The geographic flow of bank funding and access to credit: Branch networks, synergies, and local competition*

Victor Aguirregabiria[†] University of Toronto, CEPR

Robert Clark[‡] Queen's University

Hui Wang [§] Peking University

October 7, 2024

Abstract

Geographic dispersion of depositors, borrowers, and banks may prevent funding from flowing to high loan demand areas, limiting credit access. Using bank-county-year level data, we provide evidence of geographic imbalance of deposits and loans and develop a methodology for investigating the contribution to this imbalance of branch networks, market power, and scope economies. Results are based on a novel measure of imbalance and estimation of a structural model of bank competition that admits interconnections across locations and between deposit and loan markets. Counterfactual experiments show branch networks, scope economies, and local competition affect the credit flow to disadvantaged markets.

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.

*This research has benefited from the financial support of SSHRC and the National Natural Science Foundation of China. We thank Mark Egan, Carlo Regiani, and Howard Smith for insightful discussions. Helpful comments were provided by Jason Allen, Heski Bar-Isaac, Dean Corbae, Lu Han, Ali Hortaçsu, Jean-François Houde, Jakub Kastl, Nate Miller, Matt Osborne, Nicola Pavanini, Xavier Vives, and participants in seminars and conferences at the Bank of Canada, Carey Business School, Notre Dame, Princeton, University of Toronto, Vancouver School of Economics, Washington University in St. Louis, *Industrial Organization of the Financial Sector conference* at the Becker-Friedman Institute, *Jornadas de Economia Industrial*, *CEPR/JIE School and Conference on Applied Industrial Organization*, *Drexel-PFED Conference on Credit Markets and the Macroeconomy*, *NBER Conference on Financial Market Regulation*, *Barcelona Summer Forum*. We are grateful to Paul Lim, Anran Tan, Kangkai Wang, and Jingchen Zhu for excellent research assistance.

†Department of Economics, University of Toronto. 150 St. George Street, Toronto, ON, M5S 3G7, Canada, victor.aguirregabiria@utoronto.ca.

‡Department of Economics, Queen's University, Dunning Hall, 94 University Avenue, Kingston, ON, K7L 3N6, Canada, clarkr@econ.queensu.ca.

§Guanghua School of Management, Peking University, Beijing, 100871, China, jackie.wang@gsm.pku.edu.cn.

1 Introduction

An important determinant of credit provision is the availability of deposits (Jayaratne and Morgan, 2000; Ben-David et al., 2017). However, in any given region, the demand for loans may not always coincide with the availability of deposits. Geographic frictions, such as asymmetric information and transaction costs, limit the flow of funds across regions such that there can arise substantial geographic heterogeneity in access to credit and possibly even *credit deserts*. In turn, limited credit access can impact entrepreneurship levels, employment, wages, and economic growth (see, for instance, Giné and Townsend, 2004).

Wholesale liquidity markets can improve the flow of funds. Banks can buy and sell liquidity in the interbank wholesale market, although transaction costs may arise due to bank precautionary motives and liquidity hoarding (Ashcraft et al., 2011; Acharya and Merrouche, 2013). Banks may also use their branch networks to overcome geographic frictions and move liquidity from one region to another, incurring transaction costs that are likely smaller than those from the interbank market (Coase, 1937). However, two counterbalancing forces can negatively affect a bank's willingness to transfer funds between its branches: (i) economies of scope and other synergies between deposits and loans at the branch level and (ii) local market power. Scope economies may arise because clients prefer to have their deposit account and mortgage in the same bank or because a bank's cost of managing a deposit account and a loan may be smaller if they belong to the same client.[1] These and other synergies create incentives to concentrate lending activity in branches with high levels of deposits, and therefore limit the geographic flow of liquidity to markets more in need of credit. Spatial heterogeneity in local market power also negatively impacts the geographic flow of credit, affecting the extent to which changes in the marginal cost of loans are passed through to borrowers through loan interest rates.

This paper aims to provide systematic evidence on the extent to which deposits and loans are geographically imbalanced in the US banking industry and to investigate empirically the contribution of branch networks, scope economies, and local market power to this imbalance. To perform our analysis, we assemble a dataset from the US banking industry for 1998-2010, merging data at the bank-county-year level from two sources. Deposit and branch-network information is collected from the *Summary of Deposits* provided by the *Federal Deposit Insurance Corporation*. Information on lending is from the *Home Mortgage Disclosure Act* data set, which provides

[1]See for instance Kashyap et al. (2002), Mester et al. (2007), Norden and Weber (2010), and Egan et al. (2022).

information on mortgage loans.

Our first contribution is to develop an index of the imbalance between deposits and loans that captures the degree to which a depository bank transfers funds away from where deposits are collected (henceforth Imbalance Index or *II*). To do so, we adapt techniques developed in sociology and labor economics to quantify residential segregation. These measures capture the extent to which individuals from different social groups live together or apart within a given geographical area.[2] Our findings suggest that, while many banks exhibit a strong home bias with deposit and loan shares being roughly equivalent in each county where they operate, some banks transfer significant funds between geographic locations. Furthermore, we find evidence that some regions of the country have much larger shares of total deposits than they do of loans and vice versa, implying an important geographic imbalance.

We do not take a stand on whether greater imbalance is good or bad. On the one hand, greater imbalance implies markets are integrated with deposit funding flowing to those where it is most in demand. On the other, it could mean funding is being withdrawn from markets with specific characteristics (poorer, less white) and allocated to markets with different characteristics, such that some populations have limited access to credit. Our objective is to document imbalances in the geographic distribution of credit, and to gain some understanding of the role that branch networks, scope economies, and local market power play in generating these imbalances.

Investigating the factors contributing to geographic imbalance of deposits and loans requires a model that allows for interconnections across geographic locations and between deposit and loan markets such that local shocks to deposits or loans can endogenously affect the volume of loans and deposits in every local market. The second contribution of this paper is to develop and estimate a structural model of bank oligopoly competition for *both* deposits and loans in multiple geographic markets, allowing for rich interconnections. We characterize an equilibrium of this multimarket oligopoly model and propose a simple algorithm to solve it. Our approach allows us to perform counterfactual experiments that provide evidence of the effect of branch networks, scope economies, local market power, and various public policies on the geographic diffusion of funds. While other papers have pointed out the existence of economies of scope between deposits and loans at the bank level, our approach allows us to determine at which geographic level these synergies occur, focusing on the possibility that they may be local.

[2] See for instance Jahn et al. [1947], Duncan and Duncan [1955], Atkinson et al. [1970], White (1983, 1986), and Cutler et al. [1999]. More recently, they have been used by Gentzkow et al. [2019] to quantify the degree of polarization in political speech in the US.

In the model, differentiated banks sell deposit and loan products in multiple local markets (counties). The model incorporates four main variables, which may affect the demand and costs of loans and deposits in a local market. First, the number of branches the bank has in the local market may affect the marginal cost of managing deposits and loans and influence consumer awareness and willingness to pay. Second, the total amount of deposits the bank has at the national level may reduce 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. The third factor is the amount of deposits (loans) the bank has in the local market, which may increase consumer demand for loans (deposits) and reduce the bank's marginal cost of loans (deposits) due to economies of scope in managing deposits and loans. The resulting structure bears a resemblance to models of two-sided markets. Finally, the model includes the fraction of securitized loans to capture the extent to which banks move loans off their balance sheets. Securitization is especially important for non-depository institutions, so our model also allows these *shadow banks* to play a role on the loan side of the market.

The structural parameters associated with branch networks, scope economies, and local market power are fundamental for the model's predictions. Estimation must address endogeneity and simultaneity of local and total deposits and loans. Our identification approach involves controlling for a rich fixed-effects specification of the unobservables, that includes *bank* $\times$ *county*, *county* $\times$ *year*, and *bank* $\times$ *year* fixed effects. We obtain difference-in-differences transformations of the structural equations and apply instrumental variables / GMM to these transformed equations. Our identifying moment conditions require that unobservables not be serially or spatially correlated after controlling for fixed effects. We present evidence from tests that validate these restrictions.

It is important to point out that datasets containing information on interest rates for loans and deposits for all depository institutions at the county level are not available, and so we estimate the social surplus (i.e., consumers' willingness-to-pay net of banks' marginal costs) for the different deposit and loan products, as well as how social surplus 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 and require imposing weaker conditions than if we were trying to identify demand and marginal costs separately. This approach allows us to assess how market frictions –such as local market power, branching restrictions, or taxes– affect social surplus and the equilibrium provision of credit and deposits, without requiring us to disentangle their separate effects on demand and supply.

Estimation yields the following results. The number of branches in a county increases the social surplus for deposits and loans, but the effect is substantially stronger for deposits. Securitization has a strong positive impact on the amount of loans. Substantial synergies exist between deposits and loans at the bank and local market levels. Finally, the effect of a bank's total deposits on the social surplus for loans is positive and significant both economically and statistically, implying that banks' internal liquidity reduces the cost of lending.

Our structural approach allows us to evaluate the effect of market features and policies on the Imbalance Index and the value of loans, deposits, and social surplus nationally and for different counties. In a first experiment, we study the contribution of branch networks to the geographic flow of credit by imposing the restriction that banks only have access to locally generated deposits and not those earned throughout their network. Our second experiment evaluates the effects of eliminating local synergies between deposits and loans. A third experiment looks at the impact of removing local market power and obtains the model equilibrium under the condition that, in all local markets, banks act competitively with prices equal to marginal cost. In a fourth experiment, we remove shadow banks as competitors in loan markets and obtain a measure of the contribution of these institutions to the geographic distribution of credit and its evolution over time. We then present three policy experiments. In the first, we evaluate the effect of a regulation prohibiting banks from operating branch networks in multiple states, as was the case before the Riegle-Neal Act of 1994. We implement this counterfactual by dividing every multi-state bank into different independent banks, one for each state. In the second, we study the impact of introducing a deposit tax that approximates the effect of inflation taxing away the real value of nominal deposits. Finally, we evaluate the impact of a hypothetical increase in the interbank rate to reflect a contractionary policy initiated by the Federal Reserve, which allows us to explore the potential geographic non-neutrality of a standard monetary policy tool and its effect on the distribution of bank funds across counties.

Our findings suggest that branch networks are essential for spreading liquidity across the regions where a bank operates, especially in smaller, poorer, and more rural areas, significantly increasing bank-level Imbalance Index scores. On the other hand, local synergies decrease the Imbalance Index by generating a home bias and preventing funds from flowing away from where they are generated. Local market power substantially negatively affects the geographic flow of credit. Limited competition in small counties is essential in constraining the amount of credit these counties receive. Shadow banks also play an important role in credit provision, particularly in large and urban markets. In terms of policy experiments, we find that Riegle-

Neal affected lending and the flow of funds, but the impact was moderated by the limitations in cross-border lending activity. A deposit tax has a modest negative impact on lending activity, felt disproportionately by socially disadvantaged markets. Lastly, increasing the interbank rate has a significant negative impact on the fraction of funds that banks redistribute across regions. Like with the tax, the effect is stronger in disadvantaged markets

Our model builds on and extends the literature on structural models of bank competition.[3] Previous work has looked separately at the market's loan or deposit sides. Corbae and D'Erasmo (2019, 2021), Benetton (2021), Crawford et al. (2018), and Benetton et al. (2021) all focus on the loan side. Dick (2008), Ho and Ishii (2011), Honka et al. (2017), Egan et al. (2022), and Xiao (2020) estimate differentiated demand models for bank deposits. Egan et al. (2017) distinguish between insured and uninsured deposits and endogenize bank defaults and bank runs. Aguirregabiria et al. (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. Wang et al. (2022) embed simple demand models for both deposits and loans into a corporate finance model to understand the impact of various financial frictions for the transmission of monetary policy. Given their focus, their models of deposits and loans are only at the national level, only for a subset of lenders, and do not allow for synergies between the two sides of the market. Similarly, Drechsler et al. (2017) study the role of market power in the transmission of monetary policy using a Dixit-Stiglitz model of monopolistic competition. Oberfield et al. [2024] study the branch-network expansion patterns and show that banks located in markets where deposits are plentiful relative to demand for credit minimize the need for wholesale funding. We extend all of these previous studies by considering an equilibrium model for deposits and loans that allows for interconnections between these markets at the local level and for the effect of a bank's total liquidity on the costs of loans in local markets. This rich connectivity is necessary to answer the specific questions we pose here, but it is also a contribution in its own right.

We are also related to recent papers that take advantage of 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, 2017; Gilje et al., 2016; Loutskina and Strahan, 2015; Petkov, 2023; and Cortés and Strahan, 2017). Our paper complements the findings of Gilje et al. [2016] in different ways. First, our empirical analysis of the relationship between the geographic location of a bank's branches (deposits) and loans extends to all US local markets (counties). Second, we study the contribution of local market power to the geographic flow of banks' funds.

[3] For an overview of this literature, see Clark et al. [2021].

Third, our approach to identifying the effect of total deposits on local loans exploits more general sources of exogenous variation than those associated with 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.

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 updated 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 total deposits, location, 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. We focus on total deposits, including insured and uninsured, as in Mankart et al. [2020].[4,5]

Under HMDA, most institutions must disclose information on the mortgage loans they originate, refinance, or purchase in a given year. At the level of financial institution, county, and year, we have information on the number and volume of mortgage applications, mortgage loans issued, and mortgage loans subsequently securitized.[6] The type of institutions reporting to HMDA include both depository institutions and non-depository institutions – mainly Indepen-

[4]Deposits are usually assigned based on the account holder's address, branch activity, or where the account was opened (Federal Deposit Insurance Corporation, 2020). Note that survey evidence from Kiser [2002] suggests that most individuals open their main checking/deposit account around the time of their first home purchase.

[5]A small proportion of branches in the SOD data set (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.

[6]We use the information on securitization to capture differences in banks' cost of lending at the county level. Summary statistics on securitization are reported in Table 1 below.

dent Mortgage Companies (IMCs), which are commonly described as *shadow banks*.[7]

Our analysis focuses on the depository institutions reporting to HMDA, including banks and Savings Associations, that can be matched with the SOD data.[8] Focusing on depository institutions is consistent with the research questions addressed in this paper, since they rely heavily on branching and deposits to fund their loans. By contrast, shadow banks rely on wholesale funding and mortgage brokers (Rosen, 2011). Nevertheless, our structural model of demand and supply of mortgages includes competition from shadow banks. Together, these depository institutions and shadow banks represent the working sample that we use for estimation and counterfactual experiments. They account for 80% of all deposits and 94% of all mortgage loans. Other financial institutions, including the depository institutions that we cannot match with HMDA, are excluded from our analysis because we cannot assign either their deposits or loans (or both) to particular counties. Although large in number, these institutions represent a negligible fraction of lending activity and a small share of deposits. 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's (OTS) Thrift Financial Report (TFR) for saving associations. Appendix A.2 presents a detailed description of the construction of our working sample.

County-level data on socioeconomic characteristics are obtained from various products of the US Census Bureau. 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). We also use information on county-level house prices for 2742 counties from the Federal Housing Finance Agency (see Bogin and Larson, 2019), and county-level bankruptcy data from the U.S. Bankruptcy Courts.[9]

[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 the FDIC to every insured depository institution) or/and their RSSD number (assigned by the Federal Reserve to every financial institution). We match thrifts using their docket numbers. We match financial institutions supervised by the OCC through the Call Reports, which allow us to match information from SOD and HMDA.

[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.

2.2 Data features

Four features of our data and empirical approach deserve specific discussion. First, we have data on mortgage loans at the bank-county-year level but not on other forms of bank credit at the same level of disaggregation. Ideally, we would incorporate information on other types of loans, but, to our knowledge, such data are not publicly available at the bank-county-year level.[10] However, mortgage loans represent the most substantial part of bank loans and even assets. Using bank-level information from the 2010 Call Reports, Mankart et al. [2020] show mortgages account for between 62% and 72% of all bank loans, and between 38% and 45% of total bank assets, depending on bank size (with larger banks typically having smaller shares). They also report that deposits represent between 68% and 85% of total bank liabilities. These patterns hold in our sample too. Therefore, our focus on deposits and mortgages, though motivated by data availability, captures a substantial fraction of total bank liabilities and assets, respectively.

Second, our empirical focus is on stocks of deposits and flows of new loans, as opposed to either the stocks of both deposits and loans or only new deposits and new loans.[11] These are the levels at which most work studying deposit- or loan markets in isolation have considered these series. See for instance Dick [2008] and Egan et al. [2017] for deposits, and Benetton [2021] and Buchak et al. [2024a] for loans. The assumption underlying the decision is that consumers can choose in every period where to put their entire stock of deposits and where to get new loans (or where to refinance their loan). We are justified in making this assumption because switching costs are higher for loans than for deposits. While there are costs involved in moving deposits, they are typically less important than the financial penalties imposed when moving mortgage loans from one financial institution to another. That said, one problem with using stocks is that past deposit inflows, which make up today's stock of deposits, have already been used to fund prior mortgages. To address this, flows could be constructed from the SOD dataset as the net change in deposit stocks at the bank-county level by first differencing by year. However, this represents a net change – newly attracted deposits minus withdrawn deposits. This net change can be negative, which makes the deposit share at the county level difficult to calculate. It is also the case that the net change in deposits would underestimate the funds available to banks to create new loans since a fraction of existing loans are repaid (come due) every year. Therefore,

[10]Some data on small-business loans are available, but, for most of our sample period, only very large banks (i.e., those with more than $1 billion in assets) were required to reveal this lending activity.

[11]We include refinances in our sample since borrowers can move their mortgage to a new bank when they refinance, so the refinance decision looks very similar to the initial decision to get a mortgage with a lender.

we construct an *adjusted flow* measure, including net deposit changes along with those deposits freed up today as a fraction of previously funded loans is paid off. The Call Reports provide information for each bank on the fraction of loans coming due each year, and we use this to construct our measure. We have performed our analysis using new loans and adjusted deposit flows. Our findings regarding the evolution of the Imbalance Index are robust to this alternative measure. Details are provided in Appendix A.4.[12]

Third, publicly available data on interest rates of deposits and loans for all financial institutions are unavailable 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, contain potentially important measurement errors. The loan-rate data, particularly, are available only for a small set of lenders. The lack of good price data at the bank-county-year level would be an essential limitation if we wanted to estimate demand and marginal cost separately, but this is not the goal of this paper. To answer all the empirical questions in this paper, we need to estimate the value of consumers' willingness-to-pay net of banks' marginal costs (i.e., the *social surplus*) for the different deposit and loan products, as well as how net willingness-to-pay depends on different variables such as local bank branches and market power. We show that these primitives can be identified without information on prices of deposits and loans and require imposing weaker conditions than if we were trying to identify demand and marginal costs separately. One might be particularly concerned that without price data, it would be impossible to say anything meaningful about market power. Still, we provide evidence that the predictions of our structural model are consistent with theoretical implications on market power and with the stylized facts about margins and spreads presented in recent work on market power in the industry. We discuss this in more detail in Sections 3 and 6.1.

Finally, we define our markets as counties, the primary administrative divisions for most states. Markets determine the set of banks 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

[12] We have also considered the case of *stocks* of deposits and *stocks* of loans. We take advantage of the fact that, in addition to the flow of new loans at the bank-county level, we have information on stocks of loans at the bank level (i.e., aggregated across counties). To infer stock at the bank-county level, we assume that the distribution of stocks across counties is the same as the distribution of flows for the same bank. We have performed descriptive analysis using this measure, and the results are essentially unchanged. See Appendix A.4.

Huang, 2008; Gowrisankaran and Krainer, 2011; Uetake and Watanabe, 2019).

2.3 Summary statistics

Our working sample consists of all matched depository banks along with shadow banks (cases 1, 3, and 7 from Table A2), in 3,146 counties during the period 1998-2010. Table 1 presents summary statistics from this sample. The top panel provides bank-level statistics based on 61,418 bank-year observations for depository banks, and 18,552 bank-year observations for shadow banks. The bottom panel includes county-level statistics for 40,811 county-year observations.

The median number of counties where a depository bank obtains deposits from its branches is two, while the median number where a bank sells mortgage loans is eight.[13] Banks' branch networks are geographically more concentrated than the networks of counties where they provide loans. Similarly, the median number of banks providing deposit services in a county is only four, but the median number of banks selling mortgages is 85. This includes the shadow banks. The median number of counties where shadow banks actively sell loans is 31. The median Herfindahl-Hirschman index (HHI) is 3,450 for deposit markets and 655 for loan markets. A possible explanation for the different market structures is that the sunk cost of entry in a local market is higher for deposits than for loans.

Figure 1a shows the evolution of the number of banks and branches per county for depository institutions. At the start of our sample period, there were just under five banks and 20 branches per county taking deposits. These numbers increased steadily to nearly seven and 26, respectively, by 2010. The increase coincides with the rolling out of Riegle-Neal, which permitted banks to branch across state lines. Over the same period, the percentage of banks with branches in multiple states increased from less than 1% to around 7%. Figure 1b displays the evolution of the number of depository institutions and shadow banks offering loans per county. The number of depository banks making loans increases steadily until the crisis and then decreases by around 20%. From 1998 to 2010 the percentage of banks making loans in multiple states increased from 40% to 55%. The number of shadow banks increases until the crisis after which it drops sharply.

2.4 Geographic imbalance of deposits and loans

In this subsection, we present evidence of the extent to which deposits and loans are geographically imbalanced. We adapt the measures of residential segregation used in sociology and labor

[13] Table A3 in Appendix A.3 presents information on multi-state branching and lending.

Table 1: Summary Statistics

Panel A: Bank Level Statistics

Variable	Mean	S. D.	Pctile 5	Median	Pctile 95
Depository Banks: 61,418 bank-year observations					
Number of branches	15.22	113.46	1	4	34
Number of counties with deposits > 0	4.01	17.80	1	2	10
Number of counties with new loans > 0	29.78	149.39	1	8	68
Total deposits (in million $)	985	11,369	35	150	1,720
Total new loans (in million $)	186	3,149	1	12	253
Securitization rate of loans (first year) (%)	20.80	31.10	0.00	0.00	88.63
Shadow Banks: 18,552 bank-year observations					
Number of counties with new loans > 0	151.52	361.40	2	31	813
Total new loans (in million $)	742	5,280	1	79	2088
Securitization rate of loans (first year) (%)	76.75	38.14	0.00	98.97	100.00

Panel B: County Level Statistics (40,811 county-year observations)

Variable	Mean	S. D.	Pctile 5	Median	Pctile 95
Banks					
Number of branches (per county)	22.90	61.85	0	6	101
Number of banks with branches (per county)	6.04	8.24	0	4	20
Number of banks with new loans (per county)	113.69	99.70	13	85	319
" – Depository banks	44.82	39.71	6	34	124
" – Shadow banks	68.87	62.97	6	50	196
HHI market of deposits	4381	2899	1202	3450	10000
HHI market of new loans	910	879	267	655	2344
Deposits per capita (in ,000 $)	8.67	9.84	0.00	7.57	19.79
New loans per capita (in ,000 $)	3.31	4.08	0.38	2.05	10.21
Securitization rate of loans (first year) (%)	64.90	13.88	39.22	67.09	83.59
Demographics					
Income per capita (in ,000 $)	27.89	8.09	18.09	26.56	41.74
Population (in ,000 people)	93.40	301.24	3.04	25.24	396.38
Share population $\leq$ 19 (in %)	27.44	3.47	22.16	27.28	33.17
Share population $\geq$ 50 (in %)	33.27	6.33	23.37	33.00	44.23
Annual change in house price index	3.21	7.40	-7.48	3.35	13.91
Number of bankruptcy filings per year	435	1530	6	106	1790

economics to capture the dissimilarity between the geographic distributions of deposits and loans, both for individual banks and for all the banks.

