A Quantitative Model of Banking Industry Dynamics *

Dean Corbae
University of Texas
at Austin

Pablo D'Erasmo
University of Maryland
at College Park

December 8, 2010

Abstract

We develop a model of banking industry dynamics to study the relation between commercial bank market structure, business cycles, and borrower default frequencies. The model is calibrated to match a set of key aggregate and cross-sectional statistics for the U.S. banking industry. We then test the model against business cycle moments and important characteristics of banks of different sizes. As in the data, the model generates countercyclical interest rates on loans, bank failure rates, borrower default frequencies, and charge-off rates as well as procyclical loan supply and profit rates. The model is also consistent in generating the negative relation between default frequencies, loan return rates, charge-off rates and variance of returns with bank size. The model is used to study the effects of bank competition and the benefits/costs of policies to mitigate bank failure.

1 Introduction

The objective of this paper is to formulate a simple quantitative structural model of the banking industry consistent with data in order to understand the relation between market structure and risk taking by financial intermediaries. Once the underlying technological parameters are consistently estimated, we can also use the model to address important regulatory questions. We want the model to be rich enough to answer questions like those posed by Ben Bernanke "I want to be very, very clear: too big to fail is one of the biggest problems we face in this country, and we must take action to eliminate too big to fail."[1]

Banks in our environment intermediate between a large number of households who supply funds and a large number of borrowers who demand funds to undertake risky projects. By lending to a large number of borrowers, a given bank diversifies risk that any particular household may not accomplish individually. Since we will eventually estimate preference

*The authors wish to thank John Boyd, Gianni DeNicolo, Nobu Kiyotaki, Raphael Repullo, Juan Sanchez, Kenechi Ueda, and Stephen Williamson for comments on an earlier version of this paper as well as seminar participants at the Bank of England, IMF, Federal Reserve Banks of Chicago, Dallas, and St. Louis, Wharton, the University of Wisconsin, the SED, and the Workshop on Financial Systems (CFSP) at MIT.

[1]Time, December 28, 2009/January 4, 2010, p. 78.

and technology parameters, we require our model to be parsimonious. When mapping the model to data, we attempt to match both aggregate and cross-sectional statistics for the U.S. banking industry.

Our model assumes spatial differences between banks; there are "national" geographically diversified banks that may coexist in equilibrium with "regional" and very small "fringe" banks that are both restricted to a geographical area. Since we allow for regional specific shocks to the success of borrower projects, smaller banks may not be well diversified. This assumption generates ex-post differences between big and small (regional and fringe) banks. As documented in the data section, the model generates not only procyclical loan supply but also countercyclical interest rates on loans, borrower default frequencies, and bank failure rates. Since bank failure is paid for by lump sum taxes to fund deposit insurance, the model predicts countercylical taxes.

Some of the important questions to be addressed in this project are: How much does bank competition contribute to risk taking (as measured for instance by realized default frequencies)? Are crises less likely in more concentrated banking industries? What are the costs of policies to mitigate bank failure? Besides our quantitative approach, the benefit of our model relative to the existing literature is that the size distribution of banks is derived endogenously and varies over the business cycle - a fact which is evident in the data. We conduct counterfactuals to shed light on the debate about competition and bank risk taking.

Our project is most closely related to the following literature on the industrial organization of banking. Our underlying model of banking is based on Allen and Gale [3] (hereafter A-G), section IV of Boyd and De Nicolo [9] (hereafter B-D), and Martinez-Miera and Repullo [25].[2] In those models, the authors study the implications of exogenously varying the number of banks on loan supply and borrower risk taking. In fact, there is a theoretical debate between A-G (whose framework delivers that more concentration leads to more stability) and B-D (whose model delivers the result that more concentration leads to more fragility).[3] Unlike the previous frameworks, we do not exogenously fix the number of banks but instead solve for an equilibrium where banks may enter and exit so that the number of banks is endogenously determined. To keep the model simple, here we focus only on loan market competition while there is a voluminous IO literature on deposit market competition (see for example Aguirregabiria, Clark, and Wang [1]).

There is a vast empirical literature that takes up the "concentration-stability" versus "concentration-fragility" debate. For example, Beck, Demirguc-Kunt, and Levine [4] run probit regressions where the probability of a crisis depends on banking industry concentration as well as a set of controls. In their regressions a "crisis" is defined to be a significant fraction of insolvent banks (or a fraction of nonperforming loans exceeding 10%). While Beck, et. al.

[2]Another strand of literature uses the costly state verification approach of Townsend [29] either ex-ante as in Diamond [12] or ex-post as in Williamson [30] to rationalize the existence of banks. These papers all study a competitive market structure however.

[3]In particular, an exogenous increase in the number of banks in both A-G and B-D raises interest rates that banks must pay their depositors. However, in A-G those costs are passed on to borrowers since there is not loan market competition; this results in higher borrower default probabilities and ultimately lower realized profits. In B-D, on the other hand, increased loan market competition lowers interest rates on loans as well and this lowers borrower default probabilities which may ultimately raise realized profits. Which effect dominates in the B-D case is a quantitative matter.

find evidence in favor of the concentration-stability view, in general there are mixed results from this empirical work.

We address this debate using our quantitative structural model. After calibrating the model to match aggregate and cross-sectional statistics for the U.S. banking industry, we compare the types of "crisis" dependent variables that the empirical literature studies across the endogenously determined differences in market structure. We find that more concentration can lead to large increases in borrower default frequencies but since only big, well diversified banks remain, there is a decrease in bank exit rates (another "crisis" measure). Thus the model is consistent with mixed results as in the data.

We require our quantitative model to be consistent with U.S. data on market concentration. Instead of assuming perfect competition, we consider Cournot competition and apply a version of the Markov Perfect Industry equilibrium concept of Ericson and Pakes [17] augmented with a competitive fringe along the lines of Gowrisankaran and Holmes [18]. In this way, we depart from quantitative competitive models of banking such as Bernanke, et. al. [7], Carlstrom and Fuerst [10], or Diaz-Gimenez, et. al. [14], thereby allowing big banks to act strategically in the loan and deposit market. Further, dropping the competitive assumption along with our spatial restrictions generates a nontrivial size distribution of banks where both intensive and extensive margins can vary over the business cycle which is broadly consistent with data. When mapping the model to data, we use the same dataset as Kashyap and Stein [22].

The remainder of the paper is organized as follows. Section 2 documents a select set of banking data facts. Section 3 lays out a simple model environment to study bank risk taking and loan market competition. Section 4 describes a markov perfect equilibrium of that environment. Section 5 discusses how the preference and technology parameters are chosen and section 6 provides results for the simple model. Section 7 conducts three counterfactuals: (i) one experiment assesses the effects of bank competition on business failures and banking stability; (ii) another experiment assesses the effects of regulation which restricts banks to a geographical region; and (iii) a final experiment assesses the consequences of a "too big to fail" policy. Section 8 concludes and lists a set of extensions to the simple model which we are currently pursuing.

2 Some Banking Data Facts

In this section, we document the cyclical behavior of entry and exit rates, bank lending, measures of loan returns and the level of concentration in the U.S. We focus on individual commercial banks in the U.S.[4] The source for the data is the Consolidated Report of Condition and Income (known as Call Reports) that insured banks submit to the Federal Reserve each quarter.[5] We compile a large data set from 1976 to 2008 using data for the last quarter

[4]In some cases, commercial banks are part of a larger bank holding company. For example, in 2008, 1383 commercial banks (20% of the total) were part of a bank holding company. As in Kashyap and Stein [22] we focus on individual commercial banks. As they argue, there are not signficant differences in modelling each unit. The holding company is subject to limited liability protection rules with respect to the losses in any individual bank.

[5]The number of institutions and its evolution over time can be found at http://www2.fdic.gov/hsob/SelectRpt.asp?EntryTyp=10.

of each year. We follow Kashyap and Stein [22] in constructing consistent time series for our variables of interest.

One clear trend of the commercial banking industry during the last three decades is the continuous drop in the number of banking institutions. In 1980, there were approximately 14,000 institutions and this number has declined at an average of 360 per year, bringing the total number of commercial banks in 2008 to less than 7,100. This trend was a consequence of important changes in regulation that were introduced during the 1980's and 1990's (deposit deregulation in the early 1980's and the relaxation of branching restrictions later).

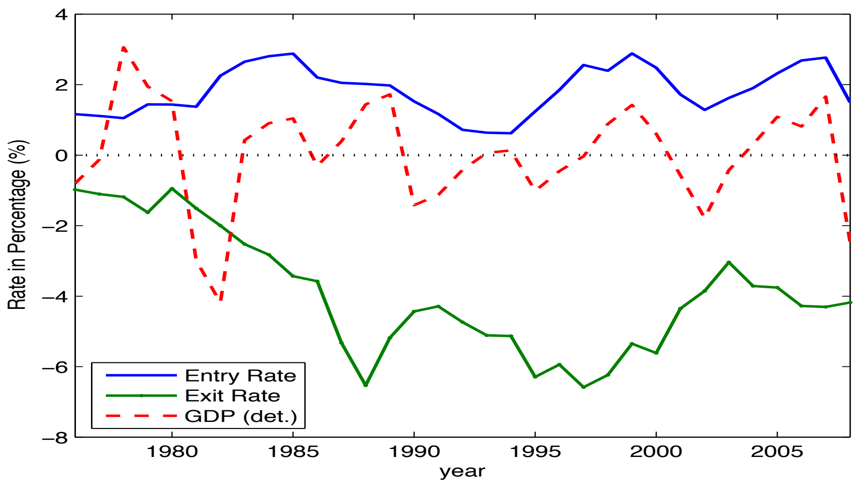
This decline in the level (or stock) of banks is evidenced by flow measures of exit, entry, and (resulting) net entry. The number of exits (including mergers and failures) and entrants expressed as a fraction of the banking population in the previous year are displayed in Figure 1. We also incorporate real detrended GDP to understand how entry and exit rates move along the cycle.[6] There was an important increase in the fraction of banks that exited starting in the 1980's and the high level continued through the late 1990's due to the aforementioned regulatory changes. The bulk of the decline was due to mergers and acquisitions. However, from 1985 through 1992, failures also contributed significantly to the decline in the number of banks. Figure 2 decomposes the exit rate into mergers and failures as well as the fraction of 'troubled' banks for a subset of the period.[7] The figure also shows that there has been a consistent flow of entry of new banks, cycling around 2%. Since 1995 the net decline in the number of institutions has trended consistently lower (except in 2008) so that the downward trend is leveling off. This recent leveling off of the trend is also documented in Table 6 of Janicki and Prescott [21].

Balance Sheet and Income Statements items can be found at https://cdr.ffiec.gov/public/.

[6] The H-P filter with parameter equal to 6.25 is used to extract the trend from real GDP data.

[7] A troubled bank, as defined by the FDIC, is a commercial bank with CAMEL rating equal to 4 or 5. CAMEL is an acronym for the six components of the regulatory rating system: **C**apital adequacy, **A**sset quality, **M**anagement, **E**arnings, **L**iquidity and market **S**ensitivity (since 1998). Banks are rated from 1 (best) to 5 (worst), and banks with rating 4 or 5 are considered 'troubled' banks (see FDIC Banking Review 2006 Vol 1). This variable is only available since 1990.

Figure 1: Bank Industry Dynamics and Business Cycle



Note: Data corresponds to commercial banks in the US. Source: FDIC, Consolidated Report of Condition and Income. Entry corresponds to new charters and conversions. Exit correspond to unassisted mergers and failures. GDP (det) correspond to real GDP detrended. The trend is extracted using the H-P filter with parameter 6.25.

Figures 1 and 2 also make clear that there was a significant amount of cyclical variation in entry and exit. The correlation of the entry and exit rates with detrended GDP is 0.25 and 0.07 respectively (if we restrict to the post-reform years - after 1990 - the correlations with detrended GDP are 0.87 and 0.11 for entry and exit rates respectively). Since exits can occur as the result of a merger, these correlations hide what we usually think of as the cyclical component of exits; failures and 'troubled' banks have a more important cyclical component than mergers. We find that the failure and 'troubled' bank rates are countercyclical while the merger rate displays a procyclical behavior. Specifically, the correlation with real detrended GDP of the failure and 'troubled' bank rates (since 1990) are -0.39 and -0.49 respectively. On the other hand, the correlation of the merger rate for the same period equals 0.28.

Figure 2: Exit Rate Decomposed

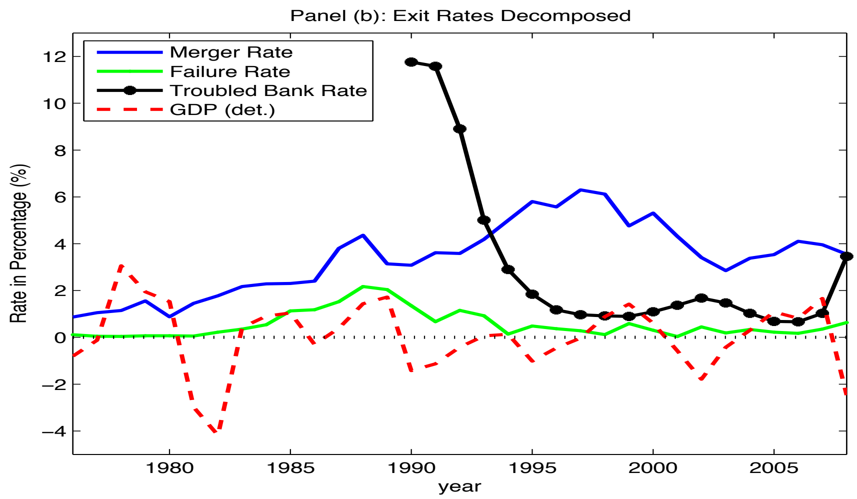


Figure 3: Loans, Deposits and Business Cycle

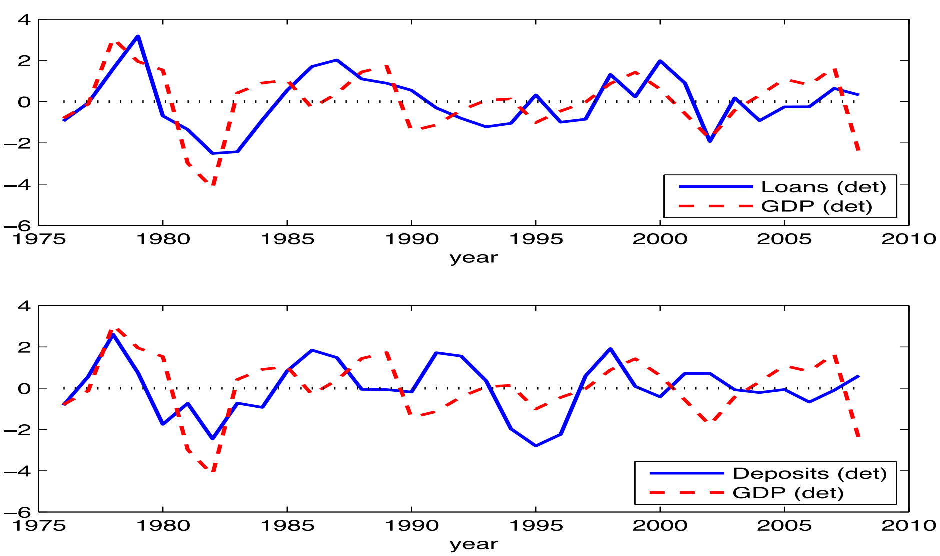


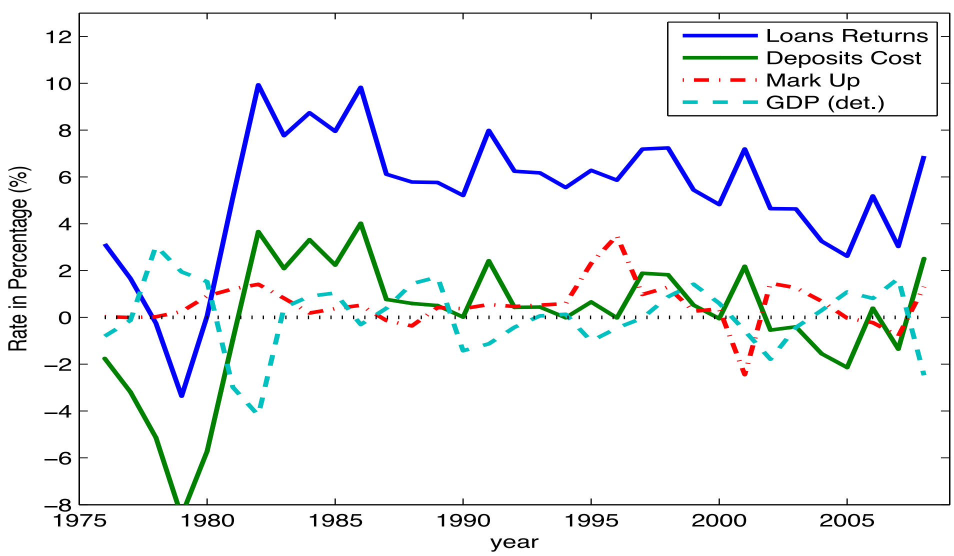
Figure 3 displays how bank lending and deposits move along the cycle. The series for loans and deposits are constructed using the individual commercial bank level data.[8] The stock

[8]The data for total loans since 1984 come from item RCON2122, total loan and leases net of unearned income. Prior to 1984, Lease Financing Receivables (RCON2165) are not included as part of total loans so the two series need to be summed to insure comparability. The Implicit GDP deflator is used to convert nominal variables into real. The H-P filter with parameter equal to 6.25 is used to extract the trend from the data.

of loans and deposits have an important cyclical component. We find that both measures of bank activity are highly procyclical where correlations with detrended GDP equal 0.58 and 0.41 for loans and deposits respectively.

