Capital Buffers in a Quantitative Model of Banking Industry Dynamics[*]

Dean Corbae[†]
University of Wisconsin - Madison and NBER

Pablo D'Erasmo[‡]
Federal Reserve Bank of Philadelphia

May 14, 2021

Abstract

We develop a model of banking industry dynamics to study the quantitative impact of regulatory policies on bank risk taking and market structure. Since our model is matched to U.S. data, we propose a market structure where big banks with market power interact with small, competitive fringe banks as well as non-bank lenders. Banks face idiosyncratic funding shocks in addition to aggregate shocks which affect the fraction of performing loans in their portfolio. A nontrivial bank size distribution arises out of endogenous entry and exit, as well as banks' buffer stock of capital. We show the model predictions are consistent with untargeted business cycle properties, the bank lending channel, and empirical studies of the role of concentration on financial stability. We find that regulatory policies can have an important impact on banking market structure, which, along with selection effects, can generate changes in allocative efficiency and stability. *JEL* Classification Numbers: E44, G21, L11.

Keywords: macroprudential policy, bank size distribution, industry dynamics with imperfect competition.

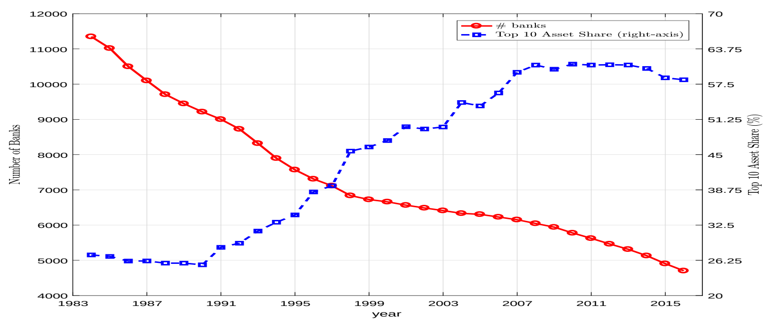
[*]We wish to thank our editor Gianluca Violante and four anonymous referees for extremely helpful comments on an earlier version of this paper. We also wish to thank John Boyd, Gianni DeNicolo, Jean-Francois Houde, Victor Rios-Rull, and Skander Van Den Heuvel, as well as seminar participants at the Federal Reserve Board, European Central Bank, European Commission, and FDIC; Central Banks of Argentina, Canada, Chile, Columbia, Mexico, Netherlands, Norway, Portugal, Spain, Sweden, Turkey, and Uruguay; the Federal Reserve Banks of Atlanta, Chicago, Cleveland, Kansas City, Minneapolis, New York, and St. Louis; the Universities of British Columbia, UCLA, Carnegie Mellon/Pittsburgh, Chicago Booth, Cologne, Colorado, Columbia, Cornell, Drexel, European University Institute, George Washington, Goethe, Indiana, London Business School, Maryland, McMaster, Michigan, Minnesota, NYU, Notre Dame, Ohio State, Penn, Penn State, Princeton, Queens, Rice, Rochester, Southern California, Tilburg, Toronto, Tsinghua; the Allied Social Science Association, CIREQ Macroeconomics Conference, Econometric Society, Sciences Po Macro Finance Conference, and Society for Economic Dynamics for helpful comments. **Disclaimer:** The views expressed in these papers are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

[†]Email: dean.corbae@wisc.edu
[‡]Email: pablo.derasmo@phil.frb.org

1 Introduction

The banking literature has focused on two main functions of bank capital. First, because of limited liability and deposit insurance, banks have an incentive to engage in risk shifting. Requiring banks to hold a minimum ratio of capital to assets and sufficient liquidity to meet funding shocks constrains the banks' ability to take risk. Second, bank capital acts like a buffer that may offset losses and save its charter value. In this paper, we develop a quantitative model of banking industry dynamics with imperfect competition and an endogenous bank size distribution to answer the following question: How much do regulatory policies enacted in the Dodd-Frank Wall Street Reform and Consumer Protection Act of 2010 affect failure rates, lending, interest rates, and market shares of large and small banks? Further, since empirical work has studied the effect of market structure on financial stability - one of the objectives of regulatory policy - we use our model to study the fixed point between policy on market structure and market structure on the efficacy of policy.

Figure 1: Number of Banks and Bank Concentration



Notes: Number refers to the number of bank holding companies and stand-alone commercial banks. Top 10 Asset share refers to the share of total assets in the hands of the top 10 banks. Source: Consolidated Reports of Condition and Income

As Figure 1 makes clear, the U.S. banking industry appears highly concentrated with the top 10 banks now comprising nearly 60% of the market share while the remaining nearly 5000 banks comprise a little over 40%. For this reason, we choose to model the industry using a dominant firm and competitive fringe framework along the lines of Gowrisankaran and Holmes [37].[1] Including endogenous entry and exit allows us to understand how regulatory policy affects market structure as well as to consider how market structure influences profitability and risk taking behavior that policymakers are attempting to regulate.

We assume that banks' deposit inflows follow a Markov process that is independently distributed across banks that vary with the business cycle. When we go to the data, we find size dependent differences in these processes; big banks have a higher level but lower variance of deposits (consistent with diversification). These processes form the capacity constraints in our Cournot model. It is important to note that even if banks only face idiosyncratic funding shocks, the presence of non-atomistic banks implies aggregate uncertainty.

Also motivated by the data, our model banks face both marginal and fixed costs of making loans with decreasing average costs across big and small banks consistent with a modest degree of increasing returns. Our data findings are consistent with a delegated monitoring model along the lines of Diamond [27].[2]

[1] In an earlier paper (Corbae and D'Erasmo [21]), we augmented the dominant firm portion of the model to include both national and regional banks with market power along the lines of Ericson and Pakes [32] with a competitive fringe under assumptions that induced a simpler state space.

[2] There are numerous empirical papers documenting the existence of scale economies in banking such as Berger

In response to variations in their cash flows, banks may choose to exit under limited liability in the event of shortfalls if their charter value is not sufficiently high. Banks whose charter value is sufficiently valuable can use their stock of net securities as a buffer or issue seasoned equity at a state-dependent cost. This induces banks in our model to undertake precautionary savings in the face of idiosyncratic shocks as in a standard consumer income fluctuations problem, but with a strategic twist, since here, big banks have loan market power. Similar to the consumer's problem, banks with a higher variance of funding shocks (i.e. the small banks) will endogenously hold a larger buffer of equity than more diversified big banks, again consistent with U.S. data.[3]

Our framework deviates from the frictionless Modigliani-Miller paradigm by including government deposit insurance and limited liability generating a moral hazard problem for banks, bankruptcy and equity issuance costs, agency conflicts between the manager and shareholders, and imperfect competition. Regulation in this environment can help mitigate bank risk taking but can have unintended consequences (e.g. raise big bank market shares).

A benefit of our structural framework is that we can conduct policy counterfactuals from newly enacted regulation in the Dodd-Frank Act. For example, in Section 7 we study a rise in level of capital requirements from 4% under Basel II to 8.5% (corresponding to the required minimum risk weighted capital requirement of 6% plus a 2.5% capital conservation buffer) motivated by changes recommended by Basel III.[4] We find that such a rise in capital requirements has the intended consequence of decreasing long run exit rates of small banks from the model's long run benchmark but also leads to a more concentrated industry since it inhibits entry. In the short run, big banks decide to strategically gain loan market share financed by issuing more equity, cutting dividends, and retaining more earnings. The net effect of higher big bank lending and lower small bank lending is a decrease in total bank lending of over 7% in the short run and nearly 9% in the long run but only a modest rise in interest rates on loans in the long run.

The modest response in interest rates (and aggregate lending) depends critically on the financing options facing borrowers if they do not choose bank loans. We model that option as do several others (e.g. Buchak, et al. [18]) through a discrete choice problem over bank and non-bank (sometimes called the "shadow banking" sector) options solved by borrowers subject to extreme value shocks. We calibrate parameters based on market share of bank and non-bank finance. Thus, our framework accounts for the impact of regulatory arbitrage associated with a change in capital and/or liquidity requirements. This margin of adjustment captures the effect of regulation on the size of the regulated banking sector.

To understand the interaction between regulatory policy and market structure, we also conduct a counterfactual where we increase the entry cost for dominant banks to a level that prevents their entry. Since our benchmark model with dominant and fringe banks nests an environment with only perfect competition, we use this counterfactual to understand the role of imperfect competition on the banking sector in Section 7. After calibrating the model of perfect competition to match U.S. banking data, we find several important differences on the efficacy of policy. For example, we find that the perfectly competitive economy generates a significantly different response in banking industry dynamics (entry and exit) to a rise in capital requirements. Specifically, entry and exit rates drop much more in the imperfectly competitive economy in response to the policy change in the long run than one with perfect competition. This leads to the intended consequence of lowering the probability of a crisis in the imperfectly competitive model but an unintended rise

and Hannan [14], Berger and Mester [15], and Wheelock and Wilson [66].

[3]It is in this sense we use the language "capital buffer" (i.e. capital in excess of the requirement).

[4]FDIC Rules and Regulations (Part 325) establishes the criteria and standards to calculate capital requirements and adequacy (see DSC Risk Management Manual of Examination Policies, FDIC, Capital (12-04)). See a full description in BIS [16].

in the probability of a crisis in the perfectly competitive case. In the long run, this translates to modest welfare gains in the imperfectly competitive case and modest welfare losses in the perfectly competitive case. The long run differences between the two models arise because with imperfect competition the big bank increases its lending to deter small bank entry resulting in a more select measure of fringe banks, which in turn leads to a much more significant reduction in deadweight losses associated with bank failure.

Related Literature

Our paper is related to the literature studying the impact of financial regulation in quantitative models of banking. One strand of literature studies dynamic bank decision problems.[5] These papers, however, cannot consider the impact of such policies on loan interest rates and the equilibrium bank size distribution.

The second strand of literature studies dynamic general equilibrium models with a representative bank under perfect competition in loan and deposit markets. Van Den Heuvel [64] was one of the first to study the welfare impact of capital requirements with perfect competition.[6] In these papers, constant returns and perfect competition imply that there is an indeterminate distribution of bank sizes, so they do not make predictions for how regulation affects banking industry market structure. Others with perfect competition assume idiosyncratic shocks which can generate an endogenous size distribution of banks.[7] In such models, big banks have no impact on lending or deposit rates and technically the failure of an individual big bank has no market impact (since it is of measure zero).

Given high concentration in the banking industry, another branch of the literature considers imperfect competition in loan and/or deposit markets. Our dynamic banking industry model builds upon the static framework of Allen and Gale [5] and Boyd and De Nicolo [17]. In those models, there is an exogenous number of banks that are Cournot competitors in the loan and/or deposit market.[8] Given both aggregate productivity and idiosyncratic funding shocks, we endogenize the number of banks by considering dynamic entry and exit decisions. While ours is one of the first quantitative models to focus on imperfect competition in loan markets, there is also an important set of papers analyzing imperfect competition in the deposit market (see for example Aguirregabiria, Clark, and Wang [3], and Egan, Hortascu, and Matvos [30]). We focus on loan markets since in the recent financial crisis, aggregate bank risk weighted asset accumulation (including loans) fell at a significantly higher rate than aggregate bank borrowings.[9]

The computation of our model is a nontrivial task. Since idiosyncratic shocks to large banks do not wash out in the aggregate, all equilibrium objects, such as value functions and prices, are a function of the infinite dimensional distribution of banks. Thus, we solve the model using an extension of the algorithm proposed by Ifrach and Weintraub [40] adapted to this environment. This entails approximating the distribution of banks by a finite number of moments.

Roadmap

[5] For example, De Nicolo et al. [26] show an inverted U-shaped relationship between capital requirements and bank lending.

[6] Among others, see Aliaga-Diaz and Olivero [4], Begenau [9], Bianchi and Bigio [10], Clerc et al. [19], Elenev, Landvoigt, and Van Nieuwerburgh [31].

[7] For example, see Rios-Rull et al. [58] and Nguyen [51].

[8] Other important papers studying the role of imperfect competition in the banking industry and regulation include Martinez-Miera and Repullo [46], Perotti and Suarez [54], Repullo [57], and Wang et al. [65].

[9] Specifically, aggregate bank risk weighted asset accumulation (including loans) grew at an annual 8.5% rate in 2006 and shrank at a −4.3% rate in 2009 at the same time that aggregate bank borrowings (including deposits) grew at an annual 10.9% rate in 2006 and shrank at a −0.4% rate in 2009.

The paper is organized as follows. Section 2 documents a set of banking data facts relevant to this paper. 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 we estimate parameters of the model to match U.S. Call Report data as well as how the model does on relevant untargeted business cycle correlations. Section 6 illustrates certain key properties of the model including its consistency with empirical studies of the competition-fragility hypothesis and the bank lending channel. Section 7 conducts policy counterfactuals put forward in the Dodd-Frank Act and studies their impact on allocative efficiency in the banking sector and welfare.