Figure 1: Number of banks and branches per county

(a) Depository banks with deposits >0 per county

(b) Depository and shadow banks with loans >0 per county

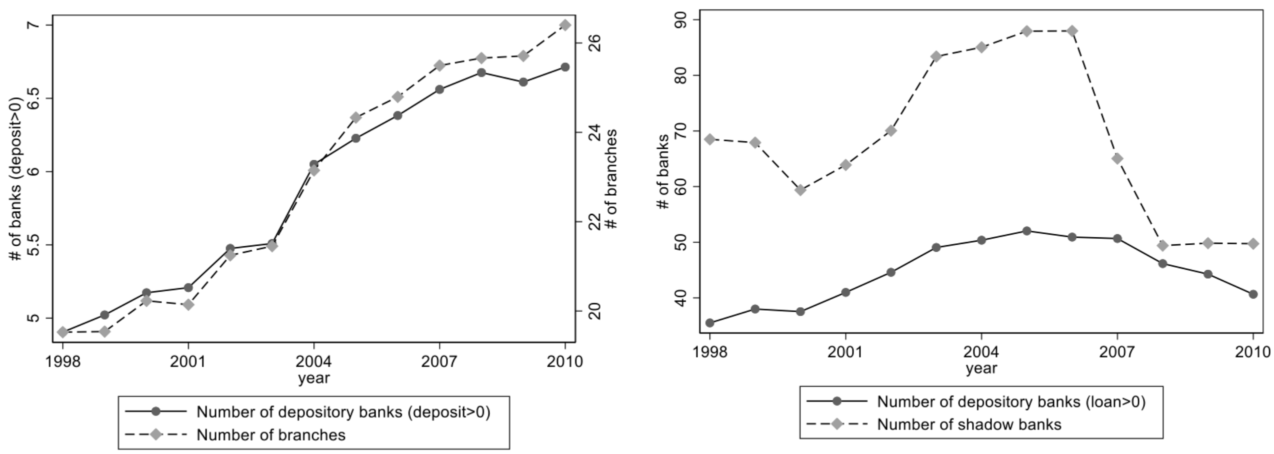
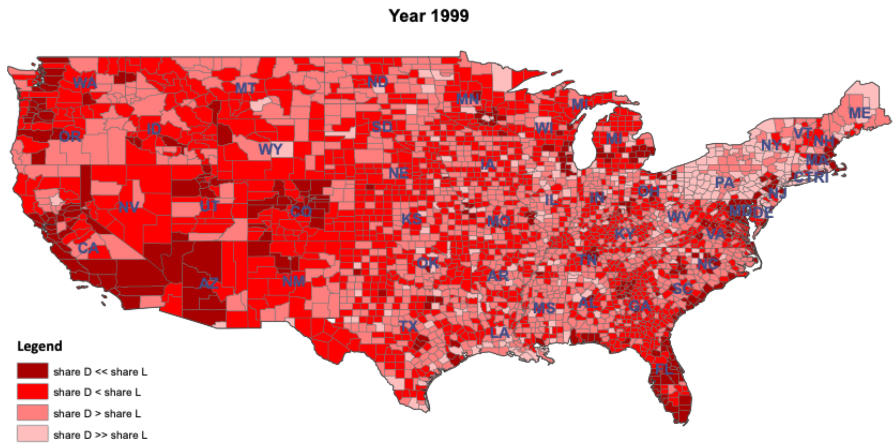


Figure 2 presents maps with the geographic distribution of counties' positions as net borrowers or net lenders. We present these maps for three different years: 1999, 2004, and 2009. We calculate the county's share of deposits for every county-year over aggregate national deposits. Similarly, we calculate the county's share of new loans over the nation's aggregate amount of new loans. 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 new mortgage loans and its share of deposits. The values of S_{L-D} provide the geographic distribution of the borrowing and lending positions of the different counties. By construction, the mean over counties of these indexes equals zero, and there are positive and negative values for net borrowing and net lending counties, respectively. The 10th, 50th, and 90th percentiles of the distribution of the index S_{L-D} are -23.8%, -0.5%, and 35.6%, respectively, as percentages of the average deposit share of a county. That is, for counties at the top 10% (bottom 10%), the difference between the share of loans, and the share of deposits exceeds (is below) the national average by more than 41 (28) percentage points. Using these cutoff points, we sort counties into four groups: (i) counties belonging to the top 10 percentiles of S_{L-D} (Share Loans >> Share Deposits); (ii) counties between the 10th and 50th percentiles (Share Loans > Share Deposits); (iii) counties between the 50th and 90th percentiles (Share Loans < Share Deposits); and (iv) counties belonging to the bottom 10 percentiles (Share Loans << Share Deposits).

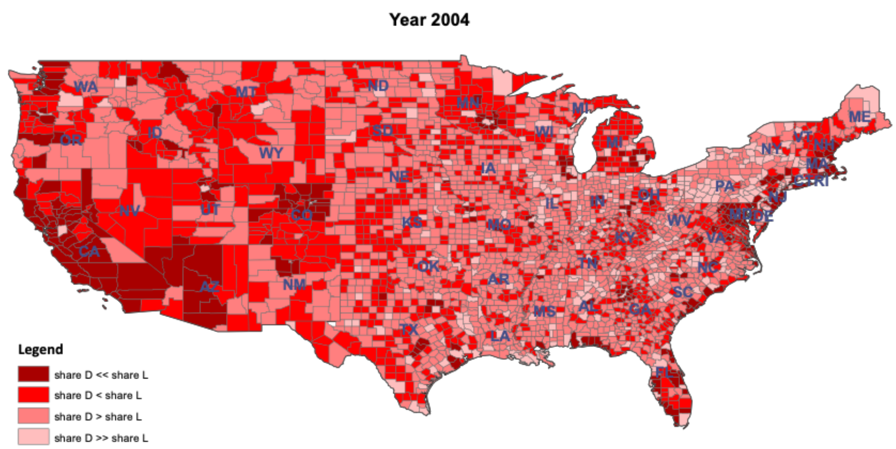
Figure 2 shows clear evidence of deposit and loan imbalances: some regions have high share of deposits, but low share of loans and vice versa. It also reveals regional patterns in net

Figure 2: Distribution of borrower/lender counties

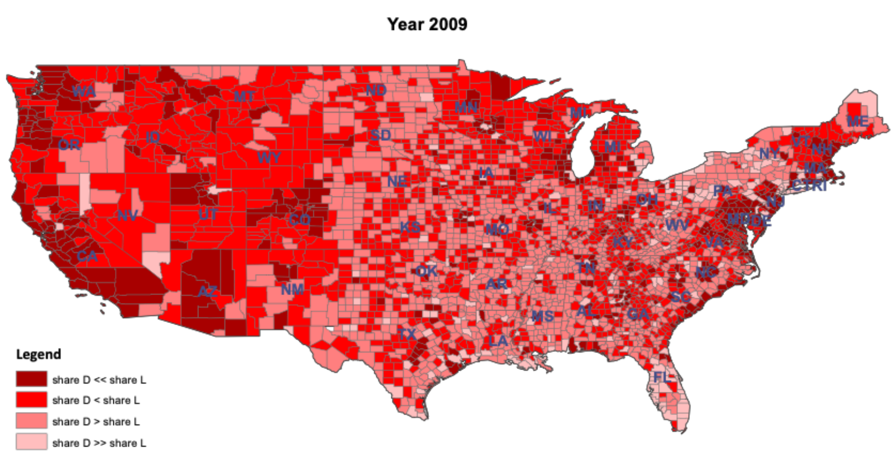
(a)



(b)



(c)



borrowing/lending positions, the most obvious of which is that counties in the interior tend to be net lenders, while those on the coasts are typically net borrowers. There have also been interesting changes related to the mortgage boom and subsequent financial crisis at the end of the decade. In 1999, several counties in California were in the bottom 10 percentiles of S_{L-D}, indicating that their share of deposits was much larger than their share of total loans. By 2004, almost all counties in the state were in the top 10 percentiles, likely reflecting the build-up of mortgage debt during the housing boom. Five years later, during the crisis, many counties had flipped again with deposit shares higher than loan shares.

Borrowing from the literature on racial geographic segregation, we consider the following index to capture the *imbalance* of deposits and loans for bank j:

$$II_{jt} \;=\; \frac{1}{2}\sum_{m=1}^{M}\left|\frac{q^d_{jmt}}{Q^d_{jt}} - \frac{q^\ell_{jmt}}{Q^\ell_{jt}}\right|, \tag{1}$$

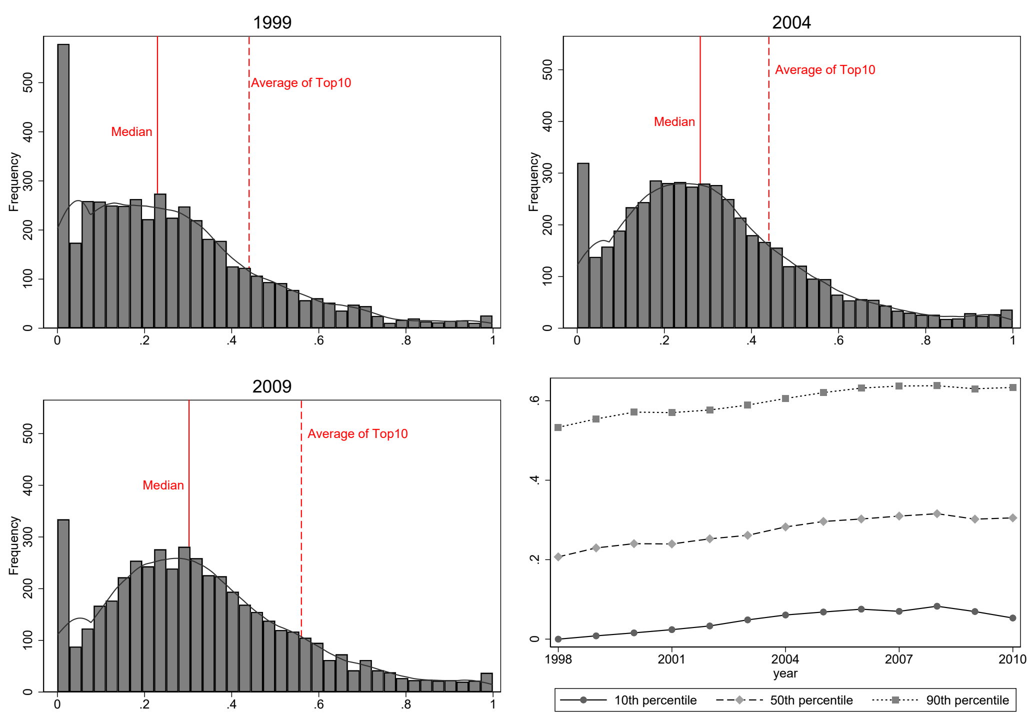
where q^d_{jmt} and q^ℓ_{jmt} represent the amount of deposits and loans, respectively, in county m and year t for bank j, and Q^d_{jt} and Q^ℓ_{jt} represent the j's total deposits and loans. This index is a measure of the imbalance of a bank's deposits and loans or a measure of the bank's home bias. An *Imbalance Index* (II) 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, an Imbalance Index equal to one means that the bank gets all its deposits in markets where it does not provide loans and sells loans only in markets where it does not have deposits.

Figure 3 presents the empirical distribution of the Imbalance Index calculated at the bank-year level.[14] We can see that while most banks are involved to some degree in transferring funds across geographic locations, some have a substantial home bias. Each year, there is a mass of banks with a score equal to zero. These are exclusively banks operating in a single county (i.e., taking deposits and making loans in a single county). At the other extreme are banks with very high Imbalance Index scores: the score is greater than 0.5 for roughly 20% of the banks. We can also see a noticeable shift to the right of the distribution over time, suggesting that more deposit funds are being distributed outside the county where they were generated (home county). A Kolmogorov–Smirnov test rejects equality of the distributions in each of the three years. Note that the distribution shifted rightward between 2004 and 2007 until the Great Financial Crisis started, at which point the trend reversed, and the distribution shifted left. The

[14]Here we present the Imbalance Index using stocks of deposits and flows of loans. In Appendix A.4, we show that the distributions of the bank-level Indexes are almost identical when using adjusted deposit flows instead. The correlation between the two Imbalance Indexes is 0.963.

median bank-level index increased from 0.23 to 0.32 between 1999 and 2010, peaking in 2008.

Figure 3: Distribution of Bank-Level Imbalance Index between Deposits and Loans



Note: The solid red vertical line represents the median of the distribution, while the dashed line shows the mean Imbalance Index for the Top 10 banks as measured by the banks' assets in each year. The bottom right panel shows the evolution of the 10th, 50th, and 90th percentile of the bank-level Imbalance Index distribution.

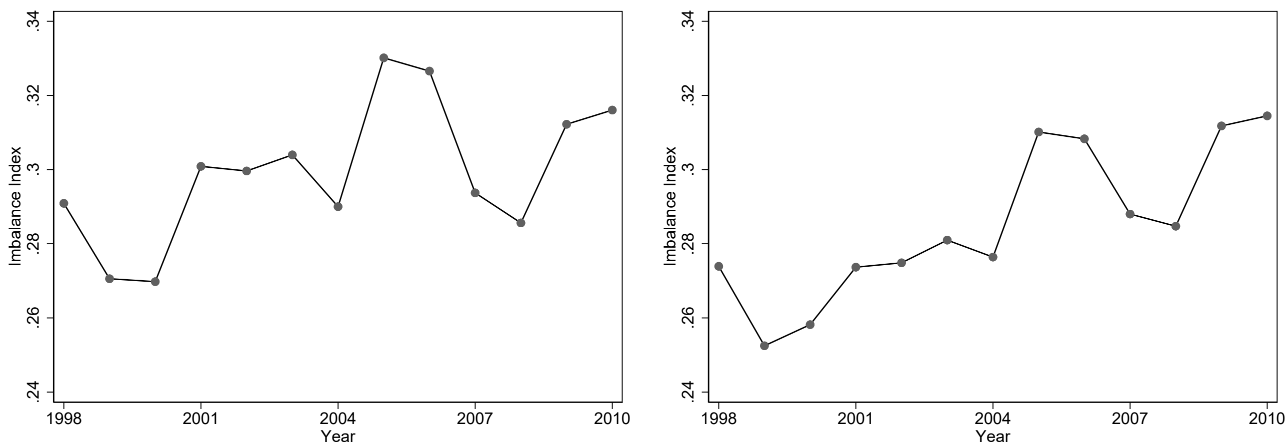
This increase over time is also noticeable in Figure 4, which presents the time series of a national-level Imbalance Index calculated using county-level observations:

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

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. It measures the imbalance of funds between geographic locations. Figure 4a presents the national Imbalance Index for the estimation sample (depository banks that can be matched with HMDA and shadow banks). It exhibits a positive trend, although the overall level of variation is not large (between 0.27 and 0.33). It peaked in 2005 before the crisis and dropped temporarily during the crisis. Figure 4b restricts attention to matched depository banks. It displays the same general trend, but is shifted down, suggesting shadow banks are important for the diffusion of credit.

Figure 4: Time Series of the National Imbalance Index

(a) Depository & shadow banks (Cases 1,3,7) (b) Depository banks (Cases 1,3)



2.4.1 Bank heterogeneity in the Imbalance Index

In this subsection, we investigate the effect of bank characteristics on the Imbalance Index. Specifically, we study the impact of bank size, as measured by number of counties in which they operate total assets, and total deposits (Figures A6, A7, and A8); geographic location of the bank's headquarters, as measured by urban vs rural (Figure A9), and by North, South, East or West (Figure A10); and socioeconomic composition of the bank's headquarter county, as measured by median income (Figure A11) and by the percentage of non-whites (Figure A12). In each case we report Kolmogorow-Smirnov tests for equality of the distributions.

Whether measured by assets, deposits, or county-network, the size of a bank has a positive effect on its Imbalance Index score, with the most significant impact coming from county presence. The rightward shift over time observed in Figure 3 is driven by the fact that the size distribution of banks is changing, with fewer small banks over time. However, even if we restrict attention to the ten largest banks in the US as measured by assets, the same pattern is present, as seen from the dashed vertical line in the figure (and from Table A7). The Imbalance Index scores for these banks are much higher than for the median sample (0.44 vs 0.23 in 1999), and they increased to 0.56 by 2009. For these larger banks, the rise in their Imbalance Index occurs because the increase in the number of counties in which they make loans outpaces the rise in the number of counties where they collect deposits.

Although initially, there was no significant difference in the Imbalance Indexes of poor vs. rich or rural vs. urban counties, by 2002, banks with headquarters in poor and rural counties began to see their Imbalance Indexes shift rightward and remain that way until the end of our sample. These results suggest that over time, banks headquartered in poorer and more rural counties are sending more of their deposits out of counties where they are raised. In contrast, the racial composition of the headquarters county has almost no effect on the Imbalance Index.

3 Model

Consider an economy with M geographic markets, indexed by $m \in \mathcal{M} = \{1, 2, ..., M\}$, and J banks, indexed by $j \in \{1, 2, ..., J\}$.[15] 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. For traditional banks, $\mathcal{M}_j^\ell$ 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., $\mathcal{M}_j^d \subseteq \mathcal{M}_j^\ell$. For shadow bank j, $\mathcal{M}_j^d$ is, by definition, empty. We take networks $\{\mathcal{M}_j^d\}_{j=1}^J$ and $\{\mathcal{M}_j^\ell\}_{j=1}^J$ as given and focus on the endogenous determination of the amounts of deposits and loans in the equilibrium of this static model of multi-market oligopoly competition.

Each local market is populated by *savers* who demand deposit products and *investors* who demand loan products. Importantly, some savers are also investors and vice versa. These products are horizontally differentiated between banks due to different product characteristics and spatial differentiation within a local market. This view of banks' services as differentiated products is in the spirit of previous papers in the literature.[16] A novel feature of our model, which is key for our analysis, is that for traditional banks, it introduces endogenous links between deposit and loan markets and between these markets at different geographic locations.

For traditional bank j, the (variable) profit function is equal to interest earnings from new loans (pre-existing loans and deposits are considered 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 in interbank wholesale markets:

$$\Pi_j \;=\; \sum_{m=1}^{M} p_{jm}^\ell \; q_{jm}^\ell + p_{jm}^d \; q_{jm}^d - C_{jm}\left(q_{jm}^\ell, q_{jm}^d\right) - \left(r_0 + c_{j0}\right) B_j, \tag{3}$$

where p_{jm}^ℓ and p_{jm}^d are the prices for loans and deposits, respectively, for bank j in market m, and q_{jm}^ℓ and q_{jm}^d are the corresponding amounts (flows) of loans and deposits. Note that, typically,

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

[16] See Degryse [1996], Schargrodsky and Sturzenegger [2000], Cohen and Mazzeo (2007 and 2010), Gowrisankaran and Krainer [2011], or Egan et al. [2017], among others.

the price for loans will be positive ($p^{\ell}_{jm} > 0$) because borrowers pay a positive interest rate to obtain a loan. In contrast, the price of deposits is typically negative ($p^{d}_{jm} < 0$) because the bank pays savers to 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 cost associated with using the interbank market. The Federal Reserve determines the interbank interest rate, which 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. This cost includes the expected cost of loan default or pre-payment, as well as the expected cost reduction associated with loan securitization. A bank's resources constraint implies that, $B_j = Q^{\ell}_j - Q^{d}_j$, 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.[17] 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 a positive sign for loans and a negative for deposits.

Given the interest rate in the interbank market, r_0, the equilibrium of our model determines the amounts of loans and deposits of every bank in every local market, and it also determines the net position of a bank in the interbank market, since $B_j = Q^{\ell}_j - Q^{d}_j$. Then, given the net positions of the private banks, the position of the Federal Reserve, represented by B_0, is also endogenously determined, such that the interbank market clears; that is, the equilibrium condition $\sum_{j=1}^{J} B_j + B_0 = 0$ is satisfied.

The profit function for shadow banks is similar in that they earn interest income collected directly or via servicing rights, but unlike traditional banks, they do not fund mortgages via deposits and do not pay interest as in equation (3). Rather, these institutions use an originate-to-distribute funding model to finance the loans they issue through securitization. We capture this funding model through our simple specification that includes securitization in C_{jm}.[18]

Section 3.1 describes the demand system for deposits and loans. Section 3.2 presents our specification of bank variable costs. In section 3.3, we derive the expression for the social surplus implied by demand and cost functions. Section 3.4 describes the model's equilibrium.

[17] We have that $B_j = S^{\ell}_j - S^{d}_j + Q^{\ell}_j - Q^{d}_j$, where S^{ℓ}_j and S^{d}_j are stocks of live pre-existing loans and deposits, respectively. However, these stocks are pre-determined and do not affect the flow of variable profits.

[18] Similar to the profit function for shadow banks specified in Buchak et al. [2024a].

3.1 Demand for deposit and loan products

(a) Demand for deposit products. There is a population of H_m^d savers in market m. Each saver has a fixed wealth that we normalize to one unit.[19] A saver has to decide whether to deposit her savings unit in a bank and, if so, in which one. Due to transportation costs, savers consider only banks with branches in their local market. In other words, banks can only attract deposits in markets in which they have branches.[20] Banks provide differentiated deposit products. The (indirect) utility for a saver from depositing her wealth in bank j in market m is (omitting the individual-saver subindex in variables u_{jm}^d and ε_{jm}^d):

$$u_{jm}^d \ = \ \mathbf{x}_{jm}^d \ \beta^d - \alpha^d \ p_{jm}^d + \xi_{jm}^d + \varepsilon_{jm}^d. \tag{4}$$

$\mathbf{x}_{jm}^d$ is a vector of characteristics of bank j (other than the deposit interest rate) and market m that are valued by depositors and observable to the researcher, such as the number of branches of bank j in the market. The vector β^d contains the marginal utilities of the characteristics $\mathbf{x}_{jm}^d$. Variable p_{jm}^d is the 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_{jm}^d represents other characteristics of bank j in market m that are observable and valuable to savers but unobservable for us as researchers. Variables ε_{jm}^d represent savers' idiosyncratic preferences, and we assume that they are independently and identically distributed across banks with type 1 extreme value distribution. The utility from the outside alternative is normalized to zero. Let $s_{jm}^d \equiv q_{jm}^d / H_m^d$ be the market share of bank j in the market for deposits at location m. The model implies that:

$$s_{jm}^d \ = \ \frac{1\left\{m \in \mathcal{M}_j^d\right\} \ \exp\left\{\mathbf{x}_{jm}^d \ \beta^d - \alpha^d \ p_{jm}^d + \xi_{jm}^d\right\}}{1 + \sum_{k=1}^{J} 1\left\{m \in \mathcal{M}_k^d\right\} \ \exp\left\{\mathbf{x}_{km}^d \ \beta^d - \alpha^d \ p_{km}^d + \xi_{km}^d\right\}}, \tag{5}$$

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

The vector of product characteristics $\mathbf{x}_{jm}^d$ includes four elements that are important for the implications of the model: (i) the number of branches, n_{jm}; (ii) the fraction of securitized loans, sec_{jm}; (iii) the bank's loans in the local market, $\ln(1 + q_{jm}^\ell)$; and (iv) the bank's total amount of deposits, $\ln(Q_j^d)$. The number of branches captures the effects of consumer transportation costs and awareness about the bank's presence. By including the bank's amount of loans in demand for deposits (and, as we will show below, the amount of deposits in the demand for loans), we try

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

[20] Honka et al. [2017] demonstrate the importance of branch presence for bank choice.

to capture, in a simple and parsimonious way, not only economies of scope and other synergies in demand for deposits and loans (i.e., one-stop banking) but also the two-sided-market nature of the banking business (see Section 4 of Vives [2016]).[21] The bank's total deposits capture consumers' concerns about the probability of default or bank run. Therefore, we have,

$$\mathbf{x}_{jm}^{d}\beta^{d} = \beta_{n}^{d}\ h(n_{jm}) + \beta_{rs}^{d}\ sec_{jm} + \beta_{\ell}^{d}\ \ln(1+q_{jm}^{\ell}) + \beta_{Q}^{d}\ \ln(Q_{j}^{d}). \tag{6}$$

$h(.)$ is a monotonic function. We use the function $s_{jm}^{d} = d_{jm}(p_{jm}^{d}, s_{jm}^{\ell}, Q_{j}^{d})$ to represent the demand for deposits, where, for notational convenience, we include explicitly as arguments the endogenous variables $(p_{jm}^{d}, s_{jm}^{\ell}, Q_{j}^{d})$.

(b) Demand for loan products. Each local market is also populated by investors/borrowers. Let H_{m}^{ℓ} be the number of new borrowers in market m. Each (new) borrower is endowed with an investment project requiring a loan unit.[22] A borrower's set of possible choices is not limited to the banks with branches in the market. Some banks sell mortgages in the market but do not have physical branches (recall that $\mathcal{M}_{j}^{d} \subseteq \mathcal{M}_{j}^{\ell}$). 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:

$$u_{jm}^{\ell} = \mathbf{x}_{jm}^{\ell}\ \beta^{\ell} - \alpha^{\ell}\ p_{jm}^{\ell} + \xi_{jm}^{\ell} + \varepsilon_{jm}^{\ell}. \tag{7}$$

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 a loan's interest rate 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:

$$s_{jm}^{\ell} = \frac{1\left\{m \in \mathcal{M}_{j}^{\ell}\right\}\ \exp\left\{\mathbf{x}_{jm}^{\ell}\ \beta^{\ell} - \alpha^{\ell}\ p_{jm}^{\ell} + \xi_{jm}^{\ell}\right\}}{1 + \sum_{k=1}^{J} 1\left\{m \in \mathcal{M}_{k}^{\ell}\right\}\ \exp\left\{\mathbf{x}_{km}^{\ell}\ \beta^{\ell} - \alpha^{\ell}\ p_{km}^{\ell} + \xi_{km}^{\ell}\right\}}. \tag{8}$$

The vector of product characteristics $\mathbf{x}_{jm}^{\ell}$ includes: (i) the number of branches, n_{jm}; (ii) the fraction of securitized loans, sec_{jm}; (iii) the bank's amount of deposits in the local market,

[21] To capture scope economies, we could consider a demand model for deposits and loans that endogenizes consumers' decisions to bundle deposits and mortgages in the same bank, as in Allen et al. [2019]. However, our dataset lacks information on bundling decisions, even in aggregate form, so our approach simplifies the way we capture this demand complementarity. Aggregating individual consumer decisions leads to the type of relationship between market shares represented in our model.