Figure 4 shows the cyclical behavior of the rate of return on loans, the rate of return on deposits, and the 'mark-up' rate (the difference between the two inclusive of other costs).[9] Rates of return on loans, deposits, and the mark-up display a highly countercyclical behavior. Their correlation with detrended GDP equals -0.41, -0.37 and -0.42 respectively.

Figure 4: Loans Return, Deposits Cost, Mark-Up and Business Cycle

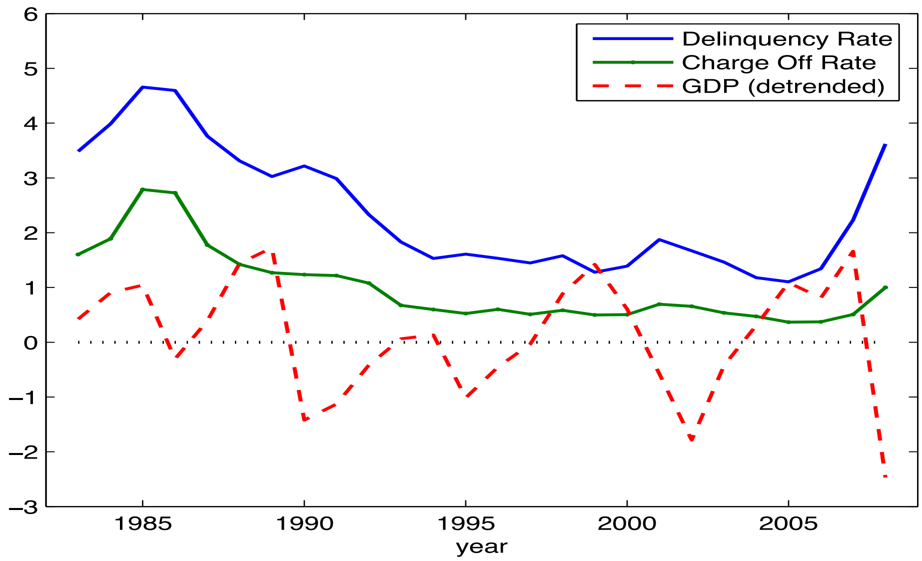


In Figure 5 we show the evolution of loan delinquency rates and charge off rates. Data is only available since 1984.[10] Delinquency rates and charge off rates are countercyclical. Their correlation with detrended GDP is -0.15 and -0.29 respectively.

[9]The rate of return on loans is defined as total income from loans divided by total loans. The rate of return on deposits is constructed as the total expense on deposits divided by the stock of deposits. Total income from loans is computed from Income Statement data (schedule RI) using variables RIAD4010 and RIAD4059. The stock of deposits is constructed from RCON 2200. Nominal returns are converted to real returns using a measure of expected inflation derived from the CPI index. In particular, we assume that inflation follows an AR(1) process and use the fitted value for each period as expected inflation.

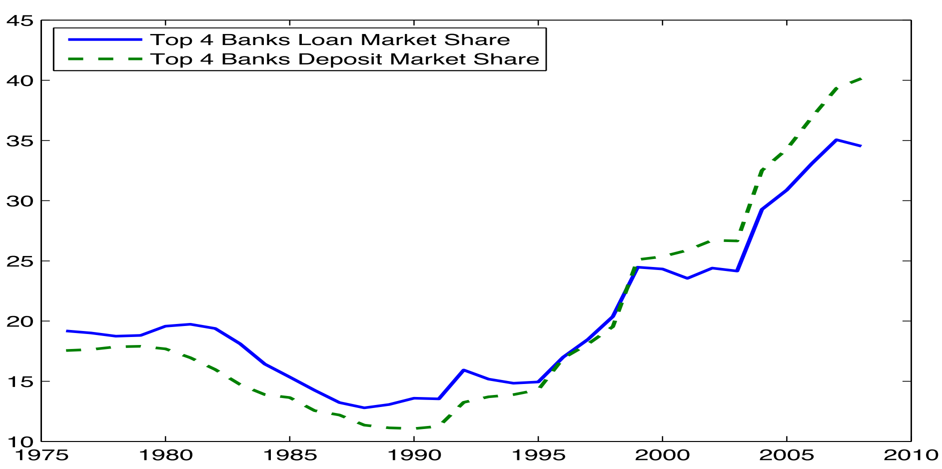
[10]The delinquency rate is the ratio of the dollar amount of a bank's delinquent loans to the dollar amount of total loans outstanding. Charge-off rates are defined as the flow of a bank's net charge-offs (gross charge-offs minus recoveries) during a year divided by the average level of its loans outstanding during the same period.

Figure 5: Loan Delinquency Rates, Charge Off Rates and Business Cycle



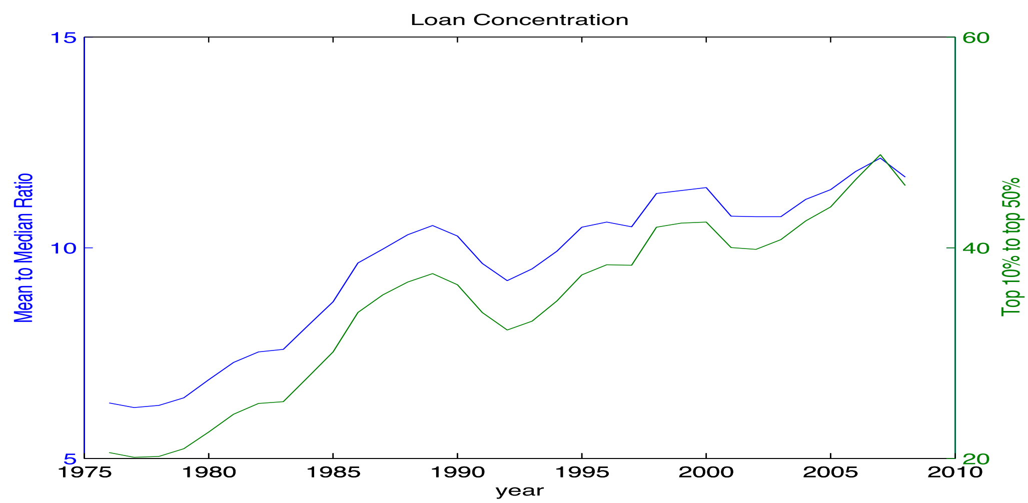
The size distribution of banks has always been skewed but the large number of bank exits (mergers and failures) that we documented above resulted in an unparalleled increase in loan and deposit concentration during the last 35 years. For example, in 1976 the four largest banks (when sorted by assets) held 19 and 18 percent of the banking industry's loans and deposits respectively while by 2008 these shares had grown to 35 and 40 percent. Figure 6 displays the trend in the share of loans and deposits in the hands of the four largest banks since 1976.

Figure 6: Increase in Concentration: Loan and Deposit Market



The increase in the degree of concentration is also evident if we look at the mean-to-median ratio and the share of loans of the top 10 percent banks to the bottom 50 percent of the asset distribution. Figure 7 displays the trend in these two measures of concentration in the loan market since the year 1976.

Figure 7: Increase in Concentration: Loan Market



The increase in concentration is also the result of considerable exit (merger and failure) and entry by banks of small size. In Table 1, we show entry and exit statistics by bank size (when sorted by assets).

Table 1: Entry and Exit Statistics by Bank Size (sorted by assets)

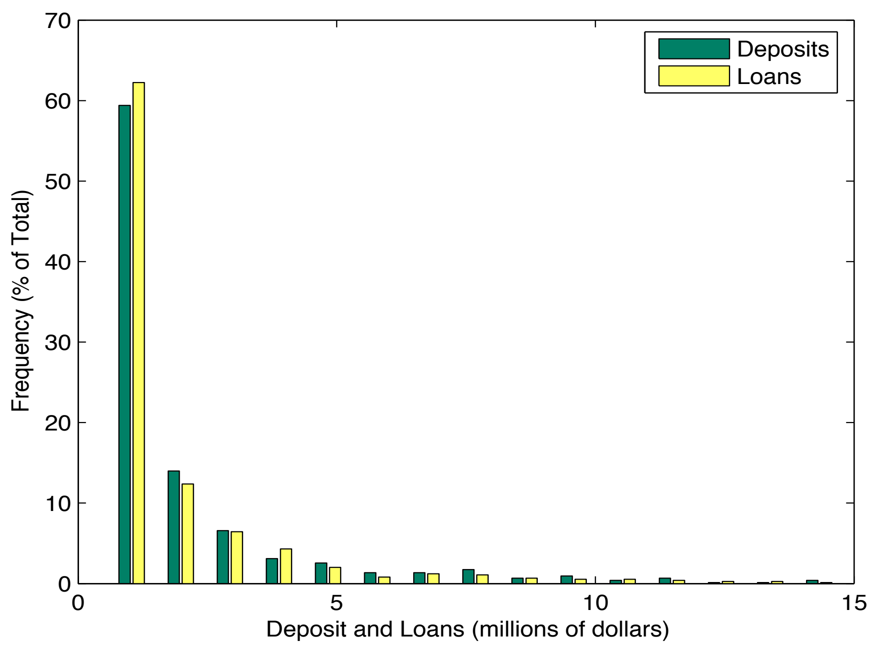
Fraction of Total x, accounted by:	x Entry	Exit	Exit by Merger	Exit by Failure
Top 4 Banks (avg.)	0.00	0.01	0.02	0.00
Top 1 % Banks (avg.)	0.27	1.07	1.57	2.88
Top 10 % Banks (avg.)	4.87	14.18	16.13	15.36
Bottom 99 % Banks (avg.)	99.73	98.93	98.43	97.12
Top 4 Banks (max.)	0.00	0.23	0.24	0.00
Top 1 % Banks (max.)	2.02	2.97	3.04	25.00
Top 10 % Banks (max.)	20.00	24.05	24.11	50.00
Bottom 99 % Banks (max.)	100.00	100.00	99.64	100.00

Note: Data corresponds to commercial banks in the US. Source: FDIC, Call and Thrift Financial Reports. Entry and Exit period extends from 1976 to 2008. Merger and Failure period is 1990 - 2007.

We note that the bulk of entry, exit, mergers and failures correspond to banks that are in the bottom 99% of the distribution. The time series average accounted for by the bottom 99% is close to 99% across all categories. We also observe that, in some periods, all entry, exit, and failures correspond to small banks (as measured by the maximum over time).

The high degree of concentration in the banking industry is the reflection of the large number of small banks and a few large banks. In Figure 8 we provide the distribution of deposits and loans for the year 2007. Given the large number of banks at the bottom of the distribution we plot only banking institutions with less than 15 million dollars in deposits (93% of the total). Banks with 1 million dollars of deposits and loans account for approximately sixty percent of the total number of banks. However, total deposits and loans in these banks make up only twenty percent of the total loans and deposits in the industry.

Figure 8: Distribution of Bank Deposits and Loans in 2008



Note: Data corresponds to commercial banks in the US.
Source: FDIC, Call and Thrift Financial Reports, Balance Sheet and Income Statement.

In Table 2, we provide measures of deposit and loan concentration for commercial banks in the US for the year 2008.[11] The table shows the high degree of concentration in deposits and loans. It is striking that the the four largest banks (measured by the C_4) hold approximately forty percent of deposits and loans and that the top 1 percent hold 77 and 75 percent of total deposits and loans respectively. We also observe a ratio of mean-to-median of around

[11]C_4 refers to the top 4 banks concentration index. The Herfindahl Index (HHI) is a measure of the size of firms in relation to the industry and often indicates the amount of competition among them. It is computed as $\sum_{n=1}^{N} s_n^2$ where s_n is market share of bank n. The Herfindahl Index ranges from $1/N$ to one, where N is the total number of firms in the industry.

10 suggesting sizeable skewness of the distribution. This high degree of inequality is also evident in the Gini coefficient of around 0.9 (recall that perfect inequality corresponds to a measure of 1). The National Herfindahl Index is between 4-5%. This is because the largest four banks have the same market share (around 10% each) and there is a large number of firms that have a very small market share (more than 95% of banking institutions have deposit and loan market shares below 1%). The national values are a lower bound since they do not consider regional market shares but are much higher than the values of HHI that one would obtain when all firms have equal market shares (i.e. 1/N and $HHI - 0.13\%$).

Table 2: Bank Deposit and Loan Concentration (in 2008)

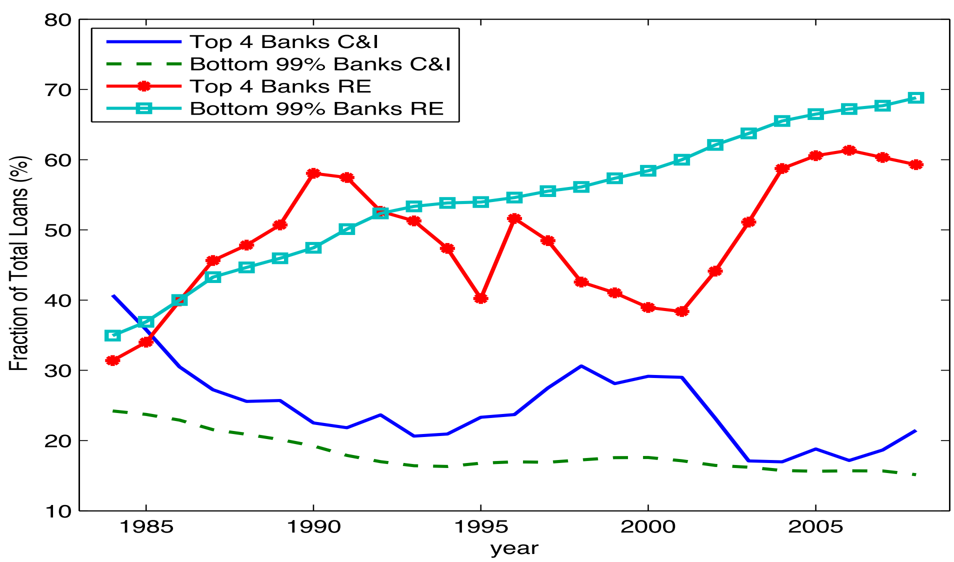
Measure	Deposits	Loans
Percentage of Total in top 4 Banks (C_4)	40.2	34.5
Percentage of Total in top 1% Banks	76.7	74.7
Percentage of Total in top 10% Banks	89.2	89.1
Ratio Mean to Median	10.5	10.6
Ratio Total Top 10% to Total Smallest 50%	36.9	40.7
Gini Coefficient	.91	.90
HHI : Herfindahl Index (National) (%)	4.9	3.8

Note: Data corresponds to commercial banks in the US. Source: FDIC, Call and Thrift Financial Reports. Total Number of Banks 7092. Top 1% banks corresponds to 71 banks. Top 10% banks corresponds to 709 banks. Bottom 50% banks corresponds to 3545 banks.