2 Banking Data Facts

In our previous paper [21], we documented a set of facts for the U.S. using data from the Consolidated Report of Condition and Income (known as Call Reports) that insured banks submit to the Federal Reserve each quarter.[10] Entry is procyclical and exit by failure is countercyclical (correlation with detrended GDP is equal to 0.61 and -0.16, respectively for the period 1984-2016). Almost all entry and exit by failure is by small banks (defined as banks in the bottom 99% of the asset distribution). Loans and deposits are procyclical (correlation with detrended GDP is equal to 0.44 and 0.18, respectively for the same period). Loan returns, margins, markups, and delinquency rates are countercyclical.

The countercyclicality of margins and markups is important. Our model provides an endogenous amplification mechanism consistent with this data fact. During a downturn, there is exit by smaller banks, which generates less competition among existing banks, raising the interest rate on loans. But since loan demand is inversely related to the interest rate, the ensuing increase in interest rates (barring government intervention) lowers loan demand even more, thereby amplifying the downturn. In this way our model provides a novel mechanism - imperfect loan market competition - to produce endogenous business cycle amplification arising from the banking sector. Further, we show in subsection 5.3 that a calibrated model with perfect competition does not share these untargeted cyclical properties.

2.1 On Imperfect Competition

As Figure 1 makes clear, the number of commercial banks in the U.S. has fallen from over 11,000 in 1984 to under 5000 in 2016 while the asset market share of the top 10 banks has grown from 27.2% in 1984 to 58.3% in 2016. Rising market shares of big banks motivate us to consider a model of the banking industry that allows for imperfect competition. Rising concentration, however, is only suggestive and not sufficient to establish market power. In this subsection, we consider other measures that are also suggestive of imperfect competition, motivating our modeling choice.

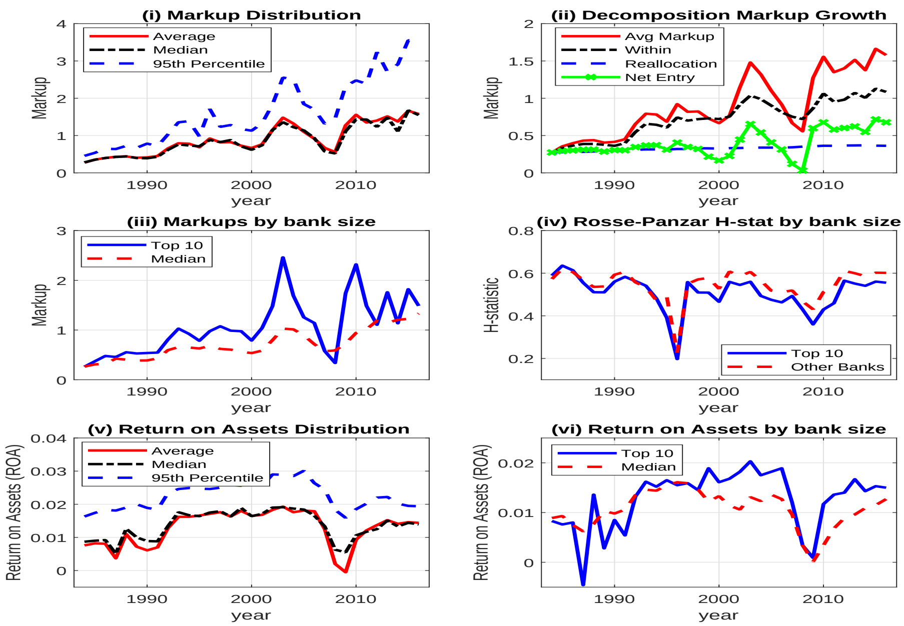
We begin by presenting evidence along the lines of De Loecker, Eeckhout, and Unger [25] using our Call Report data (the universe of U.S. commercial banks) rather than their non-banking datasets.[11] Panel (i) of Figure 2 graphs the 95th percentile of markups versus the median and

[10] Balance Sheet and Income Statements items can be found at https://cdr.ffiec.gov/public/. We group commercial banks to the bank holding company level and work with individual commercial banks when no bank holding company exists.

[11] The estimation and definition of all measures presented in Figure 2 are described in Online Appendix A-1.3. Appendix A-1.3 also presents additional evidence on the distribution of markups as well as the positive link between markups and bank size.

average markup for all U.S. commercial banks.[12] The asset-weighted average markup is calculated as $m_t = \sum_i s_i m_{it}$, where s_{it} corresponds to the asset share over total assets of bank i in period t and m_{it} to the markup of bank i in period t. There are two important features to note which are consistent with their findings for other industries. First, average markups have been rising over time. Second, the growth has been fueled by the upper tail of the distribution.

Figure 2: Markups, Passthrough, and Return on Asset Distributions



Note: Sample period: 1984-2016. Series displayed correspond to asset-weighted average markup ("Average"), median markup ("Median"), and the 95th percentile of the markup distribution. To construct the figure in panels (*i*) and (*v*) we rank each bank by its markup and return on assets and weight each bank by its asset share of total assets in the industry. Series displayed in Panel (*ii*) correspond to decomposition as presented in equation (A.1.4). Series displayed in Panels (*iii*), (*iv*), and (*vi*), correspond the evolution of (asset-weighted) average markup, Rosse-Panzar H or return on assets by bank size. We group banks in the Top 10 of the asset distribution ("Top 10"). Source: Consolidated Reports of Condition and Income

Next we present a decomposition, following the literature on resource misallocation, of the change in markups along the lines of equation (9) in De Loecker, Eeckhout, and Unger [25]. In particular, we decompose the growth of asset-weighted average markup into the increase derived from an increase in average markups ("within"), the increase derived from "reallocation" (i.e., the increase derived from growing asset shares of banks with high markups keeping markups fixed), and a final term coming from changes in markups derived from entry and exit. In Panel (*ii*) of Figure 2, we present the evolution of asset-weighted average markup, as well as three experiments based on the decomposition starting in 1984. We set the initial level in 1984 equal across experiments and then cumulatively add the changes of each of the components presented in the last equation. For example, the "within" experiment shows the evolution of markups if only the Δwithin would be allowed to change and all other components would be set to zero. Consistent with Figure 16.2 in their Appendix for the entire Finance, Insurance, and Real Estate industry (52-53), we find that most of the increase in markups is derived from the "within" component.

Here, we also choose to present the distribution of markups by bank size.[13] Panel (*iii*) in

[12]Our Panel (*i*) of Figure 2 is the analogue of Figure 3 in De Loecker, Eeckhout, and Unger [25] across all industries and the analogue of Figure 12.1 in their Appendix for the Finance and Insurance Industry Specific Markup.

[13]There is no analogue of a firm size dependent markup in De Loecker, Eeckhout, and Unger (2019).

Figure 2 presents the evolution of average (asset-weighted) markups by bank size. We compare banks in the Top 10 of the asset distribution with the average markup of the median bank. This figure establishes that big bank markups exceeded those of small banks in almost every year of the sample.

In addition, we present another typical measure of market power in the industrial organization literature. Following Shaffer [60], we estimate a measure of passthrough (the elasticity of marginal revenue with respect to factor prices). A value of 100 for the Rosse-Panzar H indicates the presence of perfect competition. Panel (iv) of Figure 2 presents the evolution of the Rosse-Panzar H by bank size over time. Notably passthrough is lower for the top 10 banks than the rest, especially after consolidation following the Riegle-Neal Act, again suggestive of imperfect competition.

As noted by De Loecker, Eeckhout, and Unger [25], before concluding whether higher markups are associated with market power, they suggest considering measures of profitability. Panel (v) of Figure 2 presents the evolution of different moments of the Return on Assets (ROA) distribution as a measure of profitability.[14] As with the distribution of markups, to construct this figure, for every year in the sample, we order the banks by the value of their ROA and weight each bank by its asset share of the entire industry. The ranking and asset shares are updated every year. Similar to the evolution of markups, the increases in the very top of the distribution drive the increase in average ROA.

Panel (vi) of Figure 2 presents the evolution of average (asset-weighted) ROA by bank size. Similar to the case of markups, we compare banks in the Top 10 of the asset distribution with the markup of the median bank (bank in percentile 50 of the asset distribution). Post Riegle-Neal (i.e. 1994), profitability as measured by ROA is typically higher for big banks than small, suggestive of market power. In conclusion, these diverse measures are suggestive of increasing market power at the national level, which motivates us to model an imperfectly competitive banking industry.[15]

2.2 Capital Buffers

Since we are interested in the effects of capital and liquidity requirements on bank lending and financial stability of big and small banks, we organize the data along those dimensions. We refer to the Top 10 banks in the asset distribution as "big" banks and we refer to the remaining banks as "small" or "fringe".[16]

Panel (i) in Figure 3 presents the evolution of the Tier 1 capital-to-risk-weighted-assets ratio for the 10 largest banks and the remaining banks when sorted by assets. In all periods, risk-weighted capital ratios are lower for large banking institutions than those for small banks.[17] The fact that capital ratios are above what regulation defines as well capitalized suggests a buffer stock motive which we model.

While Panel (i) in Figure 3 presents the cross-sectional average for big and small banks across time, the average masks the fact that some banks spend time at the constraint (and even violate the

[14] Figure 9.1 in the Appendix of De Loecker, Eeckhout, and Unger [25] shows that the return on assets and other measures of profitability are highly correlated in the data.

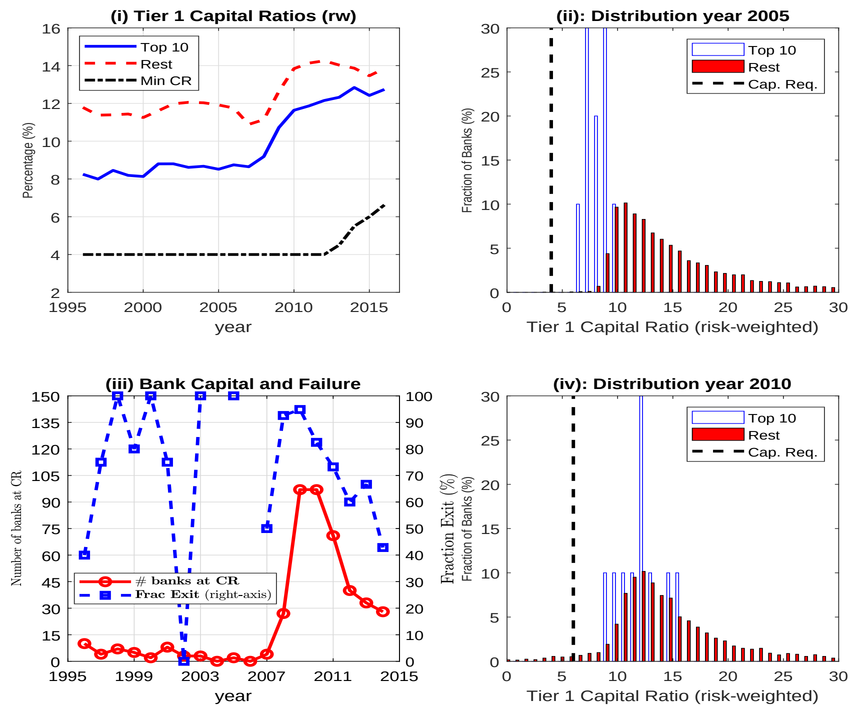
[15] We leave the question as to whether our national results for the banking industry also hold at the local level to future research. Differences between national and local Herfindahl indices for sales are documented in papers such as Rossi-Hansberg, Sarte, and Trachter [59]. Some evidence for high national *and* local concentration in the deposit market using Herfindahl indices can be found in Meyer [49].

[16] As of December 2016, most of the banks in the group of Top 10 were classified as Global Systemically Important Banks (G-SIBs). The exceptions are U.S. Bankcorp, PNC Financial Services, and Toronto Dominion Group US. While not classified as G-SIBs, these three banks are still above the $250 billion asset threshold for the size dependent policies that we study.

[17] Capital ratios based on total assets (as opposed to risk-weighted assets) present a similar pattern.

constraint). Panels (*ii*) and (*iv*) in Figure 3 plot the histogram of all banks for years 2005 and 2010, respectively. These figures make clear that large (Top 10) banks have consistently lower capital buffers than most other banks. Government assistance, private injection of equity and changes in capital regulation have induced shifts in the distribution of capital. Moreover, during the crisis, a considerable number of banks failed, merged with other institutions under distress or received government assistance. Panel (*iv*) in Figure 3 (year 2010) shows that it is possible to find banks close to or even below the minimum required. Many of these banks end up failing.

Figure 3: Distribution of Bank Capital Ratios and Relationship between Bank Capital and Failure



Note: Tier 1 Capital (rw) refers to Risk-weighted Tier 1 Capital Ratio. Averages are computed as asset-weighted averages. Min CR refers to minimum capital requirement (risk-weighted) plus capital conservation buffer for banks with less than \$50 billion in assets (all of these banks included in the "Rest" group). Banks with more than \$50 billion are required to hold additional capital since 2013. # of banks at CR refers to the number of banks with capital ratios no larger than 0.5% more than the minimum required. Fraction Exit refers to the fraction of banks within the group of banks with capital ratios no larger than 0.5% that fail during that given year or the following. The line for the Fraction Exit is missing during periods where there are no banks with capital ratios at or below the minimum required. Source: Consolidated Reports of Condition and Income

To analyze the relationship between bank failure and capital ratios, panel (*iii*) in Figure 3 shows the number of banks that are at or below the minimum risk-weighted capital required (as in panel (iv)) and the fraction of those that exit (via failure or merger the corresponding year or the year after). This figure makes clear that most banks with capital ratios close to the minimum required exit the industry. The average fraction of banks that exit conditional on being close to the minimum required is well above 70 percent.

