local market. We characterize an equilibrium of this multimarket oligopoly model and propose an algorithm to solve for an equilibrium. We estimate this structural model using data from the US banking industry for the period 1998-2010. Our novel dataset merges data at the bank-county-year level from two sources. Deposit and branch-network information are collected from the *Summary of Deposit* (SOD) data provided by the *Federal Deposit Insurance Corporation* (FDIC). Information on loans comes from the *Home Mortgage Disclosure Act* (HMDA) data set, which provides detailed information on mortgage loans.

In our model, differentiated banks sell deposit and loan products in multiple local markets (counties). The model incorporates three (endogenous) variables, which are key factors in a bank's demand and cost of loans and deposits in a local market. A first factor is the number of branches

[2]As we discuss in our Model section (section 3), these economies of scope between deposits and loans may be driven either by consumer demand (i.e., *one-stop banking*) or by variable costs. See also Kashyap, Rajan, and Stein (2002) and Egan, Lewellen, and Sunderam (2017) for models and empirical evidence on the positive synergies between banks' deposit and lending activities.

[3]Black and Strahan (2002) and Cetorelli and Strahan (2006) provide empirical evidence of how entrepreneurs and potential entrants in nonfinancial sectors face more difficult access to credit in local markets characterized by a concentrated banking sector.

the bank has in the local market. The number of branches reduces marginal costs of lending and may generate consumer awareness and willingness to pay. A second factor is the total amount of deposits the bank has at the national level, that reduces the bank's risk for liquidity shortage and the need to borrow at interbank wholesale markets. This introduces an important interconnection between local markets in a bank's operation. A third factor is the amount of deposits the bank has in the local market that increases consumer demand for loans and reduces the bank's marginal cost of a loan due to economies of scope in managing deposits and loans. These three factors are fundamental in the determination of the geographic flow of liquidity in the equilibrium of the model. The stronger the effect of local branches on the demand and cost of loans, the more concentrated are loan markets and this has a negative impact on the geographic diffusion of credit. Economies of scope between deposits and loans also reduce geographic flow of credit. In contrast, the effect of total bank deposits on local loans have a positive impact on the geographic diffusion of credit.

Our model builds on and extends the literature on structural models of bank competition. Neven and Röller (1999) estimate a model of bank oligopoly competition in deposits and loans in seven European countries. Their model assumes competition at the national level and it does not allow for multiple local markets or for economies of scope between deposits and loans. Previous studies have proposed and estimated structural equilibrium models for bank deposits as a differentiated product. Dick (2008) estimates a demand model for bank deposits, and Egan, Hortaçsu, and Matvos (2017) distinguish between insured and uninsured deposits, and endogenize bank defaults and bank runs. Our paper extends these previous studies by: (a) incorporating demand, supply, and competition in the market for bank loans; (b) allowing for economies of scope between deposits and loans, that introduces and important link between these markets at the local market level; and (c) including the effect of a bank's total liquidity on the demand and costs of deposits and loans in local markets.[4] Corbae and D'Erasmo (2013) propose and calibrate a dynamic equilibrium model of the US banking industry that incorporates Stackelberg oligopoly competition in both deposits and loans, endogenous market entry and exit, and multiple geographic markets. Our model is static and it does not endogenize bank-entry exit decisions. However, it provides a more detailed description of the geographic inter-connections between deposits and loans at the bank-county level. Aguirregabiria, Clark, and Wang (2016) estimate a model of banks' geographic location of branches, and study the role of geographic risk diversification in the configuration of bank branch networks. In the current paper, we extend this previous model by incorporating competition in

[4]In Egan, Hortaçsu & Matvos (2017), the demand for uninsured deposits of a bank depends on the bank's total liquidity. However, their model does not incorporate demand and supply of loans and how they depend on the bank's liquidity.

both loans and deposits, and inter-connections between these two markets and across geographic markets. Here we also focus on competition at the intensive margin and omit the part of the model that has to do with competition at the extensive margin, i.e., opening and closing branches, and entry and exit in loans/deposits local markets.

Three sets of structural parameters are fundamental for the predictions of the model: (a) parameters that capture the effect of the number of local branches on a bank's demand and marginal cost for deposits and loans; (b) parameters that capture economies of scope between deposits and loans at the local level; and (c) parameters that measure the effect of a bank's global deposits on the marginal cost of loans. Estimation of these parameters must address endogeneity and simultaneity issues. For the identification of the effect of local variables (i.e., number of branches and local deposits), we exploit moment conditions that combine assumptions from 'BLP instruments' in the estimation of demand of differentiated products (Berry, Levinsohn, and Pakes, 1995), restrictions on the serial correlation structure of the unobservables, and a time-to-build assumption on a bank's decision to open (close) branches in the spirit of dynamic panel data models (Arellano and Bond, 1991, Arellano and Bover, 1995, Blundell and Bond, 1999). The identification of the effect of total deposits is more challenging since this variable has only variation over time at the bank level. Our identification strategy exploits a zero covariance restriction between the local idiosyncratic shocks of far apart locations.

The estimation of the model provides the following results. The number of branches in a county increases (reduces) substantially the demand (cost) for both deposits and loans, though the effect is significantly smaller for loans. We find evidence of substantial economies of scope between deposits and loans at the level of bank and local market. The effect of a bank's total deposits on demand (cost) of loans is positive (negative) and very significant both economically and statistically. Banks' internal liquidity reduces the costs of lending.

Our structural approach allows us to evaluate factual and counterfactual policies that affect the flow of funding to those markets where deposits are scarce. We consider the following counterfactual experiments. First, we look at the effects of the consolidation of US baking industry and the geographic expansion of branch networks by looking at the counterfactual equilibrium if banks' branch networks were the ones in 1994, before Riegle-Neal act. Second, we study the effects of eliminating the home bias due to economies of scope between deposits and loans. Third, we look at the effect of eliminating county heterogeneity in local market power. Finally, we study the potential *geographic non-neutrality* of different government policies. We evaluate how a (counterfactual) tax on deposits would affect the provision of credit and, more interestingly, its geographic distribution.

We also investigate to what extent national aggregate shocks (e.g., business cycle, monetary policy) affect bank credit in a geographic non-neutral way.

We are not the first to study the relationship between retail funding and loan activity. The closest to our work is a recent set of papers that take advantage of the exogenous variation provided by the shale boom to study the extent to which banks use their branch networks to transfer funds from one local market to another (Gilje, 2012; Gilje, Loutskina, and Strahan, 2016; Loutskina and Strahan, 2015; and Petkov, 2016). Our paper complements in different ways the empirical findings by Gilje, Loutskina, and Strahan (2016). First, our empirical analysis of the relationship between the geographic location of a bank's branches (deposits) and loans extends to all the local markets (counties) in US. Second, we study the contribution of local market power to the geographic flow of banks' funds. Third, our approach for the identification of the effect of total deposits on local loans exploits more general sources of exogenous variation than those associated to local catastrophic events or discoveries of natural resources. Finally, our structural model allows us to identify the different sources of transaction costs for the flow of funding, and to perform counterfactual experiments to evaluate the effect on credit of reducing these costs.

The rest of the paper proceeds as follows. In the next section we describe the data and present descriptive evidence on the geographic dispersion of deposits and loans. In Section 3 we describe our model and in Section 4 we explain how we go about estimating it. Section 5 presents our empirical results and Section 6 describes our counterfactual experiments. Finally, Section 7 concludes.

2 Data and descriptive evidence

2.1 Data sources

We combine two data sources at the bank-county level. Branch and deposit information is collected from the Summary of Deposit (SOD) data provided by the *Federal Deposit Insurance Corporation* (FDIC). Information on mortgage loans comes from the Home Mortgage Disclosure Act (HMDA) data set.

The SOD dataset is collected on June 30th of each year and covers all depository institutions insured by the FDIC, including commercial banks and saving associations. The dataset includes information at the branch level on deposits, address, and bank affiliation. Based on the county identifier of each branch, we can construct a measure of the number of branches and total deposits for each bank in each county.[5]

[5]A small proportion of branches in the SOD dataset (around 5% of all branches) have zero recorded deposits. These might be offices in charge of loans or administrative issues. We exclude them in our analysis.

Under the HMDA, most mortgage lending institutions are required to disclose information on the mortgage loans that they originate or purchase in a given year.[6] At the level of financial institution, county, and year, we have information on the number and volume of mortgage applications, mortgage loans actually issued, and mortgage loans subsequently securitized.

The type of institutions reporting to HMDA include both depository institutions and non-depository institutions, mainly Independent Mortgage Companies (IMCs).[7] By definition, only the former, including banks and thrifts, can be matched with the SOD data.[8] Other than this matching issue, this paper focuses on depository institutions because these are the institutions that rely heavily on branching and deposits to fund their loans. By contrast, IMCs rely on wholesale funding and mortgage brokers (Rosen, 2011). Focusing on depository institutions is consistent with the research questions addressed in this paper. Nevertheless, to take into account competition in the mortgage market from non-depository institutions, we aggregate at the county-year level the total number and volume of loan mortgages from these institutions, and we use this information in our construction of market shares and in the estimation of our structural model of demand and supply of mortgages.

County level data on socioeconomic characteristics are obtained from various products of the Census Bureau. The US Census Bureau provides various data products through which we obtain detailed county level characteristics to estimate our model. Population counts by age, gender, and ethnic group are obtained from the Population Estimates. Median household income at the county level is extracted from the State and County Data Files, whereas income per capita is provided by the Bureau of Economic Analysis (BEA). Information on local business activities such as two-digit-industry level employment and number of establishments is provided by the County Business Patterns. Finally, detailed geographic information, including the area and population weighted centroid of each county, and locations of the landmarks in the US, is obtained from the Topologically Integrated Geographic Encoding and Referencing system (TIGER) dataset.

We also use information on county-level house prices for 2742 counties between 1990-2015 from the Federal Housing Finance Agency (see Bogin, Doerner and Larson, 2016), and county-level

[6] There are some geographic restrictions on loan reporting. According to the Community Reinvestment Act (CRA), large banks have to report information on all their loans regardless of the geographic location. Furthermore, regardless their size, lenders located in an MSA must report on loans originated in an MSA, though they can choose not to report loans outside MSAs. Only small lenders located outside of MSAs do not have to report. This means that the HMDA dataset may not include mortgage loans issued by small banks and originated in rural locations. However, according to the US census, about 83 percent of the population lived in an MSA region during our sample period. Therefore, HMDA provides information on the majority of residential mortgage lending activities.

[7] IMCs are for-profit lenders that are neither affiliated nor subsidiaries of banks' holding companies.

[8] We match banks in the SOD and HMDA datasets using their certificate number (provided by FDIC to every insured depository institution) or/and their RSSD number (assigned by Federal Reserve to every financial institution). We match thrifts using their docket numbers.

bankruptcy data from the U.S. Bankruptcy Courts.[9]

We derive bank-level characteristics from balance sheets and income statement information in the banks' quarterly reports provided to the different regulatory bodies: the Federal Reserve Board (FRB)'s Report on Condition and Income (Call Reports) for commercial banks, and the Office of Thrift Supervision (OTS)'s Thrift Financial Report (TFR) for saving associations.

The National Information Center records the timing of major historical events, such as renaming, merger and acquisition, and bankruptcy, of all depository institutions that ever existed in the United States. This information allows us to identify all the merger cases and the involved banks during the sample period.

There are two features of our data and empirical approach that deserve specific discussion. First, we have data on mortgage loans at the bank-county-year level but we do not have data on other forms of bank credit. Ideally, we would like to use information on other types of bank loans, but, to our knowledge, such data are not publicly available at the bank-county-year level. However, mortgage loans represent the most substantial part of bank loans, and even of bank assets. Using bank level information from the 2010 Call Reports, Mankart, Michaelides, and Pagratis (2016) show that mortgages account for between 62% and 72% of all bank loans, and between 38% and 45% of total bank assets, where the range of values captures heterogeneity in these ratios according to bank size (i.e., larger banks tend to have a smaller share of mortgage loans in total loans and assets). They also report that bank deposits represent between 68% and 85% of total bank liabilities. Therefore, our focus on deposits and mortgages, though motivated by data availability, captures a very substantial fraction of total bank liabilities and assets, respectively. Furthermore, other sorts of loans may be taken out at one location, but used to finance projects elsewhere. This would make studying the flow of funding and access to credit difficult. In contrast, mortgages are much more local.