Bergstresser [6] documents (see his Table 1) that when computed for Metropolitan Statistical Areas (MSA) the Herfindahl Index is much higher (around 20%). Those numbers are typically associated with a highly concentrated industry (values between 10-20%). If we follow the traditional approach to competition that associates more firms with more price competition and fewer firms with less-competitive behavior, these numbers can be understood as evidence in favor of an imperfectly competitive banking industry. However, an alternative view is one where firms that have higher productive efficiency have lower costs and therefore higher profits. These firms tend to do better and so naturally gain market share, which can lead to concentration. Therefore, by this logic, concentration reflects more efficient banks, not necessarily an increase in market power. For this reason, different approaches have been suggested to attempt to measure the competitive conduct of banks without explicitly using information on the number of firms in the market. One approach, known as contestability, estimates deviations from competitive pricing (i.e. the difference between marginal revenue and marginal cost). One of the most widely used contestability tests is proposed by Panzar and Rosse [27] which essentially tests that the elasticity of marginal revenue with respect to factor prices (marginal cost) is sufficiently below 1 (which is the perfect competition prediction). Using this technique, Bikker and Haaf [8] estimate the degree of competition in the banking industry for a panel of 23 (mostly developed) countries. They find that for all slices of the sample, perfect competition can be rejected convincingly, i.e. at the 99% level of confidence. In summary, taken together these measures suggest the banking industry is less than perfectly competitive.

An important trend in the loan portfolio composition of commercial banks is the increase in the fraction of loans secured by real estate and the decrease in the amount of commercial and industrial lending. In Figure 9, we document this trend for the largest 4 banks and the bottom 99% "small" banks when sorted by total loans. We compute the share of total loans that corresponds to commercial and industrial loans ($C\&I$) and real estate loans (RE) for each bank and then plot the average of these shares for each group and year.

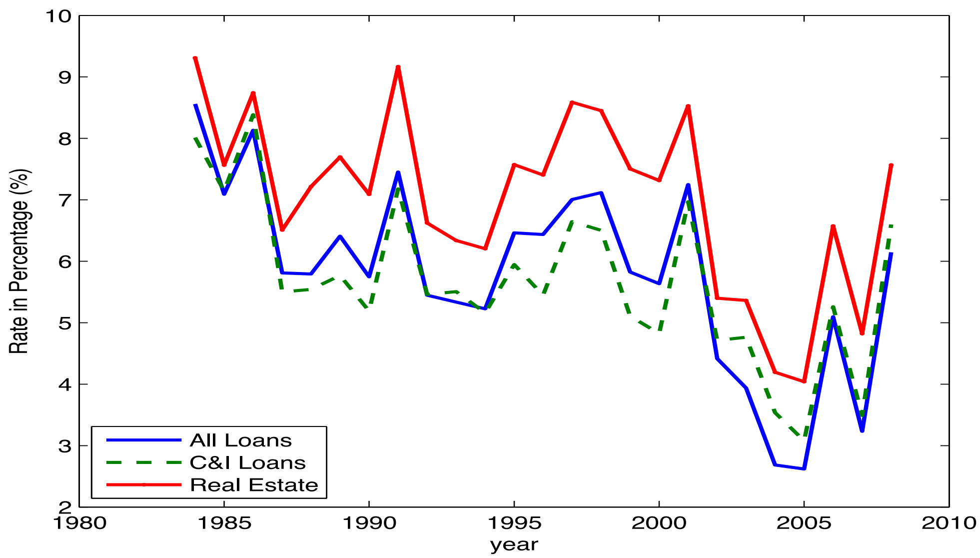
Figure 9: Loan Portfolio Composition by Bank Size (sorted by assets)



We observe that for both small and large banks, loans secured by real estate have become much more important. In the case of small banks, the share of loans secured by real estate almost doubled during this period (the share went from approximately 35 percent to around 70 percent). A similar trend is observed for the largest banks. The share of real estate loans in their portfolio increased from 40 percent to 60 percent. For this group of banks we also note a faster increase in this share during the last decade. The counterpart of the increase in real estate loans is the decrease in the share of commercial and industrial loans. We note that one of the differences between small and big banks is the portfolio composition. During the entire period loans secured by real estate constitute a more important component for small banks than for big banks. The opposite is true for commercial and industrial loans.

In Figure 10 we show that loan returns display countercyclical behavior even when dissagregated by loan type (commercial and industrial loans and real estate loans). The correlation with detrended GDP equals -0.17 and -0.09 for real estate and commercial and industrial loans respectively.

Figure 10: Loan Returns by Loan Type



In Figure 11, we analyze the evolution of loan returns by loan type and bank size. Small banks have higher returns for both real estate and commercial and industrial loans.

Figure 11: Loan Returns by Loan Type and Bank Size

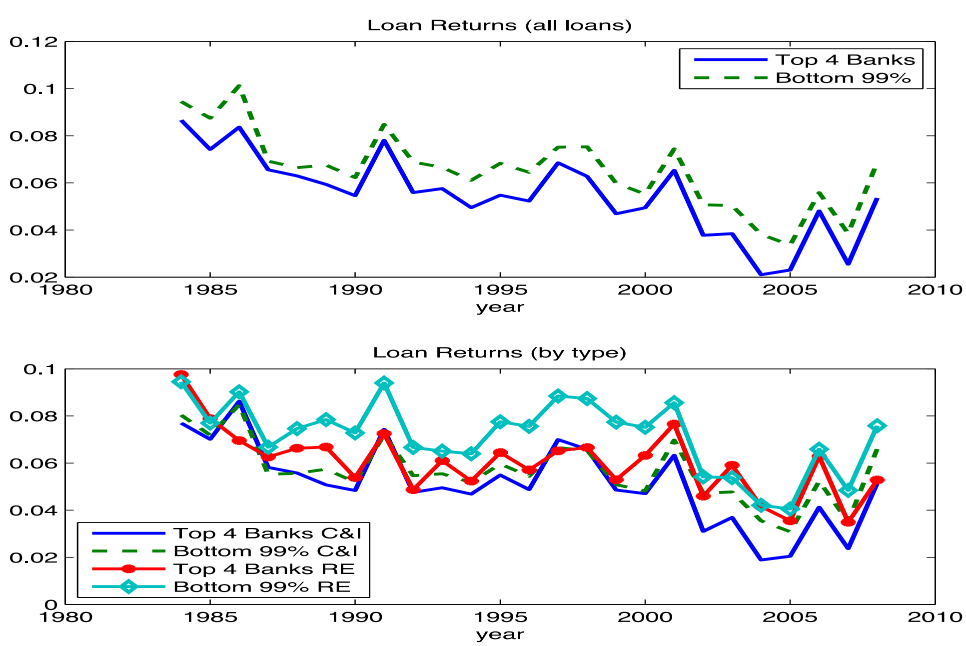


Table 3: Loan Return and Volatility by Bank Size (when sorted by assets)

Loan Returns (all loans)	Avg.	Std. Dev.	Corr. with GDP
Top 4 Banks (%)	4.93	0.75	-0.10
Top 1% Banks (%)	5.28	1.80	-0.18
Top 10% Banks (%)	5.51	1.56	-0.17
Bottom 99% Banks (%)	5.99	2.34	-0.17

Note: Data corresponds to commercial banks in the US. Source: FDIC, Call and Thrift Financial Reports. Values correspond to the time series average of the corresponding cross-sectional average and standard deviation among each group since 1984.

In Table 3 we report the average over time of the cross-sectional mean and standard deviation of loan returns by bank size since 1984. We note that smaller banks have higher returns and higher volatility of returns and a stronger negative correlation with detrended GDP. An important component of loan returns is the fraction of loans that are not repaid. For this reason, in Figures 12 and 13 we present charge-off rates and delinquency rates by bank size over time.

Independent evidence for the benefits of geographic diversification associated with bigger banks is given in Liang and Rhoades [24]. They test the hypothesis that geographic diversification lowers bank risk by regressing alternative measures of risk like the probability of bank insolvency, probability of failure, and the standard deviation of net income-to-assets on, among other controls, geographic diversification proxied by the inverse of the sum of squares of the percentage of a bank's deposits in each of the markets in which it operates. They find (see their Table 1) that the standard deviation of net-income-to-assets is significantly (both statistically and quantitatively) lower for firms that operate in a greater number of geographic markets. This will be consistent with our model.

Figure 12: Charge-Off Rates by Bank Size (when sorted by assets)

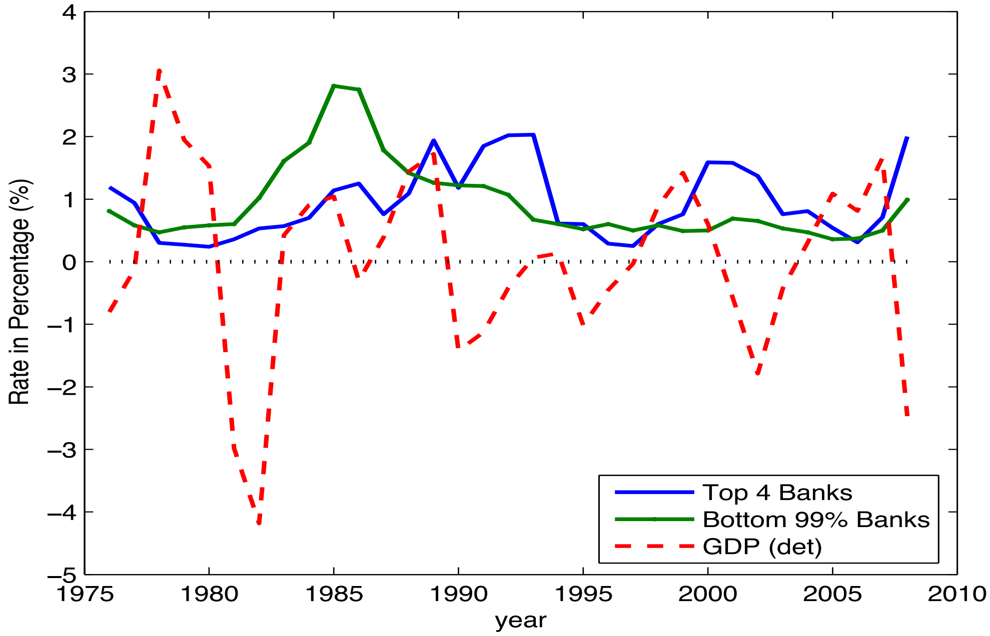
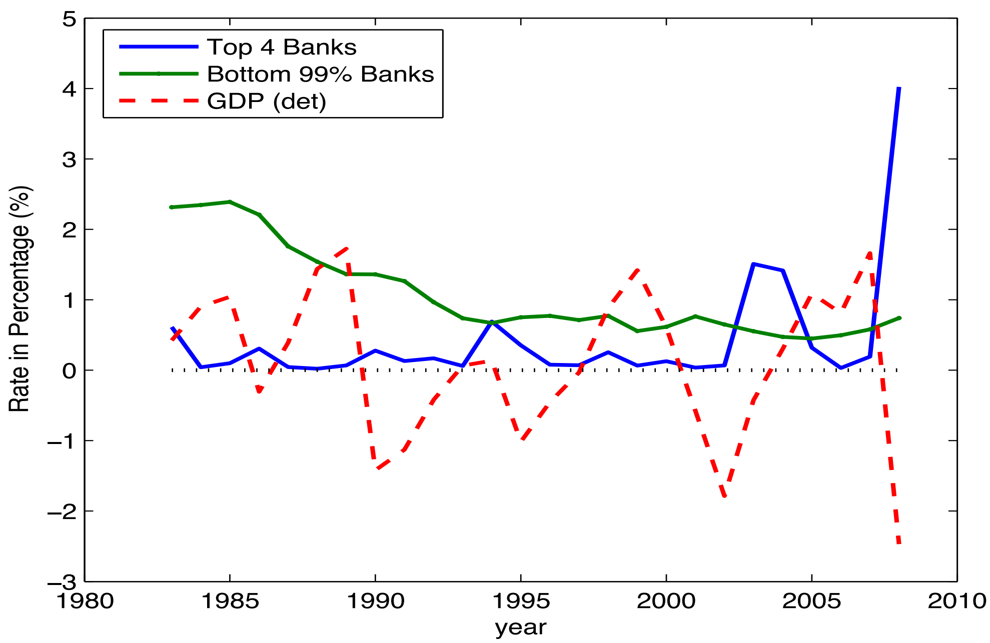


Figure 13: Delinquency Rates by Bank Size (when sorted by assets)



Both charge-off rates and delinquency rates have an important negative cyclical component. From Figure 12 we observe no clear pattern between small and big bank charge-off rates. In the early period of the sample (1976-1988), charge-off rates were higher for small banks than for big banks. However, the opposite is true after 1988. On the contrary, delinquency rates are lower for big banks than for small banks through virtually the entire sample.

In Table 4 we present the time series average of the cross-sectional average and standard deviation as well as the correlation with detrended GDP of charge-off rates and delinquency rates.

Table 4: Charge-Offs and Delinquency Rates by Bank Size (when sorted by assets)

Moment (%)	Avg.	Std. Dev.	Corr. with GDP
Charge Off Rate Top 4 Banks	0.94	0.53	-0.18
Charge Off Rate Top 1% Banks	1.08	1.25	-0.27
Charge Off Rate Top 10% Banks	0.84	1.29	-0.33
Charge Off Rate Bottom 99% Banks	0.93	1.55	0.04
Del. Rate Top 4 Banks	1.23	0.54	-0.53
Del. Rate Top 1% Banks	1.54	0.95	-0.37
Del. Rate Top 10% Banks	1.93	1.13	0.02
Del. Rate Bottom 99% Banks	2.39	1.95	0.14

Note: Data corresponds to commercial banks in the US. Source: FDIC, Call and Thrift Financial Reports. Values correspond to the time series average of the corresponding cross-sectional average and standard deviation among each group. Data on delinquency rates is available only since 1983 and corresponds to loans delinquent 90 days or more plus non accrual loans.

Banks of different sizes also differ in their sources of non-interest income and expenses. It is important to consider these differences since their relevance as a fraction of total profits has been rising the past three decades. Table 5 presents the decomposition of non interest income and expense by bank size. The bottom 99 % banks have the highest net expense to loan ratio while the lowest corresponds to the Top 1% banks.

Table 5: Non Interest Expense and Income as Fraction of Loans

	Non-Int Expense	Non-Int Income	Net Expense
Top 4 Banks	8.31	5.56	3.03
Top 1 % Banks	6.85	5.14	2.04
Top 10 % Banks	5.46	2.78	2.94
Bottom 99 % Banks	7.96	3.33	4.89

Note: Data corresponds to commercial banks in the US. Source: FDIC, Call and Thrift Financial Reports. Net expense is calculated as the difference of Non Interest Expenses and Non Interest Income. Reported values correspond to the time series average of each group.

3 Model Environment

Time is infinite. There are two regions $j \in \{e, w\}$, for instance "east" and "west". Each period, a mass B of one period lived ex-ante identical borrowers and a mass $H = 2B$ of one period lived ex-ante identical households (who are potential depositors) are born in each region.[12]

3.1 Borrowers

Borrowers in region j demand bank loans in order to fund a project. The project requires one unit of investment at the beginning of period t and returns at the end of the period:

$$\begin{cases} 1 + z_{t+1}R_t^j & \text{with prob } p^j(R_t^j, z_{t+1}, s_{t+1}) \\ 1 - \lambda & \text{with prob } [1 - p^j(R_t^j, z_{t+1}, s_{t+1})] \end{cases} \tag{1}$$

in the successful and unsuccessful states respectively. Borrower gross returns are given by $1 + z_{t+1}R_t^j$ in the successful state and by $1 - \lambda$ in the unsuccessful state. The success of a borrower's project in region j, which occurs with probability $p^j(R_t^j, z_{t+1}, s_{t+1})$, is independent across borrowers but depends on several things: the borrower's choice of technology $R_t^j \geq 0$, an aggregate technology shock at the end of the period z_{t+1}, and a regional shock s_{t+1} (the dating convention we use is that a variable which is chosen/realized at the end of the period is dated $t + 1$).

The aggregate technology shock is denoted $z_t \in \{z_b, z_g\}$ with $z_b < z_g$ (i.e. good and bad shocks). This shock evolves as a Markov process $F(z', z) = prob(z_{t+1} = z'|z_t = z)$. The regional specific shock $s_{t+1} \in \{e, w\}$ also evolves as a Markov process $G(s', s) = prob(s_{t+1} = s'|s_t = s)$ which is independent of z_{t+1}.