3 Environment

The above data facts motivate us to analyze the impact of regulatory policy on bank lending and financial stability in an imperfectly competitive banking industry. Each period, banks and non-banks intermediate between a unit mass of ex-ante identical entrepreneurs who have a profitable project which needs to be funded (the potential borrowers) and a measure $\mathcal{H} > 1$ of identical households (the potential depositors).

3.1 Entrepreneurs

Infinitely lived, risk neutral entrepreneurs demand a loan in order to fund a new project each period. Specifically, a project requires $\overline{\ell}$ units of investment in period t and returns next period (per-unit of investment):

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

in the successful and unsuccessful states, respectively. That is, borrower gross returns are given by $1 + z_{t+1}R_t$ in the successful state and by $1 - \lambda$ in the unsuccessful state. The success of a borrower's project, which occurs with probability $p(R_t, z_{t+1})$, is independent across borrowers and time *conditional* on the borrower's choice of technology $R_t \geq 0$ and an aggregate technology shock at the beginning of the following period denoted z_{t+1}. The aggregate technology shock $z_t \in Z$ evolves as a Markov process $G_z(z', z) = prob(z_{t+1} = z'|z_t = z)$. As in simple growth models with one period time-to-build, we assume the capital investment ℓ at t becomes productive in period $t + 1$ and depreciates fully after production.

We consider an environment with two types of lenders $j \in \{c, n\}$: an endogenous measure of heterogeneous commercial banks (c) and a representative non-bank (n).[18] Both types of lenders can finance productive projects by entrepreneurs but their interest rates might differ. Taking the vector of interest rates $\mathbf{r_t} = \{r_t^c, r_t^n\}$ as given, entrepreneurs decide whether they want to fund a project given their outside option and then make a discrete choice over whether to borrow from a bank or a non-bank. Once with lender type j offering a loan at interest rate r_t^j in state z_t, they choose the risk-return tradeoff of their project R_t^j where we index the choice of technology by the lender type that finances the project. Following Bucheck et al. [18], we assume that the value associated with financing the project with each type of lender is subject to an unobservable idiosyncratic shock $\boldsymbol{\epsilon_t} = \{\epsilon_t^c, \epsilon_t^n\}$ affecting the value of taking a loan from each type of lender additively. We assume that ϵ_t^j are drawn iid from a type one extreme value distribution with scale parameter $\frac{1}{\alpha}$ which we estimate from the data.

Borrowers have an outside option. In period t, they receive a perfect private signal about their reservation utility of consumption in period $t + 1$ in case they decide not to run the project. The value of this outside option is $\omega_{t+1} \in [0, \overline{\omega}]$ drawn from distribution function $\Omega(\omega_{t+1})$. These draws are i.i.d. over time. This outside option leads to a downward sloping aggregate demand for loans while the extreme value shocks determine loan demand across bank and non-bank sectors.

The entrepreneur can save $a_{E,t+1} \in \mathbb{R}_+$ with safe return r^s (that is also accessible to households) and can choose whether to retain earnings $I_{t+1} \in [0, \overline{\ell}]$ in order to finance investment.[19] When the borrower makes her choice of technology R_t^j, the aggregate shock z_{t+1} has 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, $p(R_t^j, z_{t+1})$ is assumed

[18] We choose to use the term non-bank rather than "shadow bank" following the terminology used by the Financial Stability Oversight Council to refer to the system of financial intermediaries without access to the Federal Reserve's and Federal Deposit Insurance Company's (FDIC) backstops.

[19] If, as we assume, the entrepreneur is sufficiently impatient, then she would not choose to undertake any of these alternatives. That is, provided the discount factor β_E is sufficiently low, entrepreneurs choose not to use retained earnings to finance their projects, instead choosing to eat their earnings and fund projects using one-period loans that require monitoring.

to be decreasing in R_t^j and increasing in z_{t+1}. Thus, the technology exhibits a risk-return tradeoff that varies with the business cycle. While borrowers are ex-ante identical, they are ex-post heterogeneous owing to the realization of the shocks to the return on their project. We envision borrowers as firms choosing a technology that might not succeed, or under an alternative calibration, households choosing a home that might appreciate or depreciate.

There is limited liability on the part of the borrower at the project level so that the project return net of interest payments is bounded below at zero.[20] Table 1 summarizes the risk-return tradeoff that the borrower faces. Since the choice of R_t^j is endogenous, changes in borrowing costs can affect the default frequencies on loans through a risk shifting motive. Specifically, higher interest rates may endogenously generate a pool of riskier borrowers as in Stiglitz and Weiss [62].

Table 1: Borrower's Problem (conditional on investing)

Borrower Chooses R_t^j	Receive	Pay	Probability
			$-$ $+$
Success	$(1+z_{t+1}R_t^j)\bar{\ell}$	$(1+r_t^j)(\bar{\ell}-I_t)$	$p \quad (R_t^j, \quad z_{t+1})$
Failure	$(1-\lambda)\bar{\ell}$	$\min\{(1-\lambda)\bar{\ell},(1+r_t^j)(\bar{\ell}-I_t)\}$	$1-p \quad (R_t^j, \quad z_{t+1})$

Both R_t^j and ω_{t+1} are private information to the entrepreneur, as well as the history of past borrowing and repayment by the entrepreneur (which provides the rationale for short term loans). As in Bernanke and Gertler [13], success or failure is also private information to the entrepreneur unless the loan is monitored by the lender. With one-period loans, reporting failure is a dominant strategy in the absence of monitoring provided $(1-\lambda)\bar{\ell} < (1+r_t^j)(\bar{\ell}-I_t)$, in which case loans must be monitored. Monitoring is costly as in Diamond [27].

3.2 Households

A mass $\mathcal{H}$ of infinitely lived, risk neutral households with discount factor β_H are endowed with y units of goods each period. We assume households are sufficiently patient such that they choose to exercise their savings opportunities. In particular, households have access to an exogenous safe storage technology yielding $1+r^s$ between any two periods with $r^s \geq 0$ and $\beta_H(1+r^s)=1$. They can also choose to supply their endowment to a bank or to an individual borrower. If matched with a bank, a household who deposits its endowment there receives r_t^D whether the bank succeeds or fails since we assume deposit insurance. Households can hold a fraction of the portfolio of bank stocks yielding dividends (claims to bank cash flows) and can inject equity to banks. They can also invest in shares of the representative non-bank, which gives a claim to non-bank cash-flows. They pay lump-sum taxes/transfers τ_t, which include a lump-sum tax τ_t^D used to cover deposit insurance for failing banks, and a tax (transfer if negative) for government debt service of securities τ_t^A. Finally, if a household was to match directly with an entrepreneur (i.e. directly fund an entrepreneur's project), it must compete with bank loans. Hence, the household could not expect to receive more than the bank lending rate r_t^c in successful states and must pay a monitoring cost. Since households can purchase claims to bank cash flows, and banks can more efficiently minimize costly monitoring along the lines of Diamond [27], there is no benefit to matching directly with entrepreneurs.[21]

[20] We assume entrepreneurs assets are exempt from bankruptcy for simplicity.

[21] Since in our calibrated model loans to per-capita income of households (ℓ/y) strictly exceeds one, it takes more than one household to fund a loan out of their endowment. We focus on equilibria where group deviations by households are ruled out, so that a one-shot deviation by a single household to match with an entrepreneur results

3.3 Banks

As in Diamond [27] commercial banks exist in our environment to economize on monitoring costs besides diversifying borrower idiosyncratic project risk. Given that the U.S. commercial bank distribution is characterized by a few large banks and a large number of small banks, we assume there are two types of banks: $\theta \in \{b, f\}$ for big and small - what we call "fringe" - banks, respectively. Unlike in our earlier paper (Corbae and D'Erasmo [21]), where there are multiple big banks, to keep the analysis simple, we assume there is a representative big bank if it is active (as in Gowrisankaran and Holmes [37]). If active, the big bank is a Stackelberg leader in the loan market, each period choosing a level of loans before fringe banks make their choice of loan supply. Consistent with a Cournot framework, the dominant bank understands that its choice of loan supply will influence the interest rate on loans given the best response of fringe banks. A fringe bank takes the interest rate as given when choosing its loan supply.

At the beginning of each period t after the realization of the aggregate shock z_t, the net cash flows (denoted $\pi^i_{\theta,t}$) for bank i of type θ are realized from its previous lending (denoted $\ell^i_{\theta,t}$), liquid assets (cash and securities, denoted $A^i_{\theta,t}$), and deposits (denoted $d^i_{\theta,t}$). This defines its beginning of period equity capital (or net worth) $k^i_{\theta,t}$:

$$k^i_{\theta,t} \;\; = \;\; \ell^i_{\theta,t} + A^i_{\theta,t} + \pi^i_{\theta,t} - d^i_{\theta,t}. \tag{2}$$

Equation (2) is the bank balance sheet identity where equity and deposits equal loans, securities and bank net cash holdings. At that point, a bank can choose to exit. If the bank chooses not to exit, the incumbent is randomly matched with a set of potential household depositors $\delta_{\theta,t}$. An incumbent bank then chooses a quantity $d^i_{\theta,t+1}$ of deposits to accept up to the capacity constraint $\delta_{\theta,t}$ (i.e., $d^i_{\theta,t+1} \leq \delta_{\theta,t}$ where $\delta_{\theta,t} \in \{\delta^1_\theta, \ldots, \delta^N_\theta\} \subseteq \mathbb{R}_+$) at interest rate $r^{D,i}_{\theta,t}$ that it offers each potential depositor.[22] The capacity constraint evolves according to an exogenously given Markov process $G_{\theta,t+1}(\delta_{\theta,t+1}, \delta_{\theta,t})$ with realizations which are i.i.d across banks. The value of $\delta_{\theta,t}$ for a new entrant is drawn from the probability distribution $G^e_\theta(\delta_{\theta,t})$. Differences in the volatility of funding inflows we find in the data between big and small banks provide a reason why banks of different sizes hold different size capital buffers. Since the household can always store at rate r^s, we know $r^{D,i}_{\theta,t} \geq r^s$.

Along with possible seasoned equity injections (denoted $e^i_{\theta,t} \in \mathbb{R}_+$), an incumbent bank allocates its equity capital and deposits to its asset portfolio and pays dividends (denoted $\mathcal{D}^i_{\theta,t}$).[23] We follow Title 12 "Banks and Banking" of the Code of Federal Regulations that states that proposed dividends cannot exceed a bank's net income, and restrict dividends $\mathcal{D}^i_{\theta,t} \in [0, E_t(\pi^i_{\theta,t+1})]$.[24] We assume liquid assets (e.g. U.S. treasuries) have a return equal to r^A, which the government takes as given. The incumbent bank's portfolio and dividend policy must satisfy the following constraint

$$k^i_{\theta,t} + e^i_{\theta,t} + d^i_{\theta,t+1} \geq \ell^i_{\theta,t+1} + A^i_{\theta,t+1} + \mathcal{D}^i_{\theta,t} + \zeta_\theta(e^i_{\theta,t}, z_t) + \kappa^i_\theta + c^i_\theta\left(\ell^i_{\theta,t+1}, z_t\right), \tag{3}$$

where $\zeta_\theta(e^i_{\theta,t}, z_t)$ denotes aggregate state dependent equity issuance costs and $[\kappa_\theta + c_\theta(\ell^i_{\theta,t+1}, z_t)]$ represents non-interest expenses (including monitoring costs which are a function of loans issued).

in the project not being funded, a dominated strategy which implements Diamond's arrangement.

[22] Anticipating a "recursive" formulation of the bank decision problem, certain state variables chosen in period t but paying off in period $t+1$ will be denoted q_{t+1} (e.g. deposits d_{t+1}).

[23] While less important than deposit funding, unsecured wholesale funding plays a role for some banks, especially those at the top of the asset distribution. It is beyond the scope of this paper to add an endogenous uninsured external financing decision besides the choice to issue seasoned equity that we already have on the liability side.

[24] See https://www.govinfo.gov/app/details/CFR-2011-title12-vol2/CFR-2011-title12-vol2-sec208-5.

We assume that $\zeta_\theta(0, z_t) = 0$ and $\zeta_\theta(e^i_{\theta,t}, z_t)$ is an increasing function of $e^i_{\theta,t}$ and decreasing in z_t (i.e. external financing costs are increasing in the amount of equity issued and less costly in good times).[25] Since the bank's objective is increasing in dividends $\mathcal{D}^i_{\theta,t}$, (3) will hold as an equality constraint.

As we document in Table 3 below, banks' cost structures differ across size. We assume that banks pay non-interest expenses on their loans (as in the delegated monitoring model of Diamond [27]) that differ across banks of different sizes, which we denote $c_\theta\left(\ell^i_{\theta,t+1}, z_t\right)$. Further, as in the data we assume a fixed cost κ_θ.