Second, publicly available data on interest rates of deposits and loans are not available at the bank-county-year level, or even at a more aggregate geographic level. Furthermore, the existing proprietary data on interest rates are not as clean as the quantity data on deposits and mortgage loans that we use, and they are based on geographic interpolations, and therefore, potentially important measurement errors. The loan-rate data in particular are available only for a small set of lenders. The lack of good price data at the bank-county-year level would be an important limitation if we wanted to separately estimate demand and marginal cost. However, that is not the goal of this paper. To answer all the empirical questions in this paper, we need to estimate

[9] More specifically, we use Table F 5A Business and Nonbusiness Bankruptcy County Cases Commenced, by Chapter of the Bankruptcy Code During the 12-Month Period Ending June 30, 2007.

the value of consumers willingness-to-pay net of banks' marginal costs for the different deposit and loan products, as well as how these net willingness-to-pay depends on different variables such as local bank branches. We show that these primitives can be identified without information on prices of deposits and loans.

Finally, it is necessary to comment on the fact that we define our markets to be counties, the primary administrative divisions for most states. Markets determine the set of branches that are competing with each other for consumer deposits and loans within a geographic area. Although other market definitions, such as State or Metropolitan Statistical Area, have been employed in some previous empirical studies on the US banking industry, many have considered county as their measure of geographic market (see for instance Ashcraft, 2005; Calomiris and Mason, 2003; Huang, 2008; Gowrisankaran and Krainer, 2011; and Uetake and Wanatabe, 2012).

2.2 Summary statistics

We concentrate on the period 1998-2010. Our matched sample includes 6263 banks in 3143 counties. Of these counties, 2861 have deposits in at least one year during the sample period: there are 282 counties with zero deposits at every year during the sample period. However, we observe positive amounts of mortgage loans in these counties with zero deposits. These 282 counties with no deposits but positive mortgages are rural or suburban markets where people live and make investments but where there are no bank branches. We keep all 3143 counties in our analysis. The dataset contains a total of 1,483,729 bank-county-year observations.

Table 1
Summary Statistics

Bank Level Statistics (48,531 bank-year obs.)

Variable	Mean	S. D.	Pctile 5	Median	Pctile 95
Number of branches	10.9	51.3	1.0	4.0	28.0
Number of counties with deposits > 0	3.4	10.4	1.0	2.0	9.0
Number of counties with new loans > 0	30.4	146.0	1.0	8.0	72.0
Total deposits (in million \$)	616	3,906	37	147	1,486
Total new loans (in million \$)	184	3,286	1	13	272

County Level Statistics (40,736 county-year obs.)

Variable	Mean	S. D.	Pctile 5	Median	Pctile 95
Number of branches (per county)	13.0	32.5	0.0	4.0	55.0
Number of banks with deposits > 0	4.0	6.2	0.0	2.0	14.0
Number of banks with loans > 0	36.2	34.9	4.0	26.0	107.0
HHI market of deposits	3176	2067	1111	2548	7813
HHI market of new loans	710	650	253	528	1718
Deposits per capita (in ,000 \$)	5.3	6.4	0.0	4.2	14.0
New loans per capita (in ,000 \$)	1.2	1.6	0.1	0.8	3.8
Income per capita (in ,000 \$)	27.9	8.1	18.1	26.6	41.7
Population (in ,000 people)	94.0	302.1	3.2	25.4	398.8
Share population $\leq$ 19 (in %)	27.4	3.4	22.2	27.3	33.1
Share population $\geq$ 50 (in %)	33.3	6.3	23.4	33.0	44.2
Annual change in house price index	3.0	5.7	-5.9	3.0	12.3
Number of bankruptcy filings	435	1506	6	107	1799

Table 1 presents summary statistics from our working sample. The top panel provides bank-level statistics based on 48,531 bank-year observations, and the bottom panel includes county-level statistics using 40,736 county-year observations. The median number of counties where a bank obtains deposits from its branches is only 2, while the median number of counties where a bank sells mortgage loans is 8. The branch network of a bank is geographically more concentrated than its network of counties where it provides loans. Similarly, in the panel of county-level statistics, the median number of banks providing deposit services in a county is only 2, but the median number of banks selling mortgages is 26. The median Herfindahl-Hirschman indexes (HHI) are 2548 for deposit markets (i.e., equivalent to 4 symmetric banks per market) and 528 for loan markets (i.e., equivalent to a market with 20 symmetric banks). A possible explanation of this evidence is that branches are more important to attract consumer demand for deposits than to attract demand for loans, but branches are costly to create and operate (e.g., fixed costs). Our estimation of the structural model in section 5 provides evidence supporting this explanation.

Given this level of concentration in mortgage markets, one might argue that market power is very low and should not have any impact on the geographic diffusion of credit. However, this argument does not take into account that branches can still have an impact on demand and cost of loans, and that economies of scope imply that market power in deposit markets have spillover on the mortgage markets.

Figures 1 to 4 present time series of some aggregate magnitudes from our working sample, i.e., deposits, mortgages, loans, banks, and branches) over our sample period. Figure 1 presents the evolution of the stock of deposits and the flow of new mortgage loans aggregated over all banks and counties with a yearly frequency. Both time series follow a similar pattern, with strong growth in the early 2000s followed by a decline of new mortgages and a very modest increase of deposits in the last years of the decade. Figure 2 provides evidence on the importance of mortgage loans in assets for lenders in the HMDA dataset. The median share is just below 40% at the start of our sample, rising to over 50% at the time of the financial crisis. Figure 3 provides evidence on the share of deposits in liabilities for lenders in the HMDA dataset. The median share is around 80%. Figure 4 shows the evolution of the number of banks and branches per county. At the start of our sample there were just over 7.5 banks and about 26 branches per county. These numbers increased steadily to almost 9 and over 31, respectively, by the time of the crisis, before decreasing slightly. Note that the increase from 1994 to 2009 coincides with the rolling out of Riegel Neal, which permitted banks to branch across state lines. Over the same time period the percentage of multi-state banks increased from less than 1% to around 7%.

Figure 1: Time Series of Stock of Deposits and Flow of New Mortgage Loans

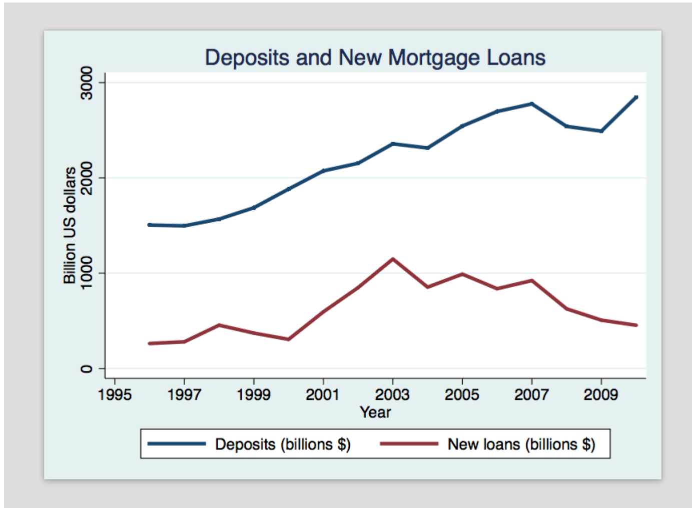


Figure 2: Share of Mortgage Loans in Total Assets

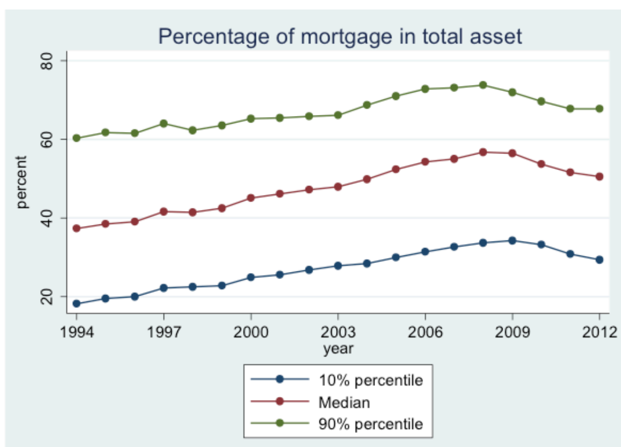


Figure 3: Share of Deposits in Liabilities

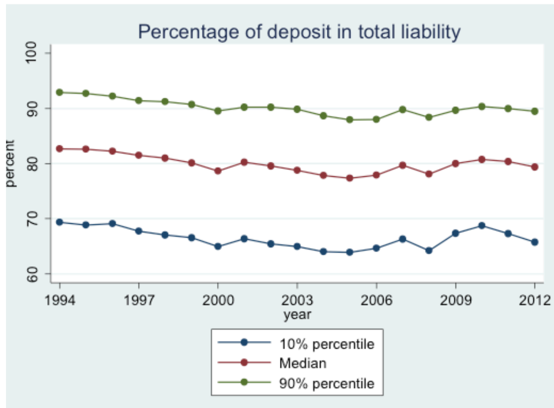
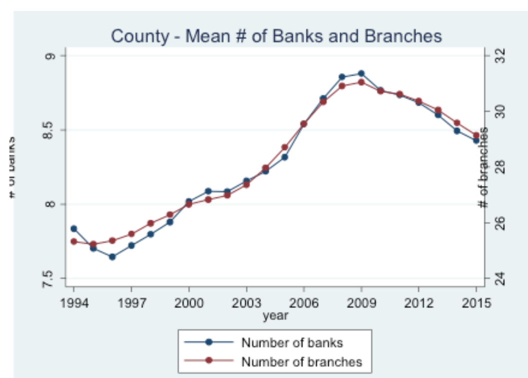


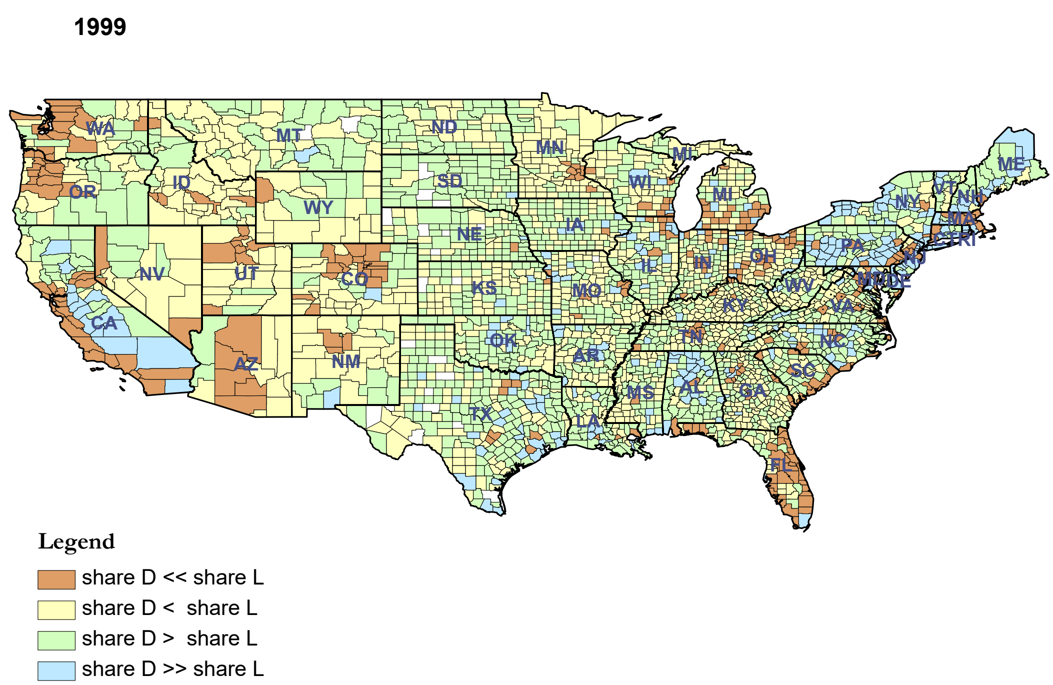
Figure 4: Number of banks and branches by county