At the beginning of the period when the borrower makes his choice of R_t both z_{t+1} and s_{t+1} have not been realized. As for the likelihood of success or failure, a borrower who chooses to run a project with higher returns has more risk of failure and there is less failure in good times. Specifically, $p^j(R_t^j, z_{t+1}, s_{t+1})$ is assumed to be decreasing in R_t^j and $p^j(R_t^j, z_g, s_{t+1}) > p^j(R_t^j, z_b, s_{t+1})$. Moreover, we assume that the borrower success probability depends positively on which region $s_{t+1} \in \{e, w\}$ receives a favorable shock. Specifically, $p^{j=s_{t+1}}(R_t^j, z_{t+1}, s_{t+1}) > p^{j\neq s_{t+1}}(R_t^j, z_{t+1}, s_{t+1})$. That is, in any period, one region has a higher likelihood of success than the other. While borrowers in a given region are ex-ante identical, they are ex-post heterogeneous owing to the realizations of the shocks to the return on their project. We envision borrowers either as firms choosing a technology which might not succeed or households choosing a house that might appreciate or depreciate.

There is limited liability on the part of the borrower. If $r_t^{L,j}$ is the interest rate on bank loans that borrowers face in region j, the borrower receives $\max\{z_{t+1}R_t^j - r_t^{L,j}, 0\}$ in the successful state and 0 in the failure state. Specifically, in the unsuccessful state he receives $1 - \lambda$ which must be relinquished to the lender.

Borrowers have an outside option (reservation utility) $\omega_t \in [\underline{\omega}, \overline{\omega}]$ drawn at the beginning of the period from distribution function $\Upsilon(\omega_t)$.

[12]The assumption $H = 2B$ is a normalization that simplifies the analysis below. Furthermore, the assumption that borrowers and depositors are one period lived is simply to restrict attention to one period loan and deposit contracts rather than to resort to anonymity as in, for instance, Carlstrom and Fuerst [10].

3.2 Depositors

All households are endowed with 1 unit of the good and have preferences denoted $u(C_t)$. All households have access to a risk free storage technology yielding $1+\overline{r}$ with $\overline{r} \geq 0$ at the end of the period. They can also choose to supply their endowment to a bank in their region or to an individual borrower. If the household deposits its endowment with a bank, they receive $r_t^{D,j}$ whether the bank succeeds or fails since we assume deposit insurance. If they match with a borrower, they are subject to the random process in (1). At the end of the period they pay lump sum taxes τ_{t+1} which are used to cover deposit insurance for failing banks.

3.3 Banks

Motivated by the data described in Section 2, we assume there are three classes of banks $\theta \in \{n, r, f\}$ for national, regional, and fringe respectively. National banks are geographically diversified in the sense that they extend loans and receive deposits in both $\{e, w\}$ regions. Regional banks are restricted to make loans and receive deposits in one geographical area (i.e. either e or w). Fringe banks are also restricted on one geographical area (i.e. either e or w). Since we allow regional specific shocks to the success of borrower projects, regional and fringe banks may not be well diversified.[13] This assumption can, in principle, generate ex-post differences in loan returns documented in the data section. A bank's type is represented by the two-tuple $(\theta, \{e, w\})$ where, for instance (r, e) denotes an eastern regional bank. [14]

We denote loans made by bank i of type (θ, j) to borrowers at the beginning of period t by $\ell_{i,t}(\theta, j)$ and accepted units of deposits by $d_{i,t}(\theta, j)$. All banks have the same linear technology for producing loans. Without an interbank market, if i is a regional or fringe bank then $\ell_{i,t}(\theta, j) \leq d_{i,t}(\theta, j)$. If i is a national bank then $\ell_{i,t}(n, e) + \ell_{i,t}(n, w) \leq d_{i,t}(n, e) + d_{i,t}(n, w)$. We assume that national and regional banks do not face any restriction on the number of deposits they can accept in their region. On the other hand, fringe banks face a capacity constraint $\bar{d}$ of available deposits. Since fringe banks take prices as given, their expected profit function is linear in the amount of loans they extend, so we need to impose this capacity constraint in order to prevent the amount of loans of a fringe bank from exceeding the total amount of deposits in the region.

The timing in the loan stage follows the standard treatment of the dominant firm model (see for example Gowrisankaran and Holmes [18]). The dominant firms, our national and regional banks, move first. They compete in a Cournot fashion and choose quantities $\ell_{i,t}(\theta, j)$ taking as given not only the reaction function of other dominant banks but also the loan supply of the competitive fringe. Each fringe bank observes the total loan supply of dominant banks and all other fringe banks (that jointly determine the loan interest rate $r_t^{L,j}$) in region j and simultaneously decide on the amount of loans to extend. Since, at a given interest rate, the production technology is linear in loans supplied, the fringe banks decision reduces simply

[13] In an interesting paper, Koeppl and MacGee [23] consider whether a model with regional banks which operate within a region with access to interbank markets can achieve the same allocation under uncertainty as a model with national banks which operate across regions.

[14] See Figure 17 for a graphical description of the regional segmentation in the model. The table above the figure also summarizes the risk-return tradeoff in the borrower's problem.

to whether to bring all their available funds to the market or not, i.e. $\ell_{i,t}(f,j) \in \{0, \overline{d}\}$.

In principle one could also have banks be Cournot competitors in the deposit market as in Boyd and DeNicolo [9]. However, since we assume that $H > 2B$ there are sufficient funds to cover all possible loans if banks offer the lowest possible deposit rate $r_t^{D,j} \geq \overline{r}$. In the future, we intend to consider the case where there are insufficient funds.

In Section 2 we documented important differences in the cost structure of banks. Based on this evidence, we assume that banks pay proportional non-interest expenses (net of non-interest income) that differ across banks of different sizes, which we denote c^θ. Here we assume that all national and regional banks face the same costs c^n and c^r, respectively. The cost for fringe bank i is denoted c_i^f which is drawn from a distribution with cdf $\Xi(c^f)$. For simplicity, we assume costs are constant over the lifespan of the bank and they are identical across regions.

There is limited liability on the part of banks. A bank that has negative profits can exit, in which case it receives value zero. We assume that if a national bank exits, it must exit both regions.

Entry costs for the creation of national and regional banks are denoted by $\kappa^n \geq \kappa^r \geq 0$.[15] We normalize the cost of creating a fringe bank to zero. Every period a large number of potential entrants M make the decision to enter the market or not. We assume that each entrant satisfies a zero expected discounted profits condition. To simplify the analysis, we assume that fringe banks can enter the market only if they have a non interest cost greater than those of incumbents. This assumption makes the computation much easier since the only relevant variable to predict the number of active fringe banks is the threshold of the active bank with the highest cost. Provided the cost of entering as a fringe bank is zero, in any given period, there are M fringe banks potentially ready to extend loans.

We denote the industry state by

$$\mu_t = \{N_t(n,\cdot), N_t(r,e), N_t(r,w), N_t(f,e), N_t(f,w)\}, \tag{2}$$

where the 5 elements of μ_t are simply the number of *active* banks by class and region. For example, if in period t, there is only one "national" and one "regional" bank in the west region, as well as 3000 fringe banks in the east and 2500 in the west, the distribution will be equal to $\mu_t = \{1, 0, 1, 3000, 2500\}$.

3.4 Information

There is asymmetric information on the part of borrowers and lenders. Only borrowers know the riskiness of the project they choose (R) and their outside option (ω). All other information is observable.

3.5 Timing

At the beginning of period t,

1. Starting from beginning of period state (μ_t, z_t, s_t), borrowers draw ω_t.

[15] As in Pakes and McGuire [26] we will assume that these costs become infinite after a certain number of firms of the given type are in the market.

2. National and regional banks choose how many loans $\ell_{i,t}(\theta, j)$ to extend and how many deposits $d_{i,t}(\theta, j)$ to accept given depositors choices.

3. Each fringe bank observes the total loan supply of dominant banks and all other fringe banks (that jointly determine the loan interest rate $r_t^{L,j}$) and simultaneously decide whether to extend loans or not. Borrowers in region j choose whether or not to undertake a project of technology R_t^j.

4. Return shocks z_{t+1} and s_{t+1} are realized, as well as idiosyncratic borrower shocks.

5. Exit and entry decisions are made in that order. Entry occurs sequentially (one bank after another).

6. Households pay taxes τ_{t+1} and consume.

4 Industry Equilibrium

Since we will use recursive methods to define an equilibrium, let any variable a_t be denoted a and a_{t+1} be denoted a'.

4.1 Borrower Decision Making

Starting in state z, borrowers take the loan interest rate $r^{L,j}$ as given and choose whether to demand a loan and if so, what technology R^j to operate. Specifically, if a borrower in region j chooses to participate, then given limited liability his problem is to solve:

$$v(r^{L,j}, z, s) = \max_{R^j} E_{z',s'|z,s}\Big[p^j(R^j, z', s')\left(z'R^j - r^{L,j}\right)\Big]. \tag{3}$$

Let $R(r^{L,j}, z, s)$ denote the borrower's decision rule that solves (3). We assume that the necessary and sufficient conditions for this problem to be well behaved are satisfied. The borrower chooses to demand a loan if

$$v(r^{L,j}, z, s) \geq \omega. \tag{4}$$

In an interior solution, the first order condition is given by

$$E_{z',s'|z,s}\Big\{\underbrace{p^j(R^j, z', s')z'}_{(+)} + \underbrace{\frac{\partial p^j(R^j, z', s')}{\partial R^j}}_{(-)}\left[z'R^j - r^{L,j}\right]\Big\} = 0 \tag{5}$$

The first term is the benefit of choosing a higher return project while the second term is the cost associated with the increased risk of failure.

To understand how bank lending rates influence the borrower's choice of risky projects, one can totally differentiate (5) with respect to $r^{L,j}$

$$\begin{aligned} 0 \;\; = \;\; & E_{z',s'|z,s}\Bigg\{\frac{\partial p^j(R^{j*}, z', s')}{\partial R^{j*}}\frac{dR^{j*}}{dr^{L,j}}z' + \frac{\partial^2 p^j(R^{j*}, z', s')}{\left(\partial R^{j*}\right)^2}\left[z'R^{j*} - r^{L,j}\right]\frac{dR^{j*}}{dr^{L,j}} \\ & + \frac{\partial p^j(R^{j*}, z', s')}{\partial R^{j*}}\left[z'\frac{dR^{j*}}{dr^{L,j}} - 1\right]\Bigg\} \end{aligned}$$

where $R^{j*} = R^j(r^{L,j}, z)$. But then

$$\frac{dR^{j*}}{dr^{L,j}} = \frac{E_{z',s'|z,s}\left[\frac{\partial p^j(R^{j*},z',s')}{\partial R^{j*}}\right]}{E_{z',s'|z,s}\left\{\frac{\partial^2 p^j(R^{j*},z',s')}{(\partial R^{j*})^2}\left[z'R^{j*} - r^{L,j}\right] + 2\frac{\partial p^j(R^{j*},z',s')}{\partial R^{j*}}z'\right\}} > 0 \tag{6}$$

since both the numerator and the denominator are strictly negative (the denominator is negative by virtue of the sufficient conditions). Thus a higher borrowing rate implies the borrower takes on more risk. Boyd and De Nicolo [9] call $\frac{dR^{j*}}{dr^{L,j}} > 0$ in (6) the "risk shifting effect". Risk neutrality and limited liability are important for this result.

An application of the envelope theorem implies

$$\frac{\partial v(r^{L,j}, z, s)}{\partial r^{L,j}} = -E_{z',s'|z,s}[p^j(R^j, z', s')] < 0. \tag{7}$$

Thus, participating borrowers are worse off the higher are borrowing rates. This has implications for the demand for loans determined by the participation constraint. In particular, since the demand for loans is given by

$$L^{d,j}(r^{L,j}, z, s) = B \cdot \int_{\underline{\omega}}^{\overline{\omega}} 1_{\{\omega \leq v(r^{L,j},z,s)\}} d\Upsilon(\omega), \tag{8}$$

then (7) implies $\frac{\partial L^{d,j}(r^{L,j},z,s)}{\partial r^{L,j}} < 0$.

4.2 Depositor Decision Making

If $r^{D,j} = \overline{r}$, then a household would be indifferent between matching with a bank and using the autarkic storage technology so we can assign such households to a bank. If it is to match directly with a borrower, the depositor must compete with banks for the borrower. Hence, in successful states, the household cannot expect to receive more than the bank lending rate $r^{L,j}$ but of course could choose to make a take-it-or-leave-it offer of their unit of a good for a return $\widehat{r} < r^{L,j}$ and hence entice a borrower to match with them rather than a bank. Given state contingent taxes $\tau(\mu, z, s, z', s')$, the household matches with a bank if possible and strictly decides to remain in autarky otherwise provided

$$\begin{aligned} U \equiv \quad & E_{z',s'|z,s}\left[u(1 + \overline{r} - \tau(\mu, z, z', s'))\right] > \\ & \max_{\widehat{r} < r^{L,j}} E_{z',s'|z,s}\Big[p^j(\widehat{R}^j, z', s')u(1 + \widehat{r} - \tau(\mu, z, s, z', s')) \\ & + (1 - p^j(\widehat{R}^j, z', s'))u\left(1 - \lambda - \tau(\mu, z, s, z', s')\right)\Big] \equiv U^E. \end{aligned} \tag{9}$$

If this condition is satisfied, then the total supply of deposits in region j is given by

$$D^{s,j} = \sum_{\theta} \sum_{i=1}^{N(\theta,j)} d_i(\theta, j) \leq H \tag{10}$$

Condition (9) makes clear the reason for a bank in our environment. By matching with a large number of borrowers, the bank can diversify the risk of project failure and this is valuable to risk averse households. It is the loan side uncertainty counterpart of a bank in Diamond and Dybvig [13].

4.3 Incumbent Bank Decision Making

An incumbent bank i of type (θ, j) chooses loans $\ell_i(\theta, j)$ in order to maximize profits and chooses whether to exit $x_i(\theta, j)$ after the realization of the aggregate shock z' and the regional shock s'.[16]

It is simple to see that no bank would ever accept more total deposits than it makes total loans.[17] Further, the deposit rate $r^{D,j} = \overline{r}$. [18] Simply put, a bank would not pay interest on deposits that it doesn't lend out and with excess supply of funds, households are forced to their reservation value associated with storage.

Let $\sigma_{-i} = (\ell_{-i}, x_{-i}, e)$ denote the industry state dependent lending, exit, and entry strategies of all other banks. Limited liability and the absence of an interbank market implies a bank will exit if its end-of-period-profits are negative. The end-of-period realized profits in state (z', s') for bank i of type (θ, j) with cost c^θ extending loans ℓ_i starting in state (μ, z, s) is given by:

$$\begin{aligned}\pi_{\ell_i(\theta,j)}(\theta, j, c^\theta, \mu, z, s, z', s'; \sigma_{-i}) \equiv \Big\{ & p^j(R^j, z', s')(1 + r^{L,j}) + (1 - p^j(R^j, z', s'))(1 - \lambda) \\ & -(1 + \overline{r}) - c^\theta \Big\} \ell_i(\theta, j).\end{aligned}$$

The first two terms represent the net return the bank receives from successful and unsuccessful projects respectively and the last terms correspond to its costs.

Differentiating with respect to ℓ_i we obtain

$$\frac{d\pi^j}{d\ell_i} = \big[\underbrace{p^j r^{L,j} - (1 - p^j)\lambda - \overline{r} - c^\theta}_{(+) \text{ or } (-)}\big] + \ell_i \Big[\underbrace{p^j}_{(+)} + \underbrace{\frac{\partial p^j}{\partial R^j} \frac{\partial R^j}{\partial r^{L,j}} (r^{L,j} + \lambda)}_{(-)}\Big] \underbrace{\frac{dr^{L,j}}{d\ell_i}}_{(-)}. \tag{11}$$

The first bracket represents the marginal change in profits coming from extending an extra unit of loans. The the second bracket corresponds to the marginal change in profits due to a bank's influence on the interest rate it faces. This term will reflect the bank's market power: for dominant banks $\frac{dr^{L,j}}{d\ell_i} < 0$ while for fringe banks $\frac{dr^{L,j}}{d\ell_i} = 0$.

The value function of a "national" incumbent bank i at the beginning of the period is given by

$$\begin{aligned}V_i(n, \cdot, \mu, z, s; \sigma_{-i}) = \max_{\{\ell_i(n,j)\}_{j=e,w}} \quad & E_{z',s'|z,s} \Big[\textstyle\sum_j \pi_{\ell_i(n,j)}(n, j, c^n, \mu, z, s, z', s'; \sigma_{-i}) \\ & + \beta V_i(n, \cdot, \mu', z', s'; \sigma_{-i}) \Big] (1 - x_i(n, \cdot))\end{aligned} \tag{12}$$

[16] In Allen and Gale (2004), banks compete Cournot in the deposit market and offer borrowers an incentive compatible loan contract that induces them to choose the project R which maximizes the bank's objective. As in Boyd and De Nicolo (2005), we assume that banks compete Cournot in the loan market and offer borrowers an incentive compatible loan contract which is consistent with the borrower's optimal decision rule.