The net cash flow of bank i of type θ after the realization of the next period's aggregate shock associated with its current lending and borrowing decisions, $\pi^i_{\theta,t+1}$, is given by

$$\pi^i_{\theta,t+1} = \left\{p(R^c_t, z_{t+1})r^c_t - (1 - p(R^c_t, z_{t+1}))\lambda\right\}\ell^i_{\theta,t+1} + r^A_t A^i_{\theta,t+1} - r^D_t d^i_{\theta,t+1}. \tag{4}$$

The first two terms represent the gross return the bank receives from successful and unsuccessful loan projects, respectively, the third term represents returns on securities, and the fourth represents interest expenses (payments on deposits).

Combining equation (2) evaluated at $t + 1$ with equation (3) with equality, we obtain that beginning-of-next-period equity (or net worth) given by

$$k^i_{\theta,t+1} = k^i_{\theta,t} + e^i_{\theta,t} + \pi^i_{\theta,t+1} - \mathcal{D}^i_{\theta,t} - \zeta_\theta(e^i_{\theta,t}, z_t) - \kappa_\theta - c_\theta\left(\ell^i_{\theta,t+1}\right). \tag{5}$$

The law of motion for net worth in (5) makes clear that retained earnings augment net worth and dividend payouts lower net worth.

Using the definition of equity in (2), when making loan, securities, and deposit decisions, bank i of type θ faces a constraint that it has to hold sufficient equity at the beginning of the next period to meet its risk weighted capital requirement in states of the world in which the bank does not exit. Since $\ell^i_{\theta,t+1}$, $A^i_{\theta,t+1}$, and $d^i_{\theta,t+1}$ are chosen in period t and $\pi^i_{\theta,t+1}$ is a function of z_{t+1}, that implies banks need to hold sufficient capital to meet the requirement for all z_{t+1} such that the bank chooses to remain active. Denoting the decision to exit by $x_{t+1} \in \{0, 1\}$ (equal to 1 when the bank exits), the capital requirement constraint is

$$\begin{aligned} k^i_{\theta,t+1} \equiv \quad & \ell^i_{\theta,t+1} + A^i_{\theta,t+1} + \pi^i_{\theta,t+1} - d^i_{\theta,t+1} \\ & \geq \varphi_{\theta,t}(w^\ell_{\theta,t}\ell^i_{\theta,t+1} + w^A_{\theta,t}(A^i_{\theta,t+1} + \pi^i_{\theta,t+1})) \;\; \forall \; z_{t+1} \text{ such that } x_{t+1} = 0, \end{aligned} \tag{6}$$

where $\varphi_{\theta,t}$ is the capital requirement and $\left(w^\ell_{\theta,t}, w^A_{\theta,t}\right)$ are risk weights associated with loans and liquid assets, respectively. We will typically take $w^\ell_{\theta,t} = 1$. Given $w^\ell_{\theta,t} > w^A_{\theta,t}$, liquid assets help relax the capital requirement constraint, but may also lower bank profitability. Lower bank profitability is the reason why the capital requirement constraint is binding in equilibrium in many macro models, while as we saw in Section 2.2 most banks hold a buffer of risk weighted capital above the regulatory constraint. The reason the constraint may only occasionally bind in our model is due to the presence of the idiosyncratic liquidity shock process $G_{\theta,t+1}(\delta_{\theta,t+1}, \delta_{\theta,t})$ and costly equity issuance $\zeta_\theta(e^i_{\theta,t}, z_t)$ in the presence of negative net cash flows, along with the desire to maintain access to the bank's charter value.

Another policy proposal is associated with bank liquidity requirements. Basel III [7] proposed that the liquidity coverage ratio, which is the stock of high-quality liquid assets (including government securities) divided by total net cash outflows over the next 30 calendar days, should exceed

[25]This "reduced form" approach to modeling equity issuance is similar to Cooley and Quadrini [20], Gomes [36], and Hennessy and Whited [38].

100% under a stress scenario. In the context of a model period being one year, this would amount to a critical value of 1/12 or roughly 8%. This is also close to the figure for reserve requirements that is bank-size dependent, anywhere from zero to 10%. For the model, we implement a liquidity requirement as:

$$\varrho_{\theta,t} d^i_{\theta,t+1} \leq A^i_{\theta,t+1} + \pi^i_{\theta,t+1}(z_{t+1} = \underline{z}), \tag{7}$$

where $\varrho_{\theta,t}$ denotes the (possibly) size and state dependent liquidity requirement and cash $\pi^i_{\theta,t+1}(z_{t+1} = \underline{z})$ is evaluated in a stress scenario.[26]

There is limited liability on the part of banks. This imposes a lower bound equal to zero in the event the bank exits. In the context of our model, limited liability implies that, upon exit, shareholders get:

$$\max\left\{k^i_{\theta,t+1} - \xi \ell^i_{\theta,t+1}, 0\right\}, \tag{8}$$

where $\xi \in [0,1]$ measures liquidation costs of an insolvent loan portfolio in the event of exit.

The objective function of the bank is to maximize the expected present discounted value of future dividends net of equity injections using the manager's discount factor which can depart from the households' discount factor β_H by the factor $\gamma \in (0,1]$:

$$E_t\left[\sum_{g=0}^{\infty} (\gamma\beta_H)^g \left(\mathcal{D}^i_{\theta,t+g} - e^i_{\theta,t+g}\right)\right]. \tag{9}$$

This introduces the possibility of agency problems through managerial myopia when $\gamma < 1$ along the lines of Acharya and Thakor [1].[27] In order to obtain a well defined distribution of banks, we need a condition that guarantees that $\gamma\beta_H(1 + r^A) < 1$, a standard assumption in models with incomplete markets as in our case. Since we assume $\beta_H(1 + r^s) = 1$ to keep the household decision problem bounded, we assume $\frac{1+r^s}{1+r^A} > \gamma$ which assures a bounded distribution over bank net worth.

Entry costs for the creation of banks are denoted by $\Upsilon_b \geq \Upsilon_f \geq 0$. Every period a large number of potential entrants make the decision of whether or not to enter the market after the realization of z_t and incumbent exit but before the realization of δ_t shocks. Entry costs correspond to the initial injection of equity into the bank subject to equity finance costs $\zeta_\theta(\Upsilon_\theta + k^i_{e,\theta,t}, z_t)$, where $k^i_{e,\theta,t}$ is the entrant's initial equity injection.

3.4 Non-Bank Lenders

A representative non-bank that discounts the future at rate β_H specializes in extending loans to entrepreneurs in a perfectly competitive market. To keep the analysis simple, the non-bank is financed with equity raised from the household sector and is not subject to limited liability. When lending to entrepreneurs non-banks face a marginal monitoring cost c_n. Like banks, the representative non-bank can diversify entrepreneurs' idiosyncratic risk but it is subject to aggregate fluctuations. Let $\pi_{n,t+1}$ denote the cash flow of the non-bank after the realization of next period's aggregate shock associated with its current lending $\ell_{n,t+1}$ given by

$$\pi_{n,t+1} = \left\{p(R^n_t, z_{t+1})r^n_t - (1 - p(R^n_t, z_{t+1}))\lambda\right\}\ell_{n,t+1}.$$

[26] Notice that an increase in $A^i_{\theta,t+1}$ and decrease in $d^i_{\theta,t+1}$ help to satisfy both the risk weighted capital requirement in (6) and the liquidity coverage ratio in (7).

[27] There are many papers on managerial myopia providing a foundation for such behavior. See for instance, Stein [61], who provides a signalling argument, or Minnick and Rosenthal [50], who provide a compensation argument.

Since the non-bank operates in a perfectly competitive market it takes the interest rate r_t^n as given. The non-bank lending and dividend/equity issuance policy $(\mathcal{D}_{n,t} - e_{n,t})$ satisfies the following flow constraint

$$\mathcal{D}_{n,t} - e_{n,t} = \pi_{n,t} + \ell_{n,t} - \ell_{n,t+1} - c_n \ell_{n,t+1}. \tag{10}$$

The objective function of the non-bank is to maximize the expected present discounted value of future cash-flows to households

$$E_t \left[\sum_{g=0}^{\infty} \beta_H^g \left(\mathcal{D}_{n,t+g} - e_{n,t+g} \right) \right], \tag{11}$$

where $\ell_{n,0} = 0$. It is important to note that the non-bank is not subject to regulatory constraints imposed on commercial banks. We assume that there is free entry into the non-bank sector, and to simplify the analysis we set the entry cost to zero.

3.5 Information

There is asymmetric information on the part of borrowers and lenders (banks and households). Only borrowers know the riskiness of the project they choose (R_t^j) and their outside option (ω_t). Success or failure of their project is only observable after observing a cost. To maintain consistency with payoffs between project choice and outside option, they receive a perfect unobservable signal about their outside option the prior period. Other information is observable.

3.6 Timing

At the beginning of period t,

1. Aggregate shock z_t and borrower opportunity ω_t are realized, which induce project returns $\pi_{E,t}^j$ for entrepreneurs funded by lender type $j \in \{c, n\}$, $\pi_{\theta,t}^i$ and $k_{\theta,t}^i$ for incumbent banks, and $\pi_{n,t}$ for non-banks.

2. Incumbents decide whether to exit and losses at failing banks determine τ_t^D.

    - Capital requirement and liquidity requirement constraints are imposed for continuing banks.

3. Potential entrants decide whether to enter, which requires an initial equity injection.

4. Funding shocks δ_t - the mass of potential depositors the bank is matched with - are realized. After observing a private signal about their future opportunities ω_{t+1}, borrowers choose whether or not $\iota_t \in \{0, 1\}$ to undertake a project requiring funding and, if so, they draw $\boldsymbol{\epsilon_t}$.

    - Those borrowers who choose to undertake a project choose the type of lender $j_t \in \{c, n\}$ and the level of technology R_t^j. All choose how much to save $a_{E,t+1}$ and retained earnings I_{t+1} implying a level of consumption.
    - The dominant bank chooses how many loans to extend, how many deposits to accept given depositors' choices, how many securities to hold, how many dividends to pay, and equity injections $(\ell_{b,t+1}^i, d_{b,t+1}^i, A_{b,t+1}^i, \mathcal{D}_{b,t}^i, e_{b,t}^i)$.

- Each fringe bank observes the total loan supply of the dominant bank ($\ell^i_{b,t+1}$) and all other fringe banks (that jointly determine the loan interest rate r^L_t) and simultaneously decide how many loans to extend, how many deposits to accept, how many securities to hold, how many dividends to pay, and equity injections ($\ell^i_{f,t+1}, d^i_{f,t+1}, A^i_{f,t+1}, \mathcal{D}^i_{f,t}, e^i_{f,t}$).
- The representative non-bank chooses how many loans to extend, how many dividends to pay, and equity injections ($\ell_{n,t+1}, \mathcal{D}_{n,t}, e_{n,t}$).
- Households pay taxes/transfers $\tau_t = \tau^D_t + \tau^A_t$ to fund deposit insurance (τ^D_t) and service government securities (τ^A_t); choose to store or deposit at a bank, how many banks and non-bank stocks to hold, equity injections for banks and non-banks; and consume.

4 Equilibrium

This section presents the equilibrium of the model. We start by describing the entrepreneur problem (which determines the demand for bank loans) and the household problem (which determines the supply of deposits and seasoned equity to banks and non-banks), followed by the bank problem.

4.1 Entrepreneur Problem

Every period, given $\mathbf{r_t} = \{r^c_t, r^n_t\}$, z_t, and the signal for ω_{t+1}, entrepreneurs choose whether ($\iota_t = 1$) or not ($\iota_t = 0$) to operate the technology. Conditional on choosing $\iota_t = 1$, entrepreneurs observe $\boldsymbol{\epsilon_t} = \{\epsilon^c_t, \epsilon^n_t\}$ and choose the type of lender $j_t \in \{c, n\}$ and technology to operate R^j_t, whether to use retained earnings $I_{t+1} \in [0, 1]$ to internally finance the project, and how much to save $a_{E,t+1} \in \mathbb{R}_+$ to solve

$$\max_{\{a_{E,t+1}, I_{t+1}, \iota_t, j_t, \{R^j_t\}_{j=c,n}\}^\infty_{t=0}} E_0 \left[\sum_{t=0}^{\infty} \beta^t_E \left\{ \begin{array}{c} (1 - \iota_{E,t})(\omega_{t+1} + I_t) + \iota_{E,t}(\pi^{j_t}_E(r^{j_t}_t, I_t, R^{j_t}_t, z_{t+1}) + \epsilon^{j_t}_t) \\ +(1 + r^s)a_{E,t} - a_{E,t+1} - I_{t+1} \end{array} \right\} \right] \tag{12}$$

where

$$\pi^j_E(r^j_t, I_t, R^j_t, z_{t+1}) = \begin{cases} \max\{0, (z_{t+1}R^j_t - r^j_t)\bar{\ell} + (1 + r^j_t)I_t\} & \text{with prob } p(R^j_t, z_{t+1}) \\ \max\{0, -(\lambda + r^j_t)\bar{\ell} + (1 + r^j_t)I_t\} & \text{with prob } 1 - p(R^j_t, z_{t+1}) \end{cases}$$

and $a_{E,0} = I_0 = 0$ and ω_0 given. The expected value of operating the technology with a loan from lender of type j is $\overline{\pi}^j_E(r^j_t, I_t, R^j_t, z_t) = E_{z_{t+1}|z_t}[\pi^j_E(r^j_t, I_t, R^j_t, z_{t+1})]$ so, the optimal choice of lender type solves

$$\Pi_E(\mathbf{r}_t, I_t, \mathbf{R}_t, z_t, \boldsymbol{\epsilon_t}) = \max_{j \in \{c,n\}} \{\overline{\pi}^j_E(r^j_t, I_t, R^j_t, z_t) + \epsilon^j_t\}. \tag{13}$$

where $\mathbf{R}_t = \{R^c_t, R^n_t\}$.