[22] In our empirical application, this will be a real estate investment.

$\ln(1+q^d_{jm})$; and (iv) the bank's total amount of deposits in all the markets, $\ln(Q^d_j)$. 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 synergies in the demand for deposits and loans 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 failure.[23] Thus, we have,

$$\mathbf{x}^\ell_{jm}\beta^\ell_m = \beta^\ell_n \; h(n_{jm}) + \beta^\ell_{rs} \; sec_{jm} + \beta^\ell_d \; \ln(1+q^d_{jm}) + \beta^\ell_Q \; \ln(Q^d_j), \tag{9}$$

We use the function $s^\ell_{jm} = \ell_{jm}(p^\ell_{jm}, s^d_{jm}, Q^d_j)$ to represent the demand for loans.

Naturally, there will be many instances where bank j's share of loans in market m is zero, and one might be concerned that these are primarily the result of a "small sample" problem arising because of a small number of potential customers in a county (see Gandhi et al., 2023 for a discussion). However, this is not the reason for zeroes in our case. Most bank-county-year observations in our dataset where loans are zero occur because they are zero in the population. That is, there are many counties where a bank does not make any loans. The observed zeroes in our market shares for loans (or deposits) are driven by banks' market entry decisions.

(c) Demand system for deposits and loans. The demand system can be represented by the equations $s^\ell_{jm} = \ell_{jm}(p^\ell_{jm}, s^d_{jm}, Q^d_j)$ and $s^d_{jm} = d_{jm}(p^d_{jm}, s^\ell_{jm}, Q^d_j)$. This system establishes links between deposits and loans within a local market and across different geographic markets. These links exist regardless of whether banks' pricing decisions internalize or not these spillover effects.

3.2 Variable cost function

We consider the following specification for the variable cost function:

$$\widetilde{C}_{jm}\left(q^\ell_{jm}, q^d_{jm}\right) \;\; = \;\; \left(\mathbf{x}^d_{jm}\;\gamma^d + \omega^d_{jm}\right)\; q^d_{jm} + \left(\mathbf{x}^\ell_{jm}\;\gamma^\ell + \omega^\ell_{jm}\right)\; q^\ell_{jm}. \tag{10}$$

Therefore, the marginal costs for deposits and loans are $c^d_{jm} \equiv \mathbf{x}^d_{jm}\;\gamma^d + \omega^d_{jm}$ and $c^\ell_{jm} \equiv \mathbf{x}^\ell_{jm}\;\gamma^\ell + \omega^\ell_{jm}$, respectively. Variables ω^ℓ_{jm} and ω^d_{jm} are unobservable to the researcher. The vector of observable variables $\mathbf{x}_{jm}$ includes the same variables as in the demand equations:

$$\begin{aligned}
\mathbf{x}^d_{jm}\;\gamma^d &= \gamma^d_n \; h(n_{jm}) + \gamma^d_{sec} \; sec_{jm} + \gamma^d_\ell \; \ln(1+q^\ell_{jm}) + \gamma^d_Q \; \ln(Q^d_j), \\
\mathbf{x}^\ell_{jm}\;\gamma^\ell &= \gamma^\ell_n \; h(n_{jm}) + \gamma^\ell_{sec} \; sec_{jm} + \gamma^\ell_d \; \ln(1+q^d_{jm}) + \gamma^\ell_Q \; \ln(Q^d_j).
\end{aligned} \tag{11}$$

The terms $\gamma^d_n \; h(n_{jm})$ and $\gamma^\ell_n \; h(n_{jm})$ 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

[23] Borrowers are concerned with bank failure because of the risk the acquiring bank may not renew their loans.

branches. The expressions $\gamma^d_\ell \ln(s^\ell_{jm})$ and $\gamma^\ell_d \ln(s^d_{jm})$ elucidate the economies of scope involved in managing deposits at the branch level. The component $\gamma^\ell_Q \ln(Q^d_j)$ delineates how the marginal cost of loans diminishes as the bank's aggregate deposit volume Q^d_j increases.

3.3 Social surplus

Let ss^ℓ_{jm} and ss^d_{jm} be the unit *social surpluses*, in dollar amounts, of the loan and deposit products of bank j in market m. The unit social surplus is the (average) consumer willingness to pay minus the unit cost. Given our specification of demand and costs, we have the following expressions for the unit social surpluses:

$$\begin{aligned} ss^\ell_{jm} &\equiv \frac{1}{\alpha^\ell}\left(\mathbf{x}^\ell_{jm}\beta^\ell + \xi^\ell_{jm}\right) - \left(\mathbf{x}^\ell_{jm}\gamma^\ell + \omega^\ell_{jm}\right) = \frac{1}{\alpha^\ell}\left(\mathbf{x}^\ell_{jm}\boldsymbol{\theta}^\ell + \eta^\ell_{jm}\right), \\ ss^d_{jm} &\equiv \frac{1}{\alpha^d}\left(\mathbf{x}^d_{jm}\beta^d + \xi^d_{jm}\right) - \left(\mathbf{x}^d_{jm}\gamma^d + \omega^d_{jm}\right) = \frac{1}{\alpha^d}\left(\mathbf{x}^d_{jm}\boldsymbol{\theta}^d + \eta^d_{jm}\right), \end{aligned} \tag{12}$$

where $\boldsymbol{\theta}^\ell \equiv \beta^\ell - \alpha^\ell\ \gamma^\ell$, $\boldsymbol{\theta}^d \equiv \beta^d - \alpha^d\ \gamma^d$, $\eta^\ell_{jm} \equiv \xi^\ell_{jm} - \alpha^\ell\ \omega^\ell_{jm}$, and $\eta^d_{jm} \equiv \xi^d_{jm} - \alpha^d\ \omega^d_{jm}$.

3.4 Bank competition and equilibrium

A key feature of our model is the rich interactions between deposit and loan markets and between these markets at different geographic locations. In our model, spillovers exist regardless of banks' pricing decisions; they are inherent to the specifications of market demand and costs described above. Any shock to local deposits changes local loan demand and marginal costs. Moreover, this change in the level of local deposits will affect bank j's total deposits, Q^d_j, which in turn cascades into broader impacts on bank j's operations across all counties in which it operates, influencing both deposits and loans.

For simplicity, we consider here a version of the model with separate Nash-Bertrand competition for deposits and loans at each local market.[24] 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.[25]

[24] Alternatively, one could also allow banks to internalize the spillover effects between loan and deposit markets, and between geographic markets when making pricing decisions. This generates an incentive to reduce price-cost margins in the two markets and positively affects market shares. Without price data, the system of equilibrium equations for market shares is the same if banks internalize the spillover effects as if they set prices to maximize local profit in each market. We have also derived the best-response pricing equations when banks internalize local and global spillover effects. The derivation of these equations is available in a previous version of this paper, https://repec.cepr.org/repec/cpr/ceprdp/DP13741.pdf.

[25] This assumption is supported by empirical evidence in Amel and Stahl [2016], who show that prices are tailored to local-market competition. They also find evidence of links in a bank's prices across markets, consistent

The marginal conditions of optimality imply the well-known pricing equations:

$$p^d_{jm} - c^d_{jm} = \frac{1}{\alpha^d(1-s^d_{jm})}, \qquad \textit{and} \qquad p^\ell_{jm} - c^\ell_{jm} = \frac{1}{\alpha^\ell(1-s^\ell_{jm})}, \tag{13}$$

where $c^d_{jm} \equiv \dfrac{\partial \widetilde{C}_{jm}}{\partial q^d_{jm}}$ and $c^\ell_{jm} \equiv \dfrac{\partial \widetilde{C}_{jm}}{\partial q^\ell_{jm}}$ represent marginal costs.

For our empirical analysis, it is convenient to write the equilibrium conditions regarding the market shares as the only endogenous variables. Let s^d_{0m} and s^ℓ_{0m} be the market shares of the outside alternative for deposits and loans in market m. The logit model implies that $\ln(s^d_{jm}/s^d_{0m}) = \mathbf{x}^d_{jm}\ \beta^d_m - \alpha^d\ p^d_{jm} + \xi^d_{jm}$ and $\ln(s^\ell_{jm}/s^\ell_{0m}) = \mathbf{x}^\ell_{jm}\ \beta^\ell_m - \alpha^\ell\ p^\ell_{jm} + \xi^\ell_{jm}$. Subbing the pricing equations into these expressions, we obtain the following system of equilibrium equations in terms of market shares:

$$\begin{aligned}
\ln\left(\frac{s^d_{jm}}{s^d_{0m}}\right) + \frac{1}{1-s^d_{jm}} &= \theta^d_n\ h(n_{jm}) + \theta^d_{rs}\ sec_{jm} + \theta^d_\ell\ \ln(1+H^\ell_m s^\ell_{jm}) + \theta^d_Q\ \ln(Q^d_j) + \eta^d_{jm}, \\
\ln\left(\frac{s^\ell_{jm}}{s^\ell_{0m}}\right) + \frac{1}{1-s^\ell_{jm}} &= \theta^\ell_n\ h(n_{jm}) + \theta^\ell_{rs}\ sec_{jm} + \theta^\ell_d\ \ln(1+H^d_m s^d_{jm}) + \theta^\ell_Q\ \ln(Q^d_j) + \eta^\ell_{jm},
\end{aligned} \tag{14}$$

where the vectors of parameters $\boldsymbol{\theta}^d$ and $\boldsymbol{\theta}^\ell$, and the unobservables η^d_{jm} and η^ℓ_{jm} are the same ones as in equation (12) for social surpluses, i.e., $\boldsymbol{\theta}^\ell \equiv \beta^\ell - \alpha^\ell\ \gamma^\ell$, $\boldsymbol{\theta}^d \equiv \beta^d - \alpha^d\ \gamma^d$, $\eta^\ell_{jm} \equiv \xi^\ell_{jm} - \alpha^\ell\ \omega^\ell_{jm}$, and $\eta^d_{jm} \equiv \xi^d_{jm} - \alpha^d\ \omega^d_{jm}$.[26]

The vector of parameters $\boldsymbol{\theta}$, together with the exogenous variables of the model, contains 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 identify demand and cost parameters separately. All our empirical results are based on the estimation of these parameters and the implementation of counterfactual experiments using the equilibrium mapping. Furthermore, the indexes $\mathbf{x}^d_{jm}\boldsymbol{\theta}^d + \eta^d_{jm}$ and $\mathbf{x}^\ell_{jm}\boldsymbol{\theta}^\ell + \eta^\ell_{jm}$ are up-to-scale measures of the social surpluses of deposits and loans, respectively, of bank j in market m. The identification of $\boldsymbol{\theta}^d$ and $\boldsymbol{\theta}^\ell$ (that we establish

with our network pricing model. However, even in the extreme scenario where a bank applies uniform pricing across all counties in which it operates, the described spillover effects persist. See also Driscoll and Judson [2013] and Hannan and Prager [2006] regarding non-uniformity of deposit pricing during our sample period, and Hendricks et al. [2024] for loan pricing.

[26] One might wonder why $\ln(Q^\ell)$ is not also included in this system of equations, since it might naturally be thought to affect lending costs. We have considered a specification including $\ln(Q^\ell)$. This has little impact on the main estimates; however, the sign on $\ln(Q^d)$ becomes negative in the loan equation, when both it and $\ln(Q^\ell)$ are included simultaneously. This is because $\ln(Q^\ell)$ and $\ln(Q^d)$ are highly correlated (correlation coefficient of 0.76). For these reasons, we omit $\ln(Q^\ell)$ from our main analysis, but we report these results in Table A9, along with results from a specification in which the collinearity problem is addressed more symmetrically via the inclusion of the sum of $\ln(Q^\ell)$ and $\ln(Q^d)$.

in Section 4) implies the identification up to scale of the social surpluses in deposit and loan markets. Though these up-to-scale surpluses are not measured in dollar amounts, we can still use them to make welfare comparisons, e.g., percentage changes in the social surplus.

The term $1/(1-s_{jm})$ in equation (14) captures market power. Removing this term is equivalent to assuming that price is equal to marginal cost and it has two direct effects on market shares in the local market: first, it reduces the market share of the outside alternative and increases total quantity; second, it reduces the variance of market shares and the degree of market concentration. In Appendix B we provide evidence that this measure of market power is consistent with the stylized facts about margins and spreads presented in recent work on the banking industry (including by Drechsler et al., 2017), and in Section 6.2, we simulate the equilibrium outcome that arises when this term is set to zero. Our findings suggest that markets become more competitive, implying that this term captures market power.

4 Identification and estimation of the structural model

The system of equations of the econometric model is:

$$\begin{aligned} y^d_{jmt} &= \mathbf{z}'_{mt}\theta^d_0 + \sum_{n=1}^{n_{\max}} \theta^d_n(n)\ \mathbb{1}(n_{jmt} \geq n) + \theta^d_{rs}\ sec_{jmt} + \theta^d_\ell\ \ln(1+q^\ell_{jm}) + \theta^d_Q\ \ln(Q^d_{jt}) + \eta^d_{jmt}, \\ y^\ell_{jmt} &= \mathbf{z}'_{mt}\theta^\ell_0 + \sum_{n=1}^{n_{\max}} \theta^\ell_n(n)\ \mathbb{1}(n_{jmt} \geq n) + \theta^\ell_{rs}\ sec_{jmt} + \theta^\ell_d\ \ln(1+q^d_{jm}) + \theta^\ell_Q\ \ln(Q^d_{jt}) + \eta^\ell_{jmt}, \end{aligned} \tag{15}$$

where $y^d_{jmt} \equiv \ln\left(s^d_{jm}/s^d_{0m}\right)+1/(1-s^d_{jm})$, $y^\ell_{jmt} \equiv \ln\left(s^\ell_{jm}/s^\ell_{0m}\right)+1/(1-s^\ell_{jm})$, and $\mathbb{1}(n_{jmt} \geq n)$ is the binary variable indicating that the number of branches n_{jmt} is greater than or equal to n, and $\mathbf{z}_{mt}$ is a vector of market characteristics capturing the relative value of the outside alternative. More specifically, $\mathbf{z}_{mt}$ captures county-level time-varying exogenous variables, including a housing price index, bankruptcy cases, income per capita, population, and age distribution.

(i) **Market size and market shares for deposits and loans**. To construct market shares we need first to construct market size variables H^d_{mt} and H^ℓ_{mt}. We postulate that market size is proportional to the total population in county m at period t, represented by variable POP_{mt}: $H^d_{mt} = \lambda^d\ POP_{mt}$ and $H^\ell_{mt} = \lambda^\ell\ POP_{mt}$ where λ^d and λ^ℓ are positive constants. The values of these coefficients are chosen such that the constructed market shares satisfy the model constraint that the sum of the market shares $\sum_j s^d_{jmt} = Q^d_{mt}/H^d_{mt}$ and $\sum_j s^\ell_{jmt} = Q^\ell_{mt}/H^\ell_{mt}$ are smaller than one for every county-year observation. More specifically, the values of these coefficients are

$\lambda^d = \max_{m,t}\left\{\frac{Q^d_{mt}}{POP_{mt}}\right\}$ and $\lambda^\ell = \max_{m,t}\left\{\frac{Q^\ell_{mt}}{POP_{mt}}\right\}$, which in our data are $\lambda^d = 382$ for deposit stocks and $\lambda^\ell = 84$ for new loans, measured in thousands of USD. Using POP_{mt} as a measure of market size and assuming that λ^d and λ^ℓ are constant across counties and over time may seem like strong restrictions. To control for measurement error, we include county $\times$ year fixed effects as explanatory variables in the model. Including these fixed effects makes our empirical results very robust to using alternative measures of market size, such as total county income.[27]

(ii) **Endogeneity.** The structural equations presented in (15) incorporate exogenous variables $\mathbf{z}_{mt}$, $\mathbb{1}(n_{jmt} \geq n)$, and sec_{jmt} alongside endogenous variables q^ℓ_{jmt}, q^d_{jmt}, and Q^d_{jt}. These endogenous variables are correlated with the error terms η^d_{jmt} and η^ℓ_{jmt} due to simultaneity. Below, we outline our assumptions aimed at addressing endogeneity.

The identification and estimation of the model are based on four assumptions: (i) a rich fixed effects specification of the unobservables; the assumption that the remaining bank-county-year transitory shocks are (ii) not correlated with the observable exogenous county characteristics, and (iii) not serially or spatially correlated, and (iv) the bank-year effects are not correlated with observable exogenous county characteristics. Assumptions ID-1 to ID-4 provide a formal description of our identifying restrictions. For the rest of this section, to avoid repetition, we present the assumptions and moment conditions using the deposit equation, but it is important to note that these same assumptions and moment conditions apply to the loans equation.

Assumption ID-1 [Fixed Effects]: The unobservables η^d_{jmt} and η^ℓ_{jmt} have the following component structure:

$$\eta^d_{jmt} = \eta^{d(1)}_{jm} + \eta^{d(2)}_{t} + \eta^{d(3)}_{mt} + \eta^{d(4)}_{jt} + \eta^{d(5)}_{jmt}. \tag{16}$$

$\eta^{d(1)}_{jm}$ denotes a bank-county fixed effect, $\eta^{d(2)}_{t}$ is a year fixed effect, $\eta^{d(3)}_{mt}$ represents a county-year fixed effect, $\eta^{d(4)}_{jt}$ indicates a bank-year fixed effect, and $\eta^{d(5)}_{jmt}$ accounts for shocks specific to each bank, county, and year combination. ■

Assumption ID-2: Regressors $\mathbf{z}_{mt}$, $\mathbb{1}(n_{jmt} \geq n)$, and sec_{jmt} are strictly exogenous with respect to shocks $\eta^{d(5)}_{jmt}$ and $\eta^{\ell(5)}_{jmt}$. For any pair of markets (m, m') and any pair of years (t, t'), we have that $(\mathbf{z}_{mt},\ \mathbb{1}(n_{jmt} \geq n),\ sec_{jmt})$ are not correlated with $\left(\eta^{d(5)}_{jm't'},\ \eta^{\ell(5)}_{jm't'}\right)$. ■

[27] We do not use the number of loan applications to measure of market size because, as explained in Agarwal et al. [2024], many prospective borrowers apply multiple times for a loan before ultimately obtaining financing or abandoning search altogether. According to their data on mortgages from a large GSEy in the US, the overall median number of applications per person is nine, and just two for those ultimately financed. Therefore, although we know the number of applications, since the HMDA data do not allow us to identify individual applicants, we cannot be sure of the number of applicants.

Assumption ID-3: Bank-county-year shocks $\eta^{d(5)}_{jmt}$ and $\eta^{\ell(5)}_{jmt}$ are not serially or spatially correlated. ■

Assumption ID-4: County characteristics in $\mathbf{z}_{mt}$ are exogenous regressors with respect to the bank-year shocks $\eta^{d(4)}_{jt}$ and $\eta^{\ell(4)}_{jt}$. For any market m and bank j, $\mathbb{E}\left(\mathbf{z}_{mt}\ \eta^{d(4)}_{jt}\right) = 0$. ■

Consider the following difference-in-difference (DiD) transformation of the structural equations in (15). First, a difference between the equations of two banks operating in the same county. This transformation eliminates the national-level effect $\eta^{d(2)}_{t}$ and the county-year effect $\eta^{d(3)}_{mt}$ from the error term. We use the $\sim$ symbol to represent this difference, e.g., $\widetilde{y}^{d}_{jmt} \equiv y^{d}_{jmt} - y^{d}_{j^*_m mt}$, where j^*_m is a baseline bank active at county m that we select to make this transformation. Second, a time difference between the equations at two consecutive periods. This transformation eliminates the bank-county fixed effect $\eta^{d(1)}_{jm}$ from the error term. We use the symbol Δ to represent this time difference transformation, e.g., $\Delta\widetilde{y}^{d}_{jmt} \equiv \widetilde{y}^{d}_{jmt} - \widetilde{y}^{d}_{jm,t-1}$. The DiD transformation is:

$$\Delta\widetilde{y}^{d}_{jmt} \;\;=\;\; \Delta\widetilde{\mathbf{x}}^{d}_{jmt}\ \theta^{d} + \Delta\widetilde{\eta}^{d(4)}_{jt} + \Delta\widetilde{\eta}^{d(5)}_{jmt}, \tag{17}$$

where $\mathbf{x}^{d}_{jmt}$ is the vector of regressors defined in equation (6).

We can also apply a third difference to eliminate the bank-year component of the error term. Let the $*$ symbol represent the difference between two counties where the bank is active, e.g., $y^{*d}_{jmt} \equiv y^{d}_{jmt} - y^{d}_{jm^*_j t}$, where m^*_j is a baseline county in the network of bank j. Therefore, we have the difference-in-difference-in-difference (DiDiD) transformation of the structural equations:

$$\Delta\widetilde{y}^{*d}_{jmt} \;\;=\;\; \Delta\widetilde{\mathbf{x}}^{*d}_{jmt}\ \theta^{d} + \Delta\widetilde{\eta}^{*d(5)}_{jmt}. \tag{18}$$

Note this DiDiD transformation removes the bank's total deposits, $\ln Q_{jt}$, from the vector of explanatory variables. Therefore, this equation cannot be used to identify parameters θ^{d}_{Q} and θ^{ℓ}_{Q}. However, as shown below, these parameters can be identified from the DiD equation.

Assumptions ID-2, ID-3, and ID-4 imply moment conditions (or valid instrumental variables) in the transformed equations. First, assumptions ID-2 and ID-3 imply the following moment conditions in the DiDiD equations:

$$\mathbb{E}\left(\left[\mathbf{z}_{mt},\ \mathbb{1}(n_{jmt} \geq n),\ sec_{jmt},\ \mathbf{x}_{jm,t-s},\ \mathbf{x}_{jm't}\right]\ \Delta\widetilde{\eta}^{*d(5)}_{jmt}\right) = 0, \tag{19}$$

for any $s \geq 2$, and m' represents any county that is a neighbor of a county contiguous to m. These moment conditions identify parameters $\theta^{d}_{n}(n)$, $\theta^{\ell}_{n}(n)$, θ^{d}_{ℓ}, and θ^{ℓ}_{d}, by combining dynamic panel models or Arellano-Bond moment conditions (Arellano and Bond, 1995) – represented by $\mathbf{x}_{jm,t-s}$ – with Hausman moment conditions (Hausman [1996]) – represented by $\mathbf{x}_{jm't}$.

Importantly, assumption ID-3 of no serial and spatial correlation in $\eta_{jmt}^{d(5)}$ and $\eta_{jmt}^{\ell(5)}$ is testable using the residuals of the estimated equations. In our empirical results, we provide evidence supporting the assumption of no serial correlation in both the loan and deposit equations. As for the spatial correlation, for the loan equation, we reject the hypothesis that shocks are not correlated with those in neighboring counties or in neighbor-of-neighbor counties. Conversely, we cannot reject this latter hypothesis for the deposit equation.

Based on the results of these tests, we consider the following instruments in the DiDiD transformation of the model. In the estimation of the deposit equation, alongside the exogenous regressors, we instrument the endogenous regressor $\ln(1 + q_{jmt}^{\ell})$ using $\ln(1 + q_{jm,t-2}^{\ell})$, and the logarithm of loans issued by bank j in all counties neighboring m's nearest neighbors. In the estimation of the loan equation using the DiDiD transformation, we instrument the endogenous regressor $\ln(1 + q_{jmt}^{d})$ using $\ln(1 + q_{jm,t-2}^{d})$ and the count of j's branches in county m lagged two periods, and we do not use spatial instruments.