[17] Suppose not and $d_i > \ell_i$. The net cost of doing so is $r^{D,j} \geq 0$ while the net gain on $d_i - \ell_i$ is zero, so it is weakly optimal not to do so.

[18] Suppose not and some bank is paying $r^{D,j} > \overline{r}$. Then the bank can lower $r^{D,j}$, still attract deposits since H=2B and make strictly higher profits, so it is strictly optimal not to do so.

subject to

$$\sum_{\theta} \sum_{i=1}^{N(\theta,j)} \ell_i(\theta, j, \mu, s, z; \sigma_{-i}) - L^{d,j}(r^{L,j}, z, s) = 0, \tag{13}$$

where $L^{d,j}(r^{L,j}, z, s)$ is given in (8) and

$$x_i(n, \cdot, \mu, z, s, z', s'; \sigma_{-i}) = \begin{cases} 1 & \text{if } \sum_{j=e,w} \pi_{\ell_i(n,j)}(n, \cdot, c^n, \mu, z, s, z', s'; \sigma_{-i}) < 0 \\ 0 & \text{if } \sum_{j=e,w} \pi_{\ell_i(n,j)}(n, \cdot, c^n, \mu, z, s, z', s'; \sigma_{-i}) \geq 0 \end{cases}. \tag{14}$$

Constraint (13), which is simply the loan market clearing condition, is imposed as a consistency condition due to the Cournot assumption whereby a national bank realizes its loan supply will influence the interest rate $r^{L,j}$. The exit decision rule in (14) is a consequence of the limited liability and no interbank loan market assumptions of the model.

The value function of a "regional" incumbent bank i in region j at the beginning of the period is given by

$$\begin{aligned} V_i(r, j, \mu, z, s; \sigma_{-i}) = \max_{\ell_i(r,j)} \quad & E_{z',s'|z,s}\Big[\pi_{\ell_i(r,j)}(r, j, c^r, \mu, z, s, z', s'; \sigma_{-i}) \\ & + \beta V_i(r, j, \mu', z', s'; \sigma_{-i})\Big](1 - x_i(r, j)) \end{aligned} \tag{15}$$

subject to (13) and where

$$x_i(r, j, \mu, z, s, z', s'; \sigma_{-i}) = \begin{cases} 1 & \text{if } \pi_{\ell_i(r,j)}(r, j, c^r, \mu, z, s, z', s'; \sigma_{-i}) < 0 \\ 0 & \text{if } \pi_{\ell_i(r,j)}(r, j, c^r, \mu, z, s, z', s'; \sigma_{-i}) \geq 0 \end{cases}. \tag{16}$$

Unlike a national bank which can transfer profits across regions, a regional bank without access to an interbank market or other assets, will exit if its profits are negative to take advantage of limited liability.

The problem of fringe bank i in region j is different from that of a dominant national or regional bank. When fringe banks make their loan supply decision, dominant banks have already made their move and since fringe banks are sufficiently small they take $r^{L,j}$ as given. As discussed following equation (11), in this case the profit function is linear in $\ell_i(f, j)$ so the quantity constraint $\ell_i(f, j) \leq \bar{d}$ will in general bind the loan decision. In particular, the value function of an incumbent fringe bank which drew cost c_i^f at entry and takes the $r^{L,j}$ which solves (13) is given by

$$\begin{aligned} V_i(f, j, c_i^f \mu, z, s; \sigma_{-i}) = \max_{\ell_i(f,j) \leq \bar{d}} \quad & E_{z',s'|z,s}\Big[\pi^j_{\ell_i(f,j)}(f, j, c_i^f, \mu, z, s, z', s'; \sigma_{-i}) \\ & + \beta V_i(f, j, c_i^f \mu', z', s'; \sigma_{-i})\Big](1 - x_i(f, j)) \end{aligned} \tag{17}$$

where

$$x_i(f, j, c_i^f, \mu, z, s, z', s'; \sigma_{-i}) = \begin{cases} 1 & \text{if } \pi_{\ell_i(f,j)}(f, j, c_i^f, \mu, z, s, z', s'; \sigma_{-i}) < 0 \\ 0 & \text{if } \pi_{\ell_i(f,j)}(f, j, c_i^f, \mu, z, s, z', s'; \sigma_{-i}) \geq 0 \end{cases}. \tag{18}$$

Since the loan interest rate is taken as given, the technology is linear in loans made and the fringe bank's decision is simply whether to bring all their available funds to the market or not, i.e. $\ell_i(f,j) \in \{0, \overline{d}\}$. Total loan supply by fringe banks in region j will be

$$L^s(f,j,\mu,z,s;\sigma_{-i}) = M\Xi\left(\overline{c}^j(\mu,z,s;\sigma_{-i})\right)\bar{d}.$$

where $\overline{c}^j(\cdot)$ denotes the highest cost such that a fringe bank will choose to offer loans in region j. More specifically, for a given loan interest rate, fringe banks will choose to offer loans whenever expected profits are greater than or equal to zero. This implies that $\overline{c}^j(\mu,z,s;\sigma_{-i})$ solves

$$E_{z',s'|z,s}\max\left\{\pi_{\bar{d}}(f,j,\overline{c}^j,\mu,z,s,z',s';\sigma_{-i}),0\right\} = 0. \tag{19}$$

The new distribution of banks after entry and exit μ' is determined by the number of active banks of type (θ,j) that remain active after the exit stage $N^x(\theta,j)$ and the number of entrants $N^e(\theta,j)$ of type (θ,j) as follows:

$$\begin{aligned}\mu' \quad &= \quad \{N^x(n,\cdot)+N^e(n,\cdot), N^x(r,e)+N^e(r,e),\\ &\quad\quad N^x(r,w)+N^e(r,w), N^x(f,e)+N^e(f,e), N^x(f,w)+N^e(f,w)\}.\end{aligned} \tag{20}$$

The number of banks of type (θ,j) in the industry after exit is given by

$$N^x(\theta,j) = \sum_{i=1}^{N(\theta,j)}\left(N(\theta,j) - x_i(\theta,j,\mu,z,s,z',s';\sigma_{-i})\right). \tag{21}$$

The number of fringe banks in region j after exit can be defined as a function of the bank with the highest cost among the survivors. More specifically, let $c^{x,j}$ be the value of c_i^f that solves $\pi_{\bar{d}}(f,j,c^{x,j},\mu,z,s,z',s';\sigma_{-i}) = 0$. Then, the number of fringe banks in region j after exit is:

$$N^x(f,j) = M\cdot\min\left\{\Xi\left(c^{x,j}(f,j,\mu,s,z,z',s';\sigma_{-i})\right), \Xi\left(\overline{c}^j(f,j,\mu,s,z,z',s';\sigma_{-i})\right)\right\}.$$

Thus, the number of fringe banks that exit in region j is $N(f,j) - N^x(f,j)$.

4.4 Entrant Bank Decision Making

In each period, new banks of type θ can enter the industry by paying the setup cost κ^θ. They will enter the industry if the net present value of entry is nonnegative. For example, taking the entry and exit decisions by other banks as given, a potential regional entrant in the west region will choose $e_i\left(r,w,\{\cdots,N^x(r,w)+N^e(r,w),\cdots\},z',s'\right) = 1$ if

$$\beta V_i(r,j,\{\cdots,N^x(r,w)+N^e(r,w)+1,\cdots\},z',s';\sigma_{-i}) - \kappa^r > 0. \tag{22}$$

4.5 Definition of Equilibrium

A pure strategy Markov Perfect Equilibrium (MPE) is a set of functions $\{v(r^{L,j},z,s)$ and $R(r^{L,j},z,s)\}$ describing borrower behavior, a set of functions $\{V_i(\theta,j,\mu,z,s;\sigma_{-i}),$

$\ell_i(\theta, j, \mu, z, s; \sigma_{-i})$, $x_i(\theta, j, \mu, z, s, z', s'; \sigma_{-i})$, and $e_i(\theta, j, \mu, z', s'; \sigma_{-i})\}$ describing bank behavior, a loan interest rate $r^{L,j}(\mu, z, s)$ for each region, a deposit interest rate $r^D = \overline{r}$, an industry state μ, a function describing the number of entrants $N^e(\theta, j, \mu, z')$, and a tax function $\tau(\mu, z, s, z', s')$ such that:

1. Given a loan interest rate $r^{L,j}$, $v(r^{L,j}, z, s)$ and $R(r^{L,j}, z, s)$ are consistent with borrower's optimization in (3) and (4).

2. For any given interest rate $r^{L,j}$, loan demand $L^{d,j}(r^{L,j}, z, s)$ is given by (8).

3. At $r^D = \overline{r}$, the household deposit participation constraint (9) is satisfied.

4. Given the loan demand function, the value of the bank $V_i(\theta, j, \mu, z, s; \sigma_{-i})$, the loan decision rules $\ell_i(\theta, j, \mu, z, s; \sigma_{-i})$, and exit rules $x_i(\theta, j, \mu, z, s, z', s'; \sigma_{-i})$, are consistent with bank optimization in (12), (14), (15), (16), (17) and (18).

5. The entry decision rules $e_i(\theta, j, \mu, z', s'; \sigma_{-i})$ are consistent with bank optimization in (22).

6. The law of motion for the industry state (20) is consistent with entry and exit decision rules.

7. The interest rate $r^{L,j}(\mu, z, s)$ is such that the loan market (13) clears. That is,

$$L^{d,j}(r^{L,j}, z, s) = B \cdot \int_{\underline{\omega}}^{\overline{\omega}} 1_{\{\omega \leq v(r^{L,j}, z, s)\}} d\Upsilon(\omega) = \sum_{\theta} \sum_{i=1}^{N(\theta,j)} \ell_i(\theta, j, \mu, s, z; \sigma_{-i} = L^{s,j}(\mu, s, z; \sigma_{-i}).$$

8. Across all states (μ, z, s, z', s'), taxes cover deposit insurance:

$$\tau(\mu, z, s, z', s') = \sum_{\theta,j} \sum_{i=1}^{N(\theta,j)} x_i(\theta, j, \mu, z, s, z', s'; \sigma_{-i}) \pi_{\ell_i(\theta,j)}(\theta, j, \mu, z, s, z', s'; \sigma_{-i}).$$

5 Calibration

We calibrate the model to match the key statistics of the U.S. banking industry described in Section 2. A model period is considered to be one year.

First, we consider the stochastic process for aggregate technology shocks $F(z', z)$. To calibrate this process, we use the NBER recession dates and create a recession indicator. More specifically, for a given year, the recession indicator takes a value equal to one if two or more quarters in that year were dated as part of a recession. The correlation of this indicator with HP filtered GDP equals -0.87. Then, we identify years where the indicator equals one with our periods of $z = z_b$ and construct a transition matrix. In particular, the maximum likelihood estimate of F_{kj}, the $(j, k)th$ element of the aggregate state transition matrix, is the ratio of the number of times the economy switched from state j to state k to the number of times the economy was observed to be in state j. We normalize the value of $z_g = 1$ and choose z_b to match the variance of detrended GDP.

We calibrate r^D using data from the banks' balance sheet. The deposit interest rate is computed as the average ratio of total interest expense over total interest bearing deposits for commercial banks in the US from 1980 to 2008.[19] The discount factor β is set to $1/(1+r^D)$.

Next we consider the stochastic process for the borrower's project. For each borrower in region j, let $y^j = \alpha z' + (1-\alpha)\varepsilon_e - bR^\psi$ where ε_e is drawn from $N(\phi(s'), \sigma_\varepsilon^2)$. The regional shock affects the mean of the idiosyncratic shock through $\phi(s') \in \{-\overline{\phi}, \overline{\phi}\}$. We assume that if $s' = j$, $\phi(s') = \overline{\phi}$ and $\phi(s') = -\overline{\phi}$ otherwise. The borrower's idiosyncratic project uncertainty is iid across agents. We define success to be the event that $y > 0$, so in states with higher z or higher ε_e success is more likely. Then

$$\begin{aligned} p^j(R, z', s') &= 1 - prob(y \leq 0 | R, z', s') \\ &= 1 - prob\left(\varepsilon_e \leq \frac{-\alpha z' + bR^\psi}{(1-\alpha)}\right) \\ &= \Phi\left(\frac{\alpha z' - bR^\psi}{(1-\alpha)}\right) \end{aligned} \tag{23}$$

where $\Phi(x)$ is a normal cumulative distribution function with mean $\phi(s')$ and variance σ_ε^2. We assume that s follows a Markov process and that the transition matrix has diagonal values equal to $\overline{G}$.

For the borrower's outside option, we assume that $\Upsilon(\omega)$ corresponds to the uniform distribution $[\underline{\omega}, \overline{\omega}]$ and set $\underline{\omega} = 0$. We let consumer's preferences be given by $u(x) = \frac{x^{1-\sigma}}{1-\sigma}$ and set $\sigma = 2$, a standard value in the macro literature. At this level of risk aversion the consumer participation constraint is satisfied. The mass of borrowers is normalized to one.

We identify the "national" bank with the top 4 banks (when sorted by assets), the "regional" banks with the top 1% banks (also when sorted by assets and excluding the top 4 banks) and the fringe banks with the bottom 99% of the bank asset distribution. Dominant bank's net non interest expenses are calibrated using the information in Table 5. The value of $c^n = 0.0303$ and $c^r = 0.0204$. We assume that c^f is distributed exponentially with location parameter equal to μ_c. Finally, we assume that $\kappa^r = \kappa^n = \kappa$ and as in Pakes and McGuire [26] we restrict the number of banks by setting the entry cost to a prohibitively high number if the number of incumbents after entry and exit exceeds a given number, in our application we choose one (i.e. there will be at most one national bank and one regional bank per region).

We are left with eleven parameters to calibrate: $\{\overline{G}, \overline{\phi}, \sigma_\varepsilon, \alpha, b, \psi, \lambda, \overline{\omega}, \kappa, \bar{d}, \mu_c\}$. Except for one data moment, we use the data for commercial banks described in Section 2 to pin down these parameters. The extra moment - the average real equity return as reported by Diebold and Yilmaz [15] - is added to shed light on the borrower's return R^*. The set of targets includes the average default frequency, borrower return, exit rate (by liquidation), loan return, charge-off rate, the relative loan returns across banks of different sizes, the relative size of delinquency rates, and the loan market share of regional and fringe banks in December of 2006 (the last observation in our sample before the crisis started).

[19] Source: FDIC, Call and Thrift Financial Reports, Balance Sheet and Income Statement (http://www2.fdic.gov/hsob/SelectRpt.asp?EntryTyp=10). The nominal interest rate is converted to a real interest rate by using the US GDP deflator (expected inflation is computed as the average of previous four quarters inflation).

Table 6 shows the calibrated parameters.

