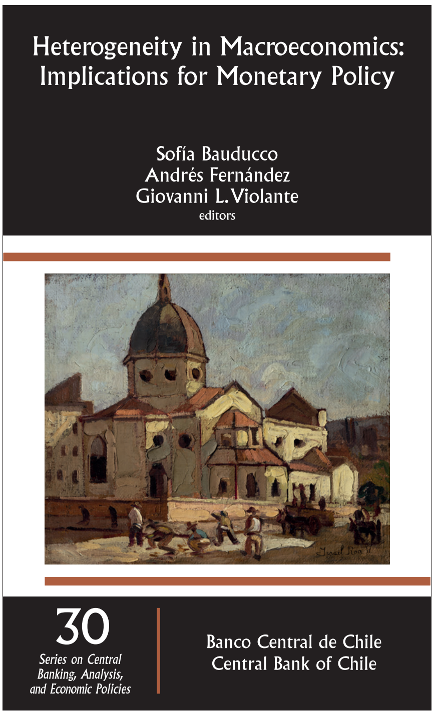
THE BANK LENDING CHANNEL ACROSS TIME AND SPACE

Dean Corbae
University of Wisconsin-Madison
National Bureau of Economic Research

Pablo D'Erasmo
Federal Reserve Bank of Philadelphia

Pages 135 – 180 | Chapter printout



The views and conclusions presented in the book are exclusively those of the authors and do not necessarily reflect the position of the Central Bank of Chile or its Board Members.

Published in Santiago, Chile
by the Central Bank of Chile
Manufactured in Chile

ISBN (print) 978-956-7421-74-9
ISBN (digital) 978-956-7421-73-2
Intellectual Property Registration 2024-A-10537

ISSN 0717-6686 (Series on Central Banking, Analysis, and Economic Policies)

THE BANK LENDING CHANNEL ACROSS TIME AND SPACE

Dean Corbae
University of Wisconsin-Madison
National Bureau of Economic Research

Pablo D'Erasmo
Federal Reserve Bank of Philadelphia

We build a model of banking industry dynamics with imperfect competition to address the following question: How does monetary policy affect lending outcomes across time, given the spatial expansion of the banking industry?

Geographic expansion of the banking industry followed from the elimination of cross-state branching restrictions begun in the McFadden Act of 1927, which permitted national banks to branch only to the same extent as state banks, thus giving the states ultimate authority. While some states permitted such cross-state branching prior to 1994, the Riegle-Neal Act removed several obstacles to banks opening branches in other states and provided a uniform set of rules regarding banking in each state.

As we document in our data in section 1, following the signing of the Riegle-Neal Act there was rapid expansion of banks crossing state lines. What is interesting is how it has translated into banking concentration. Specifically, we document that by the mid-2000s the cross-section (top 4, top 5 - 35, top 36 - 2 percent) of commercial bank cross-state deposit expansion diverged significantly. This geographic expansion coincides with the rise of U.S. bank concentration and Herfindahl indices at the national and state levels.

The authors wish to thank Jean-Francois Houde and David Moreno for helpful comments and seminar participants at the Bank of Chile. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

Heterogeneity in Macroeconomics: Implications for Monetary Policy, edited by Sofía Bauducco, Andrés Fernández, and Giovanni L. Violante, Santiago, Chile.

In section 1 we also document that, coincident with this geographic expansion, there have been important changes in the variance of deposit inflows, loan returns, and interest margins across bank sizes. An economy subject to shocks that are not perfectly correlated across space can explain these facts through geographic diversification. We also document that average costs decrease with bank size (suggestive of increasing returns to scale) which have been falling over time. These facts are consistent with a model of banking along the lines of the Diamond (1984) delegated monitoring model.[1]

Finally, section 1 examines how the bank lending channel along the lines of Kashyap and Stein (1995) and Kashyap and Stein (2000) has changed in the pre- versus post-reform data samples. We find that smaller banks exhibit greater sensitivity of their loan supply to increases in fed funds rates, consistent with a corporate finance view that smaller firms are subject to higher external finance costs. Across time we see that bigger banks have become less sensitive, while smaller banks have become more sensitive to policy hikes.

In sections 2 and 3, we build a model consistent with many of these facts. At its heart is a model of banking industry dynamics with imperfect competition as in some of our earlier work Corbae and D'Erasmo (2020) and Corbae and D'Erasmo (2021).[2] Along the lines of the 2020 paper, banks endogenously climb a size ladder consistent with higher costs to grow across space. The idea is that all banks start as state banks and Riegle-Neal lowered the cost of branching out on a regional and national basis. The equilibrium distribution of banks on the Besanko and Doraszelski (2004) ladder is solved by using the approximation techniques of Farias and others (2012).

After parameterizing the model in Section 4, we then use it to assess how geographic expansion over time affects the bank lending channel of monetary policy in Section 5. Geographic expansion has led to a skewed bank size distribution where big national banks account

1. Specifically, Diamond provides a framework where large banks arise to economize on the fixed costs of monitoring individual borrowers more efficiently than a large number of small depositors. Economies of scale in monitoring (decreasing average costs) induce size. The problem of monitoring the monitor is also solved by size; large, diversified banks can offer noncontingent (and hence incentive-compatible) deposit contracts. There are numerous empirical papers documenting the existence of scale economies in banking such as Berger and Mester (1997) or Berger and Hannan (1998). A large pool of depositors is also consistent with geographic diversification as described in Liang and Rhoades (1988).

2. A closely related paper by Aguirregabiria and others (2020) studies how geographic dispersion may prevent funding from flowing to high loan-demand areas. Also closely related is Gelman and others (2022) as well as Morelli and others (2023).

for a large share of the loan market. Monetary policy which raises the cost of external funding, e.g., a rise in fed funds, can have differential effects on the lending behavior of banks of different sizes along the lines of Kashyap and Stein (1995) distributed across different regions. Changes in monetary policy are a blunt instrument because it affects all regions instead of the affected region. As in Bellifemine and others (2022) and Wang and others (2022), we study the transmission of monetary policy in a model with bank heterogeneity and imperfect competition. We incorporate spatial differences and assess the bank lending channel across time. In particular, we conduct a counterfactual in Section 5 where we raise the cost of external finance and examine how it affects banks of different sizes across time.

1. Stylized Facts

In this section we present some data facts for a cross-section of the top 2 percent of banks across time. The data come from both the Fed's Consolidated Reports of Condition and Income for Commercial Banks, regularly called "Call Reports", which begin in 1984, and the Summary of Deposits of the Federal Deposit Insurance Corporation (FDIC), which begins in 1994.

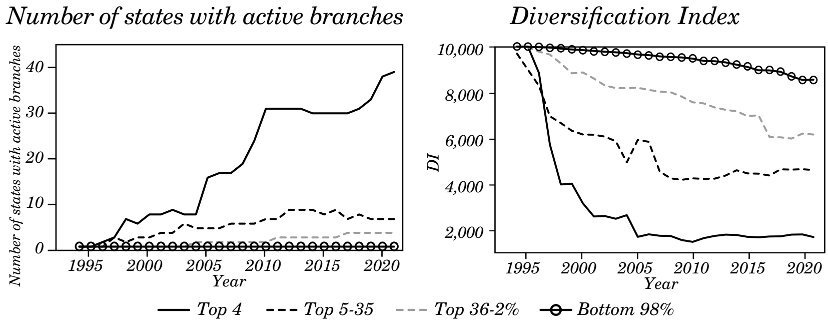
We examine geographic expansion following Riegle-Neal starting in 1994 in figure 1. The top panel graphs the number of states where a bank in each size category has an active branch. The bottom panel plots a diversification index. Let $\ell_{i,m,t}$ denote the amount of loans originated by lender i in market m in period t. Here we take m to be a state. The share of loans of lender i in state m in period t is $S_{i,m,t} = \frac{\ell_{i,m,t}}{\Sigma m \in M_{i,t} \ell_{i,m,t}}$ x 100 where $L_{i,t} = \Sigma m \in M_{i,t} \ell_{i,m,t}$ is the total amount of loans originated by lender i in period t and $M_{i,t}$ denotes the states in which lender i operates. We define a diversification index as follows:[3]

$$DI_{i,t} = \sum_{m \in M_{i,t}} s^2_{im,t}. \tag{1}$$

This index ranges between 0 and 10,000, and a smaller value indicates a more diversified lender. The bottom panel in figure 1 shows the (deposit-weighted) average of this diversification index within size categories. It is clear that there is a positive relationship between size and geographic diversification.

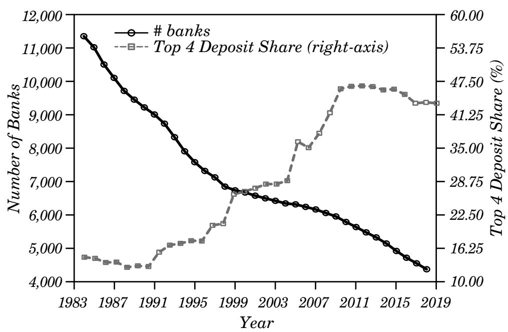
3. This measure is applied at the county level in Shin (2022).

Figure 1. Deposit Space and Size over time



Source: Source: Summary of Deposits.
Note: Banks are ranked according to deposits. Source: Summary of Deposits.

Figure 2. U.S. Banking Concentration

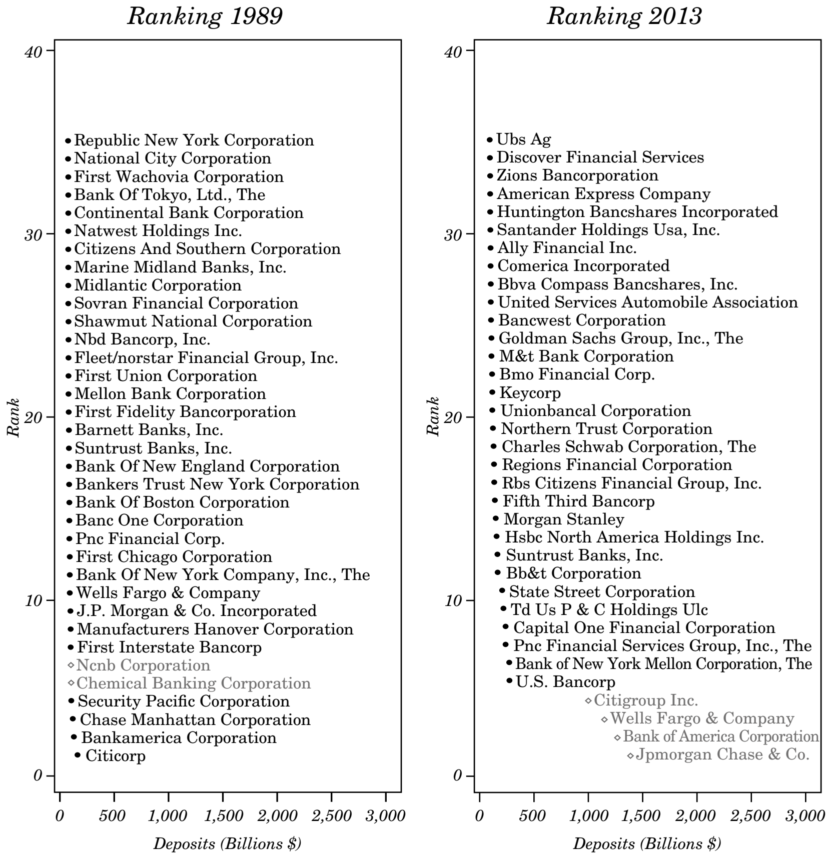


Source: Reports of Condition and Income (Call Reports).
Note: Banks are ranked according to deposits.

Figure 2 graphs the deposit market share of the top 4 banks across time. It shows that, prior to Riegle-Neal, market shares were relatively constant; then it shows a rapid transition following Riegle-Neal until the 2008 Financial Crisis, followed by another relatively constant share. Importantly, the fact that deposit shares in figure 2 are relatively constant during both the 'pre-reform' 1984–1992 and 'post-reform' 2011–2019 periods motivates our modeling choice of calibrating parameters consistent with long-run equilibria of those periods.

Figure 3 graphs how the top 35 individual banks ranked according to deposit size grew over time. Notable is the divergence of the top four from even the remaining top 5 - 35. This motivates our decision to categorize banks into three size bins (top 4, top 5 - 35, and top 36 - 2 percent).

Figure 3. Deposit Distribution Pre- and Post-Reform



Source: Reports of Condition and Income (Call Reports).
Note: Banks are ranked according to deposits. Deposits are in reported in real terms. Red bubbles identify banks that end up in the Top 4 of the distribution post-reform (2013).

As banks expand over space, they diversify shocks over their region. Here we explore how the variance of deposit inflows, loan returns, and interest margins vary across bank size and time (pre-reform 1984–1992 and the stationary portion post reform 2011–2019). Using our panel of commercial banks in the U.S., we estimate how that process of deposits evolves for bank holding companies of different sizes. We keep the same grouping convention that we described above. After controlling for firm and year fixed effects as well as a time trend, we estimate the following autoregressive process for log-deposits for bank i of type $\Theta \in \{s, r, n\}$ in period t:

$$\log\left(d^i_{\theta,t}\right) = \left(1-\rho^d_\theta\right)\overline{d}_\theta + \rho^d_\theta \log\left(d^i_{\theta,t-1}\right) + u^i_{\theta,t}, \tag{2}$$

where $d^i_{\theta,t}$ is the sum of deposits and other borrowings in period t for bank i, and $u^i_{\theta,t}$ is iid and distributed $N(0,\sigma^2_{\theta,u})$. Assuming the process is stationary, the variance of the deposit process is given by $\sigma_\theta = \frac{\sigma_{\theta,u}}{\left(1-\left(\rho^d_\theta\right)^2\right)^{1/2}}$. Since this is a dynamic model, we use the method proposed by Arellano and Bond (1991). Consistent with the evidence presented in figure 2, we estimate this process for the pre-reform period (1984–1992) and for the latest period in our sample (2009–2018).