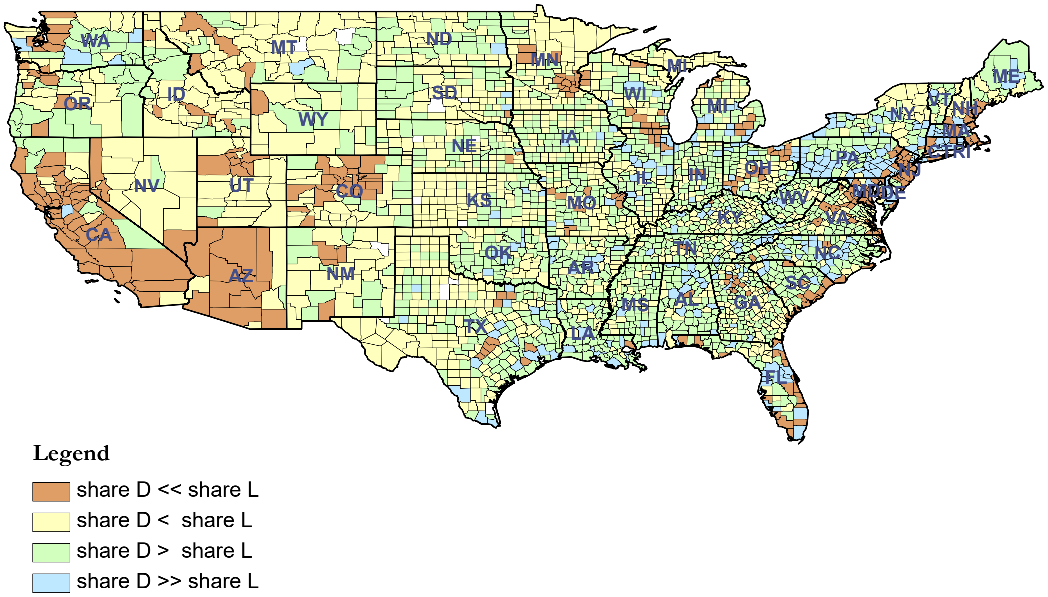
2.3 Geographic patterns of deposits and loans

Figure 5 presents maps with the geographic distribution of counties' positions as net borrowers or net lenders. We present these maps for three different years in our sample: 1999, 2004, and 2009. We first describe the construction of the statistics in these figures. For every county-year, we calculate the county's share of deposits over aggregate national deposits. Similarly, we calculate the county's share of new loans over the aggregate amount of new loans in the nation. Based on these shares, we construct at the county level the index S_{L-D} that represents the difference between the county's share of loans and its share of deposits. We sort counties into four groups: (i) counties belonging to top 10 percentiles of S_{L-D} (Share Loans >> Share Deposits); (ii) counties between the 10th and 50th percentiles of S_{L-D} (Share Loans > Share Deposits); (iii) counties between the 50th and 90th percentiles of S_{L-D} (Share Loans < Share Deposits); and (iv) counties belonging to the bottom 10 percentiles of S_{L-D} (Share Loans << Share Deposits). Figure 5 shows clear evidence of deposit and loan imbalances, with some regions having very high share of deposits, but low share of loans and vice versa. It also reveals regional patterns in the net borrowing/lending position of counties. There are also interesting changes over time that are related to the mortgage boom and the subsequent financial crisis at the end of the decade.

Figure 5. Distribution of borrower/lender counties



2004


WA
OR
CA
NV
ID
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MS
MI
IN
KY
TN
AL
OH
WV
GA
FL
SC
NC
VA
PA
NY
MD
DE
NJ
CT
RI
MA
VT
NH
ME
Legend
share D << share L
share D < share L
share D > share L
share D >> share L

2009

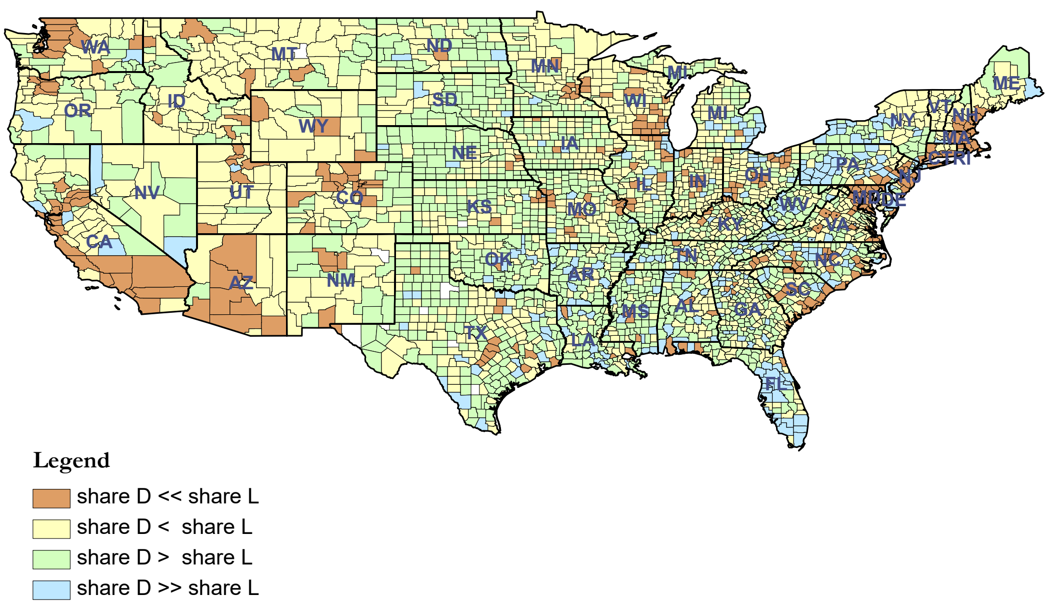

WA
OR
CA
NV
ID
MT
WY
UT
AZ
CO
NM
ND
SD
NE
KS
OK
TX
MN
IA
MO
AR
LA
WI
IL
MS
MI
IN
KY
TN
AL
OH
WV
GA
FL
SC
NC
VA
PA
NY
MD
DE
NJ
CT
RI
MA
VT
NH
ME
Legend
share D << share L
share D < share L
share D > share L
share D >> share L

Figure 6 presents the empirical distribution of *segregation indexes* calculated at the bank-year level. Borrowing from the literature on racial geographic segregation,[10] we calculate the *segregation index*: $(1/2)\sum_m |q^d_{jmt}/Q^d_{jt} - q^\ell_{jmt}/Q^\ell_{jt}|$, where q^d_{jmt} and q^ℓ_{jmt} represent the amount of deposits and loans, respectively, of bank j in county m and year t, and Q^d_{jt} and Q^ℓ_{jt} represent the bank's total amounts of deposits and loans. This index is a measure of the bank's transfer of funds between geographic locations or, alternatively, a measure of the bank's home bias. For instance, a segregation score equal to zero represents an extreme case of home bias, i.e., the bank's geographic distributions of loans and deposits are identical. At the other extreme, a segregation index equal one means that the bank gets all its deposits in markets where does not provide loans, and sells loans only in markets where does not have deposits, which is an extreme case of geographic diffusion of loans. In Figure 6 we consider the subsample of banks with at least five counties with positive values for either loans or deposits. We find that most banks are involved to some degree in the transfer of funds across geographic locations. In fact, the index is quite large (greater than 0.5) for more than one third of these banks.

Figure 7 presents the time series of a national level segregation index calculated using county level observations. This segregation index is defined as:

$$SI_t = \frac{1}{2}\sum_{m=1}^{M}\left|\frac{Q^d_{mt}}{Q^d_t} - \frac{Q^\ell_{mt}}{Q^\ell_t}\right| \tag{1}$$

where $\frac{Q^d_{mt}}{Q^d_t}$ and $\frac{Q^\ell_{mt}}{Q^\ell_t}$ are the shares of county m in the aggregate national amounts of deposits and new mortgage loans, respectively. This index measures the transfer of funds between geographic locations. The index increased substantially during the period 2000-2006, from 0.29 to 0.35. However, it experienced substantial reductions in the 2007-2008 financial crisis (from in 0.35 to 0.32) and to a lower extent at year 2000 dot-com crisis (from 0.31 to $0,29$).

[10]This type of index was first proposed by Jahn, Schmid, and Schrag (1947).

Figure 6. Segregation Indexes between Deposit and Loan Distributions: At the bank-year level

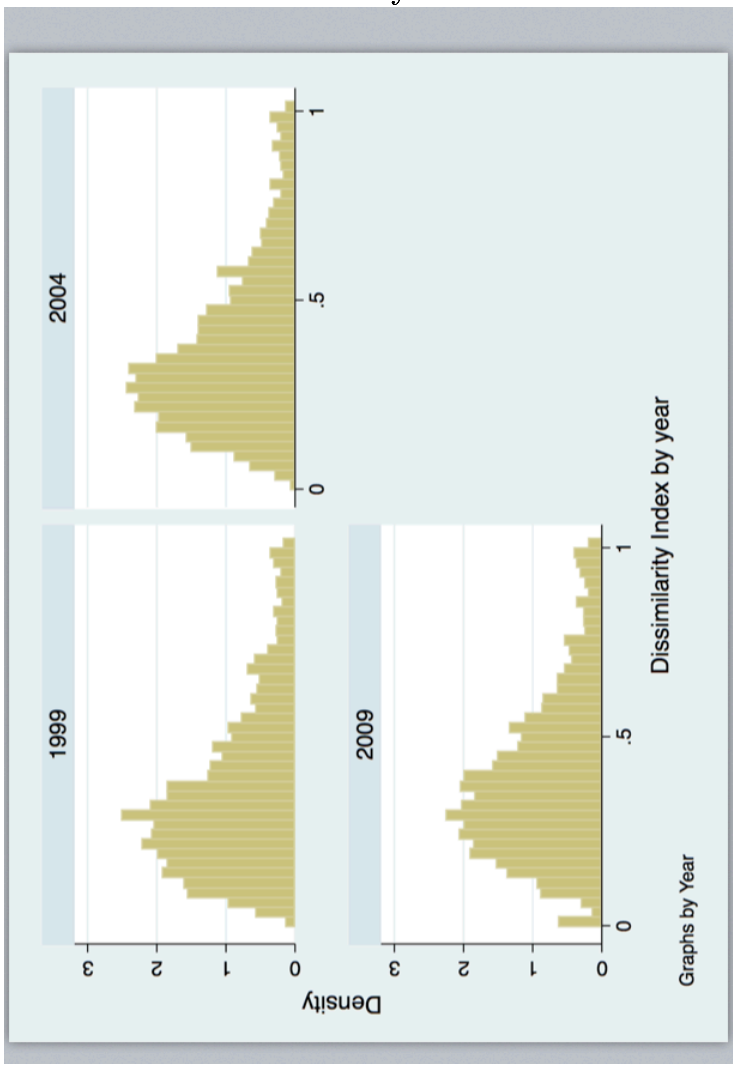
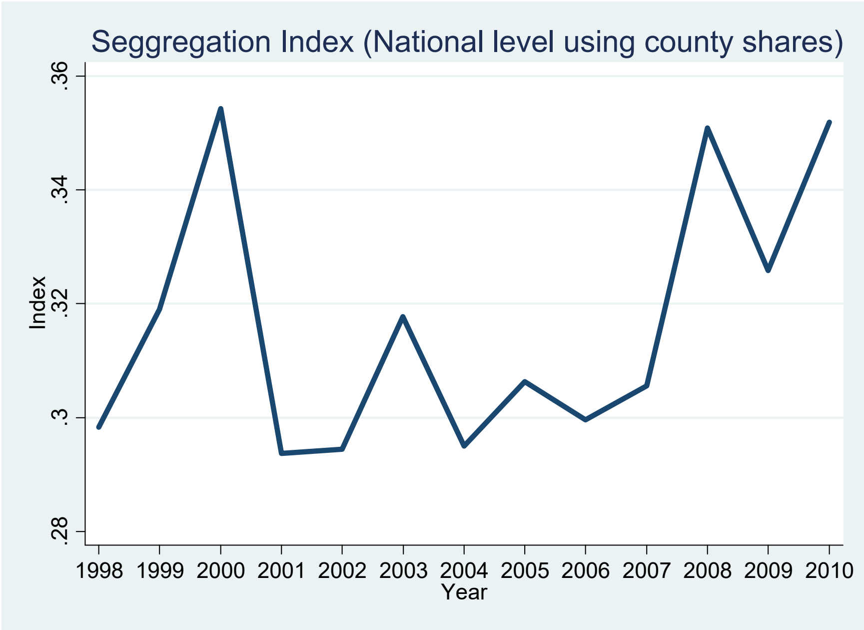


Figure 7. Time Series of the National Segregation Index



Seggregation Index (National level using county shares)
Index
.28
.3
.32
.34
.36
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
Year

3 Model

Consider an economy with M geographic markets (counties), indexed by $m \in \mathcal{M} = \{1, 2, ..., M\}$, and B banks, indexed by $j \in \{1, 2, ..., B\}$. Let $\mathcal{M}_j^d$ represent the set of markets where bank j has branches and sells deposits. Similarly, $\mathcal{M}_j^\ell$ represents the set of markets where bank j sells loans. This set of markets $\mathcal{M}_j^\ell$ includes all the markets where the bank has branches, but it may include other markets where the bank has contacts with mortgage brokers that provide clients for the bank. Therefore, $\mathcal{M}_j^\ell$ includes the set $\mathcal{M}_j^d$ but it can be larger, i.e., $\mathcal{M}_j^d \subseteq \mathcal{M}_j^\ell$.