Table 6: Model Parameters

Parameter		Value	Target
Mass of Borrowers	B	1	Normalization
Mass of Households	H	$2B$	Assumption
Depositors' Preferences	σ	2	Participation Const.
Aggregate Shock in Good State	z_g	1	Normalization
Aggregate Shock in Bad State	z_b	0.959	Std. GDP
Transition Probability	$F(z_g, z_g)$	0.85	NBER data
Transition Probability	$F(z_b, z_b)$	0.5	NBER data
Deposit Interest Rate (%)	r^D	1.12	Interest Expense
Discount Factor	β	0.99	Interest Expense
Net Non Int. Exp. Nat. Bank (%)	c^n	3.03	Net Non-Interest Expense
Net Non Int. Exp. Reg. Bank (%)	c^r	2.04	Net Non-Interest Expense
Weight Aggregate Shock	α	0.88	Default Frequency
Success Probability Parameter	b	3.77	Borrower Return
Success Probability Parameter	ψ	0.78	Bank Exit Rate
Volatility Entrep. Dist.	σ_ε	0.06	Loan Return
Loss Rate	λ	0.21	Charge off Rate
Max. Reservation Value	$\overline{\omega}$	0.22	Loan Return Top 4 to Top 1%
Regional Shock	$\overline{\phi}$	0.05	Delinq. Rate Top 4 to Top 1%
Persistence Regional Shock	$\overline{G}$	0.92	Loan Return Top 1% to Bottom 99%
Entry Cost	κ	0.29	Delinq. Rate Top 1% to Bottom 99%
Dist. Net Non Int. Exp. Fringe	μ_c	0.01	Loan Market Share Top 4 Banks
Deposit Fringe Banks	$\overline{d} \times M$	0.08	Loan Market Share Bottom 99%

Table 7 displays the targeted moments of the model and a comparison with the data. We use the following definitions to connect the model to some of the variables we presented in the data section. In particular,

- Cross-Sectional Average Default frequency (measured by the 90 day delinquency rate): $1 - p(R^*, z', s')$
- Cross-Sectional Average Borrower return: $p(R^*, z', s')(z'R^*)$
- Bank exit rate: $\frac{\sum_{\theta,j} N(\theta,j) - N^x(\theta,j)}{\sum_{\theta,j} N(\theta,j)}$
- Cross-Sectional Average Loan return: $p(R^*, z', s')r^L$
- Cross-Sectional Average Loan Charge-off rate $(1 - p(R^*, z', s'))\lambda$

Table 7: Model and Data Moments

Moment (%)	Model	Data
Default Frequency	1.95	2.39
Borrower Return	13.25	12.94
Loan Return	6.71	5.99
Charge-Off Rate	0.74	0.93
Exit Rate	1.29	1.13
Loan Return Top 4 to Top 1%	97.98	93.39
Delinq. Rate Top 4 to Top 1%	80.96	79.96
Loan Return Top 1% to Bottom 99%	98.14	88.17
Delinq. Rate Top 1% to Bottom 99%	76.72	64.32
Market Share Top 1%	41.45	40.14
Market Share Bottom 99%	24.16	25.33

The model does a good job in matching most of the targeted moments. However, it generates a 12% higher loan return and a 20% lower charge off rate than in the data.

6 Results

For the parameter values in Table 6, we find an equilibrium where national banks do not exit while regional banks and fringe banks exit in bad times (though fringe market share takes up some of the slack of the regional bank market share in bad times). In particular, we find: (i) if there is no "regional" bank in one of the regions ($N(r, j) = 0$ for $j = e$ and/or $j = w$) and $z' = z_g$ there is entry by a "regional" bank in the region with $s = j$ (this is on-the-equilibrium path); (ii) if $N(n, \cdot) = 0$, there is entry by a "national" bank (this is off-the-equilibrium path); (iii) a "regional" bank in region j exits when the regional shock changes $s = j$ and $s' \neq j$ and there is a recession $z' = z_b$; (iv) a "national" bank exits if $z = z_g$, $z' = z_b$, $s = j$, $s' \neq j$ and there is no regional bank in region j (i.e there are both aggregate and regional downturns in the region where it is the only dominant bank, but again this is also off-the-equilibrium path).

To understand the equilibrium, we first describe borrower decisions. Figure 14 shows the borrower's optimal choice of project riskiness $R^*(r^{L,j}, z, s)$ and the inverse demand function associated with $L^d(r^{L,j}, z, s)$ for region 1 (those corresponding to region 2 are similar).

Figure 14: Borrower's Project and Loan Inverse Demand

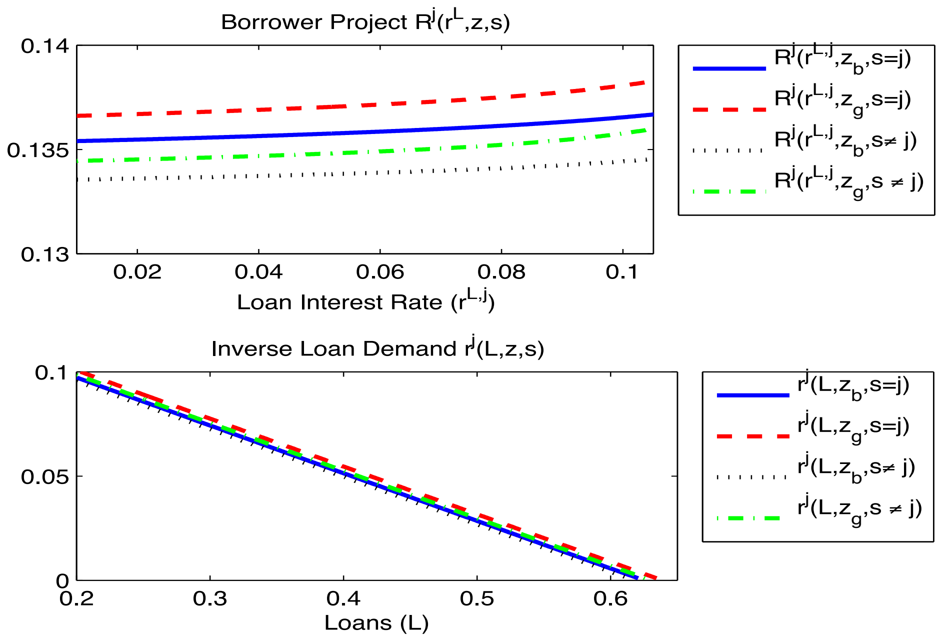


Figure 14 shows that the borrower's optimal project R is an increasing function of the loan interest rate $r^{L,j}$. Moreover, given that the value of the borrower is decreasing in $r^{L,j}$, loan demand is a decreasing function of $r^{L,j}$.

In Tables (8) to (13) and Figure 15 we provide a description of borrower and bank decision rules and their implications for loan supply, loan interest rates, borrower returns and success probabilities. Note that while these are equilibrium functions not every state is necessarily on-the-equilibrium path (starting with Table 15 we evaluate the behavior of the model on-the-equilibrium path).

Figure 15: Competitive Fringe Thresholds $\overline{c}^j$ and $c^{x,j}$

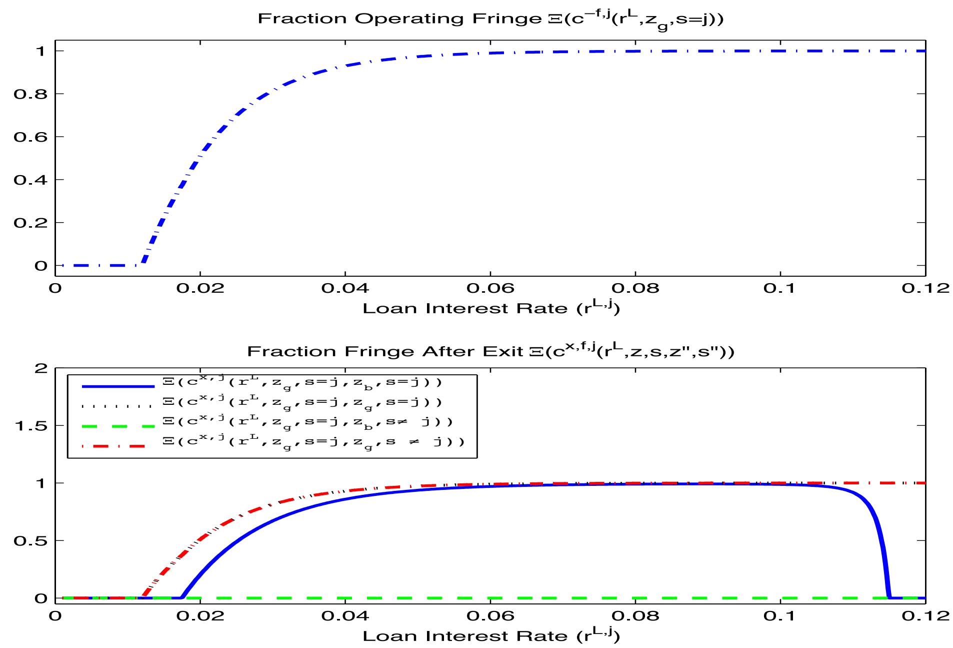


Figure 15 shows how the the fraction of active fringe banks changes with the loan interest rate for the case of $z = z_g$ and $s = j$. In the top panel, we observe the fraction of fringe banks that decide to extend loans $\Xi(\overline{c}^j)$. This fraction is increasing in the loan interest rate since expected profits are increasing in $r^{L,j}$ when $z = z_g$ and $s = j$ (i.e. the direct positive effect of $r^{L,j}$ on profits exceeds the indirect negative effect on p^j in this state for the benchmark parameters). In the bottom panel, we observe the fraction of fringe banks that survive after the exit stage $\Xi(c^{x,j})$. If the aggregate shock stays in z_g (i.e. $z' = z_g$), all the fringe banks that extended loans will remain active, $c^{x,j} = \overline{c}^j$. On the other hand, a fraction of the fringe banks will exit when $z' = z_b$ and $s' = j$ and the entire fringe sector will disappear when $z' = z_b$ and $s' \neq j$. This latter case is very unlikely (probability 0.004) since aggregate and regional shocks are highly persistent.

Table 8 provides the loan decision rules for national and regional banks. The first four columns correspond to the loans made by a national bank in region 1 and 2 respectively. The next two columns correspond to the regional bank in region 1 and the last two columns to the regional bank in region 2. We observe that banks offer more loans when $z = z_g$ than when z_b. Independent of the aggregate shock, national banks offer more loans in regions where they have more market power.

Table 8: Bank Loan Decision Rules $\ell(n, j, \mu, z, s; \sigma_{-n})$

	$\ell(n,e)$		$\ell(n,w)$		$\ell(r,e)$		$\ell(r,w)$	
μ	(z_b, e)	(z_g, e)	(z_b, e)	(z_g, e)	(z_b, e)	(z_g, e)	(z_b, e)	(z_g, e)
$\{0,1,0,\cdot\}$	-	-	-	-	0.205	0.214	-	-
$\{0,1,1,\cdot\}$	-	-	-	-	0.205	0.214	0.200	0.208
$\{1,0,0,\cdot\}$	0.182	0.192	0.178	0.187	-	-	-	-
$\{1,1,0,\cdot\}$	0.102	0.060	0.178	0.187	0.153	0.182	-	-
$\{1,1,1,\cdot\}$	0.071	0.020	0.102	0.107	0.169	0.202	0.148	0.152
	(z_b, w)	(z_g, w)	(z_b, w)	(z_g, w)	(z_b, w)	(z_g, w)	(z_b, w)	(z_g, w)
$\{0,1,0,\cdot\}$	-	-	-	-	0.200	0.208	-	-
$\{0,1,1,\cdot\}$	-	-	-	-	0.200	0.208	0.205	0.214
$\{1,0,0,\cdot\}$	0.178	0.187	0.182	0.192	-	-	-	-
$\{1,1,0,\cdot\}$	0.102	0.107	0.025	0.192	0.148	0.152	-	-
$\{1,1,1,\cdot\}$	0.102	0.107	0.071	0.020	0.148	0.152	0.169	0.202

The optimal exit rule implies that there is exit for regional banks in region j when the regional bank receives the negative regional shock ($s = j$ and $s' \neq j$) in a recession $z' = z_b$ (on-the-equilibrium path). That is,

$$x_i(r, j, \mu, z, s, z', s') = \begin{cases} 1 & \text{if } z' = z_b,\ s = j \text{ and } s' \neq j \\ 0 & \text{otherwise} \end{cases} . \tag{24}$$

Moreover, there is exit by national banks when we move into a recession ($z = z_g$ and $z' = z_b$) and a bad regional shock arrives in region j ($s' \neq j$) if there is no active regional bank in region j (off-the-equilibrium path). That is,

$$x_i(n, j, \mu, z, s, z', s') = \begin{cases} 1 & \text{if } N(r, j) = 0,\ z = z_g,\ z' = z_b,\ s = j \text{ and } s' \neq j \\ 0 & \text{otherwise} \end{cases} . \tag{25}$$

Exit occurs for a regional bank when its regional shock turns bad during a recession. This happens because borrowers take on more risk in good times and project failure is more likely in bad states. The national bank loan decision lowers realized profits of regional banks enough to induce them to exit in order to become a regional monopoly next period. To see this dynamic aspect of strategic behavior, we compare decision rules on an equilibrium path of the benchmark dynamic model versus a static economy evaluated at $\mu = \{1, 1, 1, \cdot\}$, $z = z_g$, $s = e$ in the following table.

Table 9: Dynamic vs Static Model

	Loan Decision Rules $\ell(\theta, j, \mu, z, s)$			
Model	$\ell(n, e, \cdot)$	$\ell(n, w, \cdot)$	$\ell(r, e, \cdot)$	$\ell(r, w, \cdot)$
Dynamic	0.020	0.106	0.203	0.152
Static	0.110	0.106	0.156	0.152
	Exit Rule $x(\theta, j, \mu, z, s, z' = z_b, s' = w)$			
Model	$x(n, \cdot)$		$x(r, e, \cdot)$	$x(r, w, \cdot)$
Dynamic	0		1	0
Static	1		1	0

As evident in Table 9, the national bank offers less loans in the dynamic case relative to the static case to reduce its exposure to $z' = z_b$ and $s' = w$ in order to protect its charter value. Its reduction in loans induces an increase in $r^{L,e}$ leading the borrower to increase R^e which in turn decreases the success probability p^e. The national bank's expected profits are lower but its exit probability is zero. In best response to national banks, regional banks increase their loans. This increases the success probability p^e and expected profits $E[\pi(r, e)]$ but since regional banks are not geographically diversified they still exit.

Tables 10, 11 and 12 display aggregate loan supply, the loan interest rate and borrower project as a function of the industry state μ and the aggregate state z (i.e. across market structure and the business cycle) for each region in the case that $s = e$.[20] As discussed above, not all cells in Tables 10 through 12 are on-the-equilibrium path. In particular, the equilibrium path corresponds to $\mu = \{1, 0, 0, \cdot\}$, $\mu = \{1, 1, 0, \cdot\}$ and $\mu = \{1, 1, 1, \cdot\}$.

Table 10: Loan Supply $L^{s,j}(\mu, z, s = e)$

	$L^{s,e}(\mu, z, e)$		$L^{s,w}(\mu, z, e)$	
μ	(z_b, e)	(z_g, e)	(z_b, e)	(z_g, e)
$\{0, 1, 0, \cdot\}$	0.281	0.289	-	-
$\{0, 1, 1, \cdot\}$	0.281	0.289	0.276	0.284
$\{1, 0, 0, \cdot\}$	0.257	0.267	0.254	0.262
$\{1, 1, 0, \cdot\}$	0.317	0.330	0.254	0.262
$\{1, 1, 1, \cdot\}$	0.298	0.315	0.325	0.333

[20] The entire table (where $s = w$ as well) is symmetric so we only display this case.

Table 11: Loan Interest Rate $r^{L,j}(\mu, z, s=e)$

	$r^{L,e}(\mu,z,e)$		$r^{L,w}(\mu,z,e)$	
μ	(z_b,e)	(z_g,e)	(z_b,e)	(z_g,e)
$\{0,1,0,\cdot\}$	7.87	8.01	-	-
$\{0,1,1,\cdot\}$	7.87	8.01	7.80	7.92
$\{1,0,0,\cdot\}$	8.41	8.51	8.31	8.41
$\{1,1,0,\cdot\}$	6.74	7.37	8.31	8.41
$\{1,1,1,\cdot\}$	7.08	7.81	6.69	6.78

In Tables 10 and 11 we observe that, conditional on the aggregate state z, less concentration (cases with $N(n,\cdot)+N(r,j)=2$) implies a higher loan supply and a lower loan interest rate $r^{L,j}$. This observation is consistent with Proposition 2 of Boyd and DeNicolo [9]. In particular, they consider exogenous increases in N and find that r^L declines to r^D. We also note that, conditional on the number of banks, the total loan supply is higher in good times ($z=z_g$) than in bad times ($z=z_b$). However, also conditional on the number of banks N, interest rates $r^{L,j}$ are higher in good times than in bad times and by the risk shifting effect in equation (6) the same is true for R^*.