Table 1 presents the results. The top panel provides estimates for the period prior to the passage of Riegle-Neal, which constrained bank branching to the state level (i.e. $\theta = s$), while the bottom panel provides estimates for a period of market share stability following Riegle-Neal, where some banks crossed their state borders to grew from $\theta = s$ to regional ($\theta = r$) and national ($\theta = n$) levels. Consistent with the diversification story in Diamond (1984) and empirical work such as Liang and Rhoades (1988), we find that the variance of deposit inflows σ_θ decreases as banks grow in size.

Corbae and D'Erasmo (2023) conduct the same analysis for loan returns, interest margins, and charge-off rates. Again it is apparent that the variance σ_θ of these variables decreases as banks grow in size. One interesting fact from that analysis is how close the average loan returns and margins are, which motivates our use of a Cournot-equilibrium concept with capacity constraints.

Table 1. Deposit Process Parameters

Size Group		*Pre-Reform (1984-1992)*			
Data	Model	d_θ	ρ_θ	$\sigma_{u,\theta}$	σ_θ
Top 2%	s	0.140	0.863	0.1773	0.3506

Size Group		*Post-Reform (2011-2019)*			
Data	Model	d_θ	ρ_θ	$\sigma_{u,\theta}$	σ_θ
Top 4	n	10.563	0.699	0.0306	0.0428
Top 5 - 35	r	1.000	0.764	0.0861	0.1333
Top 36 - 2%	s	0.138	0.761	0.1034	0.1595

Source: Call Reports.
Note: We study banks in the top 2 percent of the asset distribution. We group all banks ($\theta = s$) for the pre-reform period since regulation prevented them from expanding across state borders. For the post-reform period, we split this group and consider top 4 ($\theta = n$), top 5 - 35 ($\theta = r$), and top 36-2% ($\theta = s$). Average deposits are normalized to 1 for the top 5 - 35 group in the post-reform period. Average deposits d_θ is reported relative to this group.

With regard to the bank lending channel, we follow Kashyap and Stein (1995) and run the following specification:

$$\Delta y_{i,t} = \sum_{h=1}^{8}\beta_h\left(f_{t-h} - f_{t-h-1}\right) + \sum_{h=1}^{4}\alpha_h \Delta y_{i,t-h} + \gamma_j X_t + \phi x_{i,t} + a_i + \tau_t + Q_t + \in_{i,t}, \quad (3)$$

where $\Delta y_{i,t}$ denotes the growth rate of $y_{i,t}$ (loans) between quarter t and quarter $t-1$, f_{t-h} corresponds to the fed funds rate in period $t-h$, X_t captures aggregate variables (such as the inflation rate or changes in nominal GDP), $x_{i,t}$ are bank level controls that include the ratio of deposits to assets, the ratio of equity to assets, the ratio of cash and securities to assets. a_i is a bank fixed effect, τ_t is a year fixed effect, and Q_t is a quarter fixed effect.[4] Table 2 reports the value of $\Sigma^{8}_{h=1}\beta_h$ together with the p-value of the corresponding test of significance for the sum.

Table 2. Bank Lending Channel: Pre & Post by Bank Size

	Dep Var: Growth Loans	
	Pre-Reform	*Post-Reform*
Top 4	-0.2919	0.6480
Top 5 - 35	(0.22)	(0.76)
p-value	-0.2219	-1.2149
Top 36 - 2%	(0.03)	(0.07)
p-value	-0.1008	-1.3545
All (Top 2%)	(0.01)	(0.01)
p-value	-0.1212	-1.4272
ΔCPI	(0.00)	(0.00)
DeltaNGDP	yes	yes
Other Bank Controls	yes	yes
Bank FE	yes	yes
Period	84-92	11-19

Source: Reports of Condition and Income (Call Reports) and FRED, Federal Reserve Economic Data.
Note: p-values (for sum of coefficients) in parenthesis. Other bank controls include ratio of deposits to assets, the ratio of equity to assets, the ratio of cash and securities to assets.

4. See Appendix A.2 for a description of the data used to perform this analysis.

Importantly, table 2 documents that bigger banks are less sensitive to a rise in the fed funds rate than smaller banks across both pre-reform (1984–1992) and post-reform (2011–2019) periods. Specifically, the coefficients for top 4 banks are insignificant across both time periods and top 5 - 35 become insignificant at the five-percent level post reform. Interestingly, top 36 – 2 percent banks become more sensitive to rises in the fed funds rate. These results are consistent with the logic that bigger, more diversified banks have access to other sources of external funding and hence less sensitive to external funding via fed funds, as in Kashyap and Stein. Except for the smallest banks, the argument that more diversification through time makes banks less sensitive to fed funds shocks is also consistent with this idea.

In summary, expanded data analysis in Corbae and D'Erasmo (2023) documents:

1. Diversification across space grows with bank size over time as in figure 1.

2. The growth of concentration of the top 4 banks across time as in figures 2 and 3.

3. Relative stability of concentration prior to Riegle Neal in 1994 (i.e., 1984–1992) and after the Global Financial Crisis (i.e., 2011–2019), as is evident in figure 2.

4. Deviations from Zipf's law arising from growth of the right tail.

5. Rising deposit Herfindahl Indices at the national and state levels. The current average of state-level Herfindahl indices falls into the "moderately concentrated" designation by the Justice Department's Antitrust Division. The Herfindahl has grown by 80 percent from 1994. It also documents relative stability of Herfindahl indices prior to Riegle Neal in 1994 and after the Global Financial Crisis.

6. The growth in insured deposit funding and drop in variance of inflows across time by bank size, as is evident in table 1. The drop in variance associated with geographic expansion suggests geographic diversification.

7. The drop in variance of loan returns, interest margins, and charge-off rates respectively across time by bank size, again suggestive of geographic diversification.

8. The drop in average costs across time by bank size. The fact that average costs drop by bank size is consistent with increasing returns to scale at least over certain size ranges.

9. The cyclical nature of bank exit.

These data facts motivate our modeling choices:

1. Diversification and increasing returns are consistent with the delegated monitoring model of banks in Diamond (1984).

2. High levels of concentration motivate us to model the banking industry as imperfectly competitive.

3. We model bank growth with imperfect competition via a ladder along the lines of Besanko and Doraszelski (2004).

(a) We assume that pre-Riegle-Neal regulation made the cost of geographic expansion infinite while post-Riegle-Neal costs fall so that there is growth from state to regional to national consistent with figure 1.

(b) As banks grow, they expand their capacity and lower their variance of low-cost deposit inflows. We model this by bank-size-dependent Markov processes for exogenous deposit inflows consistent with the data analysis in table 1.

(c) As banks grow, they also bear lower costs of nondeposit external funding along the lines of standard models of corporate finance.

4. Banks Cournot-compete in the loan market subject to deposit capacity constraints. They also must compete with the nonbank funding sector.

5. There is endogenous bank exit across the business cycle (modeled here by the aggregate shock process). More variable interest margins and charge-off rates make smaller banks more susceptible to failure, especially in downturns.

(a) Endogenous bank exit also allows us to examine how monetary and regulatory policy can affect the bank size distribution and financial stability.

2. Model Environment

Time is discrete and there is an infinite horizon. There are two regions $j \in \{\mathcal{E}, \mathcal{W}\}$, for instance, 'east'" and 'west'. Each period, a mass B of ex-ante identical entrepreneurs who have a profitable project that needs to be funded (the potential borrowers) are born in each region. There is also a mass $H > B$ of identical households (the potential depositors) in each region that deposit their funds in the banking sector and finance banks and nonbanks $k \in \{\mathcal{B}, \mathcal{N}\}$ via equity injections, where $\mathcal{B}$ denotes traditional banks and $\mathcal{N}$ nonbanks. Financial intermediaries (banks and nonbanks) intermediate between potential borrowers and depositors.

To keep notation manageable, we let any beginning-of-period variable be denoted x and any end-of-period variable be denoted x'. Further, except where critical to understand the problem, we will not index by region j. For example, any decision rule taken in region j should be understood to depend on j.

2.1 Entrepreneurs

Ex-ante identical borrowers in region j demand loans in order to fund a risky project. The project requires one unit of investment (i.e., a loan either from a bank $k = \mathcal{B}$ or nonbank $k = \mathcal{N}$) at the beginning of period t. The entrepreneur chooses the scale R_k of the risky project in which they are investing those funds, which can be indexed on the lender type.[5] The project returns R_k at the end of the period according to:

$$\begin{cases} 1+z_j' R_k & \text{with prob } p_j\left(R_k, z_j',\right) \\ 1-\lambda & \text{with prob} \left[1-p_j\left(R_k, z_j'\right)\right] \end{cases} \tag{4}$$

in the successful and unsuccessful states, respectively. That is, borrower gross returns are by $1 + z_j' R_k$ in the successful state and by $1 - \lambda'$ in the unsuccessful state, where z_j' is a regional-specific shock and λ is the fraction lost in default. The regional shocks z_j' are assumed to be independent over time and drawn from a bivariate normal distribution $F_z(\mu_z,\sigma_z,\rho_z)$ where μ_z denotes the mean, σ_z the standard deviation, and ρ_z covariance between regions. The success of a borrower's project, which occurs with probability $p_j(R_k,z_j')$, is independent across borrowers and time conditional on the borrower's choice of technology $R_k \geq 0$ and regional shock z_j'.

As for the likelihood of success or failure, a borrower who chooses to run a project with a higher return R_k has more risk of failure. Specifically $p_j(R_k,z_j')$ is assumed to be decreasing in R_k and increasing in z_j'. Thus, the technology exhibits a risk-return trade-off. Further, since R_k is a choice variable, project returns and failure rates are endogenously determined. While borrowers are ex ante identical, they are ex-post heterogeneous owing to the realizations of the shocks to

5. Note this is the first occurrence of the notation simplification we alluded to above; in general, since risky scale is a choice variable of the entrepreneur in region j, we would denote it $R_{k,j}$, but we neglect the j subject to keep notation manageable.

the return on their project. We envision borrowers either as firms choosing a technology that might not succeed or households choosing a house that might appreciate or depreciate.

The entrepreneur makes a discrete choice over which type of financial institution to borrow from $k \in \{\mathcal{B}, \mathcal{N}\}$. Bank and nonbank interest rates on their loans to the entrepreneur can differ. Taking the vector of interest rates $r_j = \{r_{\mathcal{B},j}, r_{\mathcal{N},j}\}$ on loans as given, entrepreneurs decide whether they want to fund a project given their outside option and then make a discrete choice over whether to borrow from a bank or nonbank in their region.

Once with a lender type k offering a loan at interest rate $r_{k,j}$, the entrepreneur chooses the risk-return tradeoff of their project $R_{k,j}$. This explains why we allow project choice to depend on k; since the borrower potentially faces different rates from different lenders, they may make different risk-return project choices. Following Buchak and others (2018), we assume that the value associated with financing the project with each type of lender in region j is subject to an unobservable idiosyncratic shock $\in = \{\in_{\mathcal{B}}, \in_{\mathcal{N}}\}$ affecting the value of taking a loan from each type of lender additively. We assume that $\in_k$ are iid shocks drawn from a type-one extreme-value distribution $F_{\in}(\in;\alpha)$ with scale parameter $1/\alpha$.

Borrowers have an outside option. At the beginning of period t, they receive a realization of their reservation utility of consumption $\omega \in [0, \overline{\omega}]$ if they decide not to run the project. These draw from distribution function $\Omega(\omega)$ are i.i.d. over time and across regions. This outside option leads to a downward-sloping aggregate demand for loans, while the extreme-value shocks determine loan demand across financial institution types.

There is limited liability on the part of the borrower at the project level so that the project return net of interest payments is bounded below at zero. If $r_{k,j}$ is the interest rate on a loan that the borrower faces, the borrower receives $\max\{z_j' R_k - r_{k,j}, 0\}$ in the successful state and 0 in the failure state. Specifically, in the unsuccessful state they receive $1 - \lambda$, which must be relinquished to the lender. Table 3 summarizes the risk-return tradeoff that the borrower faces. Since the choice of R_k is endogenous, changes in borrowing costs $r_{k,j}$ can affect the default frequencies on loans through a risk-shifting motive.

Table 3. Borrower's Problem (Conditional on Investing)

Borrower Chooses R_j	*Receive*	*Pay*	*Probability*
Success	$(1+z_j' R_k)$	$(1+r_{k,j})$	$p \; (\overset{-}{R_k}, \; \overset{+}{z_j'})$
Failure	$(1-\lambda)$	$min\{(1-\lambda),(1+r_{k,j})\}$	$1-p \; (\overset{-}{R_k} \; \overset{+}{z_j'})$

Both R_k and ω are private information to the entrepreneur. As in Bernanke and Gertler (1989), success or failure is also private information to the entrepreneur unless the loan is monitored by the lender.[6] With one-period loans, since reporting failure (and hence repayment of $1-\lambda < 1 + r_{k,j}$) is a dominant strategy in the absence of monitoring, loans must be monitored. Monitoring is costly as in Diamond (1984).