If η_t is the multiplier on the non-negativity constraint on $a_{E,t+1} \geq 0$, the first order condition for $a_{E,t+1}$ is given by

$$\eta_t = 1 - \beta_E(1 + r^s). \tag{14}$$

Since we assume a sufficiently impatient entrepreneur (i.e. $\beta_E(1 + r^s) < 1$), then $a_{E,t+1} = 0$. Similarly, provided $\beta_E(1+r^s)(1+r^j) < 1$ (a condition we must verify in equilibrium), the entrepreneur chooses not to use retained earnings to fund the project (i.e. $I_{t+1} = 0$). With $I_{t+1} = 0$, reporting failure is a dominant strategy in the absence of monitoring since $(1 - \lambda) < (1 + r^j_t)$.

The solution to (13) together with the optimal choice of technology R_t^j as a function of $\mathbf{r}_t$ imply that the share of borrowers choosing a loan from a commercial bank (i.e. $j = c$) is

$$s^c(\mathbf{r}_t, z_t) = \frac{\exp(\alpha \overline{\pi}_E^c(r_t^c, 0, R_t^c(r_t^c, z_t), z_t))}{\sum_j \exp(\alpha \overline{\pi}_E^j(r_t^j, 0, R_t^j(r_t^j, z_t), z_t))}, \tag{15}$$

and the share of borrowers choosing a loan from a non-bank is $1 - s^c(\mathbf{r}_t, z_t)$. The expected value of taking out a loan evaluated at $I_t = 0$ is[28]

$$\mathbf{\Pi}_E(\mathbf{r}_t, z_t) \quad = \quad \int \Pi_E(\mathbf{r}_t, 0, \mathbf{R}_t(\mathbf{r}_t, z_t), z_t, \boldsymbol{\epsilon_t}) dF(\epsilon_t). \tag{16}$$

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

$$\frac{\partial E_{z_{t+1}|z_t} \pi_E(r^j, 0, R_t^j, z_{t+1})}{\partial r_t^j} = -E_{z_{t+1}|z_t}[p(R_t^j, z_{t+1})] < 0. \tag{17}$$

Thus, participating borrowers (i.e. those who choose to run a project rather than take the outside option) are worse off the higher is the interest rate on loans. This has implications for the aggregate demand for loans determined by the participation decision (i.e. $\omega_{t+1} \leq \mathbf{\Pi}_E$). In particular, the demand for loans is given by

$$L^d(\mathbf{r}_t, z_t) = \overline{\ell} \int_0^{\overline{\omega}} \mathbf{1}_{\{\omega_{t+1} \leq \mathbf{\Pi}_E(\mathbf{r}_t, z_t)\}} d\Omega(\omega_{t+1}), \tag{18}$$

where $\mathbf{1}_{\{.\}}$ is an indicator function that takes the value one if the argument $\{\cdot\}$ is true and zero otherwise. Then loan demand from commercial banks is given by $L^{d,c}(\mathbf{r}_t, z_t) = s^c(\mathbf{r}_t, z_t) L^d(\mathbf{r}_t, z_t)$.

In that case, everything else equal, (17) implies $\frac{\partial L^{d,c}(\mathbf{r}_t, z_t)}{\partial r_t^c} < 0$. That is, the bank loan demand curve is downward sloping. Further, bank market shares are decreasing in bank lending rates (i.e. $\frac{\partial s^c(\mathbf{r}_t, z_t)}{\partial r_t^c} < 0$) and aggregate loan demand decreases with an increase in bank lending rates (i.e. $\frac{\partial L^d(\mathbf{r}_t, z_t)}{\partial r_t^c} \leq 0$). Given the presence of nonbanks, however, restrictive regulatory policy has less of an effect on aggregate loan demand (i.e. $\frac{\partial L^{d,c}(\mathbf{r}_t, z_t)}{\partial r_t^c} \leq \frac{\partial L^d(\mathbf{r}_t, z_t)}{\partial r_t^c}$).

4.2 Household Problem

The problem of a representative household is

$$\max_{\{C_{H,t}, a_{H,t+1}, d_{H,t+1}^c, S_{t+1}^n, \{S_{\theta,t+1}^i\}_{\forall i}\}_{t=0}^{\infty}} E_0 \left[\sum_{t=0}^{\infty} \beta_H^t C_{H,t} \right] \tag{19}$$

subject to

$$\begin{aligned} & C_{H,t} + a_{h,t+1} + d_{h,t+1}^c + \sum_{\theta} \int [P_{\theta,t}^i + \mathbf{1}_{\{e_{\theta,t}^i = 1\}} (\Upsilon_\theta + k_{e,\theta,t}^i)] S_{\theta,t+1}^i di + S_{n,t+1} P_{n,t} \qquad (20) \\ = \quad & y + \sum_{\theta} \int (\mathcal{D}_{\theta,t}^i - e_{\theta,t}^i + P_{\theta,t}^i) S_{\theta,t}^i di + (1 + r^s) a_{h,t} + (1 + r_t^{D,c}) d_{h,t}^c + (\mathcal{D}_{n,t} - e_{n,t} + P_{n,t}) S_{n,t} - \tau_t, \end{aligned}$$

[28]Specifically, $\mathbf{\Pi}_E(\mathbf{r}_t, z_t) = \frac{\gamma_E}{\alpha} + \frac{1}{\alpha} \ln \left(\sum_j \exp(\alpha \Pi_E^j(r_t^j, 0, R_t^j(r_t^j, z_t), z_t)) \right)$ where γ_E is Euler's constant.

where $P^i_{\theta,t}$ and $S^i_{\theta,t+1}$ are the post-dividend stock price and stock holding of bank i of type θ, respectively, and P^n_t and S^n_{t+1} are the post-dividend stock price and stock holding of the non-bank with given initial conditions $a_{H,0} = d^c_{H,0} = S^i_{\theta,0} = S^n_0 = 0$. Given exit and entry decision rules, in cases in which a firm has exited, $P^i_{\theta,t} = 0$ on the right-hand side of the budget constraint, and, in cases in which a firm has entered, $P^i_{\theta,t} > 0$ on the left-hand side of the budget constraint.

The first order condition for bank stock holding $S^i_{\theta,t+1}$ is:

$$P^i_{\theta,t} = \beta_H E_{z_{t+1}|z_t}\left[\mathcal{D}^i_{\theta,t+1} - k^i_{\theta,t+1} + P^i_{\theta,t+1}\right], \forall i. \tag{21}$$

We will derive the expression for the equilibrium price of a bank share after we present the bank's problem. A similar condition holds for non-bank stock holding.

If banks offer the same interest rates on deposits as households can receive from their storage opportunity (i.e. $r^D_{t+1} = r^s$), then a household would be indifferent between matching with a bank and using the autarkic storage technology. In that case, any household who is matched with a bank would be willing to deposit at the insured bank. Furthermore, the first order condition for saving in the form of deposits or storage technology implies $\beta_H(1+r^s) = 1$, which we assume parametrically.

4.3 Incumbent Bank Problem

We will study equilibria which do not depend on the name i of the bank, only relevant state variables. Since we will use recursive methods to solve a bank's decision problem, let any variable $g_{\theta,t}$ be denoted g_θ and $g_{\theta,t+1}$ be denoted g'_θ. Further, we denote the cross-sectional distribution of banks or "industry state" by

$$\mu = \{\mu_b(k,\delta), \mu_f(k,\delta)\}, \tag{22}$$

where each element of μ is a measure $\mu_\theta(k,\delta)$ corresponding to *active* banks of type θ over matched deposits δ and net worth k at stage 4 in period t of our timing. The law of motion for the industry state is denoted $\mu' = F(z, \mu, \ell'_b, r^n, z', M'_e)$, where $M'_e = \{M'_{e,b}, M'_{e,f}\}$ denotes the vector of entrants of each type and the transition function F is defined explicitly below.

After being matched with δ_θ potential depositors and making them a take-it-or-leave-it deposit rate offer r^D_θ, an incumbent bank of type θ chooses loans ℓ'_θ, deposits d'_θ, and asset holdings A'_θ in order to maximize expected discounted dividends net of equity injections. Following the realization of z, banks can choose to exit setting $x_\theta = 1$ or choose to remain $x_\theta = 0$. Given the take-it-or-leave-it deposit rate offer and that the outside storage option for a household is r^s, we know in equilibrium $r^D_\theta = r^s$.

Given the Stackelberg assumption, the big bank takes into account that its loan supply affects the bank loan interest rate and that fringe banks will best respond to its loan supply. Differentiating the bank profit function π'_θ defined in (4) with respect to ℓ_θ we obtain

$$\frac{d\pi'_\theta}{d\ell'_\theta} = \underbrace{\left[pr^c - (1-p)\lambda - \frac{\partial c_\theta}{\partial \ell'_\theta}\right]}_{(+)\text{ or }(-)} + \ell'_\theta\Big[\underbrace{p}_{(+)} + \underbrace{\frac{\partial p}{\partial R^c}\frac{dR^c}{dr^c}(r^c+\lambda)}_{(-)}\Big]\underbrace{\frac{dr^c}{d\ell'_\theta}}_{(-)}. \tag{23}$$

The first bracket represents the marginal change in profits from extending an extra unit of loans. The second bracket corresponds to the marginal change in profits due to a bank's influence on the interest rate it faces. This term depends on the bank's market power; for big banks $\frac{dr^c}{d\ell'_b} < 0$ while for fringe banks $\frac{dr^c}{d\ell'_f} = 0$. Note that a change in interest rates also endogenously affects the fraction of delinquent loans faced by banks (the term $\frac{\partial p}{\partial R^c}\frac{dR^c}{dr^c} < 0$). That is, given limited liability entrepreneurs take on more risk when their financing costs rise.

Let the total supply of loans by fringe banks as a function of the aggregate state $\{z, \mu\}$ and the big bank's choice of loans ℓ'_b be given by

$$L^s_f(z, \mu, \ell'_b, r^n) = \int \ell'_f(k, \delta; z, \mu, \ell'_b, r^n)\mu_f(dk, d\delta). \tag{24}$$

The loan supply of fringe banks is a function of the big bank's loan supply ℓ'_b because fringe banks move after the big bank.

The value of an incumbent bank in period t (at stage 4) consistent with the manager's choice over $\{\{\ell'_\theta, A'_\theta, e_\theta\} \geq 0, \mathcal{D}_\theta \in [0, E[\pi'_\theta]], d'_\theta \in [0, \delta_\theta], x'_\theta \in \{0, 1\}\}$ is given by

$$\begin{aligned} V_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n) = & \max_{\{\ell'_\theta, A'_\theta, e_\theta\} \geq 0, \mathcal{D}_\theta \in [0, E[\pi'_\theta]], d'_\theta \in [0, \delta_\theta]} \Bigg\{ \mathcal{D}_\theta - e_\theta \\ & + \gamma \beta_H E_{z'|z} \left[\max_{x'_\theta \in \{0,1\}} \left\{ (1 - x'_\theta) E_{\delta'_\theta, r^{n'} | \delta_\theta, r^n} V_\theta(k'_\theta, \delta'_\theta; z', \mu', \chi', r^{n'}) + x'_\theta V^x_\theta(k'_\theta, \ell'_\theta) \right\} \right] \Bigg\} \end{aligned} \tag{25}$$

s.t.

$$\begin{aligned} k_\theta + d'_\theta + e_\theta & \geq \ell'_\theta + A'_\theta + \mathcal{D}_\theta + \zeta_\theta(e_\theta, z) + [\kappa_\theta + c_\theta\left(\ell'_\theta, z\right)] & (26) \\ k'_\theta & = \pi'_\theta + \ell'_\theta + A'_\theta - d'_\theta & (27) \\ k'_\theta & \geq \varphi_{\theta,z}(w^\ell_\theta \ell'_\theta + w^A_{\theta,z}(A'_\theta + \pi'_\theta)) \ \ \forall x'_\theta = 0 & (28) \\ \varrho_{\theta,z} d'_\theta & \leq A'_\theta + \pi'_\theta(z' = \underline{z}) & (29) \\ L^{d,c}(\mathbf{r}, z) & = \ell'_b + L^s_f(z, \mu, \ell'_b, r^n) & (30) \\ \mu' & = F(z, \mu, \ell'_b, r^n, z', M'_e), & (31) \end{aligned}$$

where the value function in (25) is defined over individual states $\{k_\theta, \delta_\theta\}$, aggregate states $\{z, \mu\}$, and $\chi = \emptyset$ if $\theta = b$ while $\chi = \ell'_b$ if $\theta = f$.[29]

Equations (26) to (29) are the bank's budget constraint, balance sheet constraint, capital requirement constraint, and liquidity coverage ratio constraint, respectively. Equation (30) is the market-clearing condition which is included since the dominant bank must take into account its impact on prices. For any given μ, $L^s_f(z, \mu, \ell'_b, r^n)$ can be thought of as a "reaction function" of fringe banks to the loan supply decision of the dominant bank. Changes in ℓ'_b affect the equilibrium interest rate through its direct effect on the aggregate loan supply (first term) but also through the effect on the loan supply of fringe banks (second term). Equation (31) corresponds to the evolution of the aggregate state to be defined below and again reflects the fact that big bank lending can have an impact on the distribution indirectly through its influence on fringe bank choices. The liquidation value of the bank for a given n'_θ and ℓ'_θ is

$$V^x_\theta(k'_\theta, \ell'_\theta) \quad = \quad \max\left\{0, k'_\theta - \xi_\theta \ell'_\theta\right\}. \tag{32}$$

The lower bound on the exit value in (32) is associated with limited liability where ξ_θ is the liquidation cost of the loan portfolio of an insolvent bank.