Table 12: Borrower's Project $R^j(\mu, z, s=e)$

	$R^e(\mu,z,e)$		$R^w(\mu,z,e)$	
μ	(z_b,e)	(z_g,e)	(z_b,e)	(z_g,e)
$\{0,1,0,\cdot\}$	13.42	13.45	-	-
$\{0,1,1,\cdot\}$	13.42	13.45	12.49	12.65
$\{1,0,0,\cdot\}$	13.50	13.45	12.49	12.65
$\{1,1,0,\cdot\}$	13.42	13.44	12.49	12.65
$\{1,1,1,\cdot\}$	13.42	13.45	12.48	12.63

Table 12 sheds light on borrower risk profiles R^j. These risk profiles depend on exogenous shocks like s and endogenous variables like $r^{L,j}$. Since there are persistent regional shocks, borrowers in each region effectively display different risk profiles. More specifically, since the probability of observing $s'=j$ (i.e. the end-of-period regional shock being good) depends on the current realization of the regional shock, there are ex-ante differences across borrowers across regions. These ex-ante differences in borrower risk profiles are reflected in the amount of loans that banks extend in each region (conditional on the level of competition and the value of the aggregate shock), the loan interest rate and finally in the overall level riskiness of their projects that borrowers choose in each region as evident in Table 12. For example, when $\mu=\{1,1,1,\cdot\}$ and $z=z_g$ and $s=e$, total loan supply in region e in Table 10 is lower (i.e. equals 0.315) than in region w (which equals 0.333) since the unlikely event that the regional shock changes to $s'=w$ exposes the national bank to a lot of risk in the e region, and hence the national bank lowers its exposure in order to maintain its high charter value. This contributes to making loan rates in the east $r^{L,e} = 7.81$ higher than $r^{L,w} = 6.78$ in Table 11 and hence leads the borrower to choose a riskier project $R^e = 13.45$ than $R^w = 12.63$ in Table 12.

In Table 13, we show the implications of loan interest rates $r^{L,j}$ and borrower project choice R^j on default frequencies $1 - p^j(R(\mu, z, s), z', s')$ across market structure and business cycle (we do not present default frequencies the table for $z' = z_g$ since they are all zero). The table shows that conditional on current state (μ, z, s), the realization of a bad aggregate shock $z' = z_b$ as well as the realization of a negative regional shock imply a higher default frequency. We also observe that more concentrated industries $N^j = 1$ have a higher default frequency across both aggregate states. Moreover, the highest default frequencies are observed at a turning point from good to bad times (i.e. from $z = z_g$ to $z' = z_b$ and $s = j$ to $s \neq j$).

Table 13: Default Freq. across Market Structure and Business Cycle (%)

			Region e		Region w	
μ	z	s	$z' = z_b, s' = e$	$z' = z_b, s' = w$	$z' = z_b, s' = e$	$z' = z_b, s' = w$
$\{0,1,0,\cdot\}$	z_b	e	0.97	25.12	-	-
$\{0,1,0,\cdot\}$	z_b	w	0.01	2.23	-	-
$\{0,1,0,\cdot\}$	z_g	e	7.66	59.49	-	-
$\{0,1,0,\cdot\}$	z_g	w	0.17	10.28	-	-
$\{0,1,1,\cdot\}$	z_b	e	0.97	25.12	2.23	0.01
$\{0,1,1,\cdot\}$	z_b	w	0.01	2.23	25.12	0.97
$\{0,1,1,\cdot\}$	z_g	e	7.66	59.49	10.28	0.17
$\{0,1,1,\cdot\}$	z_g	w	0.17	10.28	59.49	7.66
$\{1,0,0,\cdot\}$	z_b	e	1.13	27.06	2.47	0.01
$\{1,0,0,\cdot\}$	z_b	w	0.01	2.47	27.06	1.13
$\{1,0,0,\cdot\}$	z_g	e	8.78	62.31	11.51	0.21
$\{1,0,0,\cdot\}$	z_g	w	0.21	11.51	62.31	8.78
$\{1,1,0,\cdot\}$	z_b	e	0.72	21.78	2.47	0.01
$\{1,1,0,\cdot\}$	z_b	w	0.01	1.83	54.04	5.86
$\{1,1,0,\cdot\}$	z_g	e	6.53	56.23	11.51	0.21
$\{1,1,0,\cdot\}$	z_g	w	0.11	8.18	62.31	8.78
$\{1,1,1,\cdot\}$	z_b	e	0.78	22.68	1.83	0.01
$\{1,1,1,\cdot\}$	z_b	w	0.01	1.83	22.68	0.78
$\{1,1,1,\cdot\}$	z_g	e	7.27	58.41	8.18	0.11
$\{1,1,1,\cdot\}$	z_g	w	0.11	8.18	58.41	7.27

Table 14 shows the relation between the degree of bank competition measured by the number of dominant banks and important first moments for the economy. More competition (i.e. more active banks) implies a higher loan supply, a lower interest rate on loans, lower bank profit rates, and higher borrower returns. Despite lower interest rates on loans with more competition, exit rates, entry rates and default frequency display a non-linear relation with the number of dominant banks in the market.[21]

Table 14: Model Moments and Market Concentration

	$N(n,\cdot)+\sum_j N(r,j) =$		
Moment Average	1	2	3
Loan interest rate (r^L)	8.36	7.71	7.21
Loan supply	0.51	0.58	0.63
Borrower return	11.82	12.26	13.34
GDP	0.62	0.66	0.72
Exit Rate	0.01	1.26	1.06
Entry Rate	0.12	1.26	1.06
Default frequency	1.20	1.96	1.92
Bank profit rate	6.00	5.27	4.76
Loan return rate	7.91	7.23	6.73
Charge-off rate	0.62	0.79	0.80

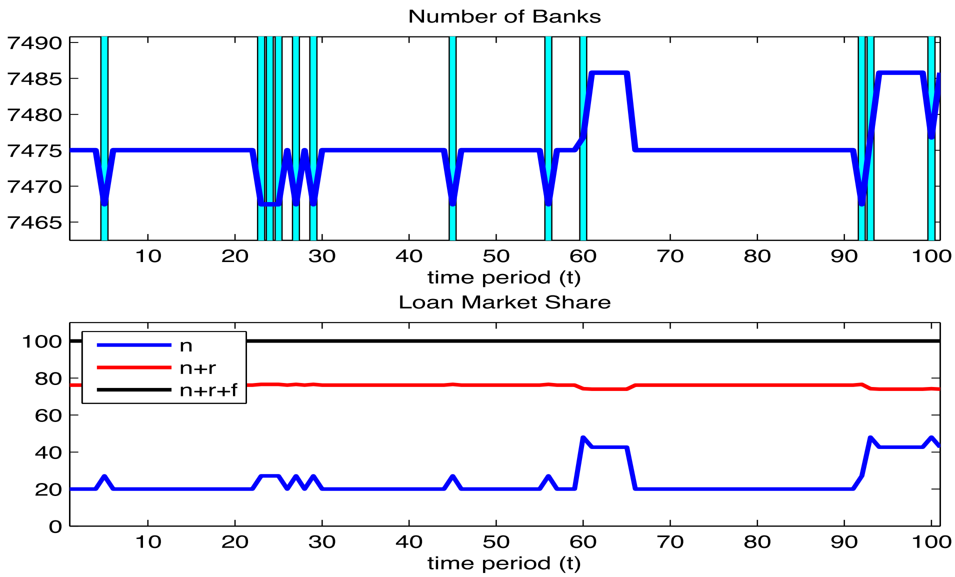
6.1 Industry Dynamics

We simulate a pseudo-panel of firms for 2000 model periods (years) and cut the last 100 periods. In Figure 16, we plot the number of firms and market shares across time noting periods of state z_b.

[21] Aggregate GDP in the model is defined as follows

$$GDP(\mu,z,s,z',s') = \sum_j L^{s,j}(\mu,z,s)\Big\{p^j(\mu,z,s,z',s')(1+z'R) + (1-p^j(\mu,z,s,z',s'))(1-\lambda)\Big\}.$$

Figure 16: Sample Path of Industry Dynamics



Note: Bars indicate periods of $z = z_b$.

As evident in Figure 16, exit is countercyclical and entry procyclical. Also, we see that the fringe accounts for a larger market share when regional banks exit. Finally, the sample exhibits periods of high concentration following recessions. We note that after the end of most recessions there is only a small increase in the number of active banks, while after a few recessions the increase in the number of banks is relatively bigger. This is due to the evolution of the regional shock. In those recessions with no change in the regional shock, there is no exit by regional banks and only a small number of fringe banks exit the market due to a reduction in profits. These fringe banks will be replaced by new fringe banks when the aggregate shock returns to $z = z_g$. On the the other hand, during recessions that happen together with a change in the regional shock, there is exit by regional banks (a dominant bank in the market). If by the time the aggregate shock returns to z_g the regional shock does not go back to its old value, the market in that region will operate only with national and fringe banks. Since national banks operate as the only dominant bank in the region, profits are higher which also induces entry by fringe banks. Hence we observe an increase in market share of national and fringe banks (see for example periods 60 to 67 of the simulation). Once the regional shock returns to its old value, there is entry by regional banks, increasing the number of dominant banks (and thus competition), reducing equilibrium profits in the given region, and generating exit by fringe banks.

6.2 Tests of the Model

We now move on to moments that the model was not calibrated to match, so that these Tables can be considered simple tests of the model. Table 15 provides the correlation between key aggregate variables with GDP.[22] We observe that, as in the data, the model generates countercylical loan interest rates, exit rates, default frequencies, loan returns and charge-off rates. Moreover, the model generates procyclical aggregate loan supply, deposit demand, entry rates and profit rates.

Table 15: Business Cycle Correlations

Variable Correlated with GDP	Model	Data
Loan Interest Rate r^L	-0.32	-0.21
Exit Rate	-0.29	-0.26
Entry Rate	0.05	0.22
Loan Supply	0.59	0.58
Deposits	0.59	0.41
Default Frequency	-0.43	-0.36
Profit Rate	0.05	0.07
Loan Return	-0.18	-0.20
Charge Off Rate	-0.32	-0.46

In Table 16 we present the main differences between "national" and "regional" banks generated on the equilibrium path by the model and compare it with the data. The model is consistent in generating lower default frequencies, loan return rates, charge off rates and variance of returns of "national" banks than "regional" banks and of "regional" banks than the competitive fringe. While the model is consistent with generating countercyclical loan returns for "national" and fringe banks, it generates procyclical loan returns for "regional" banks. Moreover, as opposed to the data, bank profit rates are lower for "national" banks than "regional" banks. In this equilibrium, "national" banks enjoy periods as the only dominant bank but in order to do so, during periods of higher competition they lower profits to induce "regional" banks to exit. Since banks exit when profits are negative and profit rates presented in the previous table are computed only for active banks, the surviving "national" banks have lower profit rates than surviving "regional" banks.

[22]We use the following dating convention in calculating correlations. Since most variables depend on z, s, z', and s' (e.g. default frequency $1 - p(R(r^L(\mu, z, s)), z', s')$), we display $corr(GDP(\mu, z, s, z', s'), k(z, z'))$.

Table 16: Model Moments by Bank Size

	National		Top 1%		Bottom 99%	
Moment Average	Model	Data	Model	Data	Model	Data
Default frequency	1.19	1.23	1.30	1.54	2.33	2.39
Bank profit rate	2.41	2.85	3.67	2.74	3.79	2.43
Loan return rate	6.29	4.94	6.51	5.28	6.66	5.99
Charge-off rate	0.75	0.94	0.70	1.08	0.70	0.93
Loan Interest rate	6.97	5.00	7.18	5.37	7.24	6.15
Variance Return	0.08	0.75	0.27	1.80	0.95	2.34
Corr(ret,gdp)	-0.22	-0.11	0.09	-0.18	-0.06	-0.17

6.3 Empirical Studies of Banking Crises, Default and Concentration

Many authors have tried to empirically estimate the relation between bank concentration, bank competition and banking system fragility and default frequency using a reduced form approach. In this section, we follow this approach using simulated data from our model to show that the model is consistent with the empirical findings. As in Beck et. al. [4], we estimate a logit model of the probability of a crisis as a function of the degree of banking industry concentration and other relevant aggregate variables. Moreover, as in Berger et. al. [5], we estimate a linear model of the aggregate default frequency as a function of banking industry concentration and other relevant controls. The banking crisis indicator takes value equal to one in periods whenever: (i) the loan default frequency is higher than 10%; (ii) deposit insurance outlays as a fraction of GDP are higher than 2%; (iii) large dominant banks are liquidated; or (iv) the exit rate is higher than two standard deviations from its mean. The concentration index corresponds to the loan market share of the national and regional banks. We use as extra regressors the growth rate of GDP and lagged growth rate of loan supply.[23] Table 17 displays the estimated coefficients and their standard errors.

[23] Beck et. al. [4] also include other controls like "economic freedom" which are outside of our model.

Table 17: Banking Crises, Default Frequencies and Concentration

Model Dependent Variable	Logit Crisis_t	Linear Default Freq._t
Concentration_t	-142.53	0.018
	(-6.61)	(10.05)
GDP growth in t	-130.39	-1.56
	(-6.38)	(-16.66)
Loan Supply Growth_{t-1}	140.66	1.31
	(5.94)	(4.28)
R^2	0.76	0.53
% Crisis Correct	64.00	-
% Correct	99.37	-

Note: $t-$statistics in parenthesis. R^2 refers to Pseudo R^2 in the logit model.

Consistent with the empirical evidence in Beck, et. al. [4], we find that banking system concentration is highly significant and negatively related to the probability of a banking crises. The results suggest that concentrated banking systems are less vulnerable to banking crises. Higher monopoly power induces periods of higher profits that prevent bank exit. This is in line with the findings of Allen and Gale [3]. Consistent with the evidence in Berger et. al. [5] we find that the relationship between concentration and loan portfolio risk is positive. This is in line with the view of Boyd and De Nicolo [9], who showed that higher concentration is associated with riskier loan portfolios.

7 Counterfactuals

7.1 On the effects of Bank Competition

Given entry costs, aggregate and regional shocks determine equilibrium entry and exit and hence the degree of industry concentration. To disentangle the effect of bank competition on risk taking and the probability of crises we run a counterfactual where entry costs κ are raised by 6% in which case "regional" banks choose not to enter the market, thus endogenously generating a more concentrated industry (inducing a market structure identical to the Gowrisankaran and Holmes [18]).

Table 18: Effects of Lower Competition

Moment	Benchmark	$\uparrow \kappa$	Change (%)
Default Frequency (%)	1.95	2.28	16.92
Exit Rate (%)	1.29	1.22	-5.43
Borrower Return (%)	13.25	13.22	-0.23
GDP	0.72	0.6	-16.67
Loan Supply	0.63	0.53	-15.87
Taxes/GDP (%)	0.04	0.03	-25.00
Loan Interest Rate (%)	6.84	8.04	17.48
Borrower Project (%)	13.57	13.59	0.15
Avg. Number Fringe Banks	7472	7497	0.33
Avg. Number Dominant Banks	2.90	1.00	-65.52

We observe that in the less competitive environment default frequencies and loan interest rates are higher, while the exit rate, entry rate, borrower return, loan supply and taxes over GDP are lower. In line with the predictions of A-G, a reduction in the level of competition reduces the exit rate (and by construction the entry rate). In this counterfactual, national banks are monopolists and operate in a region of interest rates and default frequency that avoids failure. This also increases expected profits for fringe banks reducing exit for them as well. Note that the average number of dominant banks is reduced by construction but the number of fringe banks is higher. Consistent with the predictions of B-D, a reduction in the level of competition increases the equilibrium interest rate on loans which induces borrowers to take on slightly more risk (i.e. R is 0.15% higher). This in turn, leads to an increase in default frequency by 16.9%. The increase in default frequency reduces borrower returns (i.e. pzR) by 0.2% and generates a drop in GDP and loan supply of 16.6% and 15.8% respectively. The reduction in exit rates due to lower competition reduces the amount of taxes that need to be collected to pay for deposit insurance (a 25% reduction as a fraction of GDP).

7.2 On the effects of Branching Restrictions

Important regulatory changes took place during the late eighties and early nineties in the U.S. banking industry. In 1994, the Riegle-Neal Interstate Banking and Branching Efficiency Act was passed.[24] The act allows banks to freely establish branches across state lines opening the door to the possibility of substantial geographical consolidation in the banking industry. To study the implications of branching restrictions, we study a counterfactual where we increase κ^n to 0.35, a value that prevents entry of national banks resulting in only regional and fringe banks in equilibrium. By contrasting this case with our benchmark model, we can study the benefits and costs of removing branching restrictions.

[24]The act removed the final restrictions that were in place in 1994, but the consolidation of the banking industry was a process that started during the eighties.