2.2 Households

In each region j, infinitely lived, risk-neutral households with discount factor β are endowed with one unit of the good each period. We assume households are sufficiently patient such that they choose to exercise their savings opportunities. In particular, households have access to an exogenous risk-free storage technology yielding $1 + \bar{r}$ between any two periods with $\bar{r} \geq 0$ and $\beta(1 + \bar{r}) = 1$. They can also choose to supply their endowment to a bank, a nonbank, or an individual entrepreneur. We assume that, after observing the deposit interest rate $r_{D,j}$, households who choose to deposit their earnings are randomly matched with a bank in their region at the beginning of any period t. Given deposit insurance, even if the bank fails, they receive their deposit with interest at the end of the period. Households can hold a portfolio of bank stocks yielding dividends (claims to bank cash flows) and can inject equity to banks. They can also invest in shares of the representative nonbank, which gives a claim to nonbank cash flows. They pay lump-sum taxes/transfers τ' at the end of any

6. While one interpretation of our entrepreneurs is that they are effectively one-period lived (born at the beginning of the period and dead at the end as in the OG model of Bernanke and Gertler (1989), we could have effectively modeled entrepreneurs as long-lived and added enough interperiod anonymity so that financial contracts are one-period lived as in Carlstrom and Fuerst (1997) and entrepreneurs are sufficiently impatient not to want to augment net worth.

period t which include a lump-sum tax τ'_D used to cover deposit insurance for failing banks. Finally, if a household wants to match directly with an entrepreneur (i.e., directly fund an entrepreneur's project), it must compete with bank loans. Hence, the household could not expect to receive more than the bank lending rate $r_{k,j}$ in successful states and must pay a monitoring cost. Since households can purchase claims to bank cash flows, and banks can more efficiently minimize costly monitoring along the lines of Diamond (1984), there is no benefit to matching directly with entrepreneurs.

2.3 Banks

We build a model along the lines of Ericson and Pakes (1995) where, within a region, banks Cournot-compete in a single-good market (loans) and there is endogenous entry and exit. As in Diamond (1984), banks exist in our environment to pool risk and economize on monitoring costs. We assume there are three types of banks: $\theta \in \Theta=\{\mathfrak{s},\mathfrak{r},\mathfrak{n}\}$ with size ranking $\mathfrak{s} < \mathfrak{r} < \mathfrak{n}$. We identify banks of type $\mathfrak{s}$ with small state banks and banks of type $\mathfrak{r}$ with bigger regional banks, both of which are constrained to operate in only one region. We associate banks of type $\mathfrak{n}$ with large national banks operating across regions. There can be multiple banks of each type operating and banks of all types have some degree of market power.

To save on notation, each incumbent bank with the same state variables will be treated identically. We denote loans made by such a bank of type θ in region j in period t by ℓ_θ.[7] As in Corbae and D'Erasmo (2021), bank type θ determines the mean and variance of a bank's deposits $d_\theta \in D_\theta$. In particular, banks in the model face the deposit process we estimated in equation (2). To make our definition of type consistent with the data presented in table 1, the mean of the deposit process satisfies $\overline{d}_\mathfrak{n} > \overline{d}_\mathfrak{r} > \overline{d}_\mathfrak{s}$ so that higher types have a bigger funding base. Furthermore, also consistent with the data presented in table 1, the variance of deposits satisfies $\sigma_\mathfrak{n} \leq \sigma_\mathfrak{r} \leq \sigma_\mathfrak{s}$ so that bigger banks have lower variance consistent with diversification. We discretize the continuous deposit process d_θ in equation (2) into a finite support and denote its transition matrix by $G_\theta(d'_\theta, d_\theta)$. Unlike Corbae and D'Erasmo (2021), here deposits are the only source of funding besides seasoned

7. Again, since this is a choice variable, it should be understood that ℓ_θ also depends on j and there will be places where we make that explicit.

equity. Deposits are collected at the regional level, but we assume that national banks (n) can move deposits freely across regions. Since we do not take a stand on what a 'region' is both in the model and in the estimation in equation (1), to simplify on notation, we abstract from denoting the regional origin of d_θ.

Along the lines of Besanko and Doraszelski (2004), a given bank of type θ can invest $I_\theta \in \mathbb{R}_+$ to become a larger-type bank (i.e., a small local bank can invest to become a regional bank and a medium-sized regional bank can invest to become a large national bank). One can interpret this investment technology as a reduced-form way of capturing geographic expansion in ways that can also include mergers and acquisitions. We have assumed that prior to Riegle-Neal, all banks were restricted to operate only in their home state (i.e., of type $\theta = s$). After Riegle-Neal lowers the cost of geographic expansion, any $\theta = s$ bank can then invest I_s to transit to a bigger regional type $\theta' = r$ according to the following transition function:

$$(\theta' \mid \theta = s, I_s) = \begin{cases} \dfrac{(\alpha \cdot I_s \cdot (\Delta \bar{d}_{r,s})^{-\xi}}{1+\alpha I_s \cdot (\Delta \bar{d}_{r,s})^{-\xi}} & \text{if } \theta' = r \\ \dfrac{1}{1+\alpha \cdot I_s \cdot (\Delta \bar{d}_{r,s})^{-\xi}} & \text{if } \theta' = s \end{cases}, \tag{5}$$

where the parameters $\alpha > 0$ and $\xi > 0$ measure the effectiveness of investment (at $\iota_\theta = 0$ to be precise) and $\Delta \bar{d}_{r,s} = (\bar{d}_r - \bar{d}_s) > 0$. Since banks of type s are already heterogeneous via the deposit shock process, the bigger ones may have an incentive to bear the cost of growing to r while the smaller ones may remain of type s.

After state-level banks branch out to become regional, the following transition function for a type r bank governs whether it grows to become national (n), shrinks to become state (s), or remains regional:

$$T(\theta' \mid \theta = \mathfrak{r}, I_{\mathfrak{r}}) = \begin{cases} \dfrac{(1-\delta)\alpha \cdot I_{\mathfrak{r}} \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}}{1+\alpha I_{\mathfrak{r}} \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}} & \text{if } \theta' = n \\ \dfrac{1-\delta+\delta\alpha \cdot I_{\mathfrak{r}} \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}}{1+\alpha I_{\mathfrak{r}} \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}} & \text{if } \theta' = \mathfrak{r} \\ \dfrac{\delta}{1+\alpha \cdot I_{\mathfrak{r}} \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}} & \text{if } \theta' = \mathfrak{s} \end{cases}, \quad (6)$$

where $\Delta\bar{d}_{n,\mathfrak{r}} = (\bar{d}_n - \bar{d}_{\mathfrak{r}}) > 0$.[8,9]

After the realization of θ, a given incumbent bank is randomly matched with a set of potential household depositors d_θ who receive deposit interest rate $r_{D,j}$ and then decide how many loans to extend. Regional and state banks $\theta \in \{\mathfrak{r},\mathfrak{s}\}$ can fund an amount of loans larger than its deposits by using external borrowing $a_\theta < 0$ at rate $r^a(a_\theta) > \bar{r}$. If a bank chooses an amount of loans lower than its capacity constraint, the leftover deposits a_θ can be invested in the same risk-free technology that the households have access to with return equal to $\bar{r}$. The flow constraint for such regional and small banks is $\ell_{\theta,j} + a_\theta = d_\theta$. National banks are geographically diversified in the sense that they extend loans and receive deposits in both regions $(\Sigma_j \ell_{n,j} + a_n = d_n)$. Note that, since the outside option for a household matched with a bank is to store at rate $\bar{r}$, we know that $r_{D,j} \geq \bar{r}$.

End-of-period static profits, associated with beginning-of-period deposits d_θ and lending $\ell_{\theta,j}$ in region j for an incumbent bank of type $\theta \in \{\mathfrak{s},\mathfrak{r}\}$ in industry state μ (to be described below) depends on its end-of-period state $s_j = (\mu, z'_j)$ given by

8. This specification nests Besanko and Doraszelski (2004) when $\xi_{\iota\theta} = 0$.

9. Finally, since a national bank cannot grow higher, its transition function is given by

$$T(\theta' \mid \theta = n, I_n) = \begin{cases} \dfrac{1-\delta+\alpha \cdot I_n \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}}{1+\alpha I_n \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}} & \text{if } \theta' = n \\ \dfrac{\delta}{1+\alpha \cdot I_n \cdot (\Delta\bar{d}_{n,\mathfrak{r}})^{-\xi}} & \text{if } \theta' = \mathfrak{r} \end{cases}. \quad (7)$$

We assume that a national bank that reduces its size is randomly assigned to a region $j \in \{e,w\}$ with probabilty .

$$\pi_\theta\left(d_\theta, s_j\right) = \left[p_j\left(R_{\mathcal{B}}, z'_j\right) r_{\mathcal{B},j}(\mu) - \left(1 - p_j\left(R_{\mathcal{B}}, z'_j\right)\right)\lambda\right]\ell_{\theta,j} - c_\theta\left(\ell_{\theta,j}\right) \tag{8}$$
$$+\left(\bar{r}\mathbf{1}_{\{a_\theta \geq 0\}} + r^a \mathbf{1}_{\{a_\theta < 0\}}\right) a_\theta - r_{D,j} d_\theta - c_{F,\theta},$$

where $p_j(\cdot)$ denotes the fraction of bank loans that are repaid at the end of the period in region j (an endogenous object that is consistent with the borrower's problem), $r_{\mathcal{B},j}(\cdot)$ is the Cournot-equilibrium interest rate on bank loans in region j, $c_\theta(\ell_{\theta,j})$ is the marginal cost of extending $\ell_{\theta,j}$ loans, and $C_{F,\theta}$ is the fixed operating costs. Profits for an incumbent bank of type $\theta = n$ in state $\mathbf{s} = (\mu, z'_j, z'_{-j})$ is given by

$$\pi_n\left(d_n, \mathbf{s}\right) = \sum_j \left\{\left[p_j\left(R_{\mathcal{B}}, z'_j\right) r_{\mathcal{B},j}(\mu) - \left(1 - p_j\left(R_{\mathcal{B}}, z'_j\right)\right)\lambda\right]\ell_{n,j} - c_n\left(\ell_{n,j}\right)\right\} \tag{9}$$
$$+\left(\bar{r}\mathbf{1}_{\{a_n \geq 0\}} + r^a \mathbf{1}_{\{a_n < 0\}}\right) a_n - r_{D,j} d_n - c_{F,n},$$

Given profits $\pi_\theta(d_\theta, s_j)$ for $\theta \in \{s, r\}$ and $\pi_n(d_n, s)$, banks can choose to exit. To keep the computation of the model tractable, we also incorporate a type-specific exogenous probability of exit ρ^x_θ. If a bank decides to continue, it then decides how much to invest in order to improve its capacity to collect deposits. Banks can finance investment with internal funds (π_θ) or by issuing equity e_θ whenever $I_\theta > \pi_\theta$. That is, $e_\theta = \max\{I_\theta - \pi_\theta, 0\}$. Issuing equity is costly with cost function given by $\varsigma_\theta(e_\theta)$. For tractability, unlike Corbae and D'Erasmo (2021), here we assume that banks cannot retain earnings.[10] Dividends net of equity injections for a bank of type for $\theta \in \{s, r\}$ in state (d_θ, s_j) are given by

$$\mathcal{D}_\theta\left(d_\theta, s_j\right) = \pi_\theta\left(d_\theta, s_j\right) - I_\theta - \mathbf{1}_{\{e_\theta > 0\}}\varsigma_\theta\left(e_\theta\right). \tag{10}$$

A similar equation can be written for national bank dividends $\mathcal{D}_n(d_n, s)$.

The objective function of the bank is to maximize the expected present discounted value of future dividends net of equity injections with discount factor β. It is important to note that, while deposits conditional on bank size (d_θ) are exogenous, external finance is endogenous, since bank size (θ) via investment (I_θ) and seasoned equity (e_θ) are endogenous.

10. Since one of the objectives of Corbae and D'Erasmo (2021) was to understand the role of capital buffers, we allowed for the endogenous retention of earnings which augmented a bank's capital. Here we endogenize bank size by allowing banks to invest I_θ to change θ; e.g., become bigger.

We assume there is limited liability and that incumbent banks have the option to exit after extending loans. The value of exiting for a bank of type $\theta \in \{\mathfrak{s},\mathfrak{l}\}$ is given by

$$\max\Big\{0, \zeta_\theta \Big[1 + p_j\left(R_{\mathcal{B}}, z_j'\right) r_{\mathcal{B},j}(\mu) - \left(1 - p_j\left(R_{\mathcal{B}}, z_j'\right)\lambda\right]\ell_{\theta,j} \tag{11}$$

$$-c_\theta\left(\ell_{\theta,j}\right) - c_{F,\theta} - \left(1 + r_{D,j}\right)d_\theta + \zeta_\theta 1_{\{a_n \geq 0\}}\left(1 + \overline{r}\right)a_\theta + 1_{\{a_n < 0\}}\left(1 + r^a\right)a_\theta\Big\},$$

where ζ_θ captures the recovery rate of a bank's assets at the exit stage. This induces an exit decision rule $x_\theta(d_\theta, s_j)$ for $\theta \in \{\mathfrak{s},\mathfrak{l}\}$. A similar equation can be written for the national bank.

We consider an entry process similar to Farias and others (2012). At time period t, there are a finite but large number of potential entrants. Potential entrants make entry decisions simultaneously, are short-lived, and do not consider the option of delaying entry. Entrants bear a positive entry cost κ funded by an initial equity injection by households and base their entry decision on the net present value of entering today. Entrants do not earn profits in the period they decide to enter. They appear in the following period in state $(\theta' = \mathfrak{s}, d_{\mathfrak{s}}')$ (i.e., we assume that all entrants start as a small bank), where $d_{\mathfrak{s}}'$ is drawn from $\overline{G}_{\mathfrak{s}}(d_{\mathfrak{s}})$the invariant distribution Gbar associated with $G(d_{\mathfrak{s}}', d_{\mathfrak{s}})$. We denote the number of entrants $N_{e,j}$, which is determined endogenously in equilibrium.