Second, assumptions ID-2 and ID-4 imply the following moment conditions in the DiD equations. For any (m, m', j):

$$\mathbb{E}\left(\mathbf{z}_{m't}\ \left[\Delta\widetilde{\eta}_{jt}^{d(4)} + \Delta\widetilde{\eta}_{jmt}^{d(5)}\right]\right) = 0. \tag{20}$$

These moment conditions identify θ_Q^d and θ_Q^{ℓ}. Intuitively, they imply that, after controlling for bank $\times$ county and county $\times$ year fixed effects in the error terms, we can use the exogenous socioeconomic characteristics in markets other than m where the bank is 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 $\ln Q_{jt}^d$. Socioeconomic characteristics in other markets do not have a direct effect on the structural equation for market m, satisfying an exclusion restriction. Furthermore, the model implies that these characteristics should affect the total volume of bank deposits; therefore, they are relevant instruments. This identification strategy is in the same spirit as the approaches in Gilje et al. [2016], Cortés and Strahan [2017], and Nguyen [2019].[28] We can apply this identification approach to every bank-county-year observation as long as the bank's network includes multiple counties and the county has more than one active bank.

Specifically, when estimating the parameters θ_Q^d and θ_Q^{ℓ} in the DiD transformation of deposit and loan equations, we instrument the endogenous regressor $\ln(Q_{jt}^d)$ using the logarithm of the total per capita income from all other counties where bank j operates branches in year t.

[28] Gilje et al. [2016] use shale gas discoveries in a county as exogenous shocks and study how they generate increased lending in counties connected through branch networks. Cortés and Strahan [2017] exploit exogenous variation provided by natural disasters, and Nguyen [2019] uses bank mergers. Our approach uses similar sources of local exogenous variation, but is more general since not limited to dramatic local shocks.

(iii) **Switching regressions in the loan equation.** Depository institutions face an important decision in loan markets: whether to operate with physical branches and local deposits or opt for a business model without such local infrastructure. Beyond the straightforward impact of branch numbers and deposit amounts on loan demand and costs, the rest of the parameters influencing the social surplus of loans may differ between these business models, e.g., the parameter associated with the securitization rate and the different fixed effects. Moreover, the choice between the two business models can be endogenous. Consequently, the loan equation can be framed as a switching regression model with two regimes as follows:

$$y^{\ell}_{jmt} = \begin{cases} \mathbf{x}^{\ell}_{jmt}\ \boldsymbol{\theta}^{\ell}(1) + \eta^{\ell}_{jmt}(1) & \text{if } N_{jmt} = 1 \\ \mathbf{x}^{\ell}_{jmt}\ \boldsymbol{\theta}^{\ell}(0) + \eta^{\ell}_{jmt}(0) & \text{if } N_{jmt} = 0, \end{cases} \tag{21}$$

$N_{jmt} \in \{0, 1\}$ denotes the binary decision regarding branch presence. $\boldsymbol{\theta}^{\ell}(0)$ and $\boldsymbol{\theta}^{\ell}(1)$ are two different vectors of parameters, with the elements of $\boldsymbol{\theta}^{\ell}(0)$ associated with a number of branches and amount of local deposits are constrained to be zero. We estimate the parameters of this model using a two-step control function approach *à la Heckman*, alongside GMM estimation. In the initial step, we estimate a linear probability model, with a branch presence indicator as the dependent variable, to estimate the propensity score. In the next step, we employ GMM estimation for the loan equation, integrating the propensity score as an additional regressor. This yields two sets of GMM estimates, one for each regime in the switching regression model. Identification hinges on an exclusion restriction where some regressors in the selection equation for the propensity score do not appear in the loan equation. These include dummies for branch presence in the preceding year in (i) the same county and (ii) in a neighboring county. The binary choice model incorporates all other exogenous variables and fixed effects from the loan equation as explanatory variables.

(iv) **Treatment of shadow banks.** Shadow banks are not depository institutions and so are not included in the estimation of the deposit equation. Furthermore, even if shadow banks operate through multiple establishments, these offices play a very different role than branch networks in traditional banks. The decision-making for shadow banks with regard to lending activities does not involve the dilemma of operating in a county with or without branches. Consequently, for shadow banks, the switching regression model for the loans equation simplifies to a single regime without branches. This loan equation incorporates securitization (their main source of financing) along with the rich set of fixed effects. By integrating shadow banks into our model as 'inside players', rather than merely considering them as part of the outside option, we can

investigate the potential impacts of various simulated scenarios on these entities.

5 Estimation results

Tables 2 and 3 present estimation results of the structural equation for deposits and loans, respectively. As mentioned above, the right-hand side of the equilibrium equations expressed in equation (15) represents social surplus. This surplus and the θ parameters are not measured in monetary units (dollars) but in utils. However, the dependent variables in these regressions are close to the logarithm of market shares. This enables us to effectively compare the parameters of the two equations, interpreting them as elasticities (provided the explanatory variable is also in logarithmic form) or semi-elasticities.

We report OLS Fixed-Effects (without instrumenting) and GMM (DiD and DiDiD) estimates. Compared to GMM, OLS underestimates the effect of the number of branches and the magnitude of local spillovers between deposits and loans. The main difference between the two sets of estimates is in the effect of total deposits on local loans and deposits. Statistical tests support the validity of our moment conditions/instruments.[29] For the rest of the paper, we focus on the GMM estimates.[30]

Looking first at Table 2 we see that the number of branches a bank has in a county has a substantial effect on the social surplus for a deposit product. The marginal effect of an additional branch declines with the number of branches: a second branch increases the social surplus by 67%; a third branch by 38%; a fourth branch by 29%; a fifth branch 36%; and subsequent branches by (on average) 3%. Turning to our main parameters of interest, we identify moderate economies of scope between deposits and loans: the elasticity of deposits with respect to loans is 0.10. Finally, a bank's total amount of deposits at the national level has a small effect on its social surplus at every local market where it operates.

Loan equation results are presented in Table 3. As discussed in Section 4 estimation of the loan equation involves overcoming a selection problem by using a control function approach a la Heckman. In the first step, both the number of branches in the last period and the number of branches in the neighboring counties in the last period, are strong predictors of having a branch in the local county in the current period. Their coefficients are 0.012 and 0.003, with standard errors

[29]Our GMM estimates successfully pass the Arellano-Bond (m2) serial correlation test, supporting the validity of the dynamic instruments. We also show in Appendix C that our results remain economically consistent when using different instruments sets, although the chosen specifications do not pass over-identifying restriction tests.

[30]County characteristic effects can be estimated by OLS, with all coefficients of expected sign.

Table 2: Estimation of Structural Equation for Deposits

Variable	OLS Fixed Effects		GMM DiD & DiDiD	
Number of branches				
Second branch	0.5263***	(0.0107)	0.6730***	(0.0102)
Third branch	0.2901***	(0.0086)	0.3764***	(0.0081)
Fourth branch	0.2316***	(0.0090)	0.2863***	(0.0084)
Fifth branch	0.2804***	(0.0109)	0.3552***	(0.0105)
> Fifth	0.0294***	(0.0027)	0.0315***	(0.0024)
% of loan being resold	-0.0568***	(0.0062)	-0.0408***	(0.0148)
Econ. of scope and total deposits				
log own loans in county	0.0445***	(0.0019)	0.0964***	(0.0151)
log own total deposits	0.3117***	(0.0095)	0.0346***	(0.0058)
Market characteristics				
log county income	0.0507	(0.0364)		
log county population	-0.3039***	(0.0546)		
share population age ≤ 19	2.2449***	(0.5778)		
share population age ≥ 50	1.6438***	(0.3618)		
log housing price index	0.2537***	(0.0212)		
log number of bankruptcy filings	0.0121***	(0.0044)		
log number of loan applications	-0.0267***	(0.0080)		
Bank x County Fixed Effects	YES		YES (Implicit in DiD)	
Time Dummies	YES		NO	
County x Time Dummies	NO		YES (Implicit in DiD)	
Bank x Time Dummies	NO		YES (Implicit in DiDiD)	
Number of observations	236,498		241,911	
R-square	0.9550			
S-H overid test: p-value			0.0000	
No serial correlation-m2: p-value			0.2487	
No spatial correlation: p-value			0.1769	

Note: Sample Period is 1998-2010. Robust standard errors of serial correlation and heteroscedasticity are reported in parentheses. * means p-value < 0.05; ** means p-value < 0.01; *** means p-value < 0.001.

Table 3: Estimation of Structural Equation for Loans

	Depository Banks						**Shadow Banks**	
	With branches				**W/o branches**			
	OLS		GMM		GMM		GMM	
Variable	Fixed Effects		DiD & DiDiD		DiD & DiDiD		DiDiD	
Number of branches								
Second branch	0.130***	(0.016)	0.1338***	(0.0240)				
Third branch	0.088***	(0.016)	0.0826***	(0.0165)				
Fourth branch	0.087***	(0.017)	0.0828***	(0.0155)				
Fifth branch	0.138***	(0.020)	0.1084***	(0.0181)				
> Fifth	0.020***	(0.002)	0.0101***	(0.0020)				
Securitization								
% of loans resold	1.016***	(0.018)	0.6632***	(0.0306)	0.6978***	(0.0057)	0.0613***	(0.0041)
Econ of scope and Q^d								
log own local deposits	0.190***	(0.009)	0.2963***	(0.0334)				
log own total deposits	0.291***	(0.015)	0.1727***	(0.0157)	0.4111***	(0.0104)		
Market characteristics								
log county income	0.319***	(0.071)						
log county population	-1.222***	(0.096)						
share pop. age $\leq$ $ 19	-3.592***	(0.929)						
share pop. age $\geq$ 50	-0.734	(0.667)						
log house price index	0.347***	(0.042)						
log nbr bankruptcy	-0.050***	(0.010)						
log nbr loan applications	0.670***	(0.017)						
Selection – Control function								
Propensity score	0.053	(0.374)	-0.6455*	(0.3409)	2.2301***	(0.0563)		
Propensity score square	-0.047	(0.364)	0.9942***	(0.3392)	-5.8633***	(0.5022)		
Propensity score cubic	-0.041	(0.116)	-0.3515***	(0.1052)	7.9259***	(0.8210)		
Fixed effects								
Bank x County	YES		YES		YES		YES	
Time	YES		NO		NO		NO	
County x Time	NO		YES		YES		YES	
Bank x Time	NO		YES		YES		YES	
Number of observations	194,267		196,090		797,927		2,351,846	
R-square	0.872							
Joint sign. Selection: pvalue	0.1008		0.0000		0.0000			
S-H overid test: pvalue			0.0360					
No serial correlation-m2: pvalue			0.2659					
No spatial corrrelation: pvalue			0.0000					

Note: Sample Period is 1998-2010. Robust standard errors of serial correlation and heteroscedasticity are reported in parentheses. * means p-value < 0.05; ** means p-value < 0.01; *** means p-value < 0.001.

being 0.0013 and 0.00021, respectively. We also control for all the other exogenous variables and fixed effects used in the loan equation. The goodness-of-fit in this first step is very high with an R-squared of 0.94. The control function consists of a third order polynomial in the propensity score, which is the fitted value obtained from the first step. The estimates of the control function coefficients are significant, both individually and jointly. Columns (2) and (3) present estimates from observations where depository institutions make loans and receive deposits. Columns (4) and (5) present results for observations where depository institutions make loans but do not receive deposits, while columns (5) and (6) presents results for shadow banks.

The effect of the number of branches on the social surplus of a loan product (semi-elasticity) is important but smaller than for deposits: a second branch increases social surplus by 13%; a third by 8%; a fourth by 8%; a fifth by 11%; and subsequent branches by 1%.[31] The positive and significant coefficient associated with securitization underscores the pivotal role of this funding mechanism in facilitating lending activities for traditional banks. This impact is large for these depository institutions, regardless of whether they have branches in the county or not.

We find that doubling a bank's deposits within a county implies a 30% increase in the social surplus of the bank's loans. Furthermore, a bank's total deposits at the national level substantially affect the social surplus in all local markets where it operates. A 100% surge in a bank's nationwide deposits leads to a 17% increase in the social surplus of loans if the bank has local branches and of 41% if it does not. This asymmetric effect appears intuitive since, in the absence of local deposits generated by the presence of local branches, a bank must heavily rely on the liquidity generated in other parts of its branch network to finance its local loans. These findings support the notion that a bank's internal liquidity is pivotal in facilitating lending.

It is noteworthy that the elasticity of the social surplus of loans in relation to the securitization rate differs significantly between shadow banks and depository banks. The elasticity stands at 0.06 for shadow banks, contrasting with 0.66 and 0.70 for depository banks. The substantial gap between the average securitization rates of traditional banks (21%) and shadow banks (77%) is likely a significant factor in explaining the difference in the marginal effects of securitization between these two groups of banks. In particular, there should be a considerable difference in the risk levels of the marginal securitized loans between these two types of institutions.

[31]The estimated parameters for the impact of the fifth branch appear anomalous as they disrupt the expected monotonicity in the regression function with respect to the number of branches. This apparent deviation from monotonicity arises from the linear specification for branches exceeding a certain threshold, denoted as n_{max}. It is worth noting that observations with more than five branches are scarce.

6 Counterfactual experiments

In this section we use the estimates from our model to learn more about the factors that influence the geographic flow of funds and the provision of credit, and to perform a number of policy-related experiments. We split our analysis in two parts. In Section 6.1 we present a decomposition exercise in which we alternatively shut down (i) liquidity flows from branch networks, (ii) local spillovers, (iii) local market power, (iv) and shadow banks. Results are presented in Table 4. Then, in Section 6.2 we first evaluate the impact of the Riegle-Neal Act on the flow of funds. Second, we study the consequences from the introduction of a deposit tax that approximates the impact of inflation taxing away the real value of nominal deposits. Lastly, we consider the effect of an increase in the interbank rate to reflect a contractionary policy initiated by the Federal Reserve. Findings are presented in Table 5. Additional results related to these experiments are presented in Appendix E.

For all the experiments, we use the GMM estimates for the structural parameters $\boldsymbol{\theta}$, obtain the model residuals, and then apply OLS to estimate the five different groups of fixed effects in the error terms η^d_{jmt} and η^ℓ_{jmt}. With the exception of the branch-network experiments, in conducting our analysis, we maintain the number of branches of a bank within each county at the observed value in the dataset. It seems intuitive that branch networks will evolve over time in response to the changes under consideration. Therefore, these experiments should be viewed as short-run responses to unanticipated changes in policy.

The system of equilibrium equations in (14) can be succinctly expressed as $\Psi(\mathbf{s}) = 0$, with $\mathbf{s}$ denoting the vector with the market shares of deposits and loans across all bank-counties within the networks of banks, namely $\mathcal{M}^d_j$ and $\mathcal{M}^\ell_j$. A solution to this system is an equilibrium of the model. By Brower's Theorem, an equilibrium exists, as $\Psi(.)$ is a continuous mapping, and the space of $\mathbf{s}$ is compact. To deal with multiplicity of equilibria in implementing counterfactual experiments, we apply a simple algorithm. In the first step, we compute the unique equilibrium for the vector of loan shares (s^ℓ_{jm}) in every county taking as given Q^d_j and the s^d_{jm} for every j. We then compute the unique equilibrium of deposit shares (s^d_{jm}) taking as given Q^d_j and the s^ℓ_{jm} for every j computed in the first step, and then we aggregate over counties to get Q^d_j for every j. Finally, we repeat each step such that each endogenous variable has been updated twice.[32] A nice feature of this approach is that the local equilibrium at each step is unique and easily computed using a Bisection algorithm. See Appendix D for a detailed description.

[32]We present results for two rounds of updating, but our results are similar if we update until Q^d_j converges.

We measure the effects of these counterfactual experiments by looking at the following statistics or outcome variables: (i) National Imbalance Index,[33] (ii) Median bank-level Imbalance Index (for depository institutions), (iii) Value of loans at the national level, (iv) Total social surplus in loan markets, (v) Value of loans broken down by depository and shadow banks, (vi) Value of loans across the top 100 counties and across the bottom 2500 counties ranked by loan amounts / income per capita / % urban population, (vii) Value of deposits at the national level, and (viii) Total social surplus in deposit markets.

The statistics that we present here capture a key trade-off in the geographic distribution of credit. A higher Imbalance Index implies that a larger share of bank funds is moved across counties such that credit can be used in those locations with more social surplus for loans. However, this movement of bank funds can generate not only winners, but also losers. Some counties may end up with very limited amounts of credit.

6.1 Decomposition experiments

6.1.1 Experiment 1: Eliminating liquidity flows through branch networks

First, we investigate the importance of branch networks for generating funding by considering the counterfactual equilibrium that would arise if each county in which a bank operated relied only on deposits collected from that county and not from its entire branch network. To operationalize this experiment, we set to zero all liquidity flowing into market (county) m from other markets and examine the impact on local lending and the other outcome variables. This experiment has a significant impact on banks whose operations feature counties where the amount of credit offered exceeds the locally generated deposits since, in these counties, they rely largely on liquidity from elsewhere. At the extreme, many banks make loans in counties where they have no branch presence at all, so under this experiment, access to credit would fall sharply therein.[34] Therefore, eliminating liquidity flows from throughout the branch networks should make local loans more closely related to local deposits, such that bank-level Imbalance Index scores fall.

Our findings are consistent with this. From Table 4, we can see that the median bank Imbalance Index declines substantially from 0.273 to 0.040 – that is, the proportion of funds that banks transfer between counties in which they operate declines by roughly 87 percent.

[33] In Tables 4 and 5 we present the mean national Imbalance Index over the 13 years of our sample. In Figure A13 of Appendix E, we present the full evolution across the years.

[34] It is therefore important to note that we hold fixed the amount of securitized loans, when naturally it is likely that a bank without access to funding from its deposit base would exhibit a substantially higher local securitization rate than what we observe in the data.

Table 4: Decomposition Experiments

Outcomes (13 yr avr)	Data	Exp. 1 No branch network	Exp. 2 No EOS	Exp. 3 No market power	Exp. 4 No Shadowbanks
Imbalanced Index					
National level	0.3017	0.3161(4.8%)	0.3525(16.8%)	0.2508(-16.9%)	0.2945(-2.4%)
Bank level (Median)	0.2725	0.0404(-85.2%)	0.8394(208.0%)	0.2626(-3.6%)	0.2732(0.3%)
Loans (in B$)					
Tot. value	1946	1402(-28.0%)	1366(-29.8%)	5164(165.4%)	956(-50.9%)
Tot. social surplus	2095	1484(-29.2%)	1445(-31.0%)	6203(196.1%)	991(-52.7%)
By bank types					
Depository banks	887	301(-66.1%)	263(-70.3%)	2928(230.1%)	956(7.8%)
Shadow banks	1058	1101(4.1%)	1103(4.3%)	2236(111.3%)	0(-100.0%)
By county loans					
Top 100	1109	829(-25.2%)	790(-28.8%)	2769(149.7%)	534(-51.8%)
Bottom 2500	166	104(-37.3%)	107(-35.5%)	528(218.1%)	93(-44.0%)
By county income/cap					
Top 100	543	394(-27.4%)	383(-29.5%)	1303(140.0%)	282(-48.1%)
Bottom 2500	456	313(-31.4%)	317(-30.5%)	1335(192.8%)	222(-51.3%)
By county urban%					
Top 100	769	576(-25.1%)	539(-29.9%)	1959(154.7%)	373(-51.5%)
Bottom 2500	269	174(-35.3%)	181(-32.7%)	804(198.9%)	145(-46.1%)
Deposit (in B$)					
Tot. value	4695	4172(-11.1%)	1807(-61.5%)	12948(175.8%)	4716(0.4%)
Tot. social surplus	5004	4446(-11.2%)	1870(-62.6%)	14724(194.2%)	5027(0.5%)

Overall, there is a substantial reduction in the total volume of loans, from about \$1.9 to \$1.4 trillion. Note that this decrease is driven entirely by depository institutions, who see their lending activity fall by over 66 percent. The decrease in lending is more pronounced in counties that are (i) smaller in initial lending activity, (ii) poorer, and (iii) more rural.

Note, that the large effect on lending activity is the result not only of the elimination of total liquidity, but also the subsequent negative synergies elimination entails. The total value of deposits decreases by over 11% because of economies of scope, and this decrease in turn causes lending to fall further. We also find that shadow banks actually benefit indirectly under this experiment. They experience a slight increase in lending activity because depository institutions become less attractive.

Although bank-level Imbalance Index scores fall, the national Imbalance Index actually increases. This is somewhat counterintuitive, but can be explained by the fact that, at the county level, there is a positive correlation between the total volume of deposits and the presence of banks making loans without branches. Counties with high levels of deposits are the large/rich counties to which it is profitable to extend credit even without any branches. As explained above, counties with many banks making loans without branches are affected most in this experiment, suffering larger loan volume losses. The result is that in these counties, the correlation between deposits and loans falls, causing the national Imbalance Index to increase.

Takeaway: The results from this experiment confirm the importance of branch networks for spreading liquidity across the different markets in which banks operate, especially in socially disadvantaged regions.

6.1.2 Experiment 2: Eliminating local synergies

In this experiment, we study the effects of eliminating local synergies (economies of scope) between deposits and loans. We implement this experiment by setting the parameters θ_ℓ^d and θ_d^ℓ to zero and computing the new equilibrium of the model.

Economies of scope contribute to the home bias in credit distribution, resulting in lower Imbalance Index scores. Therefore, intuitively, eliminating economies of scope should push the bank-level Imbalance Index higher. Indeed, our results suggest that the median bank-level Imbalance Index increases by 208%. The national Imbalance Index also increases by 16.8%. The overall level of credit provision falls by 29.8%, with almost all of the decrease coming from traditional banks. Since they do not have any deposit funding, shadow banks are only affected in this experiment because traditional banks become relatively less attractive. Eliminating synergies implies that the decrease in lending activity has a effect on deposit activity, such that the total value of deposits falls by 61.5%.

Our findings also suggest that eliminating synergies has a stronger effect on counties at the bottom of the tail of the loan distribution, that are poorer, and that are more rural. These counties benefit most from synergies, keeping funds for use in the markets where they are generated.

Takeaway: The results from this experiment demonstrate that economies of scope are important for generating a home bias and preventing funds from flowing away from where they are generated. This is especially true in socially disadvantaged regions.

6.1.3 Experiment 3: Removing local market power

In this experiment we obtain the equilibrium of the model under the condition that banks behave competitively in all local markets banks, charging prices equal to marginal cost. Recall from our system of equations in (15) that the left-hand-sides are defined as $y^d_{jmt} \equiv \ln\left(s^d_{jm}/s^d_{0m}\right) + 1/(1 - s^d_{jm})$ and $y^\ell_{jmt} \equiv \ln\left(s^\ell_{jm}/s^\ell_{0m}\right) + 1/(1 - s^\ell_{jm})$, where $1/(1 - s^d_{jm})$ and $1/(1 - s^\ell_{jm})$ are the price - cost margins of firm j in market m, measured in utils. The experiment involves eliminating the two margin terms.

Our findings suggest that the median bank-level Imbalance Index falls by 3.6% and that the national Imbalance Index decreases by 16.9%. The total volume of both loans and deposits increases substantially, implying that market power has a strong negative effect on social surplus. The effect is largely driven by depository institutions, but shadow banks also engage in significantly more lending as a result of the removal of local market power.[35]

Eliminating market power has a stronger effect in smaller/poorer counties, where loan volumes roughly triple. Market power is higher in these counties such that banks operating in them reduce the supply of loans relative to the level that would arise under perfect competition, and they move these funds to richer more competitive counties where they may not have branches/deposits. As a result smaller/poorer counties typically have lower loan-to-deposit ratios. Removing local market power will therefore increase the supply of loans from banks in these smaller/poorer counties, making loans closer to deposits, reducing the Imbalance Index.

Finally, it is worth pointing out that the fact that all of the results from this experiment are consistent with our expectations of the impact of shutting market power down helps to validate our structural approach and confirm that the absence of price data does not prevent us from developing a meaningful measure of market power.

Takeaway: Local market power has a strong negative effect on lending and social surplus. The effect is felt most strongly in socially disadvantaged counties.

6.1.4 Experiment 4: Removing shadow banks

Our final decomposition experiment sheds light on the role of shadow banks. We operationalize this experiment by eliminating shadow banks altogether. Doing so reduces lending by 50.9%, with the decrease being felt more strongly in high-volume and urban markets. Note from column 1 of Table 4 that shadow banks actually account for 54.4% of loan volume, but when they are

[35] At first glance, the magnitude of the changes to loans/deposits under this counterfactual may seem surprisingly large, but in fact they are quite sensible when one considers the size of deposit/loan-rate elasticities.

removed from the market, depository institutions pick up some of the slack, such that their lending activity increases by 7.8%. Because of economies of scope, deposits increase by 0.4%.