The solution to this problem provides value functions as well as bank decision rules that can be written as $\ell'_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n)$, $A'_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n)$, $\mathcal{D}_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n)$, $e_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n)$, $d'_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n)$, and $x'_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n, z')$.

[29] χ dependence captures the fact that fringe banks take as given the loan supply decision by the big bank in our Stackelberg game (not only on the equilibrium path but any arbitrary value of ℓ'_b).

Now that we have presented the problem of an incumbent bank, we can show how the price of a bank's shares and the value of a bank are related. After normalizing the number of shares of each bank to 1, the price of a share of a non-exiting bank of type θ in state $(k_\theta, \delta_\theta; z, \mu, \chi, r^n)$ after dividends have been paid and equity injected is defined by

$$P_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n) \equiv V_\theta^*(k_\theta, \delta_\theta; z, \mu, \chi, r^n) - (\mathcal{D}_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n) - e_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n)), \quad (33)$$

where $V_\theta^*(k_\theta, \delta_\theta; z, \mu, \chi, r^n)$ denotes the value of the expected discounted stream of net cash-flows using the household discount factor. Thus, substituting (33) into the household's first order condition for its stock choice in equation (21) yields

$$\begin{aligned} P_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n) &= \beta_H E_{z',\delta'|z,\delta}\Big[\mathcal{D}_\theta(k'_\theta, \delta'_\theta; z', \mu', \chi', r^{n'}) - e_\theta(k'_\theta, \delta'_\theta; z', \mu', \chi', r^{n'}) + P'_\theta(k'_\theta, \delta'_\theta; z', \mu', \chi', r^{n'})\Big] \\ &\iff V_\theta^*(k_\theta, \delta_\theta; z, \mu, \chi, r^n) - (\mathcal{D}_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n) - e_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n)) \\ &= \beta_H E_{z',\delta'|z,\delta}\Big[V_\theta^*(k'_\theta, \delta'_\theta; z', \mu', \chi', r^{n'})\Big]. \end{aligned} \quad (34)$$

Equation (34) can be re-arranged to be identical to the value of a continuing bank defined in (25) when managers' and households' preferences are aligned (i.e. when $\gamma = 1$) while $V_\theta(\cdot) \leq V_\theta^*(\cdot)$ otherwise.

4.4 Bank Entry

Next we turn to the value of entry. Both the industry state μ and the incumbent value function above in (25) are defined for stage 4 in period t of our timing. However, the entry decision is at stage 3 after exit but before the current mass of entrants $M_{e,\theta}$ is known (so that μ is not yet fully defined at that stage). Hence, we will define the entry value function in terms of stage 3 for period $t+1$. In particular, the value of entry net of entry costs for banks of type θ is given by

$$\begin{aligned} V_\theta^e(z, \mu, \ell'_b, r^n, z', M'_{e,\theta}) \equiv \max_{k'_{e,\theta}} \Big\{ &- (k'_{e,\theta} + \Upsilon_\theta)(1 + \zeta_\theta(k'_{e,\theta} + \Upsilon_\theta, z')) \\ &+ E_{\delta', r^{n'}} V_\theta(k'_{e,\theta}, \delta'_\theta, z', F(z, \mu, \ell'_b, r^n, z', M'_{e,\theta}), \chi', r^{n'}) \Big\}. \end{aligned} \quad (35)$$

Potential entrants will decide to enter if $V_\theta^e(z, \mu, \ell^{b'}, r^n, z', M'_{e,\theta}) \geq 0$. The argmax of equation (35) for those firms that enter defines the initial equity injection of banks. The mass of entrants is determined endogenously in equilibrium. Free entry implies that

$$V_\theta^e(z, \mu, \ell'_b, r^n, z', M'_{e,\theta}) \times M'_{e,\theta} = 0. \quad (36)$$

That is, in equilibrium, either the value of entry is zero, the number of entrants is zero, or both.

4.5 Evolution of the Cross-Sectional Bank Size Distribution

The distribution of banks evolves according to $\mu' = F(z, \mu, \ell'_b, r^n, z', M'_e)$ where each component is given by:

$$\begin{aligned} \mu'_\theta(k'_\theta, \delta'_\theta) = \int (1 - x'_\theta(k_\theta, \delta_\theta; z, \mu, \chi, r^n, z'))\mathbf{1}_{\{k'_\theta = k'_\theta(k_\theta, \delta_\theta, z, \mu, \chi, z'))\}} G_\theta(\delta'_\theta, \delta_\theta) d\mu_\theta(k_\theta, \delta_\theta) \\ + M'_{e,\theta}\mathbf{1}_{\{k'_\theta = k'_{e,\theta}(z, \mu, \ell'_b, r^n, z', M'_{e,\theta})\}} G_{e,\theta}(\delta'_\theta). \end{aligned} \quad (37)$$

Equation (37) makes clear how the law of motion for the distribution of banks is affected by entry and exit decisions.

4.6 Funding Deposit Insurance and Servicing Securities

The government collects lump-sum taxes (or pays transfers if negative) denoted τ that cover the cost of deposit insurance τ^D and the net proceeds of issuing securities τ^A.[30] Per-capita taxes are

$$\tau'(z, \mu, r^n, z') = \tau^{D'}(z, \mu, r^n, z') + \tau^{A'}(z, \mu, r^n, z'). \tag{38}$$

4.7 Non-Bank Problem

The representative non-bank operates in a competitive industry, so when making lending decisions it takes the loan interest rate r_t^n as given. In any state z, and taking into account that $\beta(1+r^s) = 1$, the first order condition of the non-bank with respect to ℓ_{t+1}^n is given by

$$r^s = E_{z'|z}\left[p(R^s(r^n, z), z') - (1 - p(R^n(r^n, z), z'))\lambda\right] - c_n(1 + r^s), \tag{39}$$

where $R^n(r^n, z)$ is the optimal choice of technology by the entrepreneur when taking a loan from a non-bank facing interest rate r^n. Equation (39) is one equation in one unknown which pins down the interest rate r^n of the non-bank sector as a function of z. The fact that r^n is independent of the entire distribution of banks is a form of block recursivity as in Menzio and Shi [48]. Evaluating the non-bank loan demand at this price we can determine the level of lending of the non-bank. Equation (39) also makes clear that the expected net return between a bank deposit and non-bank investment is equalized, with the spread depending on c_n. However, while depositing with a bank guarantees a risk-free return (since there is deposit insurance), equity injections in a non-bank are subject to aggregate risk.

4.8 Definition of Equilibrium

Given policy parameters $(r^s, r^A, \varphi_{\theta,z}, w_{\theta,z}^\ell, w_{\theta,z}^A, \varrho_{\theta,z})$, a *Markov Perfect Industry Equilibrium* is a set of entrepreneur financing choices $\{a_E', I', \iota, R^j\}$, a set of household saving choices $\{a_h', d_h', S_\theta', \varsigma'\}$, bank values and decisions $\{V_\theta, \ell_\theta', d_\theta', A_\theta', \mathcal{D}_\theta, e_\theta, x_\theta', V_\theta^e\}$, a cross-sectional distribution of banks μ, the mass of entrants M_θ^e, a bank loan interest rate $r^c(\mu, z)$, a non-bank loan interest rate $r^n(z)$, a deposit interest rate r^D, stock prices P_θ and P^n, and a tax function τ' such that:

1. Given $\mathbf{r} = \{r^c, r^n\}$ and r^s, $\{a_E', I', \iota, R^j\}$ are consistent with entrepreneur optimization (12) inducing an aggregate loan demand function $L^d(\mathbf{r}, z)$ in (18) and the corresponding share of commercial bank lending $s^c(\mathbf{r}, z)$ in (15).
2. Given $r^D = r^s$, P_θ, and P^s, $\{a_h', d_h', S_\theta', \varsigma^{s'}\}$ are consistent with household optimization (19)-(20) inducing a deposit matching process.
3. Given the loan demand function, $\{\ell_n', \mathcal{D}_n, e_n\}$ are consistent with non-bank optimization (39).
4. Given the loan demand function and deposit matching process, $\{V_\theta, \ell_\theta', d_\theta', A_\theta', \mathcal{D}_\theta, e_\theta, x_\theta'\}$ are consistent with bank optimization (25)-(31) inducing an aggregate loan supply function $\ell_b' + L_f^s(z, \mu, \ell_b')$ where L_f^s is defined in (24).
5. The initial equity injection rule is consistent with entrant bank optimization (35) and the free-entry condition is satisfied (36).
6. The law of motion for the industry state induces a sequence of cross-sectional distributions that are consistent with entry, exit, and asset decision rules in (37).

[30] See appendix A-2 for a detailed derivation of taxes to cover deposit insurance and servicing securities.

7. The bank interest rate $r^c(\mu, z)$ is such that the bank loan market clears. That is,

$$L^{c,d}(\mathbf{r}, z) = \ell'_b + L^s_f(z, \mu, \ell'_b, r^n).$$

8. The non-bank interest rate $r^n(\mu, z)$ is such that the non-bank loan market clears. That is,

$$L^{n,d}(\mathbf{r}, z) = \ell'_n.$$

9. Bank and non-bank stock prices satisfy versions of (33).

10. Taxes/transfers $\tau'(z, \mu, r^n, z')$ satisfy (38) for all states (z, μ, r^n, z').

5 Parameterization

In order to avoid having a low probability event like the financial crisis play a disproportionate role in our analysis, our estimation strategy is to choose model parameters to match data moments over the period 1984 to 2007. To account for the possibility of a financial crisis, we add a crisis state to our shock process that is consistent with data from 1984 to 2016. Thus, banks in our model make decisions recognizing a crisis may occur.[31]

Further, to understand the role that market structure plays in our analysis, here we also consider a separate calibration exercise where we set the entry cost for big banks (Υ_b) to infinity and recalibrate a perfectly competitive model where market power plays no role. We then assess the two models' predictions for untargeted moments in subsection 5.3.

5.1 Data Targets and Functional Forms

We use several data sources to calibrate the model. A model period is one year. Our main source for bank level variables (and aggregates derived from it) is the Consolidated Reports of Condition and Income for Commercial Banks (regularly called "call reports").[32] We aggregate commercial bank level information to the Bank Holding Company Level.[33] We also use the TFP series for the U.S. Business Sector, produced by John Fernald [33] and data provided by the Federal Deposit Insurance Corporation to identify bank failures and losses in the event of failure. Our calibration strategy involves a set of parameters taken directly from the data and a second set using Simulated Method of Moments (SMM).

We begin with the parameterization of the four stochastic processes: $F(z', z)$, $G^{\theta}_{z'}(\delta', \delta)$, $p(R, z')$, and $\Omega(\omega')$. To calibrate the stochastic process for aggregate technology shocks $F(z', z)$, we detrend the sequence of TFP using the H-P filter and estimate the following autoregressive equation $log(z') = \rho_z \log(z) + u'_z$ with $u'_z \sim N(0, \sigma^u_z)$. Once parameters ρ_z and σ^u_z are estimated, we discretize the process using the Tauchen [63] method. We set the number of grid points to four, that is $z \in Z = \{z_C, z_B, z_M, z_G\}$ (for "crisis", "bad", "median", "good"). We choose the grid in order to capture the infrequent crisis states we observe in the data. In particular, we choose z_M to match

[31] To accomplish this strategy, we simulate panels of 24 model periods and drop those panels which include a crisis when estimating the parameters for our baseline case.

[32] Source: FDIC, Call and Thrift Financial Reports, Balance Sheet and Income Statement (http://www2.fdic.gov/hsob/SelectRpt.asp?EntryTyp=10). See Appendix A-1 for a full description of the variables used in the paper.