In summary, the simple balance sheet of a bank in our environment is given by book assets equal loans (ℓ_θ), storage (a_θ), and fixed capital (κ_θ) while book liabilities equal deposits (d_θ) and equity injections (e_θ).

If all banks in a given state $(\theta, d_\theta) \in \{\Theta \times D_\theta\}$ are treated symmetrically, then the cross-sectional distribution μ specifies the number of banks across state and region. More specifically,

$$\mu = \left\{\left\{\mu_{n,*}\left(d_n\right)\right\}_{d_n \in D_n}, \left\{\mu_{\theta,j}\left(d_\theta\right)\right\}_{\theta \in \{\mathfrak{l},\mathfrak{s}\}, j \in \{e,w\}, d_\theta \in D_\theta}\right\}. \tag{12}$$

We let N denote the number of incumbent banks at time period t, that is,

$$N = \sum_{d_n \in D_n} \mu_{n,*}\left(d_n\right) + \sum_{\theta \in \{\mathfrak{l},\mathfrak{s}\}, j \in \{e,w\}, d_\theta \in D_\theta} \mu_{\theta,j}\left(d_\theta\right). \tag{13}$$

Further, the law of motion for the industry state is denoted

$$\mu' = \mathcal{H}(\mu, N_e), \tag{14}$$

where N_e denotes the number of entrants, and the transition function $\mathcal{H}$ is defined explicitly below in equation (33).

2.4 Nonbank Lenders

A representative national nonbank that discounts the future at rate β specializes in extending loans to entrepreneurs (in both regions) in a perfectly competitive market. To keep the analysis simple, the nonbank is financed with equity $e_{\mathcal{N}}$ raised from the household sector and is not subject to limited liability. When lending to entrepreneurs, nonbanks face a marginal monitoring cost $c_{\mathcal{N}}$. Like banks, the representative nonbank can diversify entrepreneurs' idiosyncratic risk, but it is subject to regional fluctuations.

Let $\pi_{\mathcal{N}}(\mathbf{s})$ denote the end-of-period profits of the nonbank after the realization of regional shocks associated with its current lending $\ell_{\mathcal{N},j}$ given by

$$\pi_{\mathcal{N}}(\mathbf{s}) = \sum_j \left\{ \left[p_j(R_{\mathcal{N}}, z_j') r_{\mathcal{N},j}(\mu) - (1 - p_j(R_{\mathcal{N}}, z_j'))\lambda \right] - c_{\mathcal{N}} \right\} \ell_{\mathcal{N},j} \tag{15}$$

subject to flow constraint $S_{\mathcal{N}} e_{\mathcal{N}} = \Sigma_j \ell_{\mathcal{N},j}$, where $(\boldsymbol{\mu}, z'_j, z'_{-j})$ and $S_{\mathcal{N}}$ are household shareholdings of the nonbank. Since the nonbank operates in a perfectly competitive market it takes the regional interest rate $r_{\mathcal{N},j}$ as given. The nonbank issues dividends according to $D_{\mathcal{N}} = \pi_{\mathcal{N}}(\mathbf{s})$.

The objective function of the nonbank is to maximize the expected present discounted value of future cash flows to households with discount factor β. We assume that there is free entry into the nonbank sector and, to simplify the analysis, we set the entry cost to zero.

2.5 Government Budget Constraint

The government collects lump-sum taxes to cover the cost of deposit insurance. Post-liquidation net transfers are given by

$$\Delta'(d_\theta, s_j) = (1 + r_{D,\theta}) d_\theta$$
$$- \zeta_\theta \left[1 + p \cdot r_{\mathcal{B}}(Z, \mu) - (1 - p)\lambda' \right] \ell_{\theta,j} - \zeta_\theta (1 + \bar{r})(d_\theta - \ell_{\theta,j}),$$

where $\zeta_\theta \leq 1$ is the post-liquidation value of the bank's asset portfolio. Aggregate taxes are given by

$$\tau'_D(s) \cdot H = \sum_{\theta, d_\theta, j} \left[\int_{\lambda'} x(d_\theta, s_j) \max\left\{0, \Delta'(d_\theta, s_j)\right\} \mu_\theta(d_\theta) df(\lambda') \right].$$

2.6 Information

There is asymmetric information on the part of borrowers and lenders (banks, nonbanks, and households). Only borrowers know the riskiness of the project they choose (R_k) and their outside option (ω). Success or failure of their project is only observable after bearing the monitoring cost. To maintain consistency with payoffs between project choice and outside option, they receive a perfect unobservable signal about their outside option at the beginning of the period. Other information is observable.

2.7 Timing

In any period t, the timing of events is as follows:

1. At the beginning of the period

(a) Bank type θ and the mass of depositors that the bank is matched with d_θ are realized given household asset decisions. That determines the industry state (i.e., cross-sectional distribution $\boldsymbol{\mu}$).

(b) After observing ω, borrowers choose whether to invest in the risky technology or to choose their outside option $\iota \in \{0,1\}$ and, if so, they draw $\in$.

(c) Those borrowers who choose to undertake a project choose the type of lender $k \in \{\mathcal{B},\mathcal{N}\}$ and the level of technology R_k.

(d) Banks and the representative nonbank choose how many loans to extend. In addition, banks choose how many deposits to accept and how many securities to acquire, and nonbanks receive their equity injections from households.

(e) The loan market is cleared determining $r_j = \{r_{\mathcal{B},j}, r_{\mathcal{N},j}\}$

2. At the end of the period, z'_j is realized:

(a) Project returns for entrepreneurs are determined.

(b) The portfolio of performing and nonperforming loans is determined via project returns and p_j resulting in a realization of $\pi_\theta(d_\theta, s_j), \pi_n(d_n, \mathbf{s})$, and $\pi_{\mathcal{N}}(\mathbf{s})$.

(c) Bank exit x_θ and entry e choices are made.

(d) Bank investment I_θ is chosen together with dividend payments and equity injections. Dividends net of equity injections for the representative nonbank are also determined.

(e) Households pay taxes τ' to fund deposit insurance and consume.

3. Equilibrium

3.1 Entrepreneur Problem

Every period, given $\boldsymbol{r}_j = \{r_{\mathcal{B},j}, r_{\mathcal{N},j}\}$ and ω, entrepreneurs located in region j choose whether ($\iota = 1$) or not ($\iota = 0$) to operate their technology. Conditional on choosing $\iota = 1$, entrepreneurs observe $\epsilon = \{\epsilon_{\mathcal{B}}, \epsilon_{\mathcal{N}}\}$ and then choose which type of lender $K \in \{\mathcal{B}, \mathcal{N}\}$to borrow from and the scale of the technology to operate R_k to solve

$$\max_{\{\iota\}} (1-\iota)\cdot\omega + \iota \cdot \boldsymbol{E}_{\epsilon}\left[\Pi_E\left(\boldsymbol{r}_j, \epsilon\right)\right], \tag{16}$$

where the value of investing (conditional on ϵ) is

$$\Pi_E\left(\boldsymbol{r}_j, \epsilon\right) = \max_{\{K, R_K\}}\Big\{1_{\{K=\mathcal{B}\}}\boldsymbol{E}_{z_j'}\left[\pi_E\left(r_{\mathcal{B},j}, R_{\mathcal{B}}, z_j'\right) + \epsilon_{\mathcal{B}}\right] + 1_{\{K=\mathcal{N}\}}\boldsymbol{E}_{z_j'}\left[\pi_E\left(r_{\mathcal{N},j}, R_{\mathcal{N}}, z_j'\right) + \epsilon_{\mathcal{N}}\right]\Big\}, \tag{17}$$

where $1_{\{\cdot\}}$ is an indicator function that takes the value one if the argument $\{\cdot\}$ is true and zero otherwise, and

$$\pi_E\left(r_{K,j}, R_K, z_j'\right) = \begin{cases} \max\left\{0, z_j' R_K - r_{K,j}\right\} & \text{with prob } p_j\left(R_K, z_j'\right) \\ \max\left\{0, -\left(\lambda + r_{K,j}\right)\right\} & \text{with prob} 1 - p_j\left(R_K, z_j'\right) \end{cases}$$

The solution to (17) implies that the share of borrowers choosing a loan from a lender of type K in region j is

$$s_{K,j}\left(\mathbf{r}_j\right) = \frac{\exp\left(\alpha \boldsymbol{E}_{z_j'}\left[\pi_E\left(r_{K,j}, R_K, z_j'\right)\right]\right)}{\sum_{\hat{K}\in\{\mathcal{B},\mathcal{N}\}}\exp\left(\alpha \boldsymbol{E}_{z_j'}\left[\pi_E\left(r_{\hat{K},j}, R_{\hat{K}}, z_j'\right)\right]\right)}. \tag{18}$$

The expected value of taking out a loan in region j is[11]

$$V_{E,j}\left(\mathbf{r}_j\right) = \int \Pi_E\left(\boldsymbol{r}_j, \epsilon\right) dF_\epsilon\left(\epsilon;\alpha\right). \tag{19}$$

If the entrepreneur undertakes the project financed by lender type K, then an application of the envelope theorem implies

$$\frac{\partial E_{z_j'}\left[\pi_E\left(r_{K,j}, R_K, z_j'\right)\right]}{\partial r_{K,j}} = -E_{z_j'}\left[p_j\left(R_K, z_j'\right)\right] < 0. \tag{20}$$

Thus, participating borrowers (i.e., those who choose to run a project rather than take the outside option) are worse off the higher the interest rate on loans is.

This has implications for the aggregate demand for loans determined by the participation decision (i.e., $\omega \leq V_{E,j}(\mathbf{r}_j)$). In particular, the total demand for loans in region j is given by

$$L_j^d\left(\mathbf{r}_j\right) = \int_0^{\overline{\omega}} \mathbf{1}_{\left\{\omega \leq V_{E,j}\left(\mathbf{r}_j\right)\right\}} \, d\Omega(\omega). \tag{21}$$

Then loan demand for commercial banks in region j is given by

$$L_{\mathcal{B},j}^d\left(\mathbf{r}_j\right) = s_{\mathcal{B},j}\left(\mathbf{r}_j\right) L_j^d\left(\mathbf{r}_j\right). \tag{22}$$

In that case, everything else equal, (20) implies $\frac{\partial L_{\mathcal{B},j}^d\left(\mathbf{r}_j\right)}{\partial r_{\mathcal{B},j}} < 0$. That is, the bank loan-demand curve is downward sloping. Furthermore, bank market shares are decreasing in bank lending rates (i.e., $\frac{\partial s_{\mathcal{B},j}\left(\mathbf{r}_j\right)}{\partial r_{\mathcal{B},j}} < 0$) and aggregate loan demand decreases with an increase in bank lending rates (i.e., $\frac{\partial L_j^d\left(\mathbf{r}_j\right)}{\partial r_{\mathcal{B},j}} \leq 0$).

3.2 Incumbent Bank Problem

As in Ericson and Pakes (1995), we consider symmetric equilibrium in the sense that all banks in the same region and individual state d_θ are treated identically. Since a bank's individual state lies in a finite set, the industry state μ is a counting measure.

11. The expected value of taking out a loan has a convenient closed form: $\frac{\gamma_E}{\alpha} + \frac{1}{\alpha} \ln\left(\Sigma_k \exp\left(\alpha E_{z_j'}\left[\pi_E\left(r_{K,j}, R_K, z_j'\right)\right]\right)\right)$ where γ_E is Euler's constant.

After being exogenously matched with d_θ potential depositors and offering them a take-it-or-leave-it deposit-rate offer $r_{D,j}$, an incumbent bank of type θ chooses loans $\ell_{\theta,j}$ in order to maximize profits. Given the outside storage option for a household is $\bar{r}$, the bank deposit rate $r_{D,j} = \bar{r}$ in all regions. In this way, we are abstracting from important deposit-side competition in order to focus on the bank lending channel. After profits are realized, banks can choose to exit setting $x_\theta = 1$ or choose to remain $x_\theta = 0$. When choosing its loan supply a small or regional bank in region j solves

$$\ell_{\theta,j}(d_\theta,\mu) = \arg\max_{\ell_{\theta,j}+a_\theta=d_\theta} E_{z_j'}\left[\pi_\theta(d_\theta,s_j)\right]. \tag{23}$$

Similarly, when choosing its loan supply across regions, a national bank solves

$$\left\{\ell_{n,j}(d_n,\mu)\right\}_{j\in\{e,w\}} = \arg\max_{\left\{\ell_{n,j}\right\}_{j\in\{e,w\}}} E_{z_j',z_{-j}'}\left[\pi_n(d_n,s)\right] \tag{24}$$

subject to

$$\sum_i \ell_{n,j} + a_n = d_n. \tag{25}$$

Given that all banks have some degree of market power, a bank takes into account that its loan supply affects the loan interest rate in its region and that other banks will best respond to its loan supply. The first-order condition for a small or regional bank in problem (23) with respect to ℓ is

$$E_{z_j'}\Bigg[\underbrace{\left(p_j r_{\mathcal{B},j}(\mu) - (1-p_j)\lambda - \frac{dc_\theta}{d\ell_{\theta,j}}\right)}_{(+)\text{ or }(-)} + \ell_{\theta,j}\Bigg(\underbrace{p_j}_{(+)} + \underbrace{\frac{\partial p_j}{\partial R_{\mathcal{B}}}\frac{\partial R_{\mathcal{B}}}{\partial r_{\mathcal{B},j}(\mu)}(r_{\mathcal{B},j}(\mu)+\lambda)\Bigg)}_{(-)} \tag{26}$$

$$-\mathbf{1}_{\{(d_\theta-\ell_{\theta,j})\geq 0\}}r - \mathbf{1}_{\{(d_\theta-\ell_{\theta,j})\geq 0\}}\mathrm{r}^a(d_\theta-\ell_{\theta,j}) = 0,$$

$$\underbrace{\frac{dr_{\mathcal{B},j}(\mu)}{d\ell_{\theta,j}}}_{(-)}\Bigg]$$

where $p_j \equiv p_j(R_{\mathcal{B}}, z_j')$. The first bracket represents the marginal change in profits from extending an extra unit of loans. The second bracket corresponds to the marginal change in profits due to a bank's influence on the interest rate it faces. This term depends on the bank's market power. A change in interest rates also endogenously affects the fraction of delinquent loans faced by banks (i.e., the term $\frac{\partial p_j}{\partial R_{\mathcal{B}}}\frac{dR_{\mathcal{B}}}{dr_{\mathcal{B},j}} < 0$). Given limited liability, entrepreneurs take on more risk when their financing costs rise. The last two terms represent the marginal cost of uninsured external borrowing for the bank. When the bank accumulates securities ($1_{\{a_\theta = d_\theta - \ell_{\theta,j} \geq 0\}}$), the marginal cost is given by the opportunity cost of the loan (what the bank could receive from storage). When the bank uses external borrowing to extend loans beyond its deposit base ($1_{\{a_\theta = d_\theta - \ell_{\theta,j} < 0\}}$), the marginal cost is given by the cost of external funds. A similar condition holds for the national bank.