Table 19: Counterfactual: Effects of Branching Restrictions

Moment	Benchmark	$\uparrow \kappa^n$	Change (%)
Default Frequency (%)	1.95	3.41	74.87
Exit Rate (%)	1.29	2.27	75.97
Borrower Return (%)	13.25	13.1	-1.13
GDP	0.72	0.60	-16.67
Loan Supply	0.63	0.53	-15.87
Taxes/GDP (%)	0.04	0.05	25.00
Loan Interest Rate (%)	6.84	8.05	17.63
Borrower Project (%)	13.57	13.63	0.44
Avg. Number Fringe Banks	7472	7488	0.21
Avg. Number Dominant Banks	2.90	1.80	-37.93

Increasing κ^n such that no national banks enter, each of the regions becomes a more concentrated market since there is at most one incumbent dominant bank each period. This results in a lower loan supply (-15.87%) and increases the loan interest rate (+17.63%) that in turn induces increases in the riskiness of the borrower's project choice (+0.44%), the default frequency (+74.87%) and the exit rate (+75.97%). The increase in loan interest rates reduces the number of entrepreneurs that choose to operate the technology resulting in a lower level of GDP (16.67% lower). Since the exit rate is higher there are higher tax collection needs to cover deposit insurance. This corresponds to a 25% increase in τ/GDP; however, this increase represents only 0.1% of GDP which represents a much smaller change than the change in the level of output or loan supply.

7.3 Too Big To Fail

We documented that the top 4 commercial banks control more than 35% of total deposits and loans. As far as we know, ours is the first structural quantitative model of banking which admits a nontrivial endogenous size distribution of banks. This makes the model suitable for analyzing changes in policies that affect banks of particular sizes.

In our benchmark economy, there exists the possibility of failure by national banks and this ends up being an off-the-equilibrium path action because national banks reduce their exposure to the region with higher risk in order to maintain their charter value. However, a policy of "too big to fail" guarantees that the the government will bail out national banks in the event of realized losses big enough to induce them to exit. Such a policy changes the ex-ante incentives of national banks since they can take on more risk guaranteed that they receive ex-post bailouts.

In this section, we compare our benchmark economy with one where there are government bailouts to national banks with negative profits. More specifically, we consider the case where if realized profits for a national bank is negative the government will cover the losses and let the bank stay in operation. The problem of a national bank becomes

$$V_i(n,\cdot,\mu,z,s;\sigma_{-i}) = \max_{\{\ell_i(n,j)\}_{j=e,w}} E_{z',s'|z,s}\Big[\sum_j \max\{0, \pi_{\ell_i(n,j)}(n,j,c^n,\mu,z,s,z',s';\sigma_{-i})\} + \beta V_i(n,\cdot,\mu',z',s';\sigma_{-i})\Big] \tag{26}$$

subject to

$$\sum_{\theta}\sum_{i=1}^{N(\theta,j)} \ell_i(\theta,j,\mu,s,z;\sigma_{-i}) - L^{d,j}(r^{L,j},z,s) = 0, \tag{27}$$

where $L^{d,j}(r^{L,j},z,s)$ is given in (8). Note that with probability one the national bank receives a bailout, so there is no exit decision and when realized profits are negative the government covers the losses.[25] These losses are paid for by taxes as in the case of the deposit insurance. In Table 20, we present the main results.

Table 20: Benchmark vs Model with National Banks Bailouts

Moment	Benchmark	Too Big to Fail	Change (%)
Default Frequency (%)	1.95	1.87	-4.10
Exit Rate (%)	1.29	1.30	0.78
Borrower Return (%)	13.25	13.20	-0.38
GDP	0.72	0.75	4.17
Loan Supply	0.63	0.66	4.76
Taxes/GDP (%)	0.04	0.04	0.00
Loan Interest Rate (%)	6.84	6.54	-4.43
Borrower Project (%)	13.57	13.54	-0.22
Avg. Number Fringe Banks	7472	7465	-0.09
Avg. Number Dominant Banks	2.90	2.80	-3.45
Unconditional Prob of Bailout	0	1.47	
Max Cost Bailout over GDP	0	1.53	

Unlike the benchmark equilibrium, we find that along the equilibrium path national banks make negative profits which introduces government bailouts when the economy heads into a recession. The unconditional probability of a government bailout equals 1.47% and it can cost up to 1.5 % of GDP. However, there are no observable changes in taxes over GDP since there is actually an increase in GDP under the counterfactual policy.

The important point is that the introduction of government bailouts induces a national bank to increase its exposure to the region with the highest risk. This “excessive” risk taking behavior is what concerns policymakers. As evident from Table (13), the highest fraction of defaults happens when we enter into a recession (i.e. $z = z_g$ and $z' = z_b$) and the regional

[25]More generally, one might think that the probability of a bailout is in $[0, 1]$ not $\{0, 1\}$, but this induces a much more complicated computational algorithm where the evolution of the banking industry depends on the realization of government bailouts.

shock changes (i.e. $s = j$ and $s' \neq j$). Thus, provided that regional shocks are persistent, borrowers in different regions have a different risk profile. A national bank will increase its exposure to risk if it increases the amount of loans to the region with the good regional shock during good times. This is precisely what happens under the too big to fail policy. In Table 21, we compare the loan decision rules for dominant banks when $\mu\{1, 1, 1, \cdot\}$, $z = z_g$ and $s = e$ in the benchmark model versus the too big to fail policy.

Table 21: Benchmark vs Too Big to Fail

	Loan Decision Rules $\ell(\theta, j, \mu, z_g, e)$			
Model	$\ell(n, e, \cdot)$	$\ell(n, w, \cdot)$	$\ell(r, e, \cdot)$	$\ell(r, w, \cdot)$
Benchmark	0.020	0.106	0.203	0.152
Too Big To Fail	0.110	0.106	0.156	0.152

Under the too big to fail policy, the national bank extends five more times loans in region e than in our benchmark economy. If $z' = z_b$ and $s' = w$ are realized, the government will effectively need to bailout the national bank (this is an on-the equilibrium-path action). Induced by the actions of national banks, regional banks, however, extend less loans than before. Interestingly, in this example, the increase in the number of loans made by national banks effectively increases the total loan supply (+ 4.76%), resulting in a lower interest rate (-4.43%), a lower borrower project (-0.22%) and a lower default probability (-4.10%). Since more projects are financed GDP increases more than 4 %.

On the other hand, since national banks make more loans, profits for regional and fringe banks are reduced. This induces them to exit more frequently resulting in a higher aggregate bank exit rate and a lower average number of dominant banks in the too big to fail economy than in our benchmark.

There is no observable change (to the second decimel place) in taxes over GDP between the two models. This is because, despite the increase in the level of taxes due to the increase in the exit rate (and hence bailouts), there is an increase in GDP.

8 Concluding Remarks

Using Call Report data from commercial banks in the U.S. from 1976 to 2008 (the same data employed by Kashyap and Stein [22]) we document that entry and exit by merger are procyclical, exit by failure is countercyclical, total loans and deposits are procyclical, loan returns and markups are countercyclical and delinquency rates and charge-offs are countercyclical. Furthermore, we show that bank concentration has been rising and that the top 4 banks have 35% of loan market share. Finally, we document important differences between small and large banks. For example, we find that smaller banks have higher returns and higher volatility of returns than large banks and that returns for smaller banks are more countercyclical than large banks.

We provide a model where "big" geographically diversified banks coexist in equilibrium with "smaller" regional and fringe banks that are restricted to a geographical area. Since we allow for regional specific shocks to the success of borrower projects, small banks (both regional and fringe) may not be well diversified. This assumption generates ex-post differences between big and small banks. As documented in the data section, the model generates not only procyclical loan supply but also countercyclical interest rates and returns on loans, bank failure rates, default frequencies, charge-off rates. Since bank failure is paid for by lump sum taxes to fund deposit insurance, the model predicts countercylical taxes. Also, the model generates differences in loan interest rates, loan returns, profit rates and default frequencies between banks of different sizes (national, regional and fringe) since large banks are able to diversify across both regions. The variance of returns is also a decreasing function of bank size but it is smaller than in the data.

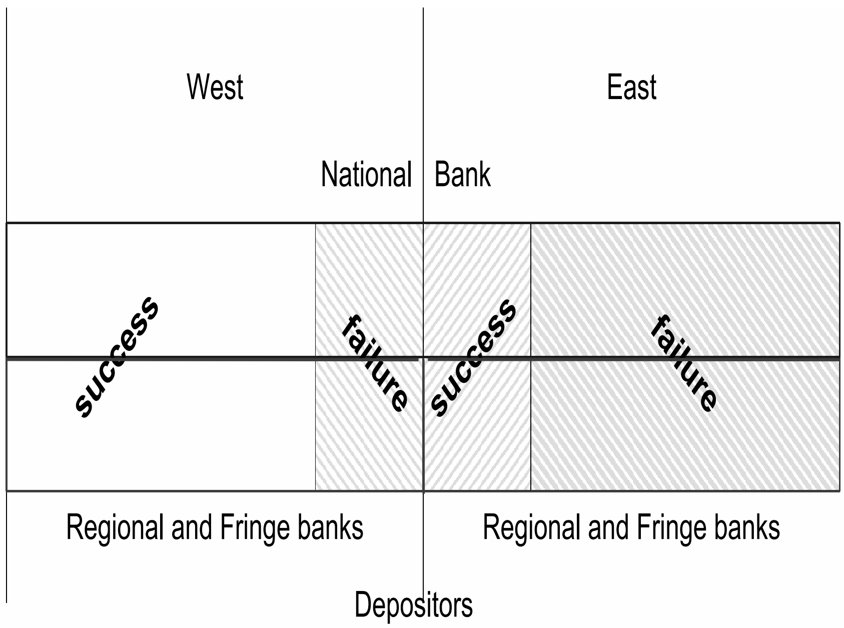
The benefit of our model relative to the existing theoretical literature is that the number of banks is derived endogenously and varies over the business cycle. To disentangle the effects of bank competition on default frequencies, borrower returns, bank exit rates and output we run a counterfactual where we increase entry costs into the banking sector to endogenously generate a more concentrated industry. As in Allen and Gale [3], we find that a reduction in the level of competition reduces bank exit. Moreover, in line with the predictions of Boyd and De Nicolo [9], less competition increases interest rates and induces borrowers to take on more risk resulting in higher default frequencies. We also show quantify the effects of a "too big to fail" policy. As expected, "too big to fail" induces big banks to extend more loans in risky states (i.e. increase their exposure), but this can induce lower interest rates and higher output on average. Increased bailout costs are actually offset by increased output in our counterfactual. Interestingly, lower exit by big banks induces more exit on average by smaller banks.

There is much work left to do. First, in order to keep the model simple and focus on the loan market, we have abstracted from deposit competition as in some other industrial organization studies (technically, this amounted to an assumption on parameters to induce an "excess supply" of depositors). In an extension we intend to add a distribution over outside options for depositors which will induce a supply of deposits which is sensitive to the deposit rate and banks will need to compete for depositors as in Allen and Gale [3]. Second, we intend to expand the bank balance sheet. In particular, currently we simply have loans on the asset side of the balance sheet. While these are the largest component (about 67%) of a bank's balance sheet, another sizeable asset (about 22%) is securities or other interbank loans. This will add another state variable to our analysis, but will allow us to start thinking about interesting policy experiments like capital requirements. [26] Further, once we have extended the bank balance sheet, we can use our model to study questions like those posed in Kashyap and Stein [22]; whether the impact of Fed policy on lending behavior is stronger for banks with less liquid balance sheets (where liquidity is measured by the ratio of securities to assets).

[26] In Corbae and D'Erasmo [11] we consider policies like capital requirements to mitigate bank failure among the competitive fringe.

Borrower chooses R^j	Receive	Pay	Probability
			$\quad\quad - \quad + \quad +$
Success	$1+z'R^j$	$1+r^{L,j}(\mu,z,s)$	$p^j \; (R^j, \; z', \; s')$
Failure	$1-\lambda$	$1-\lambda$	$1-p^j \; (R^j, \; z', \; s')$

Figure 17: Regional Segmentation



References

[1] Aguirregabiria, V., R. Clark, and H. Wang. (2010) “Bank Expansion after the Riegle-Neal Act: The Role of Diversification of Geographic Risk”, unpublished manuscript.

[2] Allen, F., E. Carletti and R. Marquez (2008) “Credit Market Competition and Capital Regulation”, unpublished manuscript.

[3] Allen, F. and D. Gale (2000) *Comparing Financial Systems,* Cambridge, MA: The MIT Press.

[4] Beck, T., A. Demirguc-Kunt, and R. Levine (2003) “Bank concentration and crises”, NBER Working Paper 9921.

[5] Berger, A. L. Klapper and R. Turk-Ariss (2008) “Banking competition and financial stability,” *Journal of Financial Services Research*, pp. 99-118.

[6] Bergstresser, D. (2004) “Market concentration and commercial bank loan portfolios ”, Harvard Business School, manuscript.

[7] Bernanke, B., M. Gertler, and S. Gilchrist (1999) “The Financial Accelerator in a Quantitative Business Cycle Framework”, in John Taylor and Michael Woodford (eds.), *The Handbook of Macroeconomics*, Amsterdam: North-Holland.

[8] Bikker, J.A. and K. Haaf (2002) “Competition, concentration and their relationship: An empirical analysis of the banking industry,” *Journal of Banking & Finance*, Vol. 26(11), p. 2191-2214.

[9] Boyd, J. and G. De Nicolo (2005) “The Theory of Bank Risk-Taking and Competition Revisited”, *Journal of Finance*, 60, p. 1329-43.

[10] Carlstrom, C. and T. Fuerst (1997) “Agency Costs, Net Worth, and Business Fluctuations: A Computable General Equilibrium Analysis”, *American Economic Review*, 87, p. 893-910.

[11] Corbae, D. and P. D’Erasmo (2010) “Preventing Small Bank Failure”, mimeo.

[12] Diamond, D. (1984) “Financial Intermediation and Delegated Monitoring”, *Review of Economic Studies*, 51, p. 393-414.

[13] Diamond, D. and P. Dybvig (1983) ”Bank Runs, Deposit Insurance, and Liquidity”, *Journal of Political Economy*, 91, p. 401–19.

[14] Diaz-Gimenez, J., E. Prescott, F. Alvarez, and T. Fitzgerald (1992) “Banking in Computable General Equilibrium Models”, *Journal of Economic Dynamics and Control*, 16, p.533-60.

[15] Diebold F. and K. Yilmaz (2009) “Measuring Financial Asset Return and Volatility Spillovers, with Application to Global Equity Markets,” *Economic Journal*, 119(534), 158-171.

[16] Ennis, H. (2001) "On the Size Distribution of Banks", Federal Reserve Bank of Richmond *Quarterly Review*, 87, p. 1-25.

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

[18] Gowrisankaran, G. and T. Holmes (2004) "Mergers and the Evolution of Industry Concentration: Results from the Dominant-Firm Model," *RAND Journal of Economics*, 35(3), p. 561-582.

[19] Hopenhayn, H. (1992) "Entry, Exit, and Firm Dynamics in Long-Run Equilibrium", *Econometrica*, 60, p. 1127-50.

[20] Huggett, M. (1993) "The risk-free rate in heterogeneous-agent incomplete-insurance economies", *Journal of Economic Dynamics and Control*, 17, p. 953-69.

[21] Janicki, H. and E. Prescott (2006) "Changes in the Size Distribution of U.S. Banks: 1960-2005", Federal Reserve Bank of Richmond *Quarterly Review*, 92, p. 291-316.

[22] Kashyap A. and J. Stein (2000) "What Do a Million Observations on Banks Say about the Transmission of Monetary Policy?" *American Economic Review*, vol. 90(3), p. 407-428.

[23] Koeppl, T. and J. MacGee (2009) "What broad banks do, and markets don't: Cross-subsidization" *European Economic Review*, vol. 53, p. 222-236.

[24] Liang, N. and S. Rhoades (1988) "Geographic Diversification and Risk in Banking" *Journal of Economics and Business*, vol. 40, p. 271-284.

[25] Martinez-Miera, D. and R. Repullo (2008) "Does Competition Reduce the Risk of Bank Failure?", CEMFI Working Paper No. 0801.

[26] Pakes, A. and P. McGuire (1994) "Computing Markov Perfect Nash Equilibrium: Numerical Implications of a Dynamic Differentiated Goods Model", *RAND Journal of Economics*, 25, p. 555-589.

[27] Panzar, J.C. and J.N. Rosse (1987) "Testing for monopoly equilibrium," *Journal of Industrial Economics*, 35, 443456.

[28] Rhoades S. (2000) "Bank Mergers and Banking Structure in the United States, 198098" *Board of Governors of the Federal Reserve System*, Staff Report 174.

[29] Townsend, R. (1979) "Optimal contracts and competitive markets with costly state verification" *Journal of Economic Theory*, 22, p. 265293.

[30] Williamson, S. (1986) "Costly Monitoring, Financial Intermediation and Equilibrium Credit Rationing", *Journal of Monetary Economics*, 18, p. 159-179.