The National Imbalance Index falls by 2.4%. This is because shadow banks are more likely to operate (make loans) in counties where depository institutions also lend heavily, such that they have relatively high loan-to-deposit ratios. Therefore, when shadow banks are removed, the distribution of loans approaches the distribution of deposits at the county level, decreasing the Imbalance Index. In Appendix E, panel (D) of Figure A13 displays the counterfactual evolution of the national Imbalance Index over the 13 years of our sample. While the average decrease resulting from this experiment is 2.4%, that actually masks important heterogeneity. In the early years of the sample, the counterfactual index is well below the factual, with a maximum decrease of 6.4% in 2001. As we approach the Great Financial Crisis, the decrease becomes smaller and actually becomes positive in 2009 and 2010.

Takeaway: Shadow banks play an important role in credit provision, especially in large and urban markets. The ability of depository banks to substitute for shadow bank credit provision is limited.

6.2 Policy Counterfactuals

6.2.1 Experiment 5: Splitting up multi-state banks (i.e., undoing Riegle-Neal)

Here, we use our model to understand the impact on the flow of funds of the Riegle-Neal Act of 1994, which allowed banks to operate branch networks in multiple states. To operationalize this experiment we divide every multi-state bank in our sample into different independent banks, one for each state. Only about 10% of depository institutions in our sample have branches in multiple states, but over 50% make loans in multiple states.[36]

Multi-state branch networks help credit flow across regions and make local lending less reliant on local deposits. A direct effect of breaking up a traditional bank's branch network is to shrink its total deposit volume since it no longer has access to any deposits generated out-of-state. This reduction in total deposits contributes to reducing the social surpluses of loans and deposits, by 19.9% and 4.5%, respectively. Furthermore, at the local level, loans become more closely related to deposits, such that the bank-level Imbalance Index falls by about 78%. For the same reasons as in Experiment 1, the national Imbalance Index increases. Finally, shadow banks actually benefit slightly under this scenario, since depository institutions become relatively less

[36]For details, see Table A3 in Appendix A.3. Note also that a significantly larger proportion of shadow banks extend their operations across state borders, with nearly 75% engaging in lending activities across multiple states.

attractive.[37]

Takeaway: Riegle-Neal impacted lending activity and the flow of funds, but the effect was moderated by the fact that cross-border lending activity without branches is somewhat limited.

Table 5: Policy Counterfactuals

Outcomes (13 yr avr)	Data	Exp. 5 Single-state networks	Exp. 6 Deposit tax	Exp. 7 Interbank rate increases 50 bps
Imbalanced Index				
National level	0.3017	0.3079(2.1%)	0.3085(2.3%)	0.2871(-4.8%)
Bank level (Median)	0.2725	0.0600(-78.0%)	0.2740(0.6%)	0.2687(-1.4%)
Loans (in B$)				
Tot. value	1946	1576(-19.0%)	1881(-3.3%)	1625(-16.5%)
Tot. social surplus	2095	1679(-19.9%)	2021(-3.5%)	1728(-17.5%)
By bank types				
Depository banks	887	490(-44.8%)	818(-7.8%)	834(-6.0%)
Shadow banks	1058	1087(2.7%)	1063(0.5%)	791(-25.2%)
By county loans				
Top 100	1109	911(-17.9%)	1075(-3.1%)	927(-16.4%)
Bottom 2500	166	132(-20.5%)	159(-4.2%)	139(-16.3%)
By county income/cap				
Top 100	543	439(-19.2%)	526(-3.1%)	458(-15.7%)
Bottom 2500	456	370(-18.9%)	439(-3.7%)	378(-17.1%)
By county urban%				
Top 100	769	628(-18.3%)	745(-3.1%)	642(-16.5%)
Bottom 2500	269	213(-20.8%)	259(-3.7%)	226(-16.0%)
Deposit (in B$)				
Tot. value	4695	4489(-4.4%)	3827(-18.5%)	8123(73.0%)
Tot. social surplus	5004	4780(-4.5%)	4054(-19.0%)	8847(76.8%)

6.2.2 Experiment 6: Tax on deposits

We evaluate how a counterfactual tax on deposits would affect the provision of credit and its geographic distribution. This experiment offers valuable insight into how exogenous changes in deposit rates affect the geographic distribution of deposits and loans. We implement this

[37] In Appendix E.2 we present results from a test of this counterfactual against pre-merger data that confirm the ability of this experiment to capture the impact of the policy.

experiment via an ad-valorem tax such that the marginal cost of deposits in every market increases by the same magnitude of τ, and social surplus in the deposit market declines by $\alpha^d\tau$. The value of τ is chosen such that social surplus in the median county – according to the empirical distribution of social surplus of deposits – declines by 20%.

Our findings suggest that the targeted reduction in deposits results in an average deposit volume decline of 18.5% for depository institutions, but a smaller decrease in lending activity, felt more strongly in smaller, poorer, and more rural counties. This implies that for traditional banks the elasticity of credit relative to deposits is roughly 0.5. These banks experience a reduction in their lending activity of 7.8% in response to a 18.5% decrease in their deposits. Note that shadow bank lending is barely affected, since they do not collect deposits, and shadow banks make up roughly half of total lending. As a result, overall, total lending activity only falls by 3.3%. This effect is felt more strongly in smaller, poorer, and more rural counties.

These changes result in only a small impact on the Imbalance Indexes. The median bank-level Imbalance Index increases slightly (by 0.6%), and the national Imbalance Index grows by 2.3%. To understand this result note that taxing deposits affects a bank's lending activity through two channels: local synergies and total deposits. Counties in which the bank collects and makes loans are affected through both channels, while counties where it only makes loans are influenced just by total deposits, such that former experience a larger decrease in lending activity. Supposing that the deposit decrease hits all of the bank's counties equally, then the fact that these two types of counties are differentially affected implies that, across the bank's network, deposits and loans will be less correlated, thereby increasing its Imbalance Index score. The same forces explain why the national Imbalance Index increases.

Takeaway: Imposition of a deposit tax has a moderate effect on lending activity. The effect is felt more strongly in socially disadvantaged markets.

6.2.3 Experiment 7: Increase the interbank rate 50 bps

Finally, we evaluate the impact of a hypothetical 50 basis point increase in the interbank rate to mimic a contractionary policy initiated by the Federal Reserve. This is a worthwhile exercise as it explores the potential geographic non-neutrality of a standard monetary policy tool and its effect on the distribution of bank funds across counties. Implementing this counterfactual is more complex, as it requires knowledge of the parameters α^d and α^ℓ. Unfortunately, as discussed, we lack data on interest rates at the bank-county level needed to estimate these parameters. To get some sense of the effect of a change in the interbank rate, for this experiment we calibrate these parameters using average loan and deposit demand-rate elasticities found in the recent literature. Full details of this calibration process are provided in the appendix.

The impact of changes in the Fed rate rate on loans, deposits, and welfare varies across regions. This variation presents itself in several ways. First, an increase in the Fed rate reduces the Imbalance Index, indicating that banks redistribute a smaller portion of their funds across counties. Notably, the Fed rate's negative impact on the Imbalance Index contrasts with the positive effects of increased market power (Exp. 3) and a deposit tax (Exp. 6). The Fed rate's asymmetric effect on loan and deposit markets suggests its influence on the distribution of bank funds differs significantly from other exogenous increases in marginal costs or price-cost margins. Second, the negative effects on loan volumes, along with associated welfare losses in loan markets, are more severe in smaller, poorer counties. Finally, as anticipated, this monetary policy is not neutral in its impact on banking competition. Shadow banks experience a more pronounced decline in activity, while certain depository institutions may actually benefit from tighter monetary conditions, as these improve the deposit side of their operations.

Takeaway: Increasing the interbank rate by 50 bps has a significant negative impact on the fraction of funds that banks redistribute across regions. The effect is felt more strongly in socially disadvantaged markets.

7 Conclusions

We use data from the SOD and HMDA for the period 1998-2010 to study the extent to which deposits and loans are segregated, and investigate factors that contribute to this imbalance. We make two main contributions. First, we adapt techniques developed in sociology and labor to measure the degree of segregation of deposits and loans. Our *Imbalance Indexes* provide information on the transfer of funds within branch networks of US banks, and across counties. Our results reveal that the majority of banks exhibit strong home bias and some regions have limited credit relative to their share of deposits. Second, we develop and estimate a structural model of bank oligopoly competition that 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. The estimated model reveals that a bank's total deposits have a significant effect on the bank's shares in loan markets. We also find evidence that is consistent with significant economies of scope between deposits and loans at the local level.

An advantage of our structural approach is that we can study counterfactual scenarios in which we adjust parameters or impose relevant policy-related restrictions. Our experiments show that multi-state branch networks contribute to the geographic flow of credit, but benefit especially larger/richer counties. Scope economies are important for generating a home bias and preventing funds from flowing away from where they are generated. Local market power, on the other hand, has a negative effect on the geographic flow of credit, with its adverse impact most pronounced in economically disadvantaged counties. Furthermore, our findings underscore the contribution of shadow banks in credit provision, particularly within large and urbanized markets. The capacity of traditional depository institutions to supplant the credit function provided by shadow banks proves to be constrained. In terms of policy experiments, we find that Riegle-Neal impacted lending activity and the flow of funds, albeit moderately due to the inherent limitations on cross-border lending activity. Additionally, introducing a deposit tax has a moderate negative effect on lending activity, with its effects disproportionately felt in socially disadvantaged markets. Lastly, increasing the interbank rate has a negative impact on the fraction of funds redistributed across regions.

Finally, it is important to point out that the US banking market continues to evolve and in the years after our sample period the number of banks and branches per county drifted back to the levels observed at the start of our sample period. This is possibly related to the use of online banking tools that make visiting a branch less important than it used to be for accessing services. In addition, as discussed, the importance of shadow banks continues to grow, as does the role of securitization. Nonetheless, branches remain important, with roughly 70,000 still operating in the US; and deposit-based funding continues to play a significant role in lending activity. The deposit share of domestic savings (non-financial sector financial assets) has actually grown from about 9% in 2000 to 13% in 2019, and informationally sensitive lending using deposit funding has stayed roughly constant over this same period at about 35% (see Buchak et al., 2024b).

A Data and Descriptive Evidence

A.1 Datasets and observation level

We rely on four main sources for the construction of our dataset. The first is the summary of deposits (SOD) collected by the Federal Deposit Insurance Corporation (FDIC), which provides county-level information on deposits outstanding (the "stock" measure) for each FDIC-insured bank.

The second is the Reports of Condition and Income (CALL) report collected by the FDIC. The CALL reports contain important bank-level characteristics, such as assets, total deposits (stock), total (mortgage) loans (stock), etc. It also provides loan turnover rate that we use to construct the "adjusted deposit flow" measures in deposits and loans described in detail in Section A.4.

Third, we make use of Home Mortgage Disclosure Act (HMDA) data collected by the Consumer Financial Protection Bureau, which provide county-level total mortgage loans issued by each reported institutions in a given year (a "flow" measure of loans).[38]

Lastly, we obtain credit union characteristics from the CALL reports (CALL-CU afterwards) collected by National Credit Union Administration (NCUA). The NCUA CALL reports are generally organized in a similar fashion as the FDIC CALL reports. However, one important feature worth mentioning is that they contain not only the mortgage loan outstanding for each credit union, but also the amount of mortgage loans newly issued in the corresponding calendar year. Similar information on loan flow is not available among the depository institutions in the FDIC CALL reports.

Table A1 summarizes the above information. The SOD and HMDA data are available annually during our study periods (1998-2010). The CALL reports are reported quarterly. In practice, we found that the June (the second quarter) CALL reports deliver the highest matching rate with the SOD and HMDA datasets each year. We therefore use the second quarter CALL reports for all the FDIC insured banks. For credit unions, we must use the December CALL data because only this quarter can give us the mortgage loans issued in the whole calendar year (the loan "flow" measure).

[38]There are geographic restrictions on loan reporting. According to the Community Reinvestment Act, large banks must report on all loans regardless of location. Regardless of size, lenders located in an MSA must report on loans originated in an MSA but can choose not to report loans outside MSAs. Only small lenders outside of MSAs do not have to report, but according to the US census, about 83% of the population lived in an MSA during our sample period. Therefore, HMDA captures most residential mortgage lending activity.

Table A1: Data sources

Dataset	Collector	Detailed Level	Variables
Summary of Deposits (SOD)	Federal Deposit Insurance Corporation (FDIC)	Institution-county level (Depository banks)	Amount of deposits (stock) Number of branches
Home Mortgage Disclosure Act Data (HDMA)	Consumer Financial Protection Bureau	Institution-county level (Depository banks & Shadow banks)	Amount of loans issued (flow) Securitization rate
Reports of Condition and Income (CALL report)	Federal Deposit Insurance Corporation (FDIC)	Institution level (Depository banks)	Amount of deposits (stock) Amount of loans (stock) Turn over rate of loans
Reports of Condition and Income for Credit Union (CU-CALL report)	National Credit Union Administration (NCUA)	Institution level (Credit Unions)	Amount of deposits (stock) Amount of loans (stock and flow) Turn over rate of loans

A.2 Construction of the working sample

Table A2 describes how all of the datasets just described are matched. The last three columns report the number of institutions, the share of deposits, and the share of loans of each case at the national level in year 2000.

Institutions in both SOD and HMDA:

The depository institutions in our working sample are those banks (case 1) and saving associations (case 3) that can be matched between SOD and HMDA. In year 2000, there are 3971 case 1 banks, accounting for 68.7% of the national deposits and 23.5% of the national loans. Meanwhile, there are 696 case 3 saving associations, accounting for 10.9% of the national deposits and 15.7% of the national loans.

Notice that case 1 banks can also be matched with the CALL report, while case 3 cannot. Matching with CALL reports is important because they contain the institutional level of loan turnover rate, which is used in constructing our "adjusted flow" measure of deposits. For case 3, we impute their turnover rate by the average turnover rate of the case 1 banks in the same county.

Institutions in HMDA but not in SOD:

Table A2: Classification of Banks & Matching Cases Between SOD, HMDA, and CALL

		Availability in Datasets				Year 2000		
Case	Inst. Category & Matching Case	HMDA	SOD	CALL	CU CALL	Nbr Banks	Share Deps	Share Loan
1	FDIC Core banks matched with HMDA	X	X	X		3971	68.7%	23.5%
2	FDIC Core banks unmatched with HMDA		X	X		5038	10.9%	2.7%
3	FDIC Savings Asso. matched with HMDA	X	X			696	10.9%	15.7%
4	FDIC Savings Asso. unmatched with HMDA		X			393	0.87%	1.8%
5	Foreign banks matched with HMDA	X		X		0	-	-
6	Foreign banks unmatched with HMDA			X		408	-	-
7	Shadow banks in HMDA	X				1047	0.0%	54.4%
8	Credit Unions in HMDA & CU CALL	X			X	1674	6.5%	1.9%
9	Credit Unions in HMDA but no CU CALL	X				11	-	-
10	Credit Unions in CU CALL but no HMDA				X	8932	2.1%	0.3%

If an institution appears in HMDA but not in SOD, and if they cannot be further matched with the CALL reports or the Credit Union Call Report, then they are labelled a shadow bank (case 7). These are mainly mortgage companies that finance their loans from sources other than deposits. In 2000, there are 1047 shadow banks, accounting for 54.4% of national loans.

Institutions that are not in SOD but are in the CALL reports are foreign banks (case 5). They are small in numbers, and also account for a tiny proportion of national deposits and loans. Therefore, they are excluded from our sample.

Credit unions (cases 8, 9) are large in number, but they are excluded from the sample for two reasons. First, their deposits and loans are only available at the institution level, not at the institution-county level. Second, they only account for a small proportion of national deposits and loans (8.6% and 2.2%, respectively, in 2000).

Institutions in SOD but not in HMDA:

Mortgage reporting for institutions smaller than a certain threshold is not compulsory. This is why most of the banks in SOD but not in HMDA (case 2 and 4 banks) tend to be small banks in assets. Although they are large in numbers (5431 institutions in year 2000), but they only account for 4.5% of total loans and 11.8% of total deposits.

Institutions not in SOD, not in HMDA:

Some foreign banks (case 6) and some credit unions (case 10) do not appear in either SOD or HMDA, but these groups are very small in terms of both deposits and loans.

Final sample: Our final working sample includes Case 1, 3, and 7 institutions. These are the institutions for which we know both their deposit distribution (Cases 1 and 3) and their loan distribution (Cases 1, 3, and 7) at the county level.

A.3 Multi-state banks

Table A3: Multi-state banks

	All	Depository Banks				Shadow banks	
year	#	#	ms_branch	ms_loan	br OR loan	#	ms_loan
1998	6123	4740	188	1861	1861	1383	936
1999	6014	4667	235	1979	1979	1347	934
2000	5892	4649	263	2026	2026	1243	912
2001	5774	4592	306	2135	2135	1182	878
2002	5856	4622	334	2294	2294	1234	923
2003	6080	4624	358	2451	2451	1456	1124
2004	6619	4906	376	2632	2632	1713	1250
2005	6615	4830	422	2662	2662	1785	1310
2006	6659	4793	441	2723	2723	1866	1371
2007	6404	4767	466	2721	2721	1637	1202
2008	6166	4789	486	2810	2810	1377	1008
2009	5913	4723	492	2657	2657	1190	890
2010	5715	4576	495	2473	2473	1139	860

A.4 Alternative Measures of the Imbalance Index – Adjusted Deposit Flows and Loan Stocks

As mentioned in Section 2.2, the loan information in HMDA refers to the mortgage loans issued in a year, and should therefore be considered a *flow* measure. On the deposit side, the SOD provides information on the deposits available on a day of the year (June 30th), and it is a *stock* variable. One problem with using stocks is that some past deposit inflows, which make up today's stock of deposits, have already been used to fund prior mortgages. To address this, deposit flows can be constructed from the SOD dataset as the net change in deposit stocks at the bank-county level by first-differencing by year. However, this represents a net change – newly attracted deposits minus withdrawn deposits.

There are two problems with using the net change. First, it can be negative, which makes it difficult to calculate the deposit share at the county level. Second, it is also the case that this net change in deposits would underestimate the funds available to banks to create new loans, since a fraction of existing loans are repaid (come due) every year.

We therefore construct what we call an *adjusted flow* measure, that includes net deposit changes along with those deposits that are freed up today as a fraction of previously funded loans are paid off. The Call Reports provide information for each bank on the fraction of loans coming due each year and we use this to construct our measure.

In our main analysis we continue to focus on stocks of deposits for the reasons mentioned in Section 2.2. However, we have performed all of our descriptive analysis using new loans and adjusted deposit flows. Our results are robust to this alternative measure of deposits – The Imbalance Index for adjusted deposit flows vs loan flows follows a similar pattern to stocks vs flows.

We also repeat the analysis with deposit stocks and replacing loan flows with loan stocks in the construction of the Imbalance Index. Again, our findings suggest that there is little change to the Imbalance Index from such a modification.

In what follows we provide a detailed explanation of the implementation of these two robustness checks, along with a discussion of the consequences for our analysis.

I. Adjusted deposit flows vs loan flows:

We start by introducing the following notation and variable definitions. For bank j in county m and year t, define:

- q^d_{jmt}: amount of deposits (stock) held by j in m as of t,

- L_{jmt}: proportion of j's deposits in m as of t used to fund mortgages throughout j's network,
- R_{jmt}: loan turnover rate for j in m in t (i.e. the proportion of loans that will mature within a year),
- DF_{jmt}: amount of deposits newly available for j in m in t (i.e. the adjusted flow of deposits).

DF_{jmt} is constructed as follows:

$$DF_{jmt} = q^d_{jmt} - q^d_{jmt-1} + R_{jmt-1} \times L_{jmt-1}.$$

This adjusted flow measure is made up of two components:

i. $q^d_{jmt} - q^d_{jmt-1}$: net change in the stock of deposits, reflecting the savings and withdrawal decisions of the local consumers.

ii. $R_{jmt-1} \times L_{jmt-1}$: loans maturing and turning over within a year – funds that should be available for the purpose of credit.

There are two simplifying assumptions we must make when constructing the adjusted deposit flow measure:

- We do not have the loan turnover rate at the bank-market level. To deal with this, we use the bank level turnover rate to proxy for the bank-market level measure (i.e., we use R_{jt-1} in place of R_{jmt-1}).
- We do not observe the amount of deposits issued as loans at the bank-county level. Instead, we infer this using observables in the dataset: the total deposit stock (Q^d_{jt}), the total loan stock (L_{jt}), and the bank-county level deposit stock (q^d_{jmt}). The first two come from the CALL reports, while the last comes from the SOD dataset. Specifically, we assume that at the bank-market level the share of deposit stock used to fund loans throughout j's network is the same as the bank-level share of deposit stock used to fund loans (i.e., $L_{jmt}/q^d_{jmt} = L_{jt}/Q^d_{jt}$). Therefore,

$$\begin{aligned} L_{jmt-1} &= L_{jt-1}/Q^d_{jt-1} \times q^d_{jmt-1} \\ &= \rho_{jt-1} \times q^d_{jmt-1}, \end{aligned}$$

where $\rho_{jt-1} = L_{jt-1}/Q^d_{jt-1}$ is the ratio between the loan (stock) and deposit (stock) at the bank level.

With the above adjustments, the adjusted deposit flow measure can be calculated as:

$$DF_{jmt} = q^d_{jmt} - q^d_{jmt-1} + R_{jt-1} \times \rho_{jt-1} \times q^d_{jmt-1}. \tag{22}$$

Data on deposits at the bank-county level come from the SOD dataset, while variables at the bank level (R and ρ) come from the CALL reports.

An additional complication arises in constructing our adjusted deposit flow measure because of mergers and the entry and exit of banks in our sample period. The construction of our adjusted flow measure requires bank-level and bank-county level information from two continuous years. Over time, there are banks that enter into or exit from the local market. Entry leads to missing values in year $t-1$, while exit leads to missing values in year t. The more complicated cases are entry or exit due to mergers, where a bank's ID should be replaced by another one.

Table A4: Entry and Exit Scenarios

				Year $t-1$	Year t	Flow calc w/o adjust.	Flow calc w/ merger adjust.	Nbr obs in 2000	% share
Scenario 0				A	A	Standard formula	Standard formula eq (22)	22.434	87
Scenario 1	Exit			A		Missing value in A	Missing value in A	322	1.2
Scenario 2	Exit through merger	2.1	B takes over A's branches	A	B	Missing value in B	Replace A's ID w/ B's in $t-1$, apply st form	1,541	6
		2.2	Within market M&A	A,B	B	Part of B's $t-1$ not counted	Replace A's ID w/ B's $t-1$, apply st formula	315	1.2
Scenario 3	B denovo enter				B	Missing value in B	B's deposits in t	1,148	4.5

Note: A and B are banks, with B being the surviving bank. Year $t-1$ and Year t columns indicate which of A and B have branches in the given county in the given year.

Table A4 lists all the possible scenarios when we compare observations at the bank-county level between year $t-1$ and t. The scenarios listed in red are those in need of modification using the merger information, otherwise we will have missing values in the deposit flow for these cases.

- ☐ In 87% of the cases, a bank exists in both years, so that the standard formula of equation (1) can be used. This is scenario 0.

- ☐ Scenario 1 describes simple exit from a county. This is infrequent and represents just 1.2% of observations.

- ☐ Most banks exit from a market because of a merger (scenario 2). Their ID in year $t-1$ needs to be replaced by the acquiring bank before applying the standard formula to the acquiring banks in year t. There are three subcases depending on whether surviving bank

B's acquisition of bank A involves taking over A's branches in counties in which it wasn't active (2.1), and in which it was (2.2).

- ☐ Scenario 3 describes denovo entry (about 4.5%). We assign the deposit stock of this bank in year t as its adjusted deposit flow in year t.

Summary statistics and results:

- ☐ **Loan turnover rate, R_{jt}:** The CALL reports provide information at the bank level on "loans and leases with a remaining maturity of one year or less." The ratio between this variable and the "total loans and leases" gives us the bank-level turnover rate for loans. The median (also the mean) turnover rate is around 30%.

- ☐ **Loan to Deposit ratio, ρ_{jt}:** The CALL reports provide information on both the loan stock and deposit stock variable. The ratio between these two can be calculated accordingly. In the year 2000, the median value of the loan/deposit ratio is 0.90. We truncate the ratio at the 5% and 95% percentile, which means that the final value of the ratio will range between 0.49 and 1.28.[39]

- ☐ **The adjustment factor, $R_{jt}\times\rho_{jt}$:** Figure A1 shows the distribution of $R_{jt}\times\rho_{jt}$ over time. The median is around 0.3, which is comforting because the resulting adjusted deposit flows are unlikely to be negative.