Changes in the loan interest rate (i.e., $\frac{dr_{\mathcal{B},j}}{d\ell_{\theta,j}}$) in (26) are derived from the market clearing condition $L^d_{\mathcal{B},j}(\mathbf{r}_j) = L^s_{\mathcal{B},j}(\mu)$, where $L^s_{\mathcal{B},j}(\mathbf{r}_j)$ is given above in (22) and $L^s_{\mathcal{B},j}(\mu)$ denotes the total supply of loans given by

$$L^s_{\mathcal{B},j}(\mu) = \sum_{\theta, d_\theta} \ell_{\theta,j}(d_\theta, \mu)\mu_{\theta,j}(d_\theta) \tag{27}$$

For a given bank distribution μ, changes in the loan supply $\ell_{\theta,j}$ of a given bank have a direct effect on the aggregate loan supply but also an indirect effect via changes in the response of its competitors.

After loans have been extended, the value of an incumbent state or regional bank $\theta \in \{s, r\}$ in region j at the exit stage 2c is

$$V_\theta(d_\theta, s_j) = \max_{x \in \{0,1\}} \left\{ V^{x=0}(d_\theta, s_j), V^{x=1}(d_\theta, s_j) \right\} \tag{28}$$

where $s_j = (\mu, z'_j)$, $V^{x=1}(d_\theta, s_j)$ is defined in equation (11) and

$$\begin{aligned} V^{x=0}(d_\theta, s_j) = \max \Big\{ & \pi_\theta(d_\theta, s_j) - I - 1_{\{\ell_\theta > 0\}} \cdot \varsigma_\theta(\ell_\theta) \\ & + \beta\rho^x_\theta E_{\theta', d'_\theta, s'_j | d_\theta, s_j}\left[V(d'_\theta, s'_j)\right] + \beta(1-\rho^x) E_{\theta', d'_\theta, s'_j | d_\theta, s_j}\left[V^{x=1}(d'_\theta, s'_j)\right] \Big\} \end{aligned} \tag{29}$$

subject to

$$\ell_\theta = \max\{I - \pi_\theta, 0\} \tag{30}$$

and the transition functions $T(\theta' \mid \theta, I)$ and $\mu' = \mathcal{H}(\mu, N_e)$. A similar problem holds for the national bank with $\pi_n(d_n, \boldsymbol{s})$ substituting for $\pi_\theta(d_\theta, s_j)$ in (29).

3.3 Bank Entry

The value of an entrant in region j net of entry costs in the industry state μ is

$$V_{e,j}(\mu) = -\kappa + \beta E\left[V_s\left(d_s', s_j'\right)\right]. \tag{31}$$

Recall that entrants do not operate in the period they enter and, consistent with the data, we assume they all start small (i.e., with $\theta = s$). Potential entrants will decide to enter if $V_{e,j}(\mu) \geq 0$. The number of entrants $N_{e,j}$ is determined endogenously in equilibrium. Free entry implies that

$$V_{e,j}(\mu) \times N_{e,j} = 0. \tag{32}$$

That is, in equilibrium, either the value of entry is zero, the number of entrants is zero, or both. The total value of entrants is given by $N_e = \sum_j N_{e,j}$.

3.4 Evolution of the Cross-Sectional Bank Size Distribution

The distribution of banks evolves according to $\mu' = \mathcal{H}(\mu, N_e)$, where each component is given by:

$$\begin{aligned} \mu'_{\theta'}(d'_{\theta}) = & \sum_{\theta \in \{\ell, s\}, j \in \{e,w\}, d_\theta \in D_\theta} \left(1 - x(d_\theta, s_j)\right)\left(1 - \rho^x_\theta\right) T\left(\theta' \mid \theta, I(d_\theta, s_j)\right) G_\theta\left(d'_\theta, d_\theta\right) \mu_\theta(d_\theta) \\ & + \sum_{d_n \in D_n} \left(1 - x(d_n, \boldsymbol{s})\right)\left(1 - \rho^x_n\right) T\left(\theta' \mid n, I(d_n, \boldsymbol{s})\right) G_n\left(d'_{\theta'}, d_n\right) \mu_n(d_n) \\ & + N_{e,j} \sum_{j, d_s} G_s(d_s) \end{aligned} \tag{33}$$

where $\overline{G}_s(d_s)$ is the distribution from which deposits for entrants are drawn. Equation (33) makes clear how the law of motion for the distribution of banks is affected by entry (N_e) and exit (X) decisions as well as the bank size investment decision (I).

3.5 Nonbank Problem

The representative nonbank operates in a competitive industry, so when making lending decisions, it takes the loan interest rate $r_{\mathcal{N},j}$ as given. Taking into account that $\beta(1+\overline{r}) = 1$, the first-order condition of the nonbank with respect to $\ell_{\mathcal{N},j}$ is given by

$$\overline{r} = E_{z_j'}\left[p\left(R_{\mathcal{N}}\left(r_{\mathcal{N},j}\right),z_j'\right)-\left(1-p\left(R_{\mathcal{N}}\left(r_{\mathcal{N},j}\right),z_j'\right)\right)\lambda\right]-c_{\mathcal{N}}, \tag{34}$$

where $R_{\mathcal{N},j}(r_{\mathcal{N}})$ is the optimal choice of technology by the entrepreneur in region j when taking a loan from a nonbank facing interest rate $r_{\mathcal{N},j}$. Equation (34) is one equation in one unknown which pins down the interest rate $r_{\mathcal{N},j}$ of the nonbank sector.[12] Evaluating the nonbank loan demand at this price we can determine the level of lending of the nonbank. Equation (34) also makes clear that the expected net return between a bank deposit and nonbank investment is equalized, with the spread depending on $c_{\mathcal{N}}$. However, while bank deposits guarantee a risk-free return (since there is deposit insurance), equity injections in a nonbank are subject to regional risk.

3.6 Definition of Equilibrium

A pure strategy Markov perfect industry equilibrium (MPIE) is:[13]

1. $\{\iota_j, K_j, R_{k,j}\}$ are consistent with entrepreneur optimization inducing an aggregate loan-demand function $L_j^d(\mathbf{r}_j)$.

2. $\{\ell_{\theta,j}, I_{\theta,j}, x_{\theta,j}, e_{\theta,j}, V_\theta\}$ are consistent with bank optimization inducing an aggregate loan-supply function $L_{\mathcal{B},j}^s$.

3. Free entry is satisfied.

4. The law of motion for the industry state $\boldsymbol{\mu}' = \mathcal{H}(\boldsymbol{\mu}, \{N_j^e\}_j)$ induces a sequence of cross-sectional distributions that are consistent with entry, exit, and investment decision rules.

5. The vector of interest rate $r_j(\boldsymbol{\mu})$ is such that the loan market clears.

6. Stock prices are consistent with bank valuation V_θ.

7. Taxes $\tau_D'(s)$ cover the cost of deposit insurance.

12. The fact that $r_{\mathcal{N},j}$ is independent of the entire distribution of banks is a form of block recursivity as in Menzio and Shi (2010).

13. See Corbae and D'Erasmo (2023) for the statement of the household problem.

4. Parameterization

When solving for the model moments, note that despite shocks d_θ which are i.i.d. across banks, the fact that banks are not of measure zero induces aggregate uncertainty.

Thus, we use the computational methods in Ifrach and Weintraub (2017).[14] In particular, the approximation methods allow for there to be strategically important (dominant) banks. We think of rising concentration as occurring between two long-run stochastic equilibria (one coinciding with pre-Riegle-Neal and one coinciding with the period following the Great Recession) that lead to different dynamics for the bank distribution following a decline in branching costs.

A model period is one year. Our main source for bank level variables (and aggregates derived from them) are the Call Reports.[15] We aggregate commercial bank level information to the bank holding company level. As discussed above, moments from the Call Report data are computed beginning in 1984, due to an overhaul of the data in that year.

Given that prior to the passing of the Riegle-Neal Interstate Banking and Branching Efficiency Act in 1994 there was only one type of bank in operation ($\mathscr{s}$ small banks), we calibrate the model to the post-reform period 2009–2018, where restrictions on opening bank branches across state lines were not in place. In the latter period, banks of three types operate $\theta \in \{\mathscr{s}, \mathscr{r}, \mathscr{n}\}$, which allows us to obtain estimates for the investment transition matrix. We focus on the top 2 percent of banks when sorted by assets and identify those in the top 4 with those with $\theta = \mathscr{n}$, those in the top 5 - 35 with $\theta = \mathscr{r}$, and those in the top 36 to 2 percent with $\theta = \mathscr{s}$.

We parameterize the stochastic process for the borrower's project as follows. For each borrower, let $y^e = a - bR + \varepsilon_e$, where ε_e is iid (across agents and time) and drawn from $N(z_j', \sigma_\varepsilon^2)$. We define success to be

14. Appendix A.1 describes the solution algorithm we use to approximate a Markov-Perfect Equilibrium.

15. Source: FDIC, Call and Thrift Financial Reports, balance sheet, and income statement items. See Appendix A.2 for a description of the data.

the event that $y^e > 0$, so in states with higher ε_e success is more likely. Then

$$\begin{aligned} p\left(R, z_j^{\prime}\right) &= 1 - \Pr\left(y^e \leq 0 \# R, z_j^{\prime}\right) \\ &= 1 - \Pr\left(\varepsilon_e \leq -a + bR\right) \\ &= \Phi_{z_j^{\prime}}\left(a - bR\right), \end{aligned}$$

where $\Phi_{z_j^{\prime}}(x)$ is a normal cumulative distribution function with mean $z_j^{\prime}$ and variance σ_{ε}^2. The stochastic process for the borrower outside option, $\Omega(\omega)$, is simply taken to be the uniform distribution $[0,\overline{\omega}]$.

We assume that the regional shock is distributed iid normal with zero mean and variance σ_z^2. We set the cross-correlation between regions ρ_z to 0.01. We discretize the regional shock and let z_j take two values $z_{j,t} \in \mathcal{Z}_j = \{z_L, z_H\}$ with $z_H > 0$ and $z_L = -z_H$.

To reduce the number of parameters to calibrate, and as we do not have enough information on the liquidation value of the assets of large banks since we do not observe liquidations in the largest category, we set $\zeta_\theta = \zeta$ and calibrate ζ using data from the FDIC. We parameterize the equity-issuance cost function $\varsigma_\theta(e_\theta) = (\varsigma_\theta^0 + \varsigma_\theta^1 e_\theta)$ where $e_\theta = \max\{0, -(\pi_\theta - \iota_\theta)\}$ and the cost of extending loans $c_\theta(\ell_\theta) = c_\theta^0 \ell_\theta + c_\theta^1 \ell_\theta^2$. We assume that the total cost of external borrowing is $r^a(a_\theta) = r_0^a a_\theta + r_1^a a_\theta^2$ when $1_{\{a_\theta<0\}}$.

As part of the calibration exercise, post reform, we estimate transition probabilities between banks of different sizes. In particular, we estimate transition matrices by counting the number of banks in each bin-year and dividing by the total number of banks of each type in a given year. We then take the time-series average of the corresponding bin for each period. For example, to compute the fraction of banks that remain in state $\mathfrak{s}$, we first count how many $\mathfrak{s}$ banks in period t are still of type $\mathfrak{s}$ in period $t + 1$. Let this number be $N_t^{\mathfrak{s},\mathfrak{s}}$. Then, evaluated at the equilibrium level of investment $I_\theta(d_\theta, s_j)$, the value in $T_t(\theta' \mid \theta, I_\theta(d_\theta, s_j))$ equals $\frac{N_t^{\mathfrak{s},\mathfrak{s}}}{N_{\mathfrak{s},t}}$ where $N_{\mathfrak{s},t}$ corresponds to all banks of type s in period t. The reported value in table 6 corresponds to the time average of $T_t(\theta' \mid \theta, I_\theta(d_\theta, s_j))$. The failure state incorporates the transition to a bank outside the top 2 percent .