We take networks $\{\mathcal{M}_j^d\}_{j=1}^B$ and $\{\mathcal{M}_j^\ell\}_{j=1}^B$ as given. One can think of these networks as being the result of a dynamic game of market entry-exit decisions with networks. The specification and estimation of such a complex game is beyond the scope of this paper. Instead, we consider these sets to be pre-determined and focus on the endogenous determination of the amounts of deposits and loans in the equilibrium of a model of multi-market oligopoly competition.

Each local market is populated by two groups of consumers: *savers* who demand deposit products, and *investors* who demand loan products. Banks sell deposit and loan products in these local markets. These products are horizontally differentiated between banks due to different product characteristics and to spatial differentiation within a local market. This view of banks' services as differentiated products is in the spirit of previous papers in the literature such as Degryse (1996), Schargrodsky and Sturzenegger (2000), Cohen and Mazzeo (2007 and 2010), Gowrisankaran and Krainer (2011), or Egan, Hortacsu, and Matvos (2017), among others. A novel feature of our model, that is key for the purposes of this paper, is that it introduces endogenous links between deposit and loan markets and between these markets at different geographic locations.

Bank j sells deposit products in every market in the set $\mathcal{M}_j^d$, and sells loan products in every market in the set $\mathcal{M}_j^\ell$.[11] The (variable) profit function of bank j is equal to interests from new loans (pre-existing loans are considered as pre-determined fixed profits), minus payments to depositors, minus costs of managing deposits and loans, and minus the costs (or returns) from the bank's activity at interbank wholesale markets:

$$\Pi_j = \sum_{m=1}^{M} p_{jm}^\ell \; q_{jm}^\ell + p_{jm}^d \; q_{jm}^d - C_{jm}\left(q_{jm}^\ell, q_{jm}^d\right) - (r_0 + c_{j0})\, B_j \tag{2}$$

where p_{jm}^ℓ and p_{jm}^d are prices for loans and deposits, respectively, for bank j in market m, and q_{jm}^ℓ and q_{jm}^d are the corresponding amounts of loans and deposits. Note that typically the price for loans will be positive ($p_{jm}^\ell > 0$) because borrowers pay a positive interest rate to obtain a loan, while the price of deposits is typically negative ($p_{jm}^d > 0$) because the bank should pay savers to

[11] For the sake of notational simplicity, we omit in this section the time subindex t.

attract their deposits. Market $m = 0$ represents the interbank wholesale market; r_0 is the interbank interest rate; B_j is the net borrowing position of bank j at the interbank market; and c_{j0} is a bank-specific transaction costs associated to using the interbank market. The interbank interest rate r_0 is determined by the Federal Reserve, and it is exogenous in this model.

The function $C_{jm}\left(q^{\ell}_{jm}, q^{d}_{jm}\right)$ represents the cost of managing deposits and loans in market m. A bank's resources constraint implies that,[12] $B_j = Q^{\ell}_j - Q^{d}_j$, where $Q^{\ell}_j \equiv \sum_{m=1}^{M} q^{\ell}_{jm}$ and $Q^{d}_j \equiv \sum_{m=1}^{M} q^{d}_{jm}$ are bank j's total new loans and deposits, respectively. Solving this restriction in the profit function, we have that $\Pi_j = \sum_{m=1}^{M} p^{\ell}_{jm}\ q^{\ell}_{jm} +\ p^{d}_{jm}\ q^{d}_{jm} -\ \widetilde{C}_{jm}\left(q^{\ell}_{jm}, q^{d}_{jm}\right)$, with $\widetilde{C}_{jm}\left(q^{\ell}_{jm}, q^{d}_{jm}\right) \equiv C_{jm}\left(q^{\ell}_{jm}, q^{d}_{jm}\right) +\ (r_0 + c_{j0})\ (q^{\ell}_{jm} - q^{d}_{jm})$. For the rest of the paper we do not include the term $(r_0 + c_{j0})\ (q^{\ell}_{jm} - q^{d}_{jm})$ explicitly in the variable cost function, but it should be understood that marginal costs include the component $r_0 + c_{j0}$ with positive sign for loans and negative for deposits.

Section 3.1 describes the demand system for deposits and loans. Section 3.2 presents our specification of bank variable costs. The equilibrium of the model is described in section 3.3. Comparative statics exercises and numerical solutions are provided in section 3.4.

3.1 Demand for deposit and loan products

(a) Demand for deposit products. There is a population of H^{d}_{m} savers in market m. Each saver has a fixed amount of wealth that we normalize to one unit.[13] A saver has to decide whether to deposit her unit of savings in a bank and at which bank. Due to transportation costs, savers consider only banks with branches in their own local market. In other words, banks can get deposits only in markets where they have branches. Banks provide differentiated deposit products. The (indirect) utility for a saver from depositing her wealth in bank j in market m is (we omit the individual-saver subindex in variables u^{d}_{jm} and ε^{d}_{jm}):

$$u^{d}_{jm} = \mathbf{x}^{d}_{jm}\ \beta^{d}_{m} - \alpha^{d}\ p^{d}_{jm} + \xi^{d}_{jm} + \varepsilon^{d}_{jm} \tag{3}$$

$\mathbf{x}^{d}_{jm}$ is a vector of characteristics of bank j (other than the deposit interest rate) that are valuable to depositors and observable to the researcher, such as the number of branches of bank j in the market, n_{jm}. The vector β^{d}_{m} contains the marginal utilities of the product characteristics $\mathbf{x}^{d}_{jm}$. These marginal utilities may vary across markets according to observable and unobservable (to the researcher) market characteristics, e.g., per capita income, age distribution, etc. Variable p^{d}_{jm} is the

[12]More precisely, we have that $B_j = S^{\ell}_j + Q^{\ell}_j - Q^{d}_j$, where S^{ℓ}_j is the stock of live pre-existing loans. However, S^{ℓ}_j is pre-determined and it does not have any effect on variable profits.

[13]See section 4 for a description of our measure of this 'unit' and of the number of consumers in the market, as well as our approach to deal with possible misspecification of these values.

price of deposit services (i.e., consumer fees minus the deposit interest rate), and α^d is the marginal utility of income. The term ξ^d_{jm} represents other characteristics of bank j in market m that are observable and valuable to savers but unobservable for us as researchers. Variable ε^d_{jm} represents savers' idiosyncratic preferences, and we assume that it is independently and identically distributed across banks with type 1 extreme value distribution. The utility from the outside alternative is normalized to zero. Let $s^d_{jm} \equiv q^d_{jm}/H^d_m$ be the market share of bank j in the market for deposits at location m. The model implies that:

$$s^d_{jm} = \frac{1\left\{m \in \mathcal{M}^d_j\right\} \;\; \exp\left\{\mathbf{x}^d_{jm} \; \beta^d_m - \alpha^d \; p^d_{jm} + \xi^d_{jm}\right\}}{1 + \sum_{k=1}^{B} 1\left\{m \in \mathcal{M}^d_k\right\} \;\; \exp\left\{\mathbf{x}^d_{km} \; \beta^d_m - \alpha^d \; p^d_{km} + \xi^d_{km}\right\}} \tag{4}$$

where $1\{.\}$ is the indicator function such that $1\left\{m \in \mathcal{M}^d_j\right\}$ is a dummy variable that indicates whether bank j has branches in market m.

The vector of product characteristics $\mathbf{x}^d_{jm}$ includes three elements that are important for the implications of the model: (i) the number of branches (n_{jm}); (ii) the bank's share of the local market for loans (s^ℓ_{jm}); and (iii) the bank's total amount of deposits (Q^d_j). The number of branches captures the effects of consumer transportation cost as well as consumer awareness about the bank's presence. The share in the local market of loans captures economies of scope at the consumer level of having deposits and loans in the same bank, i.e., one-stop banking. The bank's total deposits capture consumers' concerns for the probability of default or bank-run. Therefore, we have that,

$$\mathbf{x}^d_{jm}\beta^d_m = \beta^d_{0,m} + \beta^d_n \; h_n(n_{jm}) + \beta^d_\ell \; s^\ell_{jm} + \beta^d_Q \; h_Q(Q^d_j) \tag{5}$$

where $h_n(.)$ and $h_Q(.)$ are known monotonic functions, e.g., logarithm. We can also generalize this specification to incorporate the consumer valuation of a bank's number of branches in neighboring counties. We use the function $s^d_{jm} = d_{jm}(p^d_{jm}, s^\ell_{jm}, Q^d_j)$ to represent the demand for deposits, where, for notational convenience, we include explicitly as arguments the endogenous variables $(p^d_{jm}, s^\ell_{jm}, Q^d_j)$.

(b) Demand for loan products. Each local market is also populated by investors / borrowers. Let H^ℓ_m be the number of new borrowers in market m. Each (new) borrower is endowed with an investment project that requires 1 unit of loans. The set of possible choices that a borrower has is not limited to the banks that have branches in the market. There are banks that sell mortgages in the market but do not have physical branches, i.e., remember that $\mathcal{M}^d_j \subseteq \mathcal{M}^\ell_j$. However, borrowers may also value the geographic proximity of the bank as represented by the branches of the bank in the local market. Banks provide differentiated loan products. For a borrower located in market

m, the (indirect) utility of a loan from bank j is:

$$u_{jm}^{\ell} = \mathbf{x}_{jm}^{\ell}\ \beta_m^{\ell} - \alpha^{\ell}\ p_{jm}^{\ell} + \xi_{jm}^{\ell} + \varepsilon_{jm}^{\ell} \tag{6}$$

The variables and parameters in this utility function have a similar interpretation as in the utility for deposits presented above. Variable p_{jm}^{ℓ} represents the interest rate of a loan from bank j in market m. We also assume that the variables ε_{jm}^{ℓ} are identically distributed across banks with type 1 extreme value distribution, and the utility from the outside alternative is normalized to zero. Let $s_{jm}^{\ell} \equiv q_{jm}^{\ell}/H_m^{\ell}$ be the market share of bank j in the market for loans at location m. According to the model, we have that:

$$s_{jm}^{\ell} = \frac{1\left\{m \in \mathcal{M}_j^{\ell}\right\}\ \exp\left\{\mathbf{x}_{jm}^{\ell}\ \beta_m^{\ell} - \alpha^{\ell}\ p_{jm}^{\ell} + \xi_{jm}^{\ell}\right\}}{1 + \sum_{k=1}^{B} 1\left\{m \in \mathcal{M}_k^{\ell}\right\}\ \exp\left\{\mathbf{x}_{km}^{\ell}\ \beta_m^{\ell} - \alpha^{\ell}\ p_{km}^{\ell} + \xi_{km}^{\ell}\right\}} \tag{7}$$

As was the case for deposits, the vector of product characteristics $\mathbf{x}_{jm}^{\ell}$ includes: (i) the number of branches (n_{jm}); (ii) the bank's share of the local market for deposits (s_{jm}^{d}); and (iii) the bank's total amount of deposits in all the markets (Q_j^d). As explained above for the demand for deposits, the number of branches captures consumer transportation cost and consumer awareness, and the amount of local deposits portrays economies of scope between deposits and loans for the consumer if using the same bank. Consumers value a bank's total amount of deposits because it is related to the bank's risk of liquidity shortage and default. Thus, we have that

$$\mathbf{x}_{jm}^{\ell}\beta_m^{\ell} = \beta_{0m}^{\ell} + \beta_n^{\ell}\ h_n(n_{jm}) + \beta_d^{\ell}\ s_{jm}^{d} + \beta_Q^{\ell}\ h_Q(Q_j^d). \tag{8}$$

We use the function $s_{jm}^{\ell} = \ell_{jm}(p_{jm}^{\ell}, s_{jm}^{d}, Q_j^d)$ to represent the demand for loans.