Equation (22) does not guarantee our adjusted deposit flow measure will be positive. Fortunately, observations with negative values only account for a small proportion of the sample. The share of negative observations at the bank-county level is around 3.5%, and the share of negative observations at the bank level is around 1.3%.

Table A6 reports the summary statistics of the adjusted deposit flow, and compares it with the deposit stock and the loan flow (Total new loans). In Panel A, all variables are aggregated at the bank level. In Panel B, all variables are aggregated at the county level. The sample we use is the SOD-HMDA matched sample of depository institutions (cases 1 and 3).

From Panel A we can see that bank-level adjusted deposit flows are about a third of deposits, on average. However, despite their differences in magnitude, the deposit flow (DF) and deposit stock (DS) measures are highly correlated with each other. The correlation coefficient between

[39] All values below 0.49 or above 1.28 will be replaced by 0.49 or 1.28, respectively.

Figure A1: Adjustment factor, $R_{jt} \times \rho_{jt}$

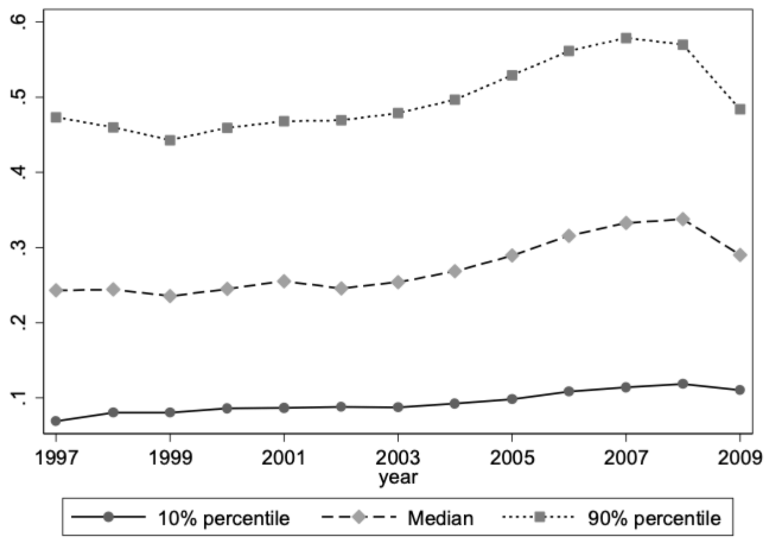


Table A5: Summary Statistics

Panel A: Bank Level Statistics: Depository Banks (61,475 bank-year obs.)

Variable	Mean	S.D.	Pctile 5	Median	Pctile 95
Total deposits (stock)	1018	11676	38	154	1799
Total deposits flow	342	4265	7	47	591
Total loan flow	188	3182	1	13	259

Panel B: County Level Statistics: Depository Banks (40,733 county-year obs.)

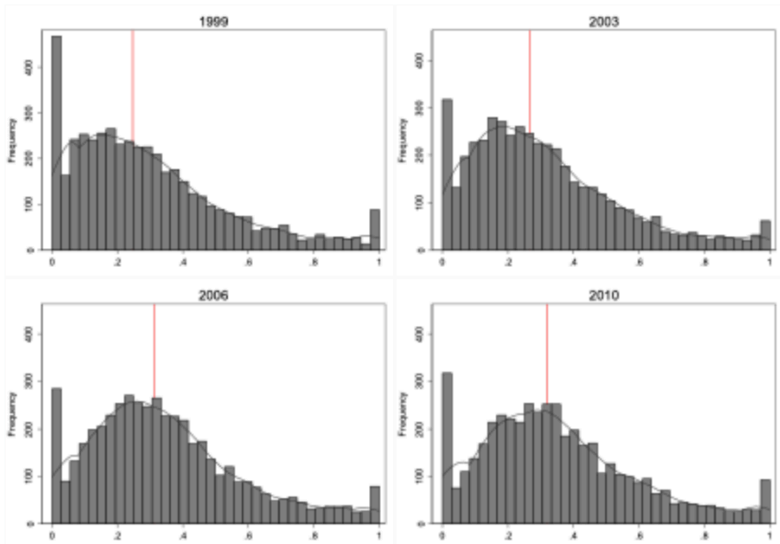
Variable	Mean	S.D.	Pctile 5	Median	Pctile 95
HHI of deposits (stock)	4321	2891	1175	3392	10000
HHI of deposits flow	4515	2902	1232	3595	10000
HHI of loan flow	1803	1346	626	1399	4357

the two is 0.86. Moreover, as can be seen from Panel B, the adjusted flow and stock measures imply very similar levels of concentration, on average.

Implications of the adjusted deposit flow for the Imbalance Index

Bank level: Using our sample of matched depository institutions, and dropping all observations with negative adjusted deposit flows, we can calculate the bank-level Imbalance Index between the distributions of adjusted deposit flows (DF) and loan flows (LF). Figure A2 of this response reports the distribution of the Imbalance Index changes over time. This figure can be directly compared to Figure 3 of the revised version of the paper. Doing so reveals that both in terms of magnitudes and evolution over time, findings are similar whether using adjusted deposit flows or deposit stocks.

Figure A2: Bank-level Imbalance Index – Adjusted Deposit flow and Loan flow Distributions



Comparison of the bank-level Imbalance Index constructed using adjusted deposit flows (DF) with the Imbalance Index using deposit stocks (DS) reveals that they are, surprisingly, very similar. For example, in year 2000, among the 4668 matched depository banks, the mean value of Imbalance Index using adjusted deposit flows is 0.304, while the mean value of Imbalance Index using deposit stocks is 0.301. Figure A13 plots II_DF against II_DS, and most of the dots are tightly located around the 45 degree line. The correlation coefficient between II_DF and II_DS is 0.944.

Table A6 reports the percentile of the Imbalance Index at the bank level using adjusted flows (panel A) and deposit stocks (panel B). Again, they all look very similar. Their trends are almost identical.

National-level Imbalance Index: Figure A4 duplicates Figure 4, which plots the National-level Imbalance Index. Panel (a) shows results for matched depository & shadow banks (cases 1, 3, and 7). The loan variables are from HMDA (i.e., loan flow measure). The dashed curve uses the deposit stock, while the solid curve uses the adjusted deposit flow measure. The Imbalance Index constructed using adjusted flows rather than stocks is higher than for stocks in every year, but the two curves follow very similar patterns. Panel (b) shows the same information, but this time just for depository institutions. Again, the Imbalance Index based on adjusted flows is everywhere greater than for stocks, but the two curves are quite similar. Based on this, we think it is reasonable to stick with our stock measure for our analysis.

Figure A3: Correlation: Imbalance Index using Adjusted Deposit Flow and Deposit Stock

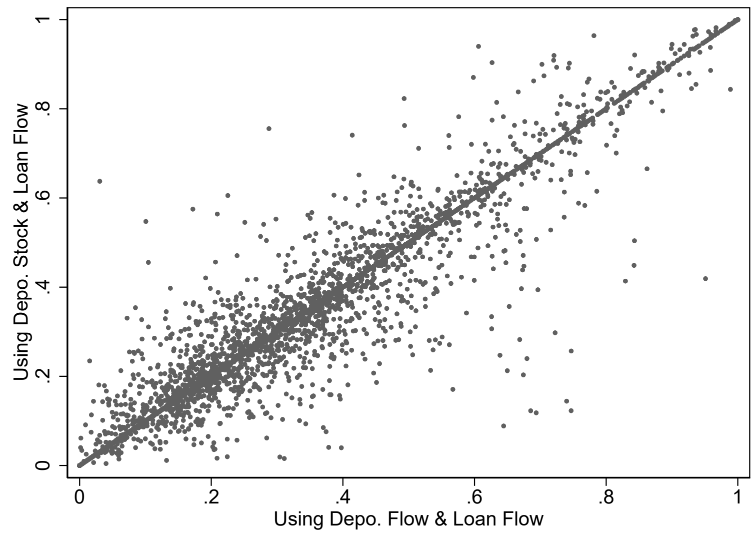


Table A6: Percentile of Imbalance Index at the Bank Level

Panel A: II using DF						Panel B: II using DS					
year	nbk	mean	10%	23%	90%	year	nbk	mean	10%	50%	90%
1998	4745	0.28	0.02	0.23	0.63	1998	4751	0.28	0.02	0.23	0.63
1999	4678	0.3	0.03	0.25	0.64	1999	4682	0.3	0.03	0.25	0.64
2000	4654	0.3	0.03	0.26	0.64	2000	4668	0.3	0.04	0.26	0.63
2001	4615	0.3	0.04	0.25	0.62	2001	4620	0.3	0.04	0.26	0.62
2002	4631	0.3	0.04	0.26	0.62	2002	4635	0.31	0.05	0.27	0.63
2003	4635	0.31	0.06	0.27	0.64	2003	4637	0.32	0.06	0.27	0.63
2004	4904	0.33	0.06	0.29	0.65	2004	4912	0.33	0.08	0.29	0.64
2005	4836	0.34	0.07	0.31	0.67	2005	4839	0.34	0.08	0.31	0.66
2006	4792	0.35	0.08	0.31	0.68	2006	4804	0.35	0.09	0.31	0.67
2007	4775	0.36	0.08	0.32	0.69	2007	4782	0.36	0.09	0.32	0.69
2008	4794	0.36	0.09	0.33	0.68	2008	4801	0.36	0.1	0.33	0.67
2009	4745	0.35	0.08	0.32	0.68	2009	4747	0.35	0.09	0.31	0.68
2010	4591	0.35	0.07	0.32	0.69	2010	4597	0.36	0.08	0.32	0.69

II. Deposit stocks vs loan stocks:

We also construct the National-level Imbalance Index using loan stocks, so that we can compare results to those derived using our main measure of loan flow (constructed from HMDA data). Mortgage-loan stocks are constructed using the CALL reports, under the assumption that stocks have the same bank-county distribution as flows from HMDA. We also construct a version based on total loans (i.e., not just mortgages), again using the CALL reports.

Note that this comparison focuses only on depository institutions, since we do not have CALL

Figure A4: Imbalance Index: Deposit stocks vs Deposit Flows

(a) Matched depository & shadow banks (b) Matched depository banks

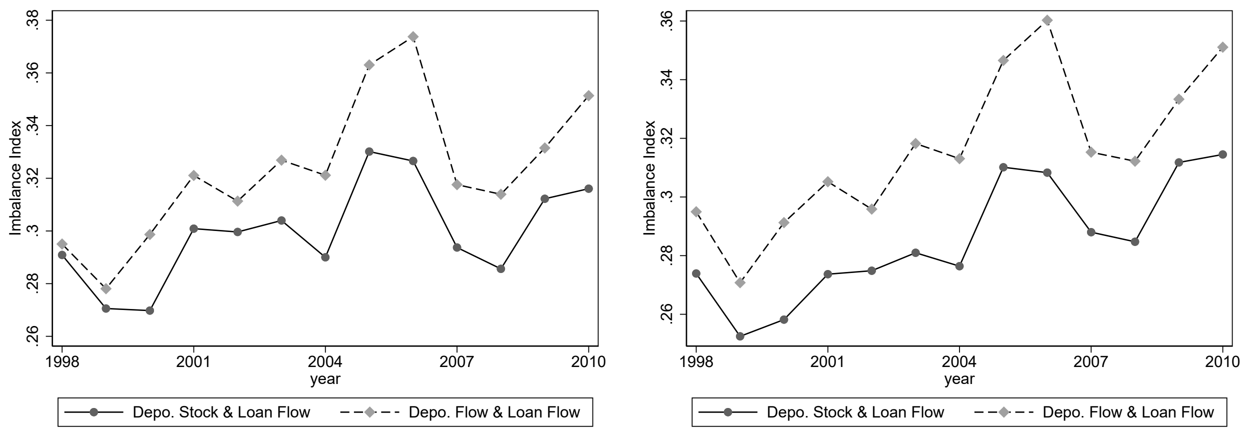
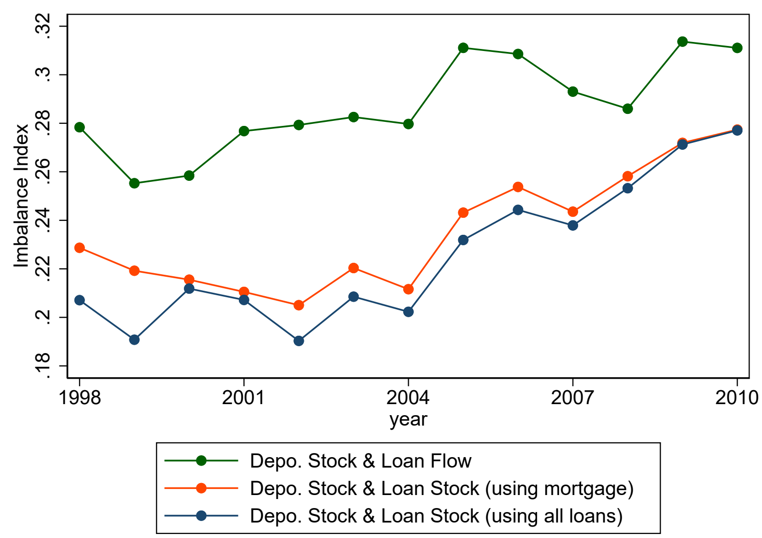


Figure A5: Imbalance Index: Loan Flow vs Loan Stock



information for shadow banks.

Our findings suggest that the Imbalance Index that uses the loan flow measure is slightly higher than the Imbalance Index constructed with loan stocks. But the overall trends of the two curves are very similar.

A.5 Bank heterogeneity in Imbalance Index

Table A7: Evolution of the Imbalance Index of the top 10 banks (by assets)

Rank	Bank Name	II
	1999	
1	CITIBANK, N. A.	0.37
2	CHASE MANHATTAN BANK	0.47
3	BANK OF AMERICA NA	0.50
4	FIRST UNION NATIONAL BANK	0.36
5	WASHINGTON MUTUAL BANK, FA	0.37
6	WELLS FARGO BANK, N. A.	0.33
7	FLEET NATIONAL BANK	0.45
8	FIRST NATIONAL BANK OF CHICAGO	0.17
9	BANKERS TRUST COMPANY	0.97
10	KEYBANK NATIONAL ASSN	0.37
Average		0.44
	2004	
1	BANK OF AMERICA NA	0.35
2	JPMORGAN CHASE BANK	0.41
3	CITIBANK NATIONAL ASSN	0.38
4	WACHOVIA BANK NATIONAL ASSN	0.41
5	WELLS FARGO BANK NA	0.51
6	WASHINGTON MUTUAL BANK FA	0.39
7	BANK ONE NATIONAL ASSN	0.70
8	FLEET NATIONAL BANK	0.37
9	U S BANK NATIONAL ASSN	0.45
10	SUNTRUST BANK	0.43
Average		0.44
	2009	
1	JPMORGAN CHASE BANK NA	0.52
2	BANK OF AMERICA NA	0.43
3	CITIBANK NATIONAL ASSN	0.75
4	WACHOVIA BANK NATIONAL ASSN	0.54
5	WELLS FARGO BANK NA	0.51
6	U S BANK NATIONAL ASSN	0.49
7	SUNTRUST BANK	0.61
8	HSBC BANK USA NATIONAL ASSN	0.66
9	BRANCH BANKING&TRUST CO	0.33
10	PNC BANK NATIONAL ASSN	0.77
Average		0.56

NOTE: Ranking based on banks' assets in each year.