We calibrate the entry cost κ by choosing the average number of entrants in each region $N_{e,j}$ that results in an average number of banks equal to 103 (the average number of banks in the top 2 percent post reform). We assume that the number of potential entrants is the same across regions. In the experiments that follow, κ is kept constant to this value and $N_{e,j}$ adjust to satisfy the equilibrium conditions. In addition, we set the origination cost for nonbanks to match the fraction of bank lending to total credit.

Table 4 presents the parameters of the model and the targets that were used. We use several moments from our panel of banks in the U.S. and the estimates of the deposit process presented in table 1. Entries above the line correspond to parameters chosen outside the model, while entries below the line correspond to parameters chosen within the model by simulated method of moments. Table 5 presents a set of data moments together with their model-generated counterparts for the post-reform period (i.e., the period used in the calibration). Moments above the line correspond to those used in the calibration procedure and those below the line are untargeted moments. In all we have 26 parameters and 26 targeted moments. Given that there is symmetry in the underlying stochastic processes and parameter values in the model, the two regions yield similar long-run averages in the tables.

Table 4. Parameters and Targets

Parameter		*Value*	*Target*
Deposit Interest Rate (%)	r^D	0.005	Avg Interest Expense Deposits
Mean Charge-off Rate	μ_λ	0.314	Avg Charge-off Rate
Exit Value Recovery	ζ	0.804	Recovery Value Bank Failures (FDIC)
Bank Discount Factor	β	0.995	$1/(1+\overline{r})$
Correlation Regional Shocks	$\rho_{\mathcal{Z}}$	0.01	regional correlation of default frequency
Measure Borrowers	B	320.0	Bank Loans to Output Ratio
Borrower Success Prob. Function	a	4.291	Avg. Borrower Return
Borrower Success Prob. Function	b	28.94	Avg. Default Frequency
Borrower Success Prob. Function	σ_e	0.107	Avg. Loan Interest Rate
Outside Option	$\overline{\omega}$	0.462	Elasticity of Loan Demand
Std. Dev Reg Shocks	σ_z	0.020	Std Dev Loan Returns
Linear Cost Loans $\mathscr{s}$	$c^0_{\mathscr{s}}$	0.001	Avg Net Mg Expense $\mathscr{s}$
Quadratic Cost Loans $\mathscr{s}$	$c^1_{\mathscr{s}}$	0.025	Elasticity Mg Expense $\mathscr{s}$
Fixed Operating Cost $\mathscr{s}$	$C_{F,\mathscr{s}}$	0.001	Fixed Cost / Loans $\mathscr{s}$
Linear Cost Loans μ	$c^0_{\mathscr{r}}$	0.001	Avg Net Mg Expense $\mathscr{r}$
Quadratic Cost Loans μ	$c^1_{\mathscr{r}}$	0.003	Elasticity Mg Expense $\mathscr{r}$
Fixed Operating Cost $\mathscr{r}$	$C_{F,\mathscr{r}}$	0.005	Fixed Cost / Loans $\mathscr{r}$
Linear Cost Loans $\mathscr{n}$	$c^0_{\mathscr{n}}$	0.003	Avg Net Mg Expense $\mathscr{n}$
Quadratic Cost Loans $\mathscr{n}$	$c^1_{\mathscr{n}}$	0.001	Elasticity Mg Expense $\mathscr{n}$
Fixed Operating Cost $\mathscr{n}$	$C_{F,\mathscr{n}}$	0.010	Fixed Cost / Loans $\mathscr{n}$
Proportional Cost Loans $\mathcal{N}$	$C_{\mathcal{N}}$	0.034	Share Bank Loans / Total Loans
Transition Probability Function	α	100.00	Loan Market Share $\mathscr{s}$
Transition Probability Function	δ	0.600	Fraction of Banks $\mathscr{s}$
Transition Probability Function	ξ	0.850	Transition $\mathscr{s}$ to $\mathscr{r}$
Fixed Equity-Issuance Costs $\mathscr{s}$	$\varsigma^0_{\mathscr{s}}$	0.001	Avg Equity Issuance $\mathscr{s}$
Proportional Equity-Issuance Costs $\mathscr{s}$	$\varsigma^1_{\mathscr{s}}$	0.050	Fract $\mathscr{s}$ Banks Issue Equity
Fixed Equity-Issuance Costs $\mathscr{s}$	$\varsigma^0_{\mathscr{r}}$	0.005	Avg Equity-Issuance *mathcal* $\mathscr{r}$
Proportional Equity-Issuance Costs μ	$\varsigma^1_{\mathscr{r}}$	0.025	Fract $\mathscr{r}$ Banks Issue Equity
Fixed Equity-Issuance Costs $\mathscr{n}$	$\varsigma^0_{\mathscr{n}}$	0.020	Avg Equity Issuance $\mathscr{n}$
Proportional Equity-Issuance Costs *n*	$\varsigma^1_{\mathscr{n}}$	0.010	Fract $\mathscr{n}$ Banks Issue Equity
Entry Cost	k	0.083	Total Number of Banks
Borrowing Cost Function	r^a_0	0.012	Spread Fed Funds to Deposit Cost
Borrowing Cost Function	r^a_1	0.008	Fed Funds / Assets

Source: Author's calculations.

Note: The entry cost is set as part of the equilibrium selection. In particular, in the baseline case, the entry cost is the one that satisfies the zero-entry condition for the value of entrants $N_{e,j} = 0.75$ (on average) that provides the best fit of the model. This entry cost is kept constant when running the experiments presented in section 5.

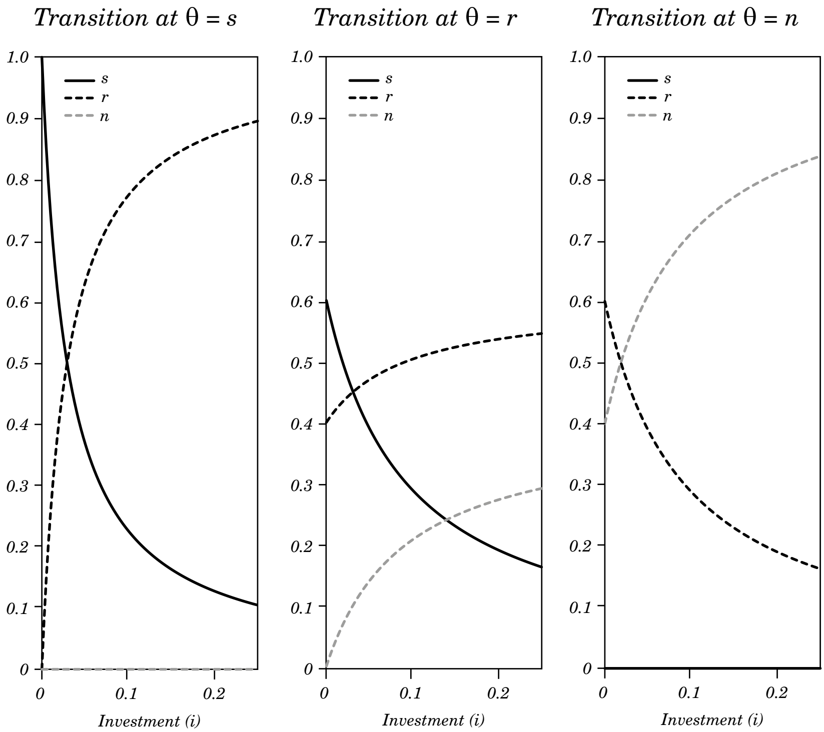
Table 5. Data & Model Moments Post Reform

Moments (%)	*Data*	*Model*
Charge-off Rate	0.685	0.39
Std Dev Charge-off Rate	0.20	0.20
Avg. Borrower Return	12.94	13.81
Avg. Default Frequency	2.09	1.25
Loan Interest Rate	2.971	2.59
Elasticity of Loan Demand	-1.1	-1.17
Avg. Net Mg Expense s	1.402	1.06
Elasticity Mg Expense s	0.875	1.83
Fixed Cost / Loans s	0.444	0.52
Avg. Net Mg Expense r	0.904	0.66
Elasticity Mg Expense r	0.940	1.69
Fixed Cost / Loans r	0.583	0.46
Avg. Net Mg Expense n	0.228	0.36
Elasticity Mg Expense n	1.05	1.08
Fixed Cost / Loans n	0.585	0.10
Loan Market Share s	16.04	20.58
Fraction of Banks s	69.35	70.80
Avg. Equity Issuance s	0.044	0.01
Fract s Banks Issue Equity	6.86	32.99
Avg. Equity Issuance r	0.04	0.00
Fract r Banks Issue Equity	4.23	0.00
Avg. Equity Issuance n	0.004	0.00
Fract n Banks Issue Equity	2.17	0.00
Bank Loans to Output Ratio	60.34	78.96
Share Bank Loans / Total Loans	50.00	78.64
Transition s to r	2.10	20.91
Spread Fed Funds to Deposit Cost	0.65	0.65
Fed Funds / Assets	2.16	15.29
Total Number of Banks	103	194.66
Exit (Failure) Rate	3.93	0.77
Deposit to Output Ratio	57.78	67.34
Markup	74.33	106.61
Avg. Net Interest Margin	4.18	2.06
Avg. Cost s	1.85	1.58
Avg. Cost r	1.487	1.12
Avg. Cost n	0.813	0.45
Fraction of Banks r	27.15	26.92
Fraction of Banks n	3.5	2.28
Loan Market Share r	50.22	44.25
Loan Market Share n	33.74	35.18
Number of Banks s	68	137.81
Number of Banks r	31	52.41
Number of Banks n	4	4.44
Deposit Market Share s	12.94	20.87
Deposit Market Share r	36.84	39.29
Deposit Market Share n	50.22	39.83

Note: Moments above the line correspond to calibration targets.

Figure 4 presents the estimated transition probability function for a given level of investment. The figure illustrates that the probability of growing (shrinking) is increasing (decreasing) in bank investment. Table 6 provides the resulting transition matrix across types evaluated at the equilibrium level of investment $I_\theta(d_\theta, s_j)$.[16] The table illustrates that failure rates in the model are decreasing in size as in the data. Further, it illustrates that size is generally persistent.

Figure 4. Calibrated Transition Probabilities (Post Reform)



16. The initial year of the post-reform period used to estimate this matrix differs from that presented in table 5. The small number of bank types we consider prevents us from using the 2009–2018 period and obtain a meaningful transition matrix.

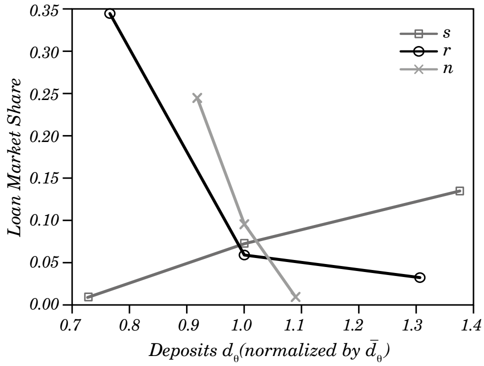
Table 6. Bank-Type Transition Matrix $T(\theta' \mid \theta, I_\theta (d_\theta, s_j))$

	Data: Post - Reform Period (1994-2019)				*Model: Post - Reform Period (1994-2019)*			
	$\theta' = s$	$\theta' = r$	$\theta' = n$	Failure	$\theta' = s$	$\theta' = r$	$\theta' = n$	Failure
Entrant	1.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
$\theta = s$	0.92	0.02	0.00	0.05	0.78	0.21	0.00	0.01
$\theta = r$	0.03	0.97	0.01	0.00	0.55	0.42	0.03	0.00
$\theta = n$	0.00	0.02	0.98	0.00	0.00	0.40	0.60	0.00

Note: We study banks in the top two percent of the asset distribution. We consider the top 4 (θ_n), the top 5 - 35 (θ_r), and the rest.

Figure 5 presents the distribution of banks for the equilibrium post reforms. The ranking of the variance of deposit inflows reflecting geographic diversification of funding shocks (i.e., highest variance for state banks and lowest for national banks) from table 1 is evident in the support of the distribution in figure 5. The market shares by bank size reflect the number of banks and the loan decisions (conditional on size). In an equilibrium where all banks extend loans equal to the amount of deposits they take, by construction, the shape of the loan distribution would derive from the shape of the deposit distribution. As deposit inflows are normally distributed, loan market shares would be distributed normal as well. Transitions from s to r and from r to n increase the number of banks with the lowest deposit value conditional on being of type r or n, as it is more likely to start with the lowest value of deposits when transitioning upwards. In addition, it is the case that in this equilibrium, most banks of type s and r extend less loans than the deposits they take (which is particularly important for the highest level of deposits), thus reducing the market share of this class of banks even further.

Figure 5. Distribution Banks Post Reform



Source: Author's calculations.

5. The Bank Lending Channel

The bank lending channel of monetary policy suggests that banks play a special role in the transmission of monetary policy. The channel works through how monetary policy effects on the cost of external funding. The corporate finance approach to the bank lending channel, as elucidated in Kashyap and Stein (1995) and Kashyap and Stein (2000), posits that larger banks are less sensitive to increases in fed funds rates since they have easier access to external funding. Thus, bigger banks lower their loan supply less than smaller banks in response to a rise in external funding costs like fed funds.

5.1 Model Mechanism

Here we describe how the bank lending channel works in the context of our model. There are three sources of external funding in our model: insured deposits d_θ at rate $\bar{r}$, fed borrowing $a_\theta = d_\theta - \ell_\theta < 0$ at rate $r^a > \bar{r}$, and equity e_θ at cost $\varsigma_\theta(e_\theta)$. Bank-type heterogeneity affects the sources of external funding in several ways.