(c) Demand system for deposits and loans. The demand system can be represented by the equations $s_{jm}^{\ell} = \ell_{jm}(p_{jm}^{\ell}, s_{jm}^{d}, Q_j^d)$ and $s_{jm}^{d} = d_{jm}(p_{jm}^{d}, s_{jm}^{\ell}, Q_j^d)$. For the moment, let us consider this demand system for a single bank, taking as given prices of loans and deposits for the rest of the banks. This system establishes links between the amount of deposits and loans in the same local market and across different geographic markets. Taking prices as given, the solution of this system of equations with respect to the markets shares $\{s_{jm}^{\ell}, s_{jm}^{d}\}$ implies the *reduced form demand system*:

$$s_{jm}^{d} = f_{jm}^{d}\left(\mathbf{p}_j^d, \mathbf{p}_j^{\ell}\right) \quad \text{and} \quad s_{jm}^{\ell} = f_{jm}^{\ell}\left(\mathbf{p}_j^d, \mathbf{p}_j^{\ell}\right) \tag{9}$$

where $\mathbf{p}_j^d$ and $\mathbf{p}_j^{\ell}$ are the vectors with bank j's interests rates for deposits and loans, respectively, at every local market where this bank is active. Loans (deposits) in a local market depend on the bank's interest rates for loans and deposits in every market where the bank operates. Therefore, the demand-price derivatives, say $\partial f_{jm}^{d}/\partial p_{jm}^{\ell}$ or $\partial f_{jm}^{\ell}/\partial p_{jm}^{d}$, incorporate *local and global multiplier effects.* For instance, taking into account that $s_{jm}^{\ell}=\ell_{jm}(p_{jm}^{\ell},s_{jm}^{d},Q_{j}^{d})$ and $s_{jm}^{d}=d_{jm}(p_{jm}^{d},s_{jm}^{\ell},Q_{j}^{d})$, we have the following system of equations:

$$\begin{aligned}
\frac{\partial f_{jm}^{\ell}}{\partial p_{jm}^{\ell}} &= \frac{\partial \ell_{jm}}{\partial p_{jm}^{\ell}}+\frac{\partial \ell_{jm}}{\partial s_{jm}^{d}}\frac{\partial f_{jm}^{d}}{\partial p_{jm}^{\ell}}+\frac{\partial \ell_{jm}}{\partial Q_{j}^{d}}\left[\sum_{m'\in\mathcal{M}_{j}^{d}}\frac{\partial f_{jm'}^{d}}{\partial p_{jm}^{\ell}}\right] \\
\frac{\partial f_{jm}^{d}}{\partial p_{jm}^{\ell}} &= \frac{\partial d_{jm}}{\partial s_{jm}^{\ell}}\frac{\partial f_{jm}^{\ell}}{\partial p_{jm}^{\ell}}+\frac{\partial d_{jm}}{\partial Q_{j}^{d}}\left[\sum_{m'\in\mathcal{M}_{j}^{d}}\frac{\partial f_{jm'}^{d}}{\partial p_{jm}^{\ell}}\right] \\
\frac{\partial f_{jm'}^{d}}{\partial p_{jm}^{\ell}} &= \frac{\partial d_{jm'}}{\partial s_{jm'}^{\ell}}\frac{\partial f_{jm'}^{\ell}}{\partial p_{jm}^{\ell}}+\frac{\partial d_{jm'}}{\partial Q_{j}^{d}}\left[\sum_{m'\in\mathcal{M}_{j}^{d}}\frac{\partial f_{jm'}^{d}}{\partial p_{jm}^{\ell}}\right] \quad \text{for } m'\neq m \\
\frac{\partial f_{jm'}^{\ell}}{\partial p_{jm}^{\ell}} &= \frac{\partial \ell_{jm'}}{\partial s_{jm'}^{d}}\frac{\partial f_{jm'}^{d}}{\partial p_{jm}^{\ell}}+\frac{\partial \ell_{jm'}}{\partial Q_{j}^{d}}\left[\sum_{m'\in\mathcal{M}_{j}^{d}}\frac{\partial f_{jm'}^{d}}{\partial p_{jm}^{\ell}}\right] \quad \text{for } m'\neq m
\end{aligned} \tag{10}$$

This is a system of linear equations in the vector of partial derivatives $\{\partial f_{jm'}^{\ell}/\partial p_{jm}^{\ell};\, \partial f_{jm'}^{d}/\partial p_{jm}^{\ell}$: for $m'\in\mathcal{M}_{j}\}$, where $\mathcal{M}_{j}\equiv\mathcal{M}_{j}^{d}\cup\mathcal{M}_{j}^{\ell}$. Solving this linear system we can obtain this vector in terms of the derivatives of the structural demand functions ℓ_{jm} and d_{jm}. The solution to this system implicitly implies the existence of local and global multiplier effects to the changes in local interest rates. More formally, let $\partial\mathbf{f}_{j,(p_{jm}^{\ell})}^{\ell}$ be the $|\mathcal{M}_{j}^{\ell}|\times 1$ vector of partial derivatives $\{\partial f_{jm'}^{\ell}/\partial p_{jm}^{\ell}:m'\in\mathcal{M}_{j}^{\ell}\}$, and similarly, let $\partial\mathbf{f}_{j,(p_{jm}^{\ell})}^{d}$ be the $|\mathcal{M}_{j}^{d}|\times 1$ vector $\{\partial f_{jm'}^{d}/\partial p_{jm}^{\ell}:m'\in\mathcal{M}_{j}^{\ell}\}$. The system of equations (10) implies the following *solution* for $[\partial\mathbf{f}_{j,(p_{jm}^{\ell})}^{\ell},\,\partial\mathbf{f}_{j,(p_{jm}^{\ell})}^{d}]$ in terms of derivatives of the structural demand functions:

$$\begin{bmatrix}\partial\mathbf{f}_{j,(p_{jm}^{\ell})}^{\ell}\\ \partial\mathbf{f}_{j,(p_{jm}^{\ell})}^{d}\end{bmatrix}=\begin{bmatrix}\mathbf{I}- & \mathbf{0}_{|\mathcal{M}_{j}^{\ell}|\times|\mathcal{M}_{j}^{\ell}|}\ ; & \mathbf{A}_{j,(p_{jm}^{\ell})}\\ & \mathbf{B}_{j,(p_{jm}^{\ell})} & \mathbf{C}_{j,(p_{jm}^{\ell})}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{i}_{|\mathcal{M}_{j}^{\ell}|}^{(m)}\ \partial\ell_{jm}/\partial p_{jm}^{\ell}\\ \mathbf{0}_{|\mathcal{M}_{j}^{d}|}\end{bmatrix} \tag{11}$$

where $\mathbf{I}$ is the identify matrix; $\mathbf{0}$ is a matrix of zeros; $\mathbf{i}_{|\mathcal{M}_{j}^{\ell}|}^{(m)}$ is a vector with 1 at the m-th element and zeroes elsewhere; and $\mathbf{A}_{j,(p_{jm}^{\ell})}$, $\mathbf{B}_{j,(p_{jm}^{\ell})}$, and $\mathbf{C}_{j,(p_{jm}^{\ell})}$ are matrices with the following definitions:

$$\begin{aligned}
\mathbf{A}_{j,(p_{jm}^{\ell})} &\equiv diag\left\{\partial\ell_{j,s^{d}}\right\}+diag\left\{\partial\ell_{j,Q^{d}}\right\}\ \mathbf{1}_{|\mathcal{M}_{j}^{\ell}|\times|\mathcal{M}_{j}^{d}|}\\
\mathbf{B}_{j,(p_{jm}^{\ell})} &\equiv diag\left\{\partial\mathbf{d}_{j,s^{\ell}}\right\}\\
\mathbf{C}_{j,(p_{jm}^{\ell})} &\equiv diag\left\{\partial\mathbf{d}_{j,Q^{d}}\right\}\ \mathbf{1}_{|\mathcal{M}_{j}^{d}|\times|\mathcal{M}_{j}^{d}|}
\end{aligned} \tag{12}$$

$diag\{\mathbf{v}\}$ is a diagonal matrix with vector $\mathbf{v}$ in the diagonal; $\mathbf{1}$ is a matrix of ones; $\partial\ell_{j,s^{d}}$ is the $|\mathcal{M}_{j}^{\ell}|\times 1$ vector with elements $\partial\ell_{jm}/\partial s_{jm}^{d}$; $\partial\ell_{j,Q^{d}}$ is the $|\mathcal{M}_{j}^{\ell}|\times 1$ vector with elements $\partial\ell_{jm}/\partial Q_{j}^{d}$; and similarly $\partial\mathbf{d}_{j,s^{\ell}}$ and $\partial\mathbf{d}_{j,Q^{d}}$ are the $|\mathcal{M}_{j}^{d}|\times 1$ vectors with elements $\partial d_{jm}/\partial s_{jm}^{\ell}$ and $\partial d_{jm}/\partial Q_{j}^{d}$, respectively.

3.2 Variable cost function

We consider the following specification for the variable cost function:

$$\begin{aligned} C_{jm}\left(q^{\ell}_{jm}, q^{d}_{jm}\right) &= \left[\omega^{d}_{jm} + \gamma^{d}_{n}\ h_{n}(n_{jm}) + \gamma^{d}_{\ell}\ s^{\ell}_{jm} + \gamma^{d}_{Q}\ h_{Q}(Q^{d}_{j})\right]\ q^{d}_{jm} \\ &+ \left[\omega^{\ell}_{jm} + \gamma^{\ell}_{n}\ h_{n}(n_{jm}) + \gamma^{\ell}_{d}\ s^{d}_{jm} + \gamma^{\ell}_{Q}\ h_{Q}(Q^{d}_{j})\right]\ q^{\ell}_{jm} \end{aligned} \tag{13}$$

The terms $\gamma^{d}_{n}\ h_{n}(n_{jm})$ and $\gamma^{\ell}_{n}\ h_{n}(n_{jm})$ portray economies of scale and scope between branches of a bank in the same market. Some costs of providing deposits and loans are shared by multiple branches. The terms $\gamma^{d}_{\ell}\ s^{\ell}_{jm}\ q^{d}_{jm}$ and $\gamma^{\ell}_{d}\ s^{d}_{jm}\ q^{\ell}_{jm}$ capture economies of scope in the management of deposits at the branch level. The component $\gamma^{\ell}_{Q}\ h_{Q}(Q^{d}_{j})$ captures how the marginal cost of loans declines with the bank's total volume of deposits Q^{d}_{j}.

3.3 Bank competition and equilibrium

A bank can charge a different interest rate for deposits (loans) at each local market. We assume that banks compete a la Nash-Bertrand. Therefore, each bank chooses its vectors of interest rates for deposits and loans, $\mathbf{p}_{j} \equiv \{p^{d}_{jm} : m \in \mathcal{M}^{d}_{j};\ p^{\ell}_{jm} : m \in \mathcal{M}^{\ell}_{j}\}$, to maximize its profit.

A marginal change in the interest rate of deposits of bank j in county m has the following effects on the bank's profit: (i) the standard marginal revenue and marginal cost effect from deposits in the same market; (ii) the indirect effect on the profits from loans in the same local market; (iii) the indirect effect on the profits from deposits in other geographic markets where the bank operates; and similarly, (iv) the indirect effect on the profits from loans in other geographic markets. That is,

$$\begin{aligned} &\underbrace{\left[q^{d}_{jm} + \left(p^{d}_{jm} - \frac{\partial C_{jm}}{\partial q^{d}_{jm}}\right)\frac{\partial f^{d}_{jm}}{\partial p^{d}_{jm}}\right]}_{(i)} + \underbrace{\left(p^{\ell}_{jm} - \frac{\partial C_{jm}}{\partial q^{\ell}_{jm}}\right)\frac{\partial f^{\ell}_{jm}}{\partial p^{d}_{jm}}}_{(ii)} \\ &+ \underbrace{\sum_{m' \neq m}\left(p^{d}_{jm'} - \frac{\partial C_{jm'}}{\partial q^{d}_{jm'}}\right)\frac{\partial f^{d}_{jm'}}{\partial p^{d}_{jm}}}_{(iii)} + \underbrace{\sum_{m' \neq m}\left(p^{\ell}_{jm'} - \frac{\partial C_{jm'}}{\partial q^{\ell}_{jm'}}\right)\frac{\partial f^{\ell}_{jm'}}{\partial p^{d}_{jm}}}_{(iv)} = 0 \end{aligned} \tag{14}$$

And we have a similar expression for the marginal conditional of optimality with respect to the interest rate of loans. This set of marginal conditions of optimality for every bank j and every geographic market m determine the equilibrium of the model.