[33] All averages from the Call Report data correspond to asset-weighted averages. That is, the average of variable x in year t equals $\sum_{i=1}^{N_t} w^i_t x^i_t$, where w^i_t is the ratio of assets of bank i in year t to total assets in year t and x_{it} is the observation of variable x for bank i in year t.

the mean of the process normalized to 1 (i.e. $z_M = 1$), select z_B ("bad times") and z_G ("good times") so they are at 1.25 standard deviations from z_M and set the value of z_C ("crisis" state) to be at 2.5 standard deviations below the mean to be consistent with the observed TFP levels during the 1982 recession and the last financial crisis (years 2008/2009).[34]

We identify "big" banks with the top 10 banks (when sorted by assets) and the fringe banks with the rest.[35] When calibrating the model with $\Upsilon_b \to \infty$ we use moments from the entire distribution of banks. As in Pakes and McGuire [53] we restrict the number of big 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, there will be at most one representative big bank and a continuum of potential fringe entrant banks. When calibrating the perfectly competitive counterfactual model (i.e., the model that sets $\Upsilon_b = \infty$) we use moments from the entire distribution of banks. For this reason we provide moments for the Top 10, the competitive fringe (i.e., all banks outside the top 10), and all banks in the following tables.

Provided that $r^A > r^s$ (which occurs in our sample period), the solution to our problem implies that the deposit capacity constraint binds in all states. Hence, we can approximate the constraint using information on deposits from our panel of commercial banks in the U.S. In particular, after controlling for firm and year fixed effects as well as a time trend, we estimate the following autoregressive process for log-deposits for bank i of type θ:

$$\log(\delta_\theta^{i\prime}) = (1-\rho_\theta^d)\upsilon_\theta^0 + \rho_\theta^d \log(\delta_\theta^i) + u_\theta^{i\prime}, \tag{40}$$

where δ_θ^i is the sum of deposits and other borrowings in the current period for bank i, and $u_\theta^{i\prime}$ is iid and distributed $N(0, \sigma_{d,\theta}^u)$.[36,37] Since this is a dynamic model we use the method proposed by Arellano and Bond [6]. Since we work with a normalization in the model (i.e., $z_M = 1$), the mean υ_θ^0 in (40) is not directly relevant. Instead, we include the mean of the finite state Markov process that depends on the aggregate state, parameterized as $\mu_\theta^d(z) = \overline{\mu}_\theta^d z^2$, as one of the parameters to be estimated via SMM.

Table 2: Deposit Process Parameters

	ρ_θ^d	$\sigma_{d,\theta}^u$	$\sigma_{d,\theta}$
Top 10 Banks	0.704	0.131	0.184
Fringe Banks	0.881	0.173	0.365
All Banks	0.882	0.173	0.366

Note: Top 10 refers to Top 10 banks when sorted by assets. Fringe Banks refers to all banks outside the top 10. Source: Consolidated Reports of Condition and Income

To keep the state space workable, we apply the method proposed by Tauchen [63] to obtain a finite state Markov representation $G_{z'}^\theta(\delta', \delta)$ to the autoregressive process in (40) with state

[34]Appendix A-1.1 presents the support and transition matrix for aggregate shocks.

[35]As of December 2016, most of the banks in the group of Top 10 were classified as Global Systemically Important Banks (G-SIBs). The exceptions are U.S. Bankcorp, PNC Financial Services, and Toronto Dominion Group US. While not classified as GSIBs, these three banks are still above the $250 billion asset threshold for the size dependent policies that we study.

[36]One could enrich this specification to include a jump process which would be a reduced form way to model random mergers that discretely increase the size of the bank.

[37]When estimating the process for $\delta_{\theta,t}^i$, to maintain consistency between our simple model balance sheet and the data, we add other borrowings to deposits. However, when estimating the process using only deposits the estimated parameters are close to that in Table 2 (see Table 7.b in Appendix A-1.4): specifically, $\sigma_{d,b} = 0.193$ and $\sigma_{d,f} = 0.346$. It is important to note, however, that our model has another source of external funding besides deposits; seasoned equity issuance which depends on the aggregate state of the economy at a higher cost than deposit funding.

dependent mean $\mu^d_\theta(z)$. We assume a 3 state Markov process for big banks and a 5 state Markov process for fringe banks. To apply Tauchen's method, we use the estimated values of equation (40) that we present in Table 2. From these estimates, we can construct the stationary variance of deposits by bank size (i.e. $\sigma_{d,\theta} = \sigma^u_{d,\theta}(1-(\rho^d_\theta)^2)^{-1/2}$) that we present in the last column of Table 2. Thus, consistent with big banks having a geographically diversified pool of funding (see Liang and Rhoades [45] and Aguirregabiria et al. [3]), big banks have less volatile funding inflows, which is one important factor explaining why they hold a smaller capital buffer in our model.

We parameterize the stochastic process for the borrower's project as follows. For each borrower, let $s = \psi_0 z' - \psi_1 R^{\psi_2} + e$, where e is iid (across entrepreneurs and time) and drawn from $N(0, \sigma^e_s)$. We define success to be the event that $s > 0$, so in states with higher z or higher e success is more likely. Then $p(R, z') = 1 - \Pr(s \leq 0 | R, z') = \Phi\left(\psi_0 z' - \psi_1 R^{\psi_2}\right)$, where $\Phi(x)$ is a normal cumulative distribution function with zero mean and standard deviation σ^e_s. The stochastic process for the borrower outside option, $\Omega(\omega)$, is simply taken to be the uniform distribution $[0, \overline{\omega}]$. We normalize the aggregate endowment of households $\mathcal{H}y$ to 1 and calibrate $\mathcal{H}$ (the mass of households) to $\mathcal{H} = 4.882$, which is consistent with a fraction of entrepreneurs in the total population ($\frac{1}{1+\mathcal{H}}$) equal to 17% that corresponds to the value reported in Quadrini [56].[38]

We estimate the marginal cost of producing a loan and the fixed cost following the empirical literature on banking (see, for example, Berger et al. [12]).[39] The marginal cost is derived from an estimate of marginal net expenses that is defined to be marginal non-interest expenses net of marginal non-interest income.[40] The estimated (asset-weighted) average of marginal non-interest expenses, marginal non-interest income, and net marginal non-interest expenses are reported in Table 3. The fixed cost κ_θ is estimated as the total cost on expenses of premises and fixed assets. The estimated (asset-weighted) average fixed cost (scaled by loans) is reported in the fourth column of Table 3. The final column of Table 3 presents our estimate of average costs for big and small banks. We find a statistically significant lower average cost for big banks than small banks. Importantly, this is consistent with increasing returns as in the delegated monitoring model of Diamond [27] and with empirical evidence on increasing returns as in, among others, Berger and Mester [15].

Table 3: Banks' Cost Structure

Moment (%)	Mg Non-Int Inc.	Mg Non-Int Exp.	Mg Net Exp. $c_\theta(\ell'_\theta, z)$	Fixed Cost $\kappa_\theta/\ell_\theta$	Avg. Cost
Top 10 Banks	4.48†	4.89†	0.41†	0.84	1.25†
Fringe Banks	2.09	3.61	1.52	0.73	2.26
All Banks	2.95	4.08	1.13	0.77	1.91

Note: Top 10 Banks refers to Top 10 banks when sorted by assets. Fringe Banks refers to all banks outside the top 10. † denotes statistically significant difference between the Top 10 and the rest. Mg Non-Int Inc. refers to marginal non-interest income, Mg Non-Int exp. to marginal non-interest expenses. Mg Net Exp. corresponds to net marginal expense and it is calculated as marginal non-interest expense minus marginal non-interest income. Fixed cost κ_θ is scaled by loans. Data correspond to commercial banks in the U.S. Source: Consolidated Reports of Condition and Income

To calibrate $r^D = r^s$ we target the average cost of funds computed as the ratio of interest expense on deposits and federal funds purchased over total deposits plus federal funds purchased,

[38] While the fraction of entrepreneurs is 17% of the total population, at the calibrated parameter values, entrepreneur output represents 39.7% of total output which is close to the estimated capital share (on average 36%) in the data. If we interpret entrepreneurs as capital owners, the model is largely in line with the data.

[39] The marginal cost estimated is also used to compute our measure of markups and the Lerner index.

[40] See Appendix A-1.2 for a detailed presentation of the estimated equation and the derivation of marginal net expenses. Marginal non-interest expenses and Marginal non-interest income are derived from trans-log functions as in the empirical literature.

which equals 0.68%. Similarly, we calibrate r^A to the ratio of interest income from safe securities over the total safe securities (net of marginal not interest expenses on securities), which results in a value equal to 1.85%.[41]

We parameterize the cost function in the model as

$$c_\theta(\ell'_\theta, z) = c_{\theta,0} z \ell'_\theta + c_{\theta,1} z (\ell'_\theta)^2. \tag{41}$$

We incorporate the estimated average marginal net expenses in our SMM procedure to help pin down the parameters of this function. We also use the estimates of the fourth column of Table 3 to pin down the fixed operating costs in the model.

The parameter λ is set to 0.397 to be consistent with the average charge-off rate that equals 0.728% in the data at the observed default frequency of 1.832%. The liquidation value of the loan portfolio $(1 - \xi_\theta)$ is estimated using data from the FDIC. We set $\xi_b = \xi_f = 0.1965$. The equity issuance cost function is parameterized as follows: $\zeta_\theta(e, z) = (\zeta_{\theta,0} e + \zeta_{\theta,1} e^2)(\frac{z_G}{z})^{\zeta_z}$. The quadratic form for equity issuance is relatively standard in the corporate finance literature (e.g. Henessy and Whited [38]) and the term $(\frac{z_G}{z})$ captures, in a parsimonious way, changes in the cost of external finance along the business cycle (consistent with the evidence presented in McLean and Zhao [47] and the "financial accelerator" literature pioneered in Bernanke and Gertler [13]). We estimate the parameters of this function by allowing equity issuance costs to depend explicitly on bank type. As our estimates show below, equity issuance costs are significantly lower for big than for small banks (much in line with the evidence described in Hughes, Mester, and Moon [39]).

In our benchmark parameterization, we use values associated with regulation in place before Basel III and the recent financial crisis. Thus we set the minimum level of the bank equity risk-weighted capital ratio for both types of banks to 4%. That is, $\varphi_{b,z} = \varphi_{f,z} = 0.04$ for all z and $w^A_{\theta,z} = 0$ for all θ and z.

Since we do not observe failure or entry by big banks during the calibration period (1984-2007), identification of Υ_b is challenging. We set the value of Υ_b to be the maximum value such that if a big bank failure were to occur a big bank would replace the failed bank immediately. The entry cost is kept constant for all our counterfactuals.

This leaves us with 23 parameters to estimate via simulated method of moments (SMM):[42]

$$\{\psi_0, \psi_1, \psi_2, \sigma^e_s, \overline{\omega}, \alpha, \ell, \gamma, \overline{\mu}^d_b, \overline{\mu}^d_f, c_{b,0}, c_{b,1}, c_{f,0}, c_{f,1}, \kappa_b, \kappa_f, \zeta_{b,0}, \zeta_{b,1}, \zeta_{f,0}, \zeta_{f,1}, \zeta_z, \Upsilon_f, c_n\}.$$

To pin down these parameters, except for a few data moments, we use the data for commercial banks described in Section 2 and in our companion paper [21]. One of the extra moments is the average real equity return (12.94%) as reported by Diebold and Yilmaz [28], added to help identify parameters associated with the borrower's return $p(R^j, z')z'R^j$.[43,44] The other moment is the elasticity of loan demand (-1.10) as estimated by Bassett, et al. [8].

As a way to determine the size of the commercial banking sector in the economy we use the deposits in commercial banks to GDP ratio and the loans and leases in commercial banks to GDP ratio, which average 38.13% and 33.71%, respectively, during our calibration period. To pin down the size of the commercial banking sector relative to the overall size of the financial sector (i.e., the

[41] The nominal interest rate is converted to a real interest rate by using CPI inflation (we use the realized inflation rate as a measure of expected inflation).

[42] Definitions to connect model moments to the data appear in Table A.1 in the Appendix.

[43] Our target lies within the estimates presented in Appendix 9 "Return on Assets" (Figure 9.1) in DeLoecker, et al. [25]. The average return on assets weighted by revenues (which would be consistent with our model) varies between approximately 8% in 1984 and 34% in 2008.

[44] As we discussed in the equilibrium section we require $\beta_E(1 + r^j)(1 + r^s) < 1$ so that there are no incentives for the entrepreneur to retain earnings. At the estimated parameters this implies that $\beta_E < 0.947$.

sum of the banking sector plus the non-bank sector), we rely on estimates from the literature. As determining the size of the non-bank sector is difficult and depends on the definition used, estimates of the ratio of bank loans to total credit (i.e., a measure that includes the non-bank sector) vary significantly from a lower bound of 35% (see Pozsar et al. [55]) to an upper bound of 65% (see Gallin [35]). We use the middle range of these estimates and target a bank loan to total credit ratio equal to 50% for our calibration period (1984-2007).