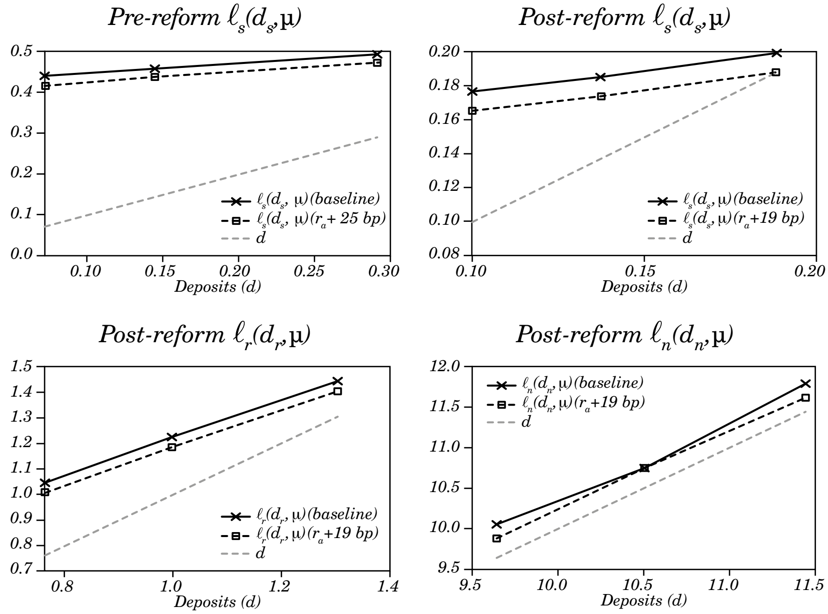
First, the type $\theta \in \{s, r, n\}$ dependent deposit process $G_\theta(d'_\theta, d_\theta)$ provides a bank with a cheap source of funds. In particular, while we assume that banks of different types face the same insured borrowing cost $\bar{r}$, its mean deposit base is increasing in size and variance is decreasing in size as in the data in table 1. This means that bigger banks have a larger source of FDIC-subsidized external funding. In our model, banks always use this cheap source of external finance first (as in a pecking order theory) when making loans, i.e., when solving (23) and (24).

Second, bank investment in type growth $I_\theta(d_\theta,s_j)$ is a function of profitability $\pi_\theta(d_\theta,s_j)$ (which is increasing in deposit base) and its type-dependent equity-issuance costs $\varsigma_\theta(e_\theta)$ in (29). There is a pecking order here as well—banks first use internal funds π_θ and then issue equity when $I > \pi_\theta$. Type $\mathfrak{s}$ banks with a high realization of deposits in a given state s_j (and hence high profits $\pi_{\mathfrak{s}}(d_{\mathfrak{s}}, s_j)$) are more likely to grow to regional types $\mathfrak{r}$. Since bigger banks face lower proportional equity-issuance costs in $\varsigma_{\mathfrak{s}}(e) > \varsigma_{\mathfrak{r}}(e) > \varsigma_{\mathfrak{n}}(e)$ as in table 4, bigger banks are more likely to stay big, enjoying low-cost deposit funding.

These two model elements provide the mechanism by which a rise in the fed funds rate may translate into more sensitivity by small banks than large banks as in Kashyap and Stein. In particular, fed fund borrowing in our model $a_\theta = d_\theta - \ell_\theta < 0$ is decreasing in d_θ for a given amount of loans ℓ_θ extended. Thus, banks with a small amount of low-cost $\bar{r}$ deposits are more exposed to borrowing at the higher fed funds rate $r^a > \bar{r}$. Thus, the cost of supplying a given amount of loans, which solves the first-order condition (26), is decreasing in deposits d_θ; since bigger banks enjoy a bigger deposit base, they are less sensitive to a rise in external borrowing via fed funds.

We illustrate this by graphing loan decision rules $\ell_\theta(d_\theta, \mu)$ for a 25 basis point rise in fed funds in the pre-reform equilibrium (top panel) and for a 19 basis point (same proportional) rise in the post-reform equilibrium (bottom panels) in figure 6. There are several things to note: (i) the difference between the decision rule ℓ and the 45-degree line illustrates fed fund borrowing, (ii) higher deposit bases lower external funding via fed funds, (iii) a rise in fed funds rates lowers loan supply more for small banks than bigger ones.

Figure 6. Lending Channel: Pre & Post-Reform Loan Decision Rules



Source: Author's calculations.

5.2 Simulation Results

We implement this policy experiment by raising the external funding cost r^a by 25 basis points in the pre-Riegle-Neal equilibrium of our model (from $r^a = 0.0125$ to $r^a = 0.015$) as well as in the post-reform, where the policy rate goes from $r^a = 0.0115$ to $r^a = 0.0134$ (the same proportional change). We evaluate the effect of this policy in the short run ($T = 1$, the response on impact, and $T = 2$, the average over two years). The starting point for the simulation is the long-run average distribution.[17]

Referring back to the first-order condition for loan choice in (26), conditional on a bank borrowing (i.e., $\mathbf{1}_{\{a_\theta<0\}}$), a rise in external funding

17. Since we focus on the short-run effects of this policy, we assume that banks do not update their beliefs about the industry state and expect to compete, in the short run, against the long-run industry from before the policy change. In Corbae and D'Erasmo (2021) we allow beliefs to be updated stochastically from the initial set of beliefs to the new long-run set of beliefs consistent with the policy change.

costs through a rise in the fed fund rate raises the marginal cost of making loans and hence lowers the 'intensive' margin of bank loans $\ell_{\theta,j}(d_\theta, \mu)$. As evident in table 7, the rise in r^a leads to lower average loans both pre and post reform. Importantly, we see larger banks are less sensitive (i.e., decrease loan supply less in response) to the rise in funding costs than smaller banks, as in Kashyap and Stein. While average loans (the intensive margin) drop, there can be changes in the distribution (the 'extensive' margin) that can induce interesting general equilibrium effects. Specifically, while aggregate loans decrease in the short run ($T = 1$) both pre and post reform, given the large post-reform interest margin at $T = 1$, we see large growth in the number of regional banks out of small. Such extensive margin changes can change the dynamics of aggregate loans and interest rates.

Table 7. Bank Lending Channel Pre & Post Reform

	Pre - Reform			*Post - Reform*		
		$r_0^a + 0.25\%$			$r_0^a + 0.19\%$	
	Baseline	T=1Δ%	T=2Δ%	Baseline	T=1Δ%	T=2Δ%
Avg. Def Freq.	1.58	-13.24	0.47	1.25	-12.84	-1.67
Loan Int. Rate	4.73	2.26	2.29	2.59	8.67	-3.44
Interest Margin	3.41	3.39	3.12	2.06	11.00	-4.25
Bank Loan Supply	82.43	-4.83	-4.90	129.06	-6.15	2.11
Bank Loans to Output	59.46	-4.42	-4.33	78.96	-4.89	1.37
Bank Loans / Total Loans	59.16	-4.27	-4.33	78.64	-4.75	1.38
Avg. Loans s	0.46	-4.59	-4.59	0.19	-5.10	-5.07
Avg. Loans r	-	-	-	1.09	-3.26	-3.95
Avg. Loans n	-	-	-	10.26	-1.09	-1.13
Loan Mkt Share s	100.00	0.00	0.00	20.52	2.21	-13.79
Loan Mkt Share r	-	-	-	44.19	-1.89	9.97
Loan Mkt Share n	-	-	-	35.28	1.09	-4.46
Total Number of Banks	178.09	-0.26	-0.26	194.66	-0.63	-0.62
N s Banks	178.09	-0.26	-0.26	137.81	1.08	-7.27
N r Banks	-	-	-	52.41	-4.82	16.91
N n Banks	-	-	-	4.44	-4.10	-1.15

Source: Author's calculations.
Note: Pre-reform $r_0^a = 1.50\%$ and Post-reform $r_0^a = 1.15\%$. The increase in the policy rate corresponds to an increase of 25 bp of in the pre-reform (a 33% increase). T denotes the number of periods used to compute reported averages.

Table 8. Bank Lending Channel Post Reform (Baseline vs No Investment Update)

		Post - Reform	
		r_0^a + 0.19%	
	Baseline	Equilibrium T = 2Δ%	Partial Equilibrium T = 2Δ%
Avg. Def Freq.	1.25	-1.67	0.88
Loan Int. Rate	2.59	-3.44	6.78
Interest Margin	2.06	-4.25	8.42
Bank Loan Supply	129.06	2.11	-4.77
Bank Loans to Output	78.96	1.37	-3.66
Bank Loans / Total Loans	78.64	1.38	-3.66
Avg. Loans s	0.19	-5.07	-5.13
Avg. Loans r	1.09	-3.95	-3.38
Avg. Loans n	10.26	-1.13	-1.12
Loan Mkt Share s	20.52	-13.79	-0.51
Loan Mkt Share r	44.19	9.97	-0.31
Loan Mkt Share n	35.28	-4.46	0.69
Total Number of Banks	194.66	-0.62	-0.63
N s Banks	137.81	-7.27	-0.13
N r Banks	52.41	16.91	-1.75
N n Banks	4.44	-1.15	-3.03

Source: Author's calculations.
Note: Pre-reform r_0^a = 1.50% and Post-reform r_0^a = 1.15%. The increase in the policy rate corresponds to an increase of 25 bp in the pre-reform (a 33% increase). T denotes the number of periods used to compute reported averages.

To attempt to decompose intensive versus extensive margin effects associated with monetary policy, table 8 shows how changes in fed funds rates affect the incentive to grow. There, we perform a counterfactual in the last column, where we compute changes two years after the increase in r^a, where we use investment decision rules from our baseline in the simulation to compare to equilibrium changes in the center column, i.e., just those from the last column of table 7.[18] Importantly, when controlling for the extensive growth margin, bank aggregate loan supply falls by nearly five percent in the counterfactual at T = 2, while aggregate bank loan supply increases by over two

18. Since investment comes at the end of any period, investment outcomes for are chosen prior to the shock so nothing is changed in for this counterfactual.

percent when using the equilibrium investment rules. This arises from growth in the number of regional banks despite the fact that average loans for each type fall. Thus, monetary policy has an impact on aggregate lending not only through the intensive margin but also via its effects on the composition of the banking industry.

6. Conclusion

In this paper we examine the consequences of geographic expansion and rising bank concentration for the bank lending channel. The model is consistent with smaller banks being more sensitive to monetary policy contractions, compatible with a corporate finance approach to banking as elucidated in Kashyap and Stein (2000). The model makes clear that monetary policy can also affect growth dynamics in the banking industry with implications for competition. Corbae and D'Erasmo (2023) provide a more general analysis of the data and implications of the model for financial stability and monetary policy.

References

Aguirregabiria, V., R. Clark, and H. Wang. 2020. "The Geographic Flow of Bank Funding and Access to Credit: Branch Networks, Local Synergies, and Competition." CEPR Discussion Paper No. DP13741. CEPR Press, Paris & London.

Arellano, M. and S. Bond. 1991. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies* 58(2): 277–97.

Bellifemine, M., R.Jamilov, and T. Monacelli. 2022. "Hbank: Monetary Policy with Heterogeneous Banks." Source? CEPR Discussion Paper No. 17129. CEPR Press, Paris & London.

Berger, A.N. and T.H. Hannan. 1998. "The Efficiency Cost of Market Power in the Banking Industry: A Test of the 'Quiet Life' and Related Hypotheses." *Review of Economics and Statistics* 80(3): 454–65.

Berger, A.N., L.F. Klapper, and R. Turk-Ariss. 2009. "Bank Competition and Financial Stability." *Journal of Financial Services Research* 35(2): 99–118.

Berger, A.N. and L.J. Mester. 1997. "Inside the Black Box: What Explains Differences in the Efficiencies of Financial Institutions?" *Journal of Banking and Finance* 21(7): 895–947.

Bernanke, B. and M. Gertler. 1989. "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review* 79(1): 14–31.

Besanko, D. and U. Doraszelski. 2004. "Capacity Dynamics and Endogenous Asymmetries in Firm Size." *RAND Journal of Economics*: 23–49.

Buchak, G., G. Matvos, T. Piskorski, and A. Seru. 2018. "Fintech, Regulatory Arbitrage, and the Rise of Shadow Banks." *Journal of Financial Economics* 130: 453–83.

Carlstrom, C. and T. Fuerst. 1997. "Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis." *American Economic Review* 87(5): 893–910.

Corbae, D. and P. D'Erasmo. 2020. "Rising Bank Concentration." *Journal of Economic Dynamics and Control*. Volume 115.

Corbae, Dean and Pablo D'Erasmo. 2021. "Capital Buffers in a Quantitative Model of Banking Industry Dynamics." Econometrica 89:2975-3023.

Corbae, D. and P. D'Erasmo. 2023. "A Quantitative Model of Banking Industry Dynamics across Time and Space." Manuscript.

Den Haan, W.J., S.W. Sumner, and G.M. Yamashiro. 2007. "Bank Loan Portfolios and the Monetary Transmission Mechanism." *Journal of Monetary Economics* 54(3): 904–924.

Diamond, D.W. 1984. "Financial Intermediation and Delegated Monitoring." *Review of Economic Studies* 51(3): 393–414.

Ericson, R. and A. Pakes. 1995. "Markov Perfect Industry Dynamics: A Framework for Empirical Work." *Review of Economic Studies* 62(1): 53–82.

Farias, V., D. Saure, and G.Y. Weintraub. 2012. "An Approximate Dynamic Programming Approach to Solving Dynamic Oligopoly Models." *The RAND Journal of Economics* 43(2): 253–82.