Using the logit structure of the demands for loans and deposits, we now develop expressions that characterize the Bertrand equilibrium and that we use for the estimation of the model parameters and for our counterfactual experiments. Under the logit specification of demand, the system of marginal conditions of optimality imply the following pricing equations:

$$\begin{aligned} p_{jm}^{\ell} &= c_{jm}^{\ell} + \Delta_j^{\ell} + \frac{1}{\alpha^{\ell}(1 - s_{jm}^{\ell})} - \frac{\beta_{\ell}^{d}}{\alpha^{d}} s_{jm}^{d} \\ p_{jm}^{d} &= c_{jm}^{d} + \Delta_j^{d} + \frac{1}{\alpha^{d}(1 - s_{jm}^{d})} - \frac{\beta_{d}^{\ell}}{\alpha^{\ell}} s_{jm}^{\ell} \end{aligned} \tag{15}$$

where c_{jm}^{ℓ} and c_{jm}^{d} represent marginal costs, and Δ_j^{ℓ} and Δ_j^{d} are terms that depend on marginal costs and demand aggregated over all the markets where bank j operates.

For our empirical analysis, it is convenient to write the equilibrium conditions in terms of the market shares as the only endogenous variables. Let s_{0m}^{d} and s_{0m}^{ℓ} be the market shares of the outside alternative for deposits and loans in market m. By definition, $s_{0m}^{d} = H_m^d - \sum_{j=1}^{B} s_{jm}^{d}$ and $s_{0m}^{\ell} = H_m^{\ell} - \sum_{j=1}^{B} s_{jm}^{\ell}$. The logit model implies that $\ln(s_{jm}^{d}/s_{0m}^{d}) = \mathbf{x}_{jm}^{d}\ \beta_m^d + \alpha^d\ p_{jm}^{d} + \xi_{jm}^{d}$. Subbing the pricing equations into this expression, we obtain the following system of equilibrium equations in terms of market shares:

$$\begin{aligned} y\left(s_{jm}^{d}, s_{0m}^{d}\right) &= \mathbf{x}_{jm}^{d}\beta_m^d - \alpha^d \left[c_{jm}^{d} + \Delta_j^{d}\right] + \frac{\alpha^d \beta_d^{\ell}}{\alpha^{\ell}} s_{jm}^{\ell} + \xi_{jm}^{d} \\ y\left(s_{jm}^{\ell}, s_{0m}^{\ell}\right) &= \mathbf{x}_{jm}^{\ell}\beta_m^{\ell} - \alpha^{\ell} \left[c_{jm}^{\ell} + \Delta_j^{\ell}\right] + \frac{\alpha^{\ell} \beta_{\ell}^{d}}{\alpha^{d}} s_{jm}^{d} + \xi_{jm}^{\ell} \end{aligned} \tag{16}$$

where, for any value of the shares (s_j, s_0), the function $y(s_j, s_0)$ is defined as $\ln\left(\frac{s_j}{s_0}\right) + \frac{1}{1 - s_j}$.

Given the structure of demand and marginal costs with $\mathbf{x}_{jm}^{d}\beta_m^d = \beta_{0,m}^d + \beta_n^d\ h_n(n_{jm}) + \beta_{\ell}^d\ q_{jm}^{\ell} + \beta_Q^d\ h_Q(Q_j^d)$ and $c_{jm}^{d} = \omega_{jm}^{d} + \gamma_n^d\ h_n(n_{jm}) + \gamma_{\ell}^d\ s_{jm}^{\ell} + \gamma_Q^d\ h_Q(Q_j^d)$, it is clear that the term $\mathbf{x}_{jm}^{d}\beta_m^d - \alpha^d\ c_{jm}^{d}$ is equal to:

$$\mathbf{x}_{jm}^{d}\beta_m^d - \alpha^d c_{jm}^{d} = \theta_n^d\ h_n(n_{jm}) + \theta_{\ell}^d\ s_{jm}^{\ell} + \theta_Q^d\ h_Q(Q_j^d) \tag{17}$$

where the θ's are structural parameters that depend on both demand and marginal cost parameters. More specifically, $\theta_n^d \equiv \beta_n^d - \alpha^d\ \gamma_n^d$, $\theta_{\ell}^d \equiv \beta_{\ell}^d - \alpha^d\ \gamma_{\ell}^d$, and $\theta_Q^d \equiv \beta_Q^d - \alpha^d\ \gamma_Q^d$, and we have similar expressions for θ_n^{ℓ}, θ_d^{ℓ}, and θ_Q^{ℓ} in the loan equation. Similarly, we have that the "error terms" in the deposit and loan equations depend on both demand and cost shocks: $\eta_{jm}^{d} \equiv \xi_{jm}^{d} - \alpha^d\ \omega_{mt}^{d}$ and $\eta_{jm}^{\ell} \equiv \xi_{jm}^{\ell} - \alpha^{\ell}\ \omega_{mt}^{\ell}$.

The parameters θ, together with the parameters $\alpha^d\beta_d^{\ell}/\alpha^{\ell}$ and $\alpha^{\ell}\beta_{\ell}^{d}/\alpha^d$ and the exogenous variables of the model, contain all the information that we need to construct the equilibrium mapping of the model and obtain an equilibrium. Given this model structure, we do not need to separately identify demand and cost parameters. All our empirical results are based on the estimation of these parameters and the implementation of counterfactual experiments using the equilibrium mapping.

4 Estimation of the structural model

4.1 Specification assumptions

We complete the specification of the econometric model with the following functional form assumptions. First, the specification of function $h_n(n_{jm})$ is completely nonparametric. Therefore, with some abuse of notation, we use the parameters $\{\theta_n^d(n_j) : \text{for } n_j = 1, 2, ..., n^{\max}\}$ to represent the terms $\theta_n^d\ h_n(n_j)$. Similarly, we define the parameters $\{\theta_n^\ell(n_j) : \text{for } n_j = 1, 2, ..., n^{\max}\}$ in the equation for loans. We consider a logarithmic specification for the function h_Q, i.e., $h_Q(Q_j^d) = \ln(Q_j^d)$. We define the parameters $\widetilde{\theta}_\ell^d \equiv \theta_\ell^d + \alpha^d \beta_d^\ell / \alpha^\ell$ and $\widetilde{\theta}_d^\ell \equiv \theta_d^\ell + \alpha^\ell \beta_\ell^d / \alpha^d$ that capture economies between deposits and loans at the bank-county level.

For the unobservables η_{jmt}^d and η_{jmt}^ℓ, we consider the following component structure specification:

$$\eta_{jmt}^d = \eta_{jm}^{d(1)} + \eta_t^{d(2)} + \eta_{jt}^{d(3)} + \eta_{jmt}^{d(4)} \tag{18}$$

$\eta_{jm}^{d(1)}$ represents bank-county fixed effects (i.e., bank-county dummies); $\eta_t^{d(2)}$ is an aggregate national shock (i.e., time dummies); $\eta_{jt}^{d(3)}$ represents an bank-specific national level shock; and $\eta_{jmt}^{d(4)}$ is a bank-county-specific shock. The error term in the loan equation has the same structure.

Given these specification assumptions, the equations of the econometric model are:

$$\begin{aligned} y_{jmt}^d &= \mathbf{z}_{mt}'\ \delta^d + \sum_{n=1}^{n_{\max}} \theta_n^d(n)\ 1_{jmt}(n) + \widetilde{\theta}_\ell^d\ s_{jmt}^\ell + \theta_Q^d\ \ln(Q_{jt}^d) + \eta_{jmt}^d \\ y_{jmt}^\ell &= \mathbf{z}_{mt}'\ \delta^\ell + \sum_{n=1}^{n_{\max}} \theta_n^\ell(n)\ 1_{jmt}(n) + \widetilde{\theta}_d^\ell\ s_{jmt}^d + \theta_Q^\ell\ \ln(Q_{jt}^d) + \eta_{jmt}^\ell \end{aligned} \tag{19}$$

where $y_{jmt}^d \equiv y\left(s_{jmt}^d, s_{0mt}^d\right)$, $y_{jmt}^\ell \equiv y\left(s_{jmt}^\ell, s_{0mt}^\ell\right)$, $1_{jmt}(n) \in \{0,1\}$ is the binary variable that indicates that the number of branches n_{jmt} is equal to n, and $\mathbf{z}_{mt}$ is a vector of market characteristics that captures the relative value of the outside alternative. More specifically, $\mathbf{z}_{mt}$ includes a housing price index and its growth, bankruptcy cases, population density, and age distribution.

Market size and market shares for deposits and loans. To construct market shares we need first to construct market size variables H_{mt}^d and H_{mt}^ℓ. We have used the following approach. First, we postulate that the the market sizes H_{mt}^d and H_{mt}^ℓ are proportional to the total population in county

m at period t:[14]

$$H^d_{mt} = \lambda^d \ POP_{mt} \quad \text{and} \quad H^\ell_{mt} = \lambda^\ell \ POP_{mt}$$

where λ^d and λ^ℓ are positive constants and POP_{mt} is total population in county m at period t. Constants λ^d and λ^ℓ are chosen such that the the constructed market shares satisfy the model constraint that the sum of the market shares $\sum_{m=1}^M s^d_{jmt} = Q^d_{mt}/H^d_{mt}$ and $\sum_{m=1}^M s^\ell_{jmt} = Q^\ell_{mt}/H^\ell_{mt}$ are always smaller than one. More specifically, $\lambda^d = \max_{m,t} \left\{ \frac{Q^d_{mt}}{POP_{mt}} \right\}$ and $\lambda^\ell = \max_{m,t} \left\{ \frac{Q^\ell_{mt}}{POP_{mt}} \right\}$. These constants are $\lambda^d = 547$ (thousand USD) and $\lambda^\ell = 103$ (thousand USD).

Admittedly, using POP_{mt} as a measure of market size, and assuming that λ^d and λ^ℓ are constant across counties and over time, seem like strong restrictions. To control for measurement error in this measure of market size, we include socioeconomic characteristics at the county-level as explanatory variables in the model. The effect of these socioeconomic characteristics in the model can be interpreted as corrections for market size.

4.2 Identification restrictions

In the structural equations in (19), regressors s^ℓ_{jmt}, s^d_{jmt}, and $\ln(Q^d_{jt})$ are endogenous variables of the model, and therefore they are correlated with the error terms η^d_{jmt} and η^ℓ_{jmt} because of simultaneity. Furthermore, though the number branches n_{jmt} is not an endogenous variable in our structural model, we expect this variable to depend also on the supply and demand shocks in deposits and loan markets. Therefore, the number of branches is also an endogenous variable in the econometric model.

Assumptions ID-1 and ID-2 describe the restrictions that we impose in the model to deal with endogeneity and to identify the θ parameters.

Assumption ID-1: Market characteristics $\mathbf{z}_{mt}$ are strictly exogenous regressors with respect to the bank-county-specific shocks $\eta^{d(4)}_{jmt}$ and $\eta^{\ell(4)}_{jmt}$, i.e., for any pair of markets (m, m') and any pair of years (t, t'), we have that $\mathbb{E}\left(\mathbf{z}_{mt} \ \eta^{d(4)}_{jm't'}\right) = 0$ and $\mathbb{E}\left(\mathbf{z}_{mt} \ \eta^{\ell(4)}_{jm't'}\right) = 0$. $\blacksquare$

Assumption ID-2: The market specific component of market characteristics, i.e., $\mathbf{z}^*_{mt} \equiv \mathbf{z}_{mt} - \mathbb{E}(\mathbf{z}_{mt}|t)$, are not correlated with the bank-specific national level shocks $\eta^{d(3)}_{jt}$ and $\eta^{\ell(3)}_{jt}$, i.e., for any market m and bank j, we have that $\mathbb{E}\left(\mathbf{z}^*_{mt} \ \eta^{d(3)}_{jt}\right) = 0$ and $\mathbb{E}\left(\mathbf{z}^*_{mt} \ \eta^{\ell(3)}_{jt}\right) = 0$. $\blacksquare$

Assumption ID-3: The bank-county-specific shocks $\eta^{d(4)}_{jmt}$ and $\eta^{\ell(4)}_{jmt}$ are not serially correlated. $\blacksquare$

[14] We have also tried total county income, instead of county population. Our empirical results are very robust to this.