Figure A6: Breakdown of imbalance index by bank size – size measured by counties

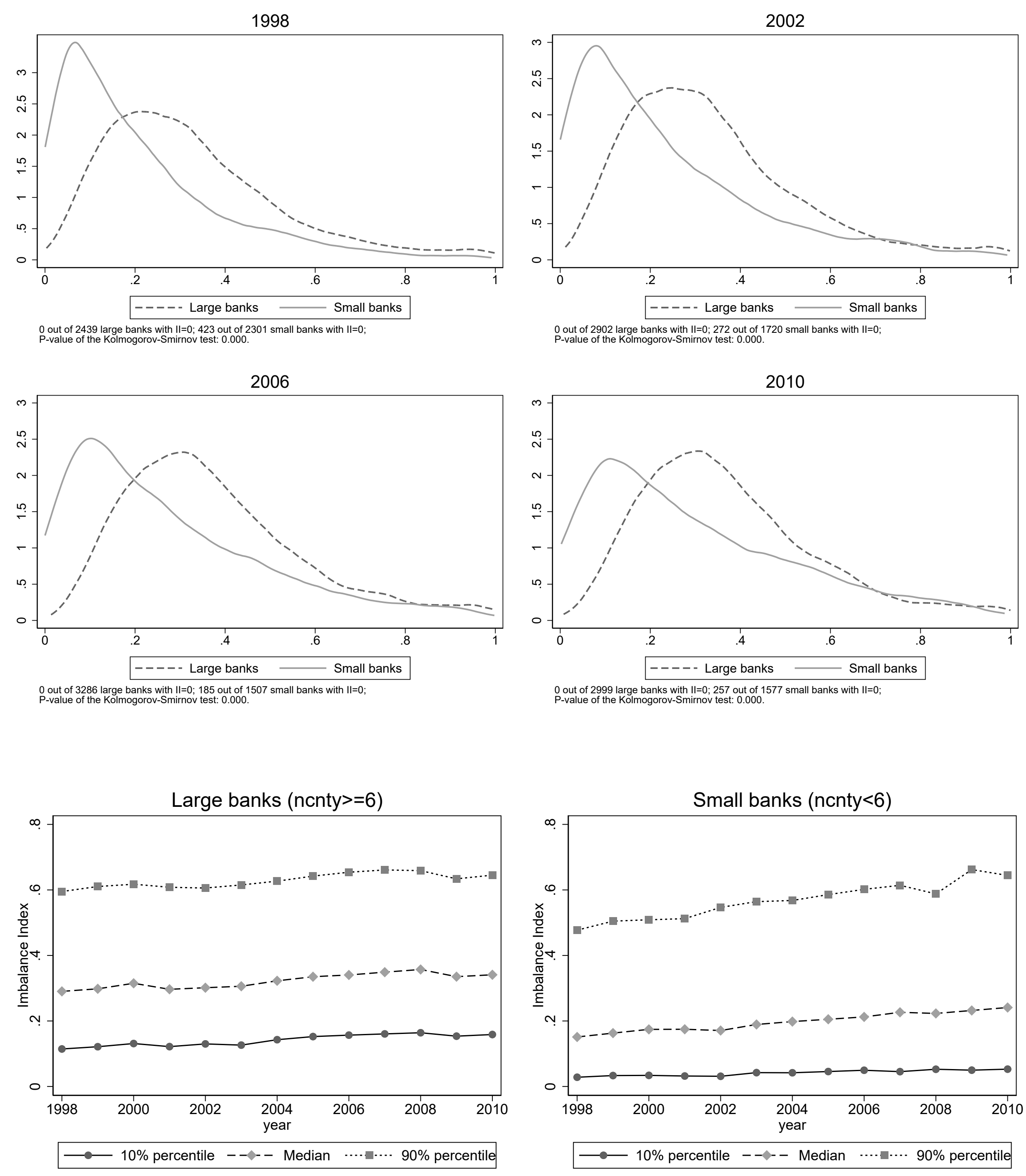


Figure A7: Breakdown of imbalance index by bank size – size measured by assets

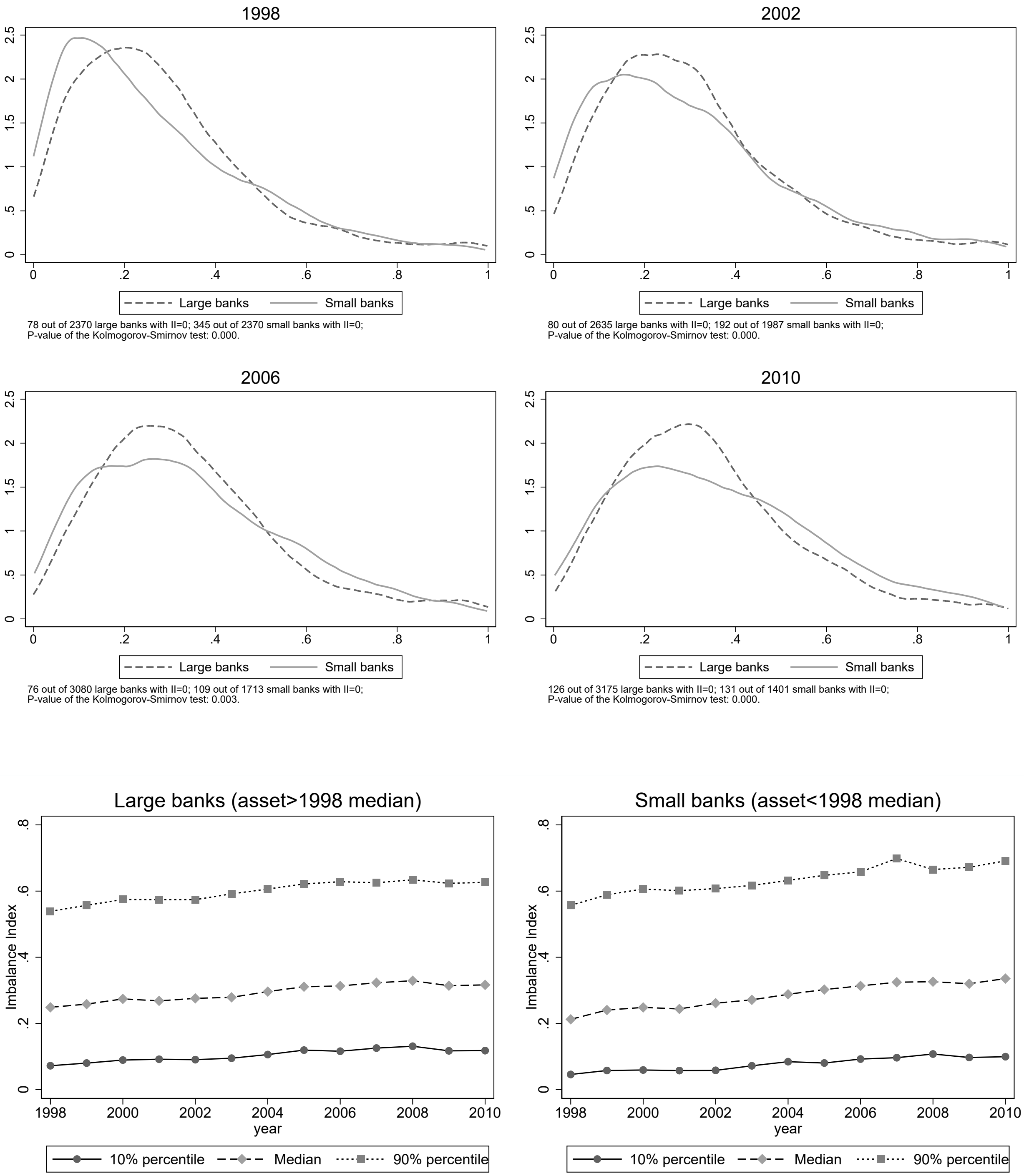


Figure A8: Breakdown of imbalance index by bank size – size measured by deposits

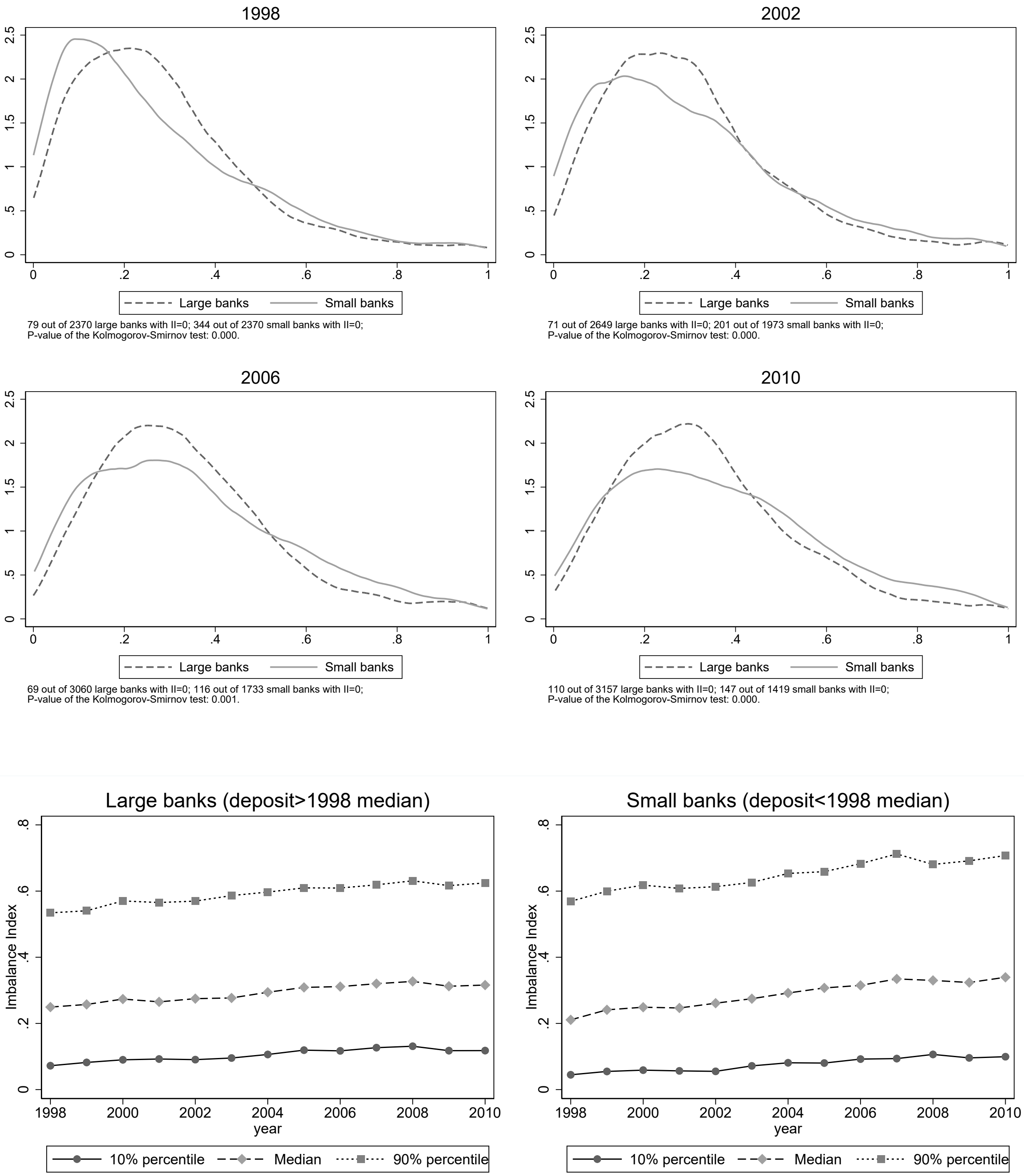
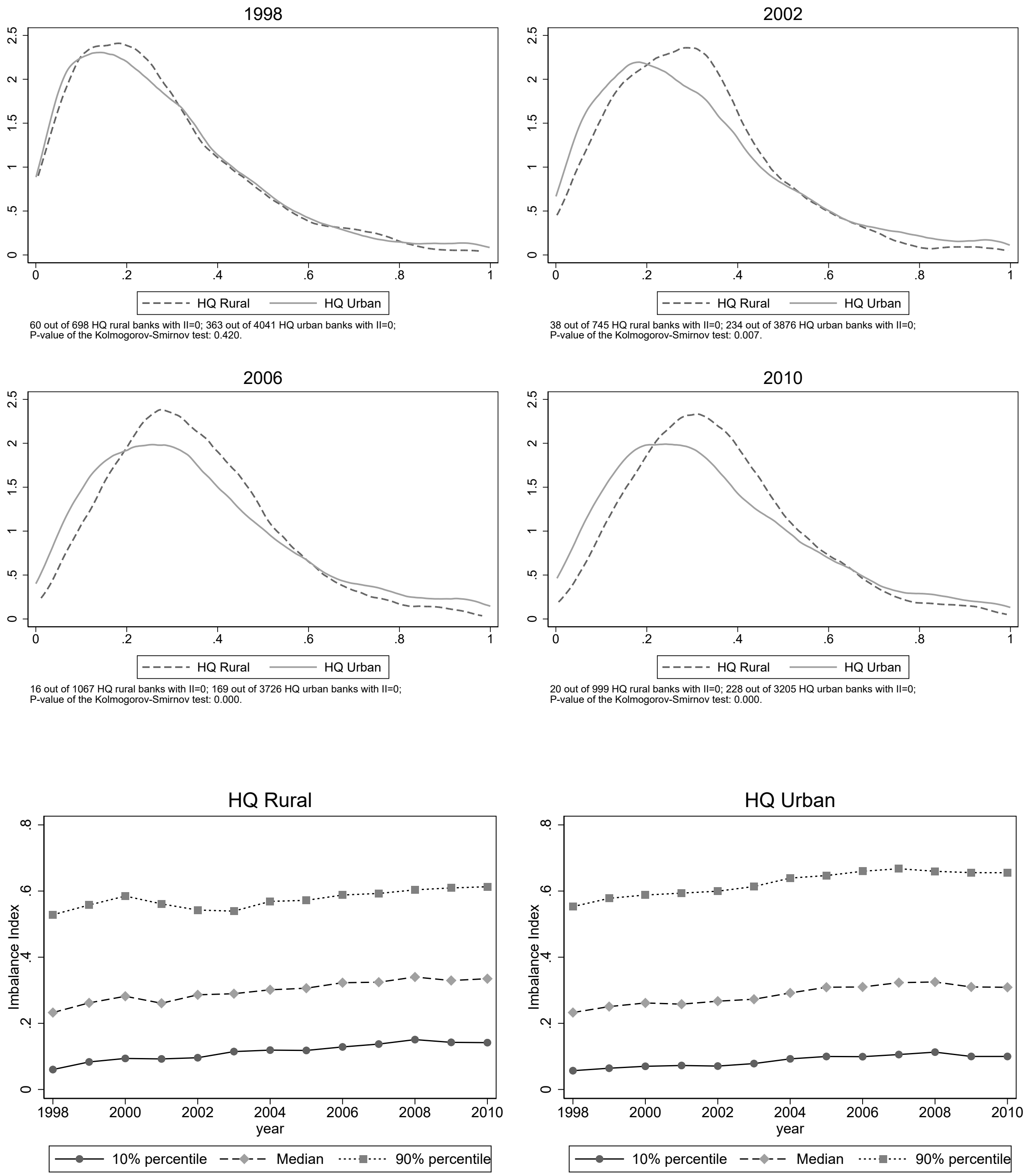


Figure A9: Breakdown of imbalance index by location – urban vs rural



A "rural county" is defined as rural population share is larger than 50%, based on population census in 2010. Reference: https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/2010-urban-rural.html

Figure A10: Breakdown of imbalance index by location – North, South, East, West

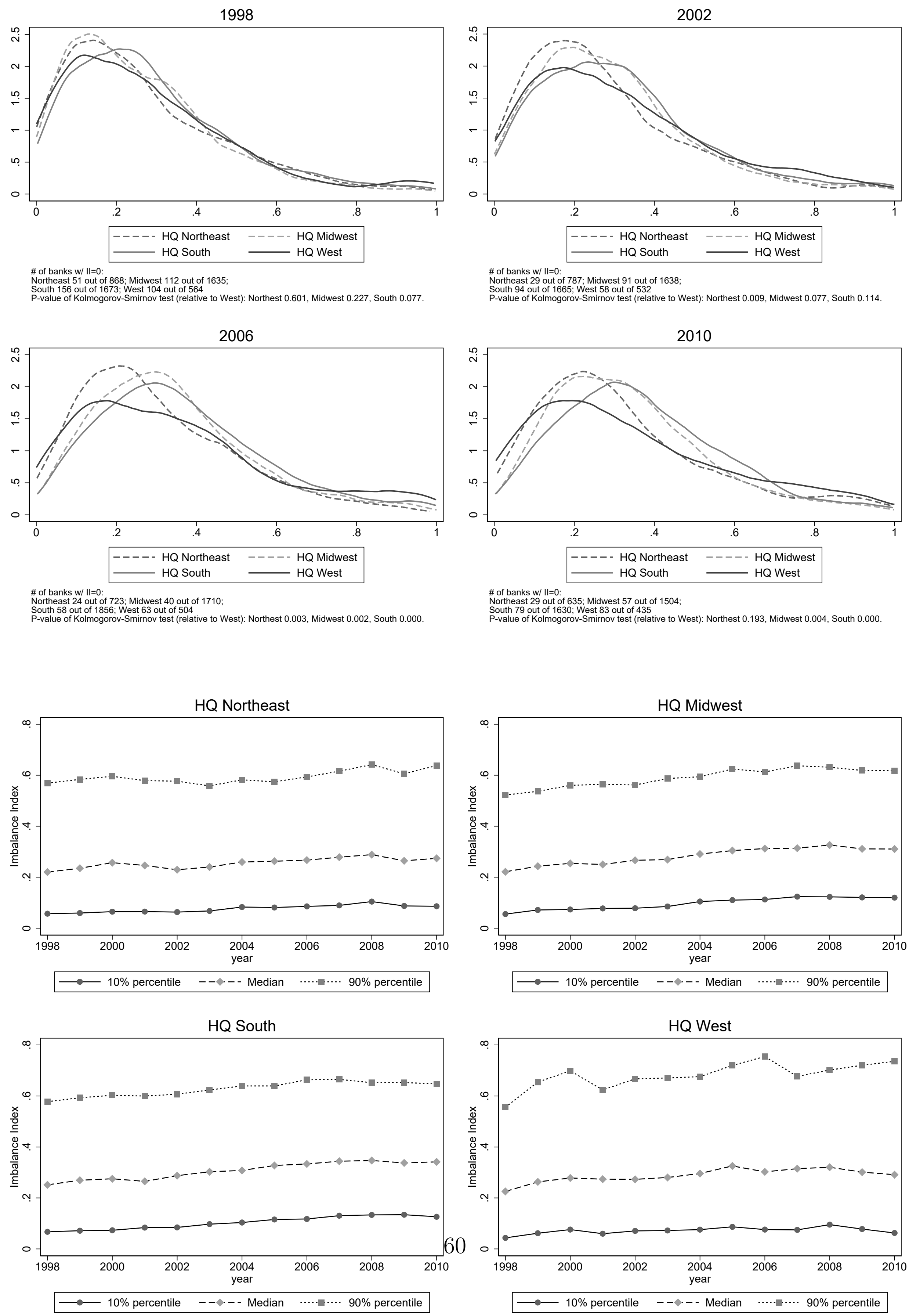


Figure A11: Breakdown of imbalance index by HQ-county income level (cutoff = 1998 median income)

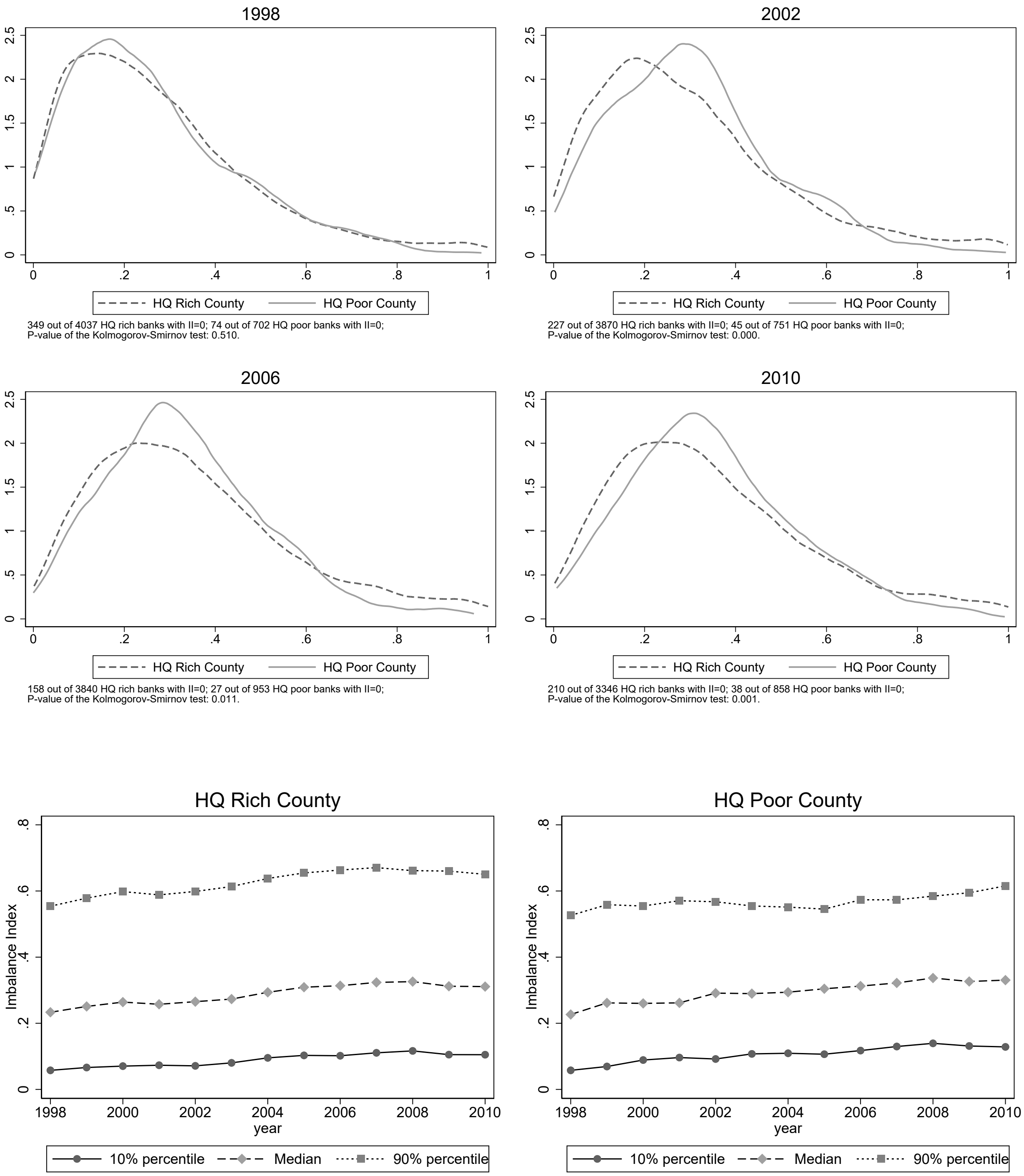
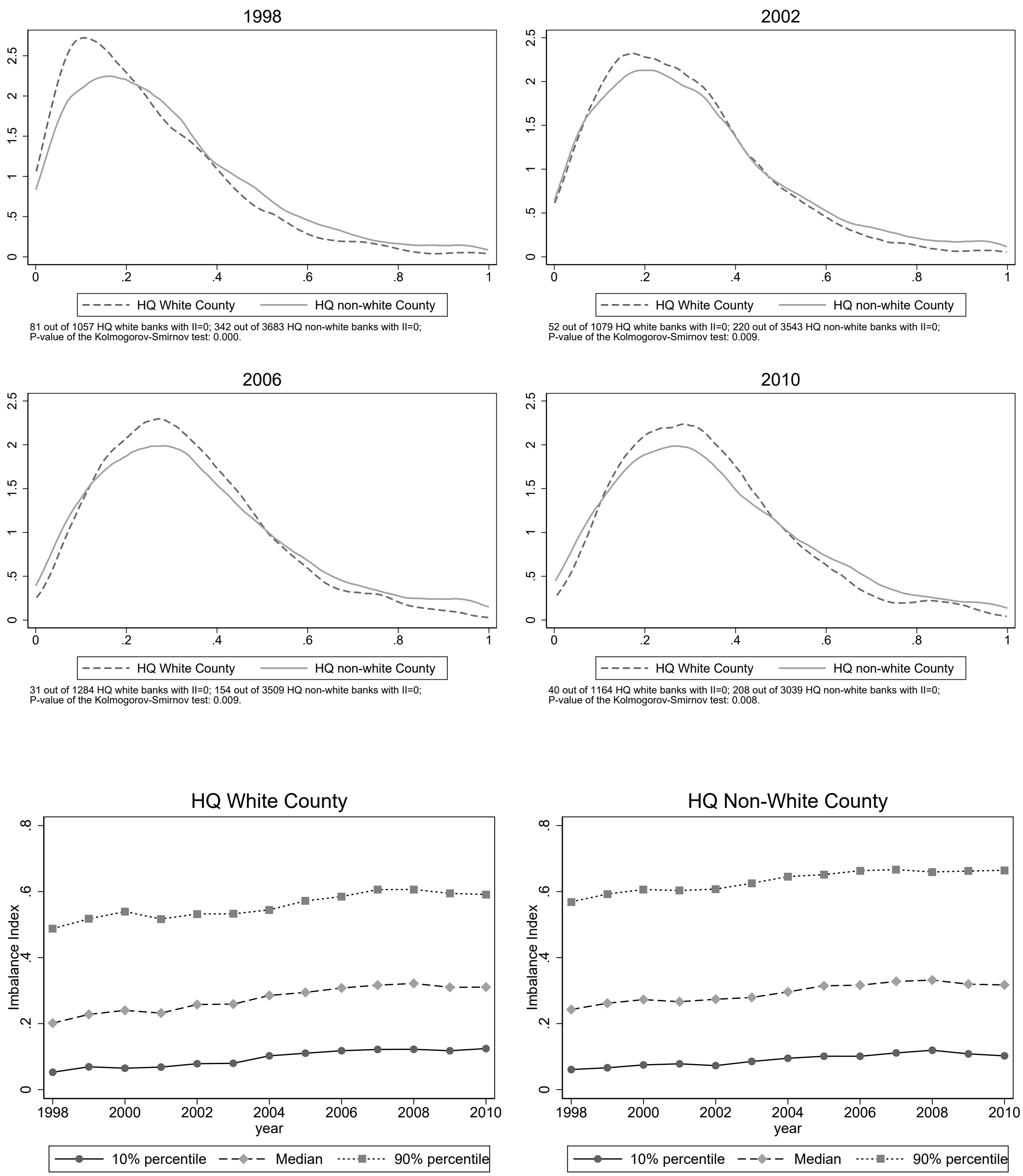


Figure A12: Breakdown of imbalance index by HQ-county white or non-white (cutoff = 1998 median white%)



B Markup predictions and stylized facts about margins

The predictions of our model are consistent with the stylized facts about margins and spreads presented in recent work by Drechsler et al. (DSS) in their 2017 QJE paper.

The main finding in DSS is that when the Fed funds rate rises, deposit spreads rise and the amount of deposit declines, and that these effects are stronger in more concentrated markets. Using our structural model we can measure market power in the local deposit market (i.e., the deposit spread (margin) at the bank-county-year level) by $1/(1-s_{jmt})$. Then, using this measure we can investigate how it changes when the Fed funds rate rises, and how this varies depending on market concentration.

Our findings are presented below in Table A8 and confirm these results. Market power in local deposit markets is positively correlated with the Fed funds rate (as measured by T-bills), and this correlations is stronger if the HHI of deposits in the local market is higher.

Table A8: Market power in local deposit markets

DV: $1/(1-s_j)$	(1)	(2)	(3)	(4)
T-bill (log)	0.002***	0.003***	0.016***	
	(0.001)	(0.000)	(0.002)	
HHI deposit (log)		0.318***	0.313***	0.197***
		(0.001)	(0.001)	(0.004)
T-bill x HHI (log)			0.008***	0.014***
			(0.001)	(0.001)
Constant	1.195***	1.714***	1.707***	1.538***
	(0.001)	(0.002)	(0.002)	(0.006)
County FE				Y
Bank-year FE				Y
Observations	225,513	225,513	225,513	197,620
R-squared	0.000	0.249	0.249	0.522

Note: Robust standard errors of serial correlation and heteroscedasticity are reported in parentheses. * means p-value < 0.05; ** means p-value < 0.01; *** means p-value < 0.001.

C Additional estimation results

Table A9: Alternative specifications of the loan equation

Variables	OLS Fixed Effects				GMM DiD & DiDiD			
Number of branches								
Second branch	0.117***	(0.012)	0.123***	(0.013)	0.1036***	(0.0231)	0.1586***	(0.0231)
Third branch	0.085***	(0.011)	0.078***	(0.013)	0.0675***	(0.0158)	0.0946***	(0.0161)
Fourth branch	0.080***	(0.012)	0.062***	(0.014)	0.0666***	(0.0144)	0.0793***	(0.0152)
Fifth branch	0.093***	(0.013)	0.092***	(0.015)	0.0785***	(0.0168)	0.0999***	(0.0174)
> Fifth	0.008***	(0.002)	0.009***	(0.002)	0.0043**	(0.0020)	0.0079***	(0.0023)
Securitization								
% of loan resold	0.158***	(0.011)	0.352***	(0.013)	0.2207***	(0.0244)	0.3480***	(0.0281)
Econ of scope, Qd, and Ql								
log own local deposit	0.210***	(0.008)	0.098***	(0.007)	0.3599***	(0.0319)	0.2094***	(0.0320)
log own total deposit	-0.432***	(0.011)			-0.6726***	(0.0209)		
log own total loans	0.871***	(0.006)			0.7639***	(0.0159)		
log(total depo + total loan)			0.640***	(0.005)			0.2799***	(0.0092)
Market characteristics								
log county income	-0.025	(0.052)	0.036	(0.058)				
log county population	-0.778***	(0.068)	-0.936***	(0.077)				
share pop. age< 19	-4.381***	(0.700)	-3.519***	(0.788)				
share pop. age >50	-1.767***	(0.477)	-1.377***	(0.531)				
log house price index	0.350***	(0.031)	0.184***	(0.034)				
log nbr bankuptcy	-0.021***	(0.008)	-0.022***	(0.008)				
log nbr loan applications	0.448***	(0.013)	0.499***	(0.014)				
Selection - Control function								
Propensity score	0.569	(0.429)	4.402***	(0.491)	-1.0385***	(0.3011)	0.3256	(0.3591)
Propensity score square	0.069	(0.506)	-5.287***	(0.582)	1.4218***	(0.2901)	0.5093	(0.3689)
Propensity score cubic	-0.223	(0.196)	1.626***	(0.227)	-0.4558***	(0.0903)	-0.3255***	(0.1209)
Fixed effects								
Bank x County	YES		YES		YES		YES	
Time	YES		YES		NO		NO	
Country x Time	NO		NO		YES		YES	
Bank x Time	NO		NO		YES		YES	
Number of observations	194,267		194,267		196,090		196,090	
R-square	0.926		0.915					

Note: Robust standard errors of serial correlation and heteroscedasticity are reported in parentheses. * means p-value < 0.05; ** means p-value < 0.01; *** means p-value < 0.001.

Tables A10 and A11 present results demonstrating the robustness of our findings to different combinations of time series and spatial instruments. In addition, we have also considered specification in which we incorporate BLP-type instruments and lagged-by-3-period values into the set of instruments used in the main specifications. The main findings remain consistent.

Table A10: Alternative set of instruments for EOS: Deposit equation

Variables	Spec. 1		Spec. 2		Spec. 3	
Number of branches						
Second branch	0.6803***	(0.0089)	0.6792***	(0.0087)	0.6673***	(0.0090)
Third branch	0.3801***	(0.0076)	0.3785***	(0.0074)	0.3722***	(0.0077)
Fourth branch	0.2904***	(0.0079)	0.2889***	(0.0078)	0.2825***	(0.0081)
Fifth branch	0.3645***	(0.0100)	0.3649***	(0.0097)	0.3525***	(0.0100)
> Fifth	0.0335***	(0.0024)	0.0336***	(0.0023)	0.0310***	(0.0022)
Securitization						
% of loan resold	-0.0354***	(0.0116)	-0.0448***	(0.0114)	-0.0573***	(0.0114)
Econ of scope and Qd						
log own local loans	0.0692***	(0.0102)	0.0762***	(0.0099)	0.1146***	(0.0101)
log own total deposit	0.0452***	(0.0049)	0.0497***	(0.0048)	0.0378***	(0.0050)
Number of observations	241,911		241,911		241,911	

Note: All specifications include Bank x County, County x Time, and Bank x Time fixed effects. The instrumental variables used for the EOS in each specification are as follows. Spec.1, value of loans in the neighbouring counties, and value of loans in the neighbours of the neighbouring counties; Spec. 2, value of loans in the neighbouring counties, and value of loans in the neighbouring counties lagged by 2 periods; Spec. 3, value of loans in the neighbouring counties, and value of loans lagged by 2 periods.