Gelman, M., I. Goldstein, and A. MacKinlay. 2022. "Bank Diversification and Lending Resiliency." Available at SSRN 4147790.

Ifrach, B. and G.Y. Weintraub. 2017. "A Framework for Dynamic Oligopoly in Concentrated Industries." *Review of Economic Studies* 84(3): 1106–50.

Kashyap, A.K. and J.C. Stein. 1995. "The Impact of Monetary Policy on Bank Balance Sheets." In *Carnegie-Rochester Conference Series on Public Policy* 42: 151–95.

Kashyap, A.K. and J.C. Stein. 2000. "What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?" *American Economic Review* 90(3): 407–28.

Liang, N. and S.A. Rhoades. 1988. "Geographic Diversification and Risk in Banking." *Journal of Economics and Business* 40(4): 271–84.

Menzio, G. and S. Shi. 2010. "Block Recursive Equilibria for Stochastic Models of Search on the Job." *Journal of Economic Theory* 145(4): 1453–94.

Morelli, J., M. Moretti, and V. Venkateswaran. 2023. "Geographical Expansion in U.S. Banking: A Structural Evaluation." Manuscript.

Shin, Y. 2022. "Deposit Market Competition and Equity Capital Issuance." Manuscript.

Wang, Y., T.M. Whited, Y. Wu, and K. Xiao. 2022. "Bank Market Power and Monetary Policy Transmission: Evidence from a Structural Estimation." *Journal of Finance* 77(4): 2093–141.

Weintraub, G.Y., C.L. Benkard, and B. Van Roy. 2008. "Markov Perfect Industry Dynamics with Many Firms." *Econometrica* 76(6): 1375–411.

Appendix A

A.1 Solution Algorithm

The analysis of Markov-Perfect Equilibrium with imperfect competition is generally limited to industries with just a few firms, less realistic than the number of banks we consider in this paper. The main restriction is that, since firms have market power, their decision rules are a function of the decision rules of all their competitors. Even if one were to restrict to symmetric strategies in which decision rules become a function of the industry state as we do, the number of industry states to be considered quickly becomes very large.

For this reason, we solve the model by adapting the approach in Farias and others (2012) to an environment with aggregate and regional shocks. The algorithm approximates a Markov-Perfect Equilibrium by assuming that firms, at each time, make decisions based on their own state and the average industry state (conditional on a set of finite moments) that prevail in equilibrium. This reduces the computational cost considerably since firms' decision rules are not explicitly a function of the sequence of industry states, but rather a function of the long-run average distribution. The results in Weintraub and others (2008) and Farias and others (2012) establish conditions under which this approximation works well asymptotically.

In our application, we approximate the industry state by assuming that it equals the average cross-sectional distribution. That is, when maximizing profits, banks choose the optimal level of loans, deposits, and securities competing against the long-run average distribution of banks. We denote that distribution by $\bar{\mu}$. We quantitatively show that in our setup, where banks do not accumulate assets and regional shocks are i.i.d., this assumption approximates the observed distribution very well. Note that, given that we know the investment and exit decision rules for each bank type, we do not need to approximate the transition from $\bar{\mu}$ to $\bar{\mu}'$, i.e., unlike the Krusell-Smith method. Instead, we simply apply the transition operator $\mu' = \mathcal{H}(\bar{\mu}, N_e)$ in equation (33) from the text.

To find an equilibrium we perform the following steps:

1. Solve the problem of the entrepreneur (16)-(17) and derive the total loan-demand function (21). Given that the extreme-value distribution implies bank and nonbank market shares given in (18), we can calculate bank loan demand as in (22).

2. Solve the problem of the nonbank (34) to obtain the residual loan demand for bank loans.

3. Set tolerances $\in^{\ell}, \in^{I}, \in^{x}, \in^{e}$, and $\in^{\mu}$ to small values. Start with a number of entrants $N^{e,g}$ where iteration $g = 0$ is an initial guess.

4. Guess an investment decision rule $I^h(\cdot)$ and an exit decision rule $X^h(\cdot)$, where iteration $h = 0$ is an initial guess.

5. Using $N^{e,g}, I^h(\cdot)$, and $X^h(\cdot)$ and a large sequence of shocks $\{z_{j,t}, z_{-j,t}\}_{t=1}^{T}$, simulate the distribution of banks $\{\mu_t\}_{t=1}^{T}$. Discard the initial 250 periods and compute the average industry state $\overline{\mu}^h$ by taking the average of the observed distribution.[19]

6. Obtain an equilibrium in the loan market:

a. Guess a loan decision rule $\ell^k(\cdot)$ where iteration $k = 0$ is an initial guess.

b. For each $\{\theta, d\}$, given that the industry state $\overline{\mu}^h$ and $\ell^k(\cdot)$ determines the loan-supply function of a bank's competitors, obtain the best response $\ell^{k+1}(\cdot)$ by maximizing profits in equation (23).

c. Compute $\Delta^{\ell} = \|\ell^{k+1}(\cdot) - \ell^k(\cdot)\|$.

d. If $\Delta^{\ell} < \in^{\ell}$, an equilibrium in the loan market has been found, so continue to the next step. If not, return to step with the updated loan decision rule $\ell^{k+1}(\cdot)$.

7. Solve the bank problem to obtain investment and exit rules:

a. For each $\{\theta, d, s_j\}$, solve the bank problem in (29) to obtain $I^{h+1}(\cdot)$ and in (28) with $V^{x=1}(\cdot)$ given in (11) to obtain $X^{h+1}(\cdot)$.

b. Using $I^{h+1}(\cdot)$ and $X^{h+1}(\cdot)$, compute a new long-run industry state $\overline{\mu}^{h+1}$ using the transition operator in equation (33).

c. Compute $\Delta^I = I^{h+1}(\cdot) - I^h(\cdot)\|$, $\Delta^x = \|x^{h+1}(\cdot) - x^h(\cdot)\|$, and $\Delta^{\mu} = \|\overline{\mu}^{h+1} - \overline{\mu}^h\|$.

d. If $\Delta^I < \in^I$, $\Delta^x < \in^x$, and $\Delta^{\mu} < \in^{\mu}$ continue to the next step. If not, return to step with the updated industry state $\overline{\mu}^{h+1}$.

8. Obtain the value of an entrant (net of entry costs) $V^e(\overline{\mu}^{h+1})$ in equation (31). If $\|V^e(\overline{\mu}^{h+1})\| < \in^e$, an equilibrium has been found. If not, update the number of entrants $N^{e,g+1}$ and return to step 5 with the updated number of entrants. The update of $N^{e,g}$ is done taking into account the value of $V^e(\overline{\mu}^{h+1})$. If $V^e(\overline{\mu}^{h+1}) > 0$, set $N^{e,g+1} > N^{e,g}$. If $V^e(\overline{\mu}^{h+1}) < 0$, set $N^{e,g+1} < N^{e,g}$.

19. Note that, to simulate the distribution, you need an initial distribution. We assume that the distribution in period 1 equals one with a number $N^{e,g}$ of θ_3 banks. As we discard the initial periods to compute the average distribution, the selection of this initial distribution is not quantitatively relevant.

9. A final check on the equilibrium is how well the 'average' industry (conditional on Z) approximates the observed distribution along the equilibrium path. We compute the average distance between the observed distribution $\{\mu_t\}(t{=}1)^T$ and the average distribution μ and the values are small.

While the algorithm just described has been proven to converge, we also experimented with a slightly modified version, where we evaluate the value of the entrant for many possible values of the number of entrants and define an equilibrium as one where the condition in point 8 is satisfied. This modified version of the algorithm is more costly computationally but robust.

A.2 Data Description

As in Corbae and D'Erasmo (2021), we compile a large panel of banks from 1984 to 2019 using data for the last quarter of each year.[20] The source for the data is the Call Reports that banks submit to the Federal Reserve each quarter.[21] Call Report data are available for all banks regulated by the Federal Reserve System, the FDIC, and the Comptroller of the Currency. All financial data are on an individual-bank basis.

We consolidate individual commercial banks to the bank holding company level and retain those bank holding companies and commercial banks (if there is not top holder) for which the share of assets allocated to commercial banking (including depository trust companies, credit card companies with commercial bank charters, private banks, development banks, limited charter banks, and foreign banks) is higher than 25 percent. We follow Kashyap and Stein (2000) and Den Haan and others (2007) in constructing consistent time series for our variables of interest. Finally, we only include banks located within the fifty states and the District of Columbia. In addition to information from the Call Reports, we identify bank failures using public data from the FDIC.[22] We also identify mergers and acquisitions

20. There was a major overhaul to the Call Report format in 1984. Since 1984, banks are, in general, required to provide more detailed data concerning assets and liabilities. Due to changes in definitions and the creation of new variables after 1984, some of the variables are only available after this date.

21. Balance sheet and income statement items can be found at the FFIEC website.

22. Data is available at the FDIC website.

using the Transformation Table on the FFIEC website. We identify 'events' where the acquired and acquiring firms are commonly owned in some form before the acquisition (i.e., the listed merger is only a corporate reorganization) and discard these events from the merger sample.

To deflate balance-sheet and income-statement variables, we use the CPI index. When we report weighted aggregate time series, we use the asset market share as the weight. To control for the effect of a small number of outliers, when constructing the loan returns, cost of funds, charge-off rates, and related series, we eliminate observations in the top and bottom 1 percent of the distribution of each variable. We also control for the effects of bank entry, exit, and mergers by not considering the initial period, the final period, or the merger period (if relevant) of any given bank.

To analyze the bank lending channel, we follow Kashyap and Stein (1995) and extend our annual data to quarterly frequency. As before, we work with data from 1984 to 2019. We discard observations for banks that are involved in a merger, de novo banks in the period they enter, as well as the final period of banks that fail. Variables are defined as in Corbae and D'Erasmo (2021). We follow Kashyap and Stein (1995) and, for each of our size categories, we run the following specification (in panel fixed-effects form)

$$\Delta y_{i,t} = \sum_{h=1}^{8} \beta_h (f_{t-h} - f_{t-h-1}) + \sum_{h=1}^{4} \alpha_h \Delta y_{i,t-h} + \gamma_j X_t + \phi x_{i,t} + a_i + \tau_t + Q_t + \in_{i,t}, \quad \text{(A.2.1)}$$

where $\Delta y_{i,t}$ denotes the growth rate of $y_{i,t}$ (total loans net of C&I loans) between quarter t and quarter $t-1$, f_{t-h} corresponds to the fed funds rate in period $t-h$, X_t captures aggregate variables (such as the inflation rate or changes in nominal GDP), $x_{i,t}$ bank level controls that include the ratio of deposits to assets and the ratio of cash and securities to assets. a_i is a bank fixed effect, τ_t is a year fixed effect, and Q_t is a quarter fixed effect. The fed funds come from FRED and correspond to the end-of-quarter value.

A.3 Cost Estimation

We estimate the marginal cost of producing a loan $c_{\theta}(\ell_{\theta,t})$ and the fixed cost $C_{F,\theta}$ following the empirical literature on banking.[23,24] The marginal cost is derived from an estimate of marginal net expenses that is defined to be marginal noninterest expenses net of marginal noninterest income. Marginal noninterest expenses for bank j are derived from the following trans-log cost function:

$$\begin{aligned}\log NIE_t^j = {} & g_1\log(w_t^j) + h_1\log(\ell_t^j) + g_2\log(q_t^j) + g_3\log(w_t^j)^2 \\ & + h_2[\log(\ell_t^j)]^2 + g_4\log(q_t^j)^2 + h_3\log(\ell_t^j)\log(q_t^j) + h_4\log(\ell_t^j)\log(W_t^j) \\ & + g_5\log(q_t^j)\log(W_t^j) + g_6^1 t + g_6^2 t^2 + g_{8,t} + g_9^j + \in_t^j,\end{aligned} \tag{A.3.2}$$

where $NIE_{\theta,t}^j$ is noninterest expenses (calculated as total expenses minus the interest expense on deposits, the interest expense on fed funds purchased, and expenses on premises and fixed assets), g_9^j is a bank fixed effect, W_t^j corresponds to input prices (labor expenses), ℓ_t^j corresponds to real loans (one of the two bank j's outputs), q_t^j represents safe securities (the second bank output), the t regressor refers to a time trend, and $k_{8,t}$ refers to time fixed effects. We estimate this equation by panel fixed effects with robust standard errors clustered by bank.[25] Noninterest marginal expenses are then computed as:

$$\text{Mg Non-Int Exp.} \equiv \frac{\partial NIE_t^j}{\partial \ell_t^j} = \frac{NIE_t^j}{\ell_t^j}\left[h_1 + 2h_2\log\left(\ell_t^j\right) + h_3\log\left(q_{jt}\right) + h_4\log\left(w_t^j\right)\right]. \tag{A.3.3}$$

Marginal noninterest income (Mg NonInt Inc.) is estimated by using an equation similar to equation (A.3.2) (without input prices), where the left-hand side corresponds to total noninterest income. Net marginal expenses (Net Exp.) are computed as the difference between marginal noninterest expenses and marginal noninterest income. The

23. See, for example, Berger and others (2009) and our previous paper Corbae and D'Erasmo (2021).

24. The marginal cost estimated is also used to compute our measure of markups and the Lerner Index.

25. We eliminate bank-year observations in which the bank organization is involved in a merger or the bank is flagged as being an entrant or a failing bank. We only use banks with three or more observations in the sample.

fixed cost $C_{F,\theta}$ is estimated as the total cost on expenses of premises and fixed assets. Table 5 presents the estimated average net expense, the fixed cost, as well as the average cost by bank size.