These three assumptions provide moment conditions that identify all the θ parameters. We can distinguish two sets of moment restrictions. Consider a bank that is active in at least two markets, say m and m^*. Then, we can construct a *difference-in-difference* transformation of the equations in (19): a difference between the equations for markets m and m^*, that removes the components $\eta_t^{d(2)}$ and $\eta_{jt}^{d(3)}$ from the error term; a time difference between the equations at periods t and $t-1$, that eliminates the bank-county fixed effects $\eta_{jm}^{d(1)}$. The transformed equations become:

$$
\begin{aligned}
\Delta \widetilde{y}_{jmt}^{d} &= \Delta \widetilde{\mathbf{z}}_{mt}' \ \delta^{d} + \sum_{n=1}^{n_{\max}} \theta_n^d(n) \ \Delta \widetilde{1}_{jmt}(n) + \widetilde{\theta}_{\ell}^{d} \ \Delta \widetilde{s}_{jmt}^{\ell} + \Delta \widetilde{\eta}_{jmt}^{d(4)} \\
\Delta \widetilde{y}_{jmt}^{\ell} &= \Delta \widetilde{\mathbf{z}}_{mt}' \ \delta^{\ell} + \sum_{n=1}^{n_{\max}} \theta_n^{\ell}(n) \ \Delta \widetilde{1}_{jmt}(n) + \widetilde{\theta}_{d}^{\ell} \ \Delta \widetilde{s}_{jmt}^{d} + \Delta \widetilde{\eta}_{jmt}^{\ell(4)}
\end{aligned}
\tag{20}
$$

where the $\sim$ symbol represents the difference between markets transformation, e.g., $\widetilde{y}_{jmt}^{d} \equiv y_{jmt}^{d} - y_{jm^*t}^{d}$, and Δ represents the time difference transformation, e.g., $\Delta \widetilde{y}_{jmt}^{d} \equiv [y_{jmt}^{d} - y_{jm^*t}^{d}] - [y_{jm,t-1}^{d} - y_{jm^*,t-1}^{d}]$. Note that the difference between markets also eliminates the effect of the aggregate volume of deposits, $\ln(Q_{jt}^{d})$, such that parameters θ_Q^{ℓ} and θ_Q^{d} cannot be identified from these equations. However, Assumptions ID-1 and ID-3 imply the following moment conditions:

$$
\begin{aligned}
\mathbb{E}\left(\begin{bmatrix} \mathbf{z}_{mt} \\ \mathbf{x}_{jm,t-s} \end{bmatrix} \ \Delta \widetilde{\eta}_{jmt}^{d(4)} \right) &= 0 \\
\mathbb{E}\left(\begin{bmatrix} \mathbf{z}_{mt} \\ \mathbf{x}_{jm,t-s} \end{bmatrix} \ \Delta \widetilde{\eta}_{jmt}^{\ell(4)} \right) &= 0
\end{aligned}
\tag{21}
$$

for $s \geq 2$, and where $\mathbf{x}_{jmt}$ represents the vector of bank-county specific endogenous variables, $\{y_{jmt}^{d}$, y_{jmt}^{ℓ}, s_{jmt}^{d}, s_{jmt}^{ℓ}, $n_{jmt}\}$. These moment conditions identify the parameters δ^d, δ^{ℓ}, $\theta_n^d(n)$, $\theta_n^{\ell}(n)$, $\widetilde{\theta}_{\ell}^{d}$, and $\widetilde{\theta}_{d}^{\ell}$.

Consider now another type of differences-in-differences transformation of the equations in (19): we still apply the time difference between the equations at periods t and $t-1$ to eliminate bank-county fixed effects $\eta_{jm}^{d(1)}$; and we apply difference with respect to the average equation at period t, i.e., $y_{jmt}^{d} - \mathbb{E}(y_{jmt}^{d}|t)$, that eliminates the aggregate national level shocks $\eta_t^{d(2)}$ from the error term. The transformed equations become:

$$
\begin{aligned}
\Delta y_{jmt}^{*d} &= \Delta \mathbf{z}_{mt}^{*\prime} \ \gamma^{d} + \sum_{n=1}^{n_{\max}} \theta_n^d(n) \ \Delta 1_{jmt}^{*}(n) + \widetilde{\theta}_{\ell}^{d} \ \Delta s_{jmt}^{*\ell} + \theta_Q^{d} \ \Delta \ln(Q_{jt}^{*d}) + \left[\Delta \eta_{jt}^{*d(3)} + \Delta \eta_{jmt}^{*d(4)} \right] \\
\Delta y_{jmt}^{*\ell} &= \Delta \mathbf{z}_{mt}^{*\prime} \ \gamma^{\ell} + \sum_{n=1}^{n_{\max}} \theta_n^{\ell}(n) \ \Delta 1_{jmt}^{*}(n) + \widetilde{\theta}_{d}^{\ell} \ \Delta s_{jmt}^{*d} + \theta_Q^{\ell} \ \Delta \ln(Q_{jt}^{*d}) + \left[\Delta \eta_{jt}^{*\ell(3)} + \Delta \eta_{jmt}^{*\ell(4)} \right]
\end{aligned}
\tag{22}
$$

where the star $*$ represents the difference with respect to the national mean, e.g., $y_{jmt}^{*d} \equiv y_{jmt}^{d} - \overline{y}_{t}^{d}$, and $\overline{y}_{t}^{d} \equiv (NM_t)^{-1} \sum_j \sum_m y_{jmt}^{d}$. Assumption ID-1 and ID-2 imply the following moment conditions:

$$\begin{array}{rcl}\mathbb{E}\left(\left[\sum_{m\in\mathcal{M}_{jt}^{d}}\mathbf{z}_{mt}^{*}\right]\left[\Delta\eta_{jt}^{*d(3)}+\Delta\eta_{jmt}^{*d(4)}\right]\right) & = & 0\\ \mathbb{E}\left(\left[\sum_{m\in\mathcal{M}_{jt}^{d}}\mathbf{z}_{mt}^{*}\right]\left[\Delta\eta_{jt}^{*\ell(3)}+\Delta\eta_{jmt}^{*\ell(4)}\right]\right) & = & 0\end{array} \tag{23}$$

These moment conditions identify the parameters θ_Q^d and θ_Q^ℓ. Intuitively, these moment conditions imply that we can use the exogenous characteristics at markets other than m where the bank in active in the deposit market, i.e., $\{\mathbf{z}_{m't}^*$ for $m' \neq m$ with $m' \in \mathcal{M}_{jt}^d\}$, to instrument the total amount of deposits $\ln(Q_{jt}^d)$. The characteristics at other markets do not have a direct effect in the structural equation for market m, i.e., they satisfy a exclusion restriction. By Assumption ID-2, they are not correlated with the error term $\Delta\eta_{jt}^{*d(3)} + \Delta\eta_{jmt}^{*d(4)}$, i.e., they are valid instruments. Furthermore, the model implies that these characteristics should have an effect on the total volume of deposits of the bank, i.e., they are relevant instruments.

We jointly estimate all the parameters of the model using a GMM estimator in the spirit of those in the dynamic panel data literature (Arellano And Bond, 1991, Arellano and Bover, 1995, and Blundell and Bond, 1999). We apply a two-step optimal GMM estimator and obtain standard errors robust of heterocedasticity and serial correlation.

5 Estimation results

Table 2 presents the results from the GMM estimation of deposit and loan equilibrium equations. As shown in Table 1 above, banks provide loans in many more counties than they obtain deposits. As a result, the number of observations in the estimation of the loan equations is almost ten times the number of observations in the estimation of the deposit equation, i.e., $1,224,465$ versus $132,265$. In general, the estimated parameters are quite precise.

By construction, the right-hand side of the equilibrium equations expressed in (19) represents consumer willingness-to-pay net of marginal cost. In fact, it is equal to the social value of the products at the bank-county-year level, relative to the value of the outside alternative. For convenience, we refer to these values as the *net willingness-to-pay* (or *net-wtp*). The parameters θ capture the causal effect of different variables on the *net-wtp.*

Unfortunately, the net-wtp and the θ parameters are not measured in monetary units (dollars) but in utils. Furthermore, the θ parameters are not directly comparable because they are measured in different util units since the variance of extreme value unobservables can be different in the demands for loans and deposits, i.e., α^d and α^ℓ can be different.

Nevertheless, the dependent variables in the left-hand-side of the equilibrium equations are very close to logarithm of county-level market shares, $\ln(s_{jmt}^d)$ and $\ln(s_{jmt}^\ell)$. Therefore, we make some comparisons between the θ parameters of the two equations by interpreting these parameters as elasticities (if the explanatory variable is also in logarithms) or semi-elasticities.

(i) Number of branches. The number of branches in the county has a very substantial effect on the net-wtp for a deposit product. The marginal effect of an additional branch declines with the number of branches: a second branch increases the net-wtp / log-share by 51%; a third branch by 26%; a fourth branch by 18%; a fifth branch 20%; and subsequent branches by (on average) 4%. The effect of the number of branches on the net-wtp / log-share of a loan product is also important, but substantially smaller than for deposits: a second branch increases the net-wtp / log-share by 19%; a third branch by 8%; a fourth branch by 8%; a fifth branch 7%; and subsequent branches do not have any additional effect on loans.

In the data and in our model, a bank needs at least one branch in the county to obtain deposits. That is not case for loans. Therefore, we can identify the effect of the first branch on the net-wtp for a loan product. The estimate is 101%, i.e., the first branch doubles the demand for a loan product. This effect, though substantial, also shows that the demand for loan products at banks without branches is still half of the demand at banks with one branch.

(ii) Economies of scope between deposits and loans at the county level. We identify significant economies of scope between deposits and loans. Doubling the amount of deposits of a bank in a county implies an 8% increase in the net-wtp / market share of the bank's loans in the same market. The elasticity of deposits with respect to loans is smaller but still significant, i.e., 2%.

(iii) Effect of total deposits. A bank's amount of deposits at the national level has a very substantial effect on the bank's net-wtp / log-share of product loans at every local market where it operates: a 100% increase in a bank's total deposits implies a 32% increase in the market share for loans at every county. This provides strong evidence that banks' internal liquidity facilitates lending.

(iv) County characteristics. Income per-capita, the housing price index, and the number of bankruptcy filings all have substantial effects on the value of a loan product relative to the outside alternative. The effect of the housing price index, with an elasticity of 1.08, is particularly important. As expected, bankruptcy filings have a negative and significant effect, with an elasticity of -0.08.

Table 2
GMM Estimation of Equilibrium equations
Sample Period: 1998-2010

Variable	(1) Deposits	(3) Loans
Number of branches		
First branch ($1\{n_{jmt} \geq 1\}$)	-	1.0116***
		(0.1150)
Second branch ($1\{n_{jmt} \geq 2\}$)	0.5076***	0.1907***
	(0.0144)	(0.0199)
Third branch ($1\{n_{jmt} \geq 3\}$)	0.2564***	0.0871***
	(0.0112)	(0.0211)
Fourth branch ($1\{n_{jmt} \geq 4\}$)	0.1813***	0.0843***
	(0.0136)	(0.0230)
Fifth branch ($1\{n_{jmt} \geq 5\}$)	0.2048***	0.0691***
	(0.0136)	(0.0257)
# of branches in county above 5th	0.0531***	0.0039
	(0.0040)	(0.0035)
Econ. of scope and total depo		
log own loans in county	0.0198***	-
	(0.0017)	
log own deposits in county	-	0.0778***
		(0.0086)
log total own deposits	0.5025***	0.3248***
	(0.0180)	(0.0040)
Market characteristics		
log County Income	0.1708***	0.3662***
	(0.0536)	(0.0438)
log County Population	-0.5360***	-0.2197***
	(0.0914)	(0.0627)
Share Population age $\leq$ 19	2.6846***	-4.4841***
	(0.6751)	(0.5448)
Share Population age $\geq$ 50	2.9298***	-2.6891***
	(0.4919)	(0.3445)
log housing price index	0.2531***	1.0820***
	(0.0283)	(0.0228)
log number of bankruptcy filings	0.0102*	-0.0804***
	(0.0059)	(0.0069)
Bank $\times$ County Fixed Effects	YES	YES
Time Dummies	YES	YES
Number of observations	132,265	1,224,465
Hansen-Sargan (p-value)	0.3678	0.2122

Note: * means p-value < 0.05; ** means p-value < 0.01; *** means p-value< 0.001

6 Counterfactual experiments

Using the estimated model, we implement counterfactual experiments to measure the effects home-bias, branch networks, and local market power on the geographic distribution of deposits and loans, and the geographic distribution of consumer welfare, i.e., which counties are better off and which are worse off. The following is a description of the experiments in terms of the structural parameters.