Table A11: Alternative set of instruments for EOS: Loan equation

Variables	Spec. 1		Spec. 2		Spec. 3	
Number of branches						
Second branch	0.0957***	(0.0298)	0.0945***	(0.0296)	0.1433**	(0.0618)
Third branch	0.0619***	(0.0191)	0.0611***	(0.0190)	0.0879**	(0.0355)
Fourth branch	0.0692***	(0.0167)	0.0694***	(0.0167)	0.0866***	(0.0285)
Fifth branch	0.0899***	(0.0201)	0.0889***	(0.0200)	0.1133***	(0.0348)
$>$ Fifth	0.0092***	(0.0022)	0.0094***	(0.0022)	0.0105***	(0.0036)
Securitization						
% of loan resold	0.6623***	(0.0304)	0.6629***	(0.0304)	0.6634***	(0.0307)
Econ of scope and Qd						
log own local deposit	0.3571***	(0.0437)	0.3614***	(0.0432)	0.2820***	(0.0907)
log own total deposit	0.1672***	(0.0157)	0.1582***	(0.0119)	0.1737***	(0.0148)
Selection - Control function						
Propensity score	-0.6647**	(0.3373)	-0.6403*	(0.3441)	-0.6357*	(0.3553)
Propensity score square	0.9122***	(0.3288)	0.8727***	(0.3330)	1.0109***	(0.3568)
Propensity score cubic	-0.3001***	(0.1004)	-0.2863***	(0.1003)	-0.3630***	(0.1255)
Number of observations	196,090		196,090		196,090	

Note: All specifications include Bank x County, County x Time, and Bank x Time fixed effects. The instrumental variables used for the EOS in each specification are as follows. Spec.1, value of deposits lagged by 2 periods; Spec. 2, value of deposits lagged by 2 periods, and value of deposits in the neighours of the neighbouring counties; Spec. 3, number of branches lagged by 2 periods, and value of deposits in the neighbouring counties lagged by 2 periods.

D Description of an equilibrium and algorithm

Let $\mathbf{s}$ be the $2MJ \times 1$ vector of local market shares for deposits and loans for every bank and county. An equilibrium is a vector $\mathbf{s}$ that satisfies the following system of equations:

$$\begin{cases} (A) & s^d_{jm} = 0, \ \forall\ (j,m), m \notin \mathcal{M}^d_j \\ (B) & s^\ell_{jm} = 0, \ \forall\ (j,m), m \notin \mathcal{M}^\ell_j \\ (C) & s^d_{jm} = \left(1 - \sum_{i=1}^{J} s^d_{im}\right) \exp\left\{ e^d_{jm} - \dfrac{1}{1 - s^d_{jm}} + \theta^d_\ell \ \ln(1 + H^\ell_m s^\ell_{jm}) + \theta^d_Q \ \ln Q^d_j \right\}, \ \forall\ (j,m), m \in \mathcal{M}^d_j \\ (D) & s^\ell_{jm} = \left(1 - \sum_{i=1}^{J} s^\ell_{im}\right) \exp\left\{ e^\ell_{jm} - \dfrac{1}{1 - s^\ell_{jm}} + \theta^\ell_d \ \ln(1 + H^d_m s^d_{jm}) + \theta^\ell_Q \ \ln Q^d_j \right\}, \ \forall\ (j,m), m \in \mathcal{M}^\ell_j \\ (E) & Q^d_j = \sum_{j=1}^{J} s^d_{jm}\ H_m, \ \forall\ j \end{cases}$$

where e^d_{jm} and e^ℓ_{jm} are exogenous terms that we define in the paper. For the description of the algorithms, it is convenient to represent this system of equations in the following compact form:

$$\begin{cases} (A') & \mathbf{s}_m = f_m\left(\mathbf{s}_m, \mathbf{Q}^d\right), \ \forall\ m = 1, 2, ...M \\ (B') & \mathbf{Q}^d = F\left(\mathbf{s}\right) \end{cases}$$

The system of equations in (A') represents equations (A) to (D) for market m, and f_m is a vector-valued function with $2J$ elements. The system of equations in (B') is simply the system in (E) in vector form, and F is a vector-valued function with J elements.

The system of equations (A') possesses two properties that significantly ease the computation of an equilibrium. First, for a fixed value of the vector $\mathbf{Q}^d$, solving for the market shares in (A') is separable across counties. This means that the system of $2JM$ equations and unknowns can be divided into M separate systems, each with a dimension of $2J$. Second, for fixed $\mathbf{Q}^d$ and deposit shares s^d_{jm}, there exists a unique vector of loan shares that solves the system (A'). Likewise, for fixed $\mathbf{Q}^d$ and loan shares s^ℓ_{jm}, there is a unique vector of deposit shares that solves the system (A').

Taking these properties into account, our algorithm for computing an equilibrium follows these steps:

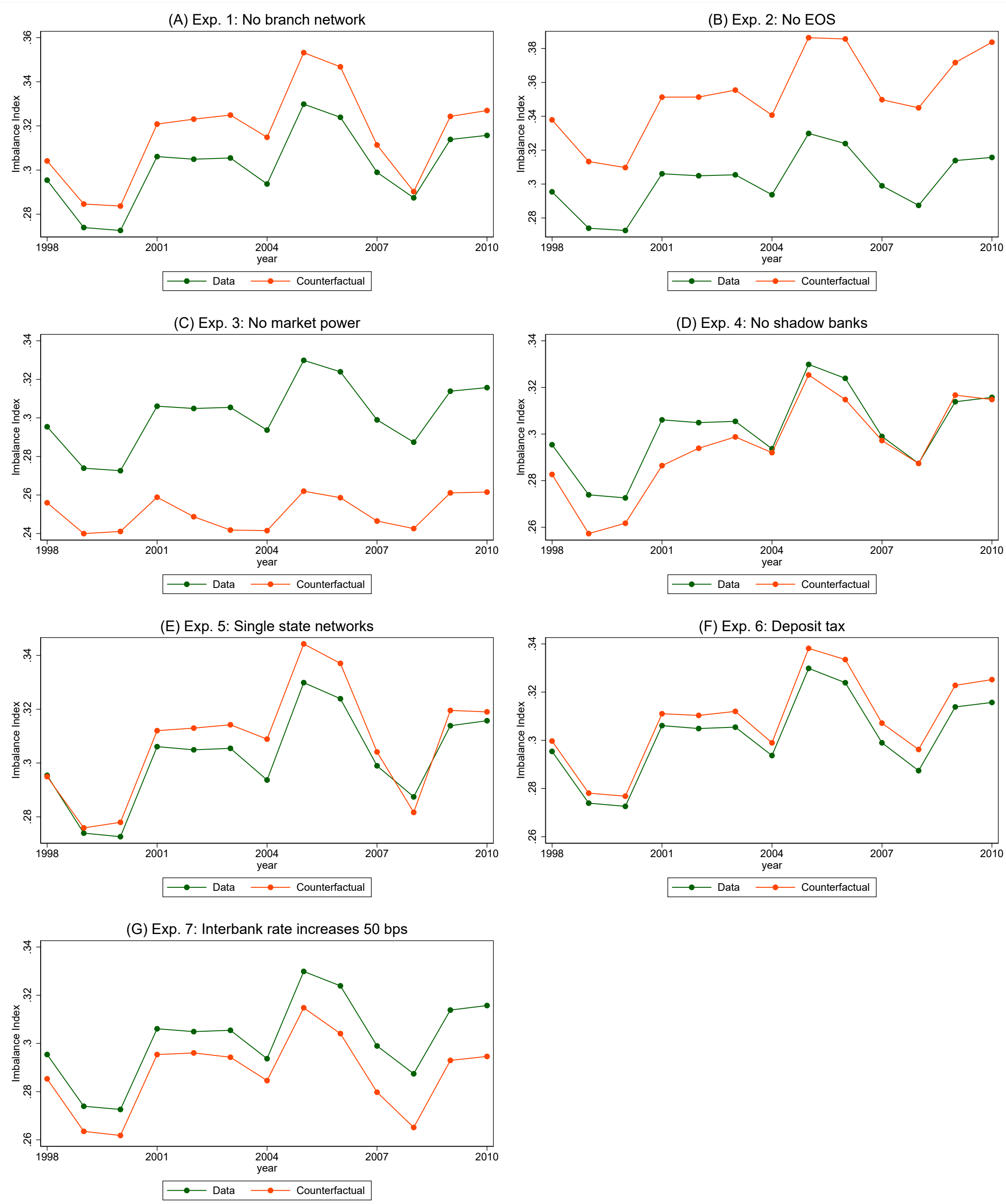
1. Initialization. We start by initializing the vector $\mathbf{Q}^d$. Typically, in most counterfactual experiments, this initial value is set to the banks' total deposits observed in the data. However, if the counterfactuals involve changes to branch networks, we adjust this initial value accordingly.
2. Iterative Process. During each iteration, we execute the following four steps sequentially:
   - Step 1. Compute the unique equilibrium for the vector of loan shares s^ℓ_{jm} in each county, given Q^d_j and local deposit shares s^d_{jm} for every j. This unique local equilibrium is computed using a Bisection algorithm.
   - Step 2. Compute the unique equilibrium of deposit shares s^d_{jm}, given Q^d_j and the local loan shares s^ℓ_{jm} computed in Step 1. We again use a Bisection algorithm for this computation.
   - Step 3. Using the market shares from Step 2, aggregate over counties to obtain Q^d_j for each j, thereby updating the vector $\mathbf{Q}^d$.
   - Step 4. Check for convergence by calculating the distance between the values of the vector $\mathbf{Q}^d$ at the beginning and the end of the iteration.

This process can be repeated until the algorithm converges.

E Additional Counterfactual results

E.1 Counterfactual National Imbalance Index

Figure A13: Evolution of the National Imbalance Index – Counterfactual scenarios



E.2 Validation of the Riegle-Neal Counterfactual

Here we test our Riegle-Neal counterfactual against the data. To do so we adopt the following strategy. We look for mergers involving two single-state banks.[40] Specifically, we are interested in cases where in period t bank A operates only in state 1 and acquires bank B that operates only in state 2, such that in period $t+1$, the survivor A' operates in states 1 and 2. There are 71 such cases throughout our sample period.

We then run our Counterfactual in which we effectively undo Riegle-Neal to see how outcomes change as we split A' into separate single-state banks A and B in $t+1$.

The first rows of Panels A and B of Table A12 present the Imbalance Index for the counterfactual unmerged survivor and non-survivor banks respectively.

Next, following the referee's suggestion we compare these counterfactual separate entities from $t+1$ to the data for A and B in t.[41] The Imbalance Index scores for A and B in period t are presented in the second rows of Panels A and B of Table A12, respectively.

Finally, the third rows of Panels A and B present statistics on the differences between the counterfactual Imbalance Index in $t+1$ and the Imbalance Index constructed using data from t, for survivor and non-survivor banks respectively.

Importantly, we would not expect the counterfactual outcomes for A and B from $t+1$ to be identical to data outcomes for A and B from t, since naturally there could have been some other changes in the market. However, should outcomes be similar, then this would validate our approach.[42]

Results:

1. For both survivor and non-survivor banks, the mean Imbalance Index score of the year $t+1$ counterfactual is very close to the mean Imbalance Index score of in the data in year t, as are their 25, 50, and 75 percentiles.

2. The difference between these two measures (the third row of each panel) is also very small at the mean, and the corresponding percentiles.

Overall, we believe that this test validates our Riegle-Neal counterfactual and allows us to learn about the importance of the ability to expand across states for the flow of credit.

[40] That is, banks operating branches in only one state although they may make loans in multiple states.

[41] The referee had suggested using data from before 1998, but in fact any merger that takes place across state lines is the result of Riegle-Neal and so we do not need to go back to pre-1998. This is useful, since not many banks had actually taken advantage of the expansion opportunities afforded them by Riegle-Neal by this time (see Aguirregabiria et al. 2016).

[42] The reason we look at expansion through mergers, instead of expansion through denovo entry is that, for the latter cases, there is no bank entity in state 2 in year t. By contrast, for expansion through mergers, for survivors, we can compare their CF II in t+1 in state 1 with their real II in t (Panel A of Table A12); and for non-survivor, we can compare the survivor's CF II in t+1 in state 2 with their real II in t (Panel B of Table A12).

Table A12: Test of Riegle-Neal Counterfactual: Mergers between Two Single-state Banks

	# obv	Mean	Std	p25	p50	p75
Panel A: Survivor bank						
CF II in t+1	71	0.304	0.170	0.180	0.292	0.403
Real II in t	71	0.314	0.159	0.174	0.309	0.427
$CF_{t+1} - Real_t$	71	-0.009	0.102	-0.046	0.001	0.042
$Diff/Real_t$	71	3.4%	51.7%	-18.8%	0.7%	16.2%
Panel B: Non-survivor bank						
CF II in t+1	71	0.302	0.221	0.124	0.300	0.471
Real II in t	71	0.302	0.226	0.138	0.272	0.512
$CF_{t+1} - Real_t$	71	0.000	0.223	-0.129	-0.020	0.129
$Diff/Real_t$	71	68.3%	370.3%	-41.6%	-10.9%	33.2%

E.3 Details on the increase in interbank rate counterfactual

In our model, the interbank rate r_0 affects both the marginal cost of loans and deposits through the bank's balance position in the interbank market, B_j. Specifically, r_0 increases the marginal cost of loans while decreasing the marginal cost of deposits. As a result, r_0 enters our equilibrium equations additively through the residual terms in the social surpluses for loans and deposits, where it reduces the social surplus of loans and increases the social surplus of deposits.

$$\eta^{\ell}_{jmt} = -\alpha^{\ell}\, r_0 + \dots$$

$$\eta^{d}_{jmt} = \alpha^{d}\, r_0 + \dots$$

Importantly, the impact of r_0 on the equilibrium equations is determined by the magnitude of parameters α^{ℓ} and α^{d}.

Our main challenge in implementing a counterfactual experiment involving changes to the interbank interest rate (Experiment 7) is that it requires knowledge of parameters α^{d} and α^{ℓ}. Since we do not have access to the necessary bank-county level interest rate data to estimate these parameters directly, we have instead calibrated them so that our model aligns with the average demand-rate elasticities for loans and deposits reported in the literature.

Let $Elast^{\ell}_{jmt}$ and $Elast^{d}_{jmt}$ represent the absolute values of the demand-rate elasticities for loans and deposits, respectively, at the bank-county-year level. According to our model, the following relationships hold:

$$Elast^{\ell}_{jmt} = \alpha^{\ell}\, (1 - s^{\ell}_{jmt})\, p^{\ell}_{jmt}$$

$$Elast^{d}_{jmt} = \alpha^{d}\, (1 - s^{d}_{jmt})\, p^{d}_{jmt}$$

Let $\overline{Elast}^{\ell}$ and $\overline{Elast}^{d}$ represent the weighted averages of these elasticities over banks, counties, and time. Combining the above equations with the definition of these weighted averages, we obtain:

$$\overline{Elast}^{\ell} = \alpha^{\ell}\; \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{m=1}^{M} \frac{q^{\ell}_{jmt}}{Q^{\ell}}\; (1 - s^{\ell}_{jmt})\; \overline{p}^{\ell}_{jt}$$

$$\overline{Elast}^{d} = \alpha^{d}\; \sum_{t=1}^{T} \sum_{j=1}^{J} \sum_{m=1}^{M} \frac{q^{d}_{jmt}}{Q^{d}}\; (1 - s^{d}_{jmt})\; \overline{p}^{d}_{jt}$$

Here, Q^{ℓ} and Q^{d} represent the total sums of loans and deposits, respectively, across all counties, banks, and years in our sample. Our calibration of parameters α^{ℓ} and α^{d} consists of the values solving these two equations. For the bank-level average interest rates, $\overline{p}^{\ell}_{jt}$ and $\overline{p}^{d}_{jt}$, we use data from Call Reports. For the average elasticities, we use values of 3.75 for loans and 2.33 for deposits, based on estimates from Dick [2008] for deposits, and Corbae and D'Erasmo [2021] and Buchak et al. [2024a] for loans.[43]

Using this approach, we calibrate α^{d} to be 1.23; and α^{ℓ} to be 0.64.

[43] Corbae and D'Erasmo [2021] estimate a loan elasticity of 1.01, while Buchak et al. [2024a] estimate it at 6.5. Given that our model, sample period, and market definition differ from those in these papers, we use the midpoint of these estimates.

References

V. Acharya and O. Merrouche. Precautionary hoarding of liquidity and interbank markets: Evidence from the subprime crisis. Review of Finance, 17(1):107–160, 2013.

S. Agarwal, J. Grigsby, A. Hortaçsu, G. Matvos, A. Seru, and V. Yao. Searching for approval. Econometrica, 92:1195–1231, 2024.

V. Aguirregabiria, R. Clark, and H. Wang. Diversification of geographic risk in retail bank networks: Evidence from bank expansion after the Riegle-Neal Act. The RAND Journal of Economics, 47(3):529–572, 2016.

J. Allen, R. Clark, and J. Houde. Search frictions and market power in negotiated price markets. Journal of Political Economy, 127(4):1550–1598, 2019.

D. Amel and J. Stahl. Pricing behavior of banks in local banking markets: Evidence from survey data of deposit rates. Working paper, 2016.

M. Arellano and S. Bond. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Review of Economic Studies, 58(2):277–297, 1995.

A. Ashcraft, J. McAndrews, and D. Skeie. Precautionary reserves and the interbank market. Journal of Money, Credit and Banking, 43:311–348, 2011.

A. B. Atkinson et al. On the measurement of inequality. Journal of Economic Theory, 2(3):244–263, 1970.

I. Ben-David, A. Palvia, and C. Spatt. Banks' internal capital markets and deposit rates. Journal of Financial and Quantitative Analysis, 52(5):1797–1826, 2017.

M. Benetton. Leverage regulation and market structure: An empirical model of the UK mortgage market. Journal of Finance, 76:2997–3053, 2021.

M. Benetton, A. Gavazza, and P. Surico. Mortgage pricing and monetary policy. Bank of England Working Paper, 2021.

W. D. Bogin, A. and W. Larson. Local house price dynamics: New indices and stylized facts. Real Estate Economics, 47(2):365–398, 2019.

G. Buchak, G. Matvos, T. Piskorski, and A. Seru. Beyond the balance sheet model of banking: Implications for bank regulation and monetary policy. Journal of Political Economy, 132:616–693, 2024a.

G. Buchak, G. Matvos, T. Piskorski, and A. Seru. The secular decline of bank balance sheet lending. 2024b.

R. Clark, J. F. Houde, and J. Kastl. The industrial organization of financial markets. In K. Ho, A. Hortaçsu, and A. Lizzeri, editors, Handbook of Industrial Organization, vol. V. Elsevier, 2021.

R. H. Coase. The nature of the firm. Economica, 4(16):386–405, 1937.

A. Cohen and M. Mazzeo. Market structure and competition among retail depository institutions. Review of Economics and Statistics, 89(1):60–74, 2007.

A. Cohen and M. Mazzeo. Investment strategies and market structure: An empirical analysis of bank branching decisions. Journal of Financial Services Research, 38(1):1–21, 2010.

D. Corbae and P. D'Erasmo. Capital requirements in a quantitative model of banking industry dynamics. Technical report, National Bureau of Economic Research, 2019.

D. Corbae and P. D'Erasmo. Capital buffers in a quantitative model of banking industry dynamics. Econometrica, 89:2975–3023, 2021.

K. Cortés and P. Strahan. Tracing out capital flows: How financially integrated banks respond to natural disasters. Journal of Financial Economics, 125(1):182–199, 2017.

G. S. Crawford, N. Pavanini, and F. Schivardi. Asymmetric information and imperfect competition in lending markets. American Economic Review, 108(7):1659–1701, 2018.

D. M. Cutler, E. L. Glaeser, and J. L. Vigdor. The rise and decline of the American ghetto. Journal of Political Economy, 107(3):455–506, 1999.

H. Degryse. On the interaction between vertical and horizontal product differentiation: An application to banking. Journal of Industrial Economics, 44(2):169–186, 1996.

A. Dick. Demand estimation and consumer welfare in the banking industry. Journal of Banking & Finance, 32 (8):1661–1676, 2008.

I. Drechsler, A. Savov, and P. Schnabl. The deposits channel of monetary policy. The Quarterly Journal of Economics, 132(4):1819–1876, 2017.

J. Driscoll and R. A. Judson. Sticky deposit rates. Finance and Economics Discussion Series 2013-80, Board of Governors of the Federal Reserve System (U.S.), 2013. URL https://EconPapers.repec.org/RePEc:fip:fedgfe:2013-80.

O. D. Duncan and B. Duncan. A methodological analysis of segregation indexes. American Sociological Review, 20(2):210–217, 1955.

M. Egan, A. Hortaçsu, and G. Matvos. Deposit competition and financial fragility: Evidence from the US banking sector. American Economic Review, 107(1):169–216, 2017.

M. Egan, S. Lewellen, and A. Sunderam. The cross section of bank value. Review of Financial Studies, 35: 2101–2143, 2022.

Federal Deposit Insurance Corporation. Summary of deposits reporting instructions. Federal Deposit Insurance Corporation (FDIC) Report, 2020. URL https://www.fdic.gov/regulations/resources/call/sod/2020-06-sod-instructions.pdf.

A. Gandhi, Z. Lu, and X. Shi. Estimating demand for differentiated products with zeroes in market share data. Quantitative Economics, 14:381–418, 2023.

M. Gentzkow, J. M. Shapiro, and M. Taddy. Measuring group differences in high-dimensional choices: Method and application to congressional speech. Econometrica, 87(4):1307–1340, 2019.

E. Gilje. Does local access to finance matter? Evidence from U.S. oil and natural gas shale booms. Management Science, 65(1):1–18, 2017.

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

X. Giné and R. M. Townsend. Evaluation of financial liberalization: A general equilibrium model with constrained occupation choice. Journal of Development Economics, 74(2):269–307, 2004.

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

T. H. Hannan and R. A. Prager. Multimarket bank pricing: An empirical investigation of deposit interest rates. Journal of Economics and Business, 58(3):256–272, 2006. ISSN 0148-6195. doi: https://doi.org/10.1016/j.jeconbus.2005.07.004. URL https://www.sciencedirect.com/science/article/pii/S0148619506000051.

J. Hausman. Valuation of New Goods Under Perfect and Imperfect Competition,. National Bureau of Economic Research, Chicago, 1996.

K. Hendricks, J. Houde, and D. Raising. Asymmetric information in the wholesale market for mortgages: The case of Ginnie Mae loans. Working paper, 2024.

K. Ho and J. Ishii. Location and competition in retail banking. International Journal of Industrial Organization, 29(5):537–546, 2011.

E. Honka, A. Hortaçsu, and M. A. Vitorino. Advertising, consumer awareness, and choice: Evidence from the US banking industry. The RAND Journal of Economics, 48(3):611–646, 2017.

R. Huang. The effect of monetary tightening on local banks. 2008.

J. Jahn, C. F. Schmid, and C. Schrag. The measurement of ecological segregation. American Sociological Review, 12(3):293–303, 1947.

J. Jayaratne and D. P. Morgan. Capital market frictions and deposit constraints at banks. Journal of Money, Credit and Banking, pages 74–92, 2000.

A. K. Kashyap, R. Rajan, and J. C. Stein. Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. The Journal of Finance, 57(1):33–73, 2002.

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

E. Loutskina and P. Strahan. Financial integration, housing, and economic volatility. Journal of Financial Economics, 115(1):25–41, 2015.

J. Mankart, A. Michaelides, and S. Pagratis. Bank capital buffers in a dynamic model. Financial Management, 49:473–502, 2020.

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

H.-L. Nguyen. Are credit markets still local? Evidence from bank branch closings. American Economic Journal: Applied Economics, 11(1):1–32, 2019.

L. Norden and M. Weber. Funding modes of german banks: structural changes and their implications. Journal of Financial Services Research, 38(2):69–93, 2010.

E. Oberfield, E. Rossi-Hansberg, N. Trachter, and D. Wenning. Banks in space, 2024. NBER Working paper 32256.

I. Petkov. Small business lending and the bank-branch network. Journal of Financial Stability, 64, 2023.

R. Rosen. Competition in mortgage markets: The effect of lender type on loan characteristics. Federal Reserve Bank of Chicago Economic Perspectives, 35(1):2–21, 2011.

E. Schargrodsky and F. Sturzenegger. Banking regulation and competition with product differentiation. Journal of Development Economics, 63(1):85–111, 2000.

K. Uetake and Y. Watanabe. Entry by merger: Estimates from a two-sided matching model with externalities. Available at SSRN 2188581, 2019.

X. Vives. Competition and Stability in Banking. Princeton University Press, 2016.

Y. Wang, T. M. Whited, Y. Wu, and K. Xiao. Bank market power and monetary policy transmission: Evidence from a structural estimation. Journal of Finance, 77:2093–2141, 2022.

M. J. White. The measurement of spatial segregation. American Journal of Sociology, 88(5):1008–1018, 1983.

M. J. White. Segregation and diversity measures in population distribution. Population Index, pages 198–221, 1986.

K. Xiao. Monetary transmission through shadow banks. The Review of Financial Studies, 33(6):2379–2420, 2020.