Experiment 1. First, we look at the importance of geographic expansion of branch networks across states by looking at the counterfactual equilibrium if banks could not operate in multiple states, as was the case prior to the Riegle-Neal Act of 1994. We keep the structural parameters and all the exogenous variables (observable and unobservable) at their observed/estimated values, including the local number of branches. However, each bank in our sample with branches in multiple states is divided into as many "counterfactual banks" as the number of states where this bank operates. Note that the main effect of this counterfactual is through the effect of total deposits on demand-cost of loans and deposits, i.e., $\theta_Q^d \ln(Q_{jt}^d)$ and $\theta_Q^\ell \ln(Q_{jt}^d)$.

Experiment 2. Second, we study the effects of eliminating the home bias due to economies of scope between deposits and loans. In this experiment, we set the parameters θ_ℓ^d and θ_d^ℓ to zero and compute the new equilibrium of the model. We are more interested in the effect of local economies of scope in reducing the geographic diffusion of credit than on their effect of increasing the net-wtp for loans and deposits. Therefore, we compensate this effect by increasing the constant terms in the two regressions such that the mean value of the net-wtps, over all bank-county-year observations and evaluated at the sample observations, is the same as in the estimated model.

Experiment 3. Third, we look at the effect of eliminating county heterogeneity in local market power. We impose the restriction that all the banks have the same market power in every market. We implement this experiment by setting $\theta_n^d = \theta_n^\ell = 0$, $\theta_d^\ell = \theta_\ell^d = 0$, and $\eta_{jm}^{d(1)} = \eta_{jm}^{\ell(1)} = \eta_{jmt}^{d(4)} = \eta_{jmt}^{\ell(4)} = 0$ for every market m. As in experiment 2, we compensate this effect by increasing the constant terms.

Experiment 4. We evaluate how a counterfactual tax on deposits would affect the provision of credit and, more interestingly, its geographic distribution. We implement this experiment by reducing by 20% (i.e., 20% ad valorem tax) the constant term in the equation for the net-wtp for deposit products.

Experiment 5. Finally, we investigate to what extent national aggregate shocks (e.g., business cycle, monetary policy) affect bank credit in a geographic non-neutral way. We implement this experiment by setting to zero the national level aggregate shocks in the equations for deposits and loans: $\eta_t^d = \eta_t^\ell = 0$ at every year t.

Table 3 presents results from these counterfactual experiments. We measure the effects of these counterfactuals by looking at two outcome variables: (a) the aggregate segregation index, SI_t, that we have defined in equation (1) and whose evolution we presented in Figure 7; and (b) inequality in counties' share of credit. For share of credit, we rank counties according to their share in the national amount of loans, and we obtain the share of credit of the bottom 2500 counties, that represent approximately 80% of all the counties and 22% of US population, and of the top 100 counties, that account for 3% of all the counties and 40% of the US population.

Table 3
Counterfactual Experiments

Outcome Variable	**Data**	**Exp. 1** No m.s.n.	**Exp. 2** No ec.s.	**Exp. 3** No l.m.p.	**Exp. 4** Tax dep.	**Exp. 5** No agg.s.
Segregation index	0.32	0.24	0.39	0.52	0.29	0.28
Bottom 2500 counties: share of credit	9%	10%	11%	14%	12%	7%
Top 100 counties: share of credit	58%	49%	52%	45%	49%	61%

Experiment 1: Remove multi-state branch networks ("No m.s.n.": No mult-state networks).
Experiment 2: Remove Economies of scope ("No ec.s.": No economies of scope).
Experiment 3: Remove local market power ("No l.m.p.": No local market power).
Experiment 4: 20% taxt on deposits ("Tax dep.": Tax on deposits).
Experiment 5: Removes aggregate national shocks. ("No ag.s.": No aggregate shocks).

7 Conclusions

The purpose of this paper is to provide systematic evidence on the extent to which banks' branch networks can reduce geographic frictions to the flow of funding to locations where credit is in greater demand. We focus on three empirical questions: how did the deregulation that allowed banks to expand geographically in the 1990s and 2000s affect the geographic flow of bank funds?; how important is the 'home bias' generated by economies of scope between deposits and loans?; and what is the contribution of local market power to the geographic distribution of bank credit?

Our results are based on the estimation of a structural model of bank oligopoly competition for deposits and loans in multiple geographic markets. The equilibrium of the model allows for rich interconnections across geographic locations and between deposit and loan markets such that local shocks in demand for deposits or loans can affect endogenously the volume of loans and deposits in every local market. We estimate this structural model using data from the US banking industry for the period 1998-2010. The estimated model shows that a bank's total deposits has a very significant effect on the bank's market shares in loan markets. We also find evidence that is consistent with significant economies of scope between deposits and loans at the local level. Counterfactual experiments show that these economies of scope generate a substantial home-bias in the utilization of funds. Local market power has also a significant negative effect on the geographic flow of credit.

References

[1] Acharya, V., and O. Merrouche (2012): "Precautionary hoarding of liquidity and interbank markets: Evidence from the subprime crisis," *Review of Finance*, 22, 1-59.

[2] Aguirregabiria, V., R. Clark, and H. Wang (2016): "Diversification of Geographic Risk in Retail Bank Networks: Evidence from Bank Expansion after the Riegle-Neal Act," *RAND Journal of Economics*, 47, 529-572.

[3] Akkus, O., J. Cookson, and A. Hortacsu (2012): "The Determinants of Bank Mergers: A Revealed Preference Analysis," manuscript. University of Chicago.

[4] Anderson, S., De Palma, A. and Thisse, J.-F. (1992): "Discrete Choice Theory of Product Differentiation," Cambridge, MA: MIT Press.

[5] Arellano, M., S. Bond (1991): "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The Review of Economic Studies*, 58(2), 277-297.

[6] Arellano, M., and O. Bover (1995): "Another look at the instrumental variable estimation of error-components models," *Journal of Econometrics*, 68(1), 29-51.

[7] Ashcraft, A., J. McAndrews, and D. Skeie (2011): "Precautionary reserves and the interbank market," *Journal of Money, Credit and Banking*, 43, 311-348.

[8] Berry, S., J. Levinsohn, and A. Pakes (1995): "Automobile Prices in Market Equilibrium," *Econometrica*, 63, 841-890.

[9] Blundell, R., and S. Bond (1998): "Initial conditions and moment restrictions in dynamic panel data models," *Journal of Econometrics*, 87(1), 115-143.

[10] Bogin, A., W. Doerner, and W. Larson (2016): *Local house price dynamics: New indices and stylized facts*, FHFA Staff Working Paper Series.

[11] Brevoort, K., and J. Wolken (2009): "Does Distance Matter in Banking?" chapter 3 in book "The Changing Geography of Banking and Finance," P. Alessandrini, M. Fratianni, and A. Zazzaro, Alberto (eds.) Springer.

[12] Coase, R. (1937): "The Nature of the Firm," *Economica*, 4(16), 386-405.

[13] Cohen, A., and M. Mazzeo (2007): "Market Structure and Competition among Retail Depository Institutions," Review of Economics and Statistics, 89, 60-74.

[14] Cohen, A., and M. Mazzeo (2010): "Investment Strategies and Market Structure: An Empirical Analysis of Bank Branching Decisions," Journal of Financial Services Research 38:1, 1-21.

[15] Corbae, D., and P. D'Erasmo (2013): "A Quantitative Model of Banking Industry Dynamics," manuscript. University of Wisconsin-Madison.

[16] Corbae, D., and P. D'Erasmo (2017): "Capital Requirements in a Quantitative Model of Banking Industry Dynamics," working paper.

[17] Cortés, K. and P. Strahan (2014): "Tracing out Capital Flows: How Financially Integrated Banks Respond to Natural Disasters," manuscript. Federal Reserve Bank of Cleveland.

[18] Degryse, H. (1996): "On the interaction between vertical and horizontal product differentiation: An application to banking," The Journal of Industrial Economics, 169-186.

[19] Dick, A. (2008): "Demand estimation and consumer welfare in the banking industry," *Journal of Banking and Finance*, 32(8), 1661-1676.

[20] Egan, M., A. Hortaçsu, and G. Matvos (2017): "Deposit Competition and Financial Fragility: Evidence from the US Banking Sector," *American Economic Review*, 107(1), 169-216.

[21] Egan, M., S. Lewellen, and A. Sunderam (2017): "The Cross Section of Bank Value" working paper

[22] Gilje, E. (2012): "Does Local Access to Finance Matter? Evidence from U.S. Oil and Natural Gas Shale Booms," Management Science (Forthcoming).

[23] Gilje, E., E. Loutskina, and P. Strahan (2016): "Exporting liquidity: Branch banking and financial integration," Journal of Finance, 71, 1159-1184.

[24] Gine, X. and R. Townsend (2004): "Evaluation of financial liberalization: a general equilibrium model with constrained occupation choice," Journal of Development Economics, 74, 269-307.

[25] Gowrisankaran, G., and J. Krainer (2011): "Entry and pricing in a differentiated products industry: Evidence from the ATM market," The RAND Journal of Economics, 42(1), 1-22.

[26] Ho, K. and J. Ishii (2011): "Location and Competition in Retail Banking," *International Journal of Industrial Organization*, 29, 537-546.

[27] Jahn, J. A., C. F. Schmid, and C. Schrag (1947): "The Measurement of Ecological Segregation?," *American Sociological Review*, 103, 293-303.

[28] Jayaratne, J. and D. Morgan (2000): "Capital Market Frictions and Deposit Constraints at Banks," *Journal of Money, Credit and Banking*, 32, 74-92.

[29] Kashyap, A., R. Rajan, and J. Stein (2002) "Banks as Liquidity Providers: An Explanation for the Coexistence of Lending and Deposit-Taking," *Journal of Finance*, 57 33-73

[30] Loutskina, E. and P. Strahan (2015): "Financial integration, housing, and economic volatility," *Journal of Financial Economics*, 115, 25-41.

[31] Mankart, J., A. Michaelides and S. Pagratis (2016): "Bank Capital Buffers in a Dynamic Model," working paper.

[32] Neven, D. and L.-H. Roller (1999): "An aggregate structural model of competition in the European banking industry," *International Journal of Industrial Organization*, 17, 1059-1074.

[33] Nguyen, H.-L. (2015): "Do Bank Branches Still Matter? The Effect of Closings on Local Economic Outcomes." manuscript.

[34] Papi, L., E. Sarno, and A. Zazzaro (2015): "The Geographical Network of Bank Organizations: Issues and Evidence for Italy," forthcoming in Handbook of Geographies of Money and Finance, R. Martin and J. Pollard (eds), E. Elgar Publisher.

[35] Petersen M,. and R. Rajan (2002): "Does Distance Still Matter? The Information Revolution in Small Business Lending," *Journal of Finance*, 57(6), 2533-2570.

[36] Petkov, I. (2015): "Small Business Lending and the Bank-Branch Network," manuscript, BC.

[37] Rosen, Richard J. (2011): "Competition in Mortgage Markets: The Effect of Lender Type on Loan Characteristics," *Federal Reserve Bank of Chicago Economic Perspectives*, 35 (1), 2-21.

[38] Schargrodsky, E., & Sturzenegger, F. (2000): "Banking regulation and competition with product differentiation," *Journal of Development Economics*, 63(1), 85-111.