The set of targets from commercial bank data includes the loan interest margin (4.68%) that is defined as the difference between the estimated loan interest rate and the cost of deposits, the standard deviation of the default frequency (0.98%), the standard deviation of bank credit to output ratio (0.43%), the loan default frequency (1.83%), marginal net expenses and fixed costs by bank size (as reported in Table 3), equity issuance over assets by bank size (0.02% and 0.11% for big and fringe banks, respectively), the bank failure and entry rate (1.07% and 1.65%, respectively), and the dividend to asset ratio by bank size (0.61% and 0.63% for big and fringe banks, respectively).

We also target some important components of the balance sheet of banks. While the balance sheet in our model is fairly rich and considers the most important pieces of its empirical counterpart such as loans, cash and securities, deposits and equity, in order to connect the model's balance sheet with the one in the data that contains several additional items, we proceed as follows in Table 4. We identify loans in the model with the reported value for risk-weighted assets. Since there are two assets in the model, loans (risky assets) and cash and securities (safe assets), once we determine the ratio of risk-weighted assets to total assets (loans to assets in the model) by bank size, the ratio of cash and securities can be obtained as a residual.[45] One of the main counterfactuals in the paper evaluates changes in capital regulation, so we target the risk-weighted Tier 1 capital ratio by bank size (equity to loans in the model). The risk-weighted capital ratio together with risk-weighted assets to total asset ratio imply the equity to asset ratio and the ratio of deposits to total assets in the model. While our simplified balance sheet lumps insured deposits together with other borrowings in Deposits & Borrowings to Total Assets in Table 4, insured deposits and borrowings with little to no risk make up between 77% and 92% of that total.[46]

Table 4: Banks' Balance Sheet by Size

Assets	Top 10	Fringe	All Banks
Cash & Securities	21.55	26.16	24.15
Loans (risk-weighted assets)	78.45	73.83	75.85
Liabilities			
Deposits & Borrowings	93.36	91.06	92.10
Equity	6.64	8.94	7.90
Capital Ratio (risk-weighted)	8.48	12.10	10.42

Note: All variables except capital ratio (risk-weighted) are reported as the ratio to total assets. Data correspond to commercial banks in the U.S. Source: Consolidated Reports of Condition and Income

The computation of the model is a nontrivial task. We solve the model using an extension of the algorithm proposed by Ifrach and Weintraub [40] adapted to our environment. This entails approximating the distribution of banks by a finite number of moments.[47] In short, the solution

[45] Tables A.2.a, A.2.b, and A.3 in Appendix A-1 describe the variables used for the balance sheet classification fully.

[46] Specifically, out of the average of total deposits and borrowings by U.S. commercial banks during our calibration period, insured deposits represent 77.4%, federal funds purchased 7.9%, other borrowed money 7.1% (which includes borrowings from Federal Home Loan banks collateralized by mortgages and mortgage related assets), trading liabilities 3.4%, and other liabilities 4.1%.

[47] Appendix A-3 describes in detail the algorithm we use to compute an approximate Markov perfect industry

entails keeping track of all the states of the dominant player (i.e., the big bank), aggregate productivity z, and approximating the evolution of the distribution of fringe banks using a finite set of moments (specifically, the mass of fringe banks denoted $\mathcal{M}$ and average liabilities of fringe banks denoted $\mathcal{K}$). It is clear from (39) that r^n is a function only of z. Consistent with the methods in Ifrach and Weintraub [40], the continuation value for fringe banks is evaluated using the equilibrium loan decision of the big bank (i.e. $V_f(k_f, \delta_f; z, \mu, \ell_b'(k_b, \delta_b; z, \mu, r^n(z)), r^n(z)))$.[48] Thus, the approximation implies that the state space of the big bank is $\{k_b, \delta_b, z, \mathcal{K}, \mathcal{M}\}$ and the state space of any fringe bank is $\{k_f, \delta_f, k_b, \delta_b, z, \mathcal{K}, \mathcal{M}\}$.

5.2 Estimation Results

Tables 5.a and 5.b show the estimated parameters, corresponding to those chosen outside and inside the model, respectively, and Table 6 provides the moments generated by our baseline model with imperfect competition and a perfectly competitive model relative to the data. As a fraction of assets, fringe banks issue more equity than the big bank (as in the data), despite significantly higher equity issuance costs for fringe banks than for the big bank (consistent with Hughes and Mester [39]). As in Berger and Mester [15] and the theory behind Diamond [27], our imperfectly competitive model delivers the observed increasing returns as average costs for big banks are smaller than those of fringe banks (which can be seen from Table 6 as the sum of net non-interest expenses plus fixed costs over loans equals 1.69 and 3.18 for big and fringe banks, respectively). Since all these estimated costs play a role in determining lending and external finance decisions, they influence the bank's capital buffer, which in turn influences stability. We find that the model with imperfect competition slightly under-predicts bank failure rates while the model with perfect competition over-predicts it.

Table 5.a: Model Parameters (chosen outside the model)

Parameter		Imperfect Competition	Perfect Competition	Target
Autocorrel. z	ρ_z	0.299	0.299	TFP US (Fernald/SanFran Fed)
Std. Dev. Error	σ_z^u	0.010	0.010	TFP US (Fernald/SanFran Fed)
Crisis state	z_C	0.976	0.976	TFP US (Fernald/SanFran Fed)
Mass Households	$\mathcal{H}$	4.882	4.882	Fraction Entrepreneurs
Household Endowment	y	0.205	0.205	Normalization ($\mathcal{H}y = 1$)
Deposit interest rate (%)	$r^s = r^d = \frac{1}{\beta_H} - 1$	0.688	0.688	Int. expense
Securities Return (%)	r^a	1.851	1.851	Return Safe Securities
Charge-off rate	λ	0.397	0.397	Charge off rate
Autocorrel. Deposits (b, f)	ρ_θ^d	$(0.70, 0.881)$	$(\cdot, 0.882)$	Deposit Process (top 10,fringe)
Std. dev. Error (b, f)	$\sigma_{d,\theta}^u$	$(0.13, 0.1726)$	$(\cdot, 0.1727)$	Deposit Process (top 10,fringe)
% loss in exit	ξ	0.1965	0.1965	Recovery Failures (FDIC)
Cap requirement weights (b, f)	$((w_\theta^\ell, w_\theta^A))$	$((1, 0), (1, 0))$	$(\cdot, (1, 0))$	Basel II Capital Regulation
Capital requirement f bank	(φ_θ)	$(0.04, 0.04)$	$(\cdot, 0.04)$	Basel II Capital Regulation

equilibrium. The approximation method is also similar to that in Krusell and Smith [44].

[48] The recursive formulation captures all possible deviations, not only contemporaneous but also future deviations. As the big bank evaluates different values of ℓ_b' it takes into account that this results in contemporaneous changes in the interest rate on loans that feeds directly into the loan and asset decisions of fringe banks, but also into the evolution of the industry (via profitability and entry/exit) that will possibly induce further changes in its loan supply.

Table 5.b: Model Parameters (chosen inside the model)

Parameter		Imperfect Competition	Perfect Competition	Target
Weight agg. shock	ψ_0	4.743	4.743	Std. dev. Def. frequency
Success prob. param.	ψ_1	26.313	25.450	Borrower Return
Success prob. param.	ψ_2	0.922	0.922	Interest Margin
Volatility borrower's dist.	σ_s^e	0.107	0.106	Default freq.
Max. reservation value	$\overline{\omega}$	0.509	0.462	Elasticity loan demand
Variance Entrep. Pref. Shock	α	98.25	98.25	Std. dev. Bank Credit to Output
Loan size	$\overline{\ell}$	3.50	3.50	Bank credit to output ratio
Discount Factor Manager	γ	0.981	0.981	Loan mkt share fringe
Avg. deposits (b, f)	μ_θ^d	$(0.372, 0.162)$	$(\cdot, 0.267)$	Deposits to output ratio Deposit mkt share fringe
Cost fun linear (b, f)	$c_{\theta,0}$	$(0.001, 0.001)$	$(\cdot, 0.0004)$	Net non-int exp. (top 10,fringe)
Cost fun quadratic (b, f)	$c_{\theta,1}$	$(0.017, 0.063)$	$(\cdot, 0.043)$	Cap. ratio (risk-weighted) (top 10,fringe)
Fixed cost (b, f)	κ_θ	$(0.001, 0.003)$	$(\cdot, 0.003)$	Fixed cost over loans (top 10,fringe)
Linear Equity Issue Cost (b, f)	$\zeta_{\theta,0}$	$(0.025, 0.200)$	$(\cdot, 0.100)$	Div. to asset ratio (top 10,fringe)
Quadratic Equity Issue Cost (b, f)	$\zeta_{\theta,1}$	$(0.100, 4.500)$	$(\cdot, 4.500)$	Equity Issuance/Assets (top 10,fringe)
Cyclic Equity Issue Cost	ζ_z	4.00	4.00	Loans to asset ratio fringe
Entry Cost (b, f)	Υ_θ	$(0.250, 0.008)$	$(\cdot, 0.057)$	Bank failure rate (top 10,fringe)
Cost function non-banks	c_n	0.0326	0.0318	Bank credit to total credit Loans to asset ratio top 10

Note: Our estimate of α implies a variance of $\frac{1}{6} \cdot \left(\frac{\pi}{\alpha}\right)^2 = 0.00017$ of the Gumbel distribution.

Table 6: Targeted Data and Model Moments

Long-Run Averages 1984-2007	Imperfect Competition		Perfect Competition	
Moment (%)	Data	Model	Data	Model
Capital ratio (risk-weighted) (b, f)	$(8.46, 12.10)$	$(10.51, 11.28)$	$(\cdot, 10.42)$	$(\cdot, 11.81)$
Bank failure rate (b, f)	$(0.0, 1.07)$	$(0.0, 0.83)$	$(\cdot, 1.07)$	$(\cdot, 1.49)$
Default freq.	1.83	1.38	1.83	1.35
Interest Margin	4.68	3.98	4.68	3.89
Loans to asset ratio (b, f)	$(78.45, 73.84)$	$(93.87, 74.71)$	$(\cdot, 75.85)$	$(\cdot, 70.67)$
Deposit mkt share fringe	65.91	75.63	100.00	100.00
Net non-int exp. (b, f)	$(0.41, 1.52)$	$(1.35, 1.83)$	$(\cdot, 1.13)$	$(\cdot, 1.77)$
Fixed cost over loans (b, f)	$(0.84, 0.73)$	$(0.34, 1.53)$	$(\cdot, 0.77)$	$(\cdot, 1.06)$
Div/Assets ratio (b, f)	$(0.61, 0.63)$	$(2.23, 0.57)$	$(\cdot, 0.62)$	$(\cdot, 1.05)$
E. I./Assets (b, f)	$(0.03, 0.11)$	$(0.00, 0.04)$	$(\cdot, 0.08)$	$(\cdot, 0.12)$
Loan mkt share fringe	64.29	70.93	100.00	100.00
$L^{s,c}$ to output ratio	33.72	31.17	33.72	32.27
$L^{s,c}$ to total credit	50.00	53.68	50.00	52.89
Deposits to output ratio	38.14	36.01	38.14	42.26
Borrower Return	12.94	14.53	12.94	15.06
Elasticity loan demand	-1.10	-1.24	-1.10	-1.15
Std. dev. Bank Credit to Output	0.43	0.92	0.43	2.49
Std. dev. Def. frequency	0.98	1.39	0.98	1.39

Note: (b, f) refers to big and fringe, respectively. In the baseline model, we identify "big" banks with the top 10 banks (when sorted by assets) and the fringe banks with the rest. The distribution of banks in the model with $\Upsilon_b \to \infty$ (i.e., Perfect Competition) is composed only by fringe banks and targets correspond to (asset-weighted) averages from the full set of banks.

5.3 Untargeted Moments: Business Cycle Correlations

Since one objective of regulation is to promote stability, we now consider the cyclical properties of the model. Since parameters of the model were not chosen to match data on business cycle correlations, the results in Table 7 can be considered a simple qualitative consistency test of the model.[49] We observe that, as in the data, the baseline model with imperfect competition generates countercyclical exit rates, default frequencies, interest margins, and markups as well as procyclical entry rates, lending, assets, deposits, and equity.[50] Importantly, the perfectly competitive model generates a counterfactual correlation of entry rates, loan supply, equity, interest margins, and markups.

Table 7: Business Cycle Correlations

Business Cycle Correlations with Commercial Bank Variables	Data	Imperfect Competition	Perfect Competition
Exit Rate	-0.12	-0.21	-0.06
Entry Rate	0.70	0.10	-0.02
Default Frequency	-0.65	-0.58	-0.67
Loans	0.54	0.46	-0.14
Total Assets	0.32	0.46	0.15
Deposits	0.29	0.36	0.43
Equity	0.18	0.91	-0.07
Interest Margins	-0.36	-0.43	0.16
Markups	-0.57	-0.07	0.44

In both models, the countercyclicality of exit rates follows from the countercyclicality of chargeoffs and the procyclicality of bank profits. As banks fail, the reduction in bank lending (and competition) together with the positive relation between bank costs and productivity will lead to an increase in markups if the reduction in bank lending and costs is larger than the deterioration in aggregate conditions (inducing a reduction loan demand and increase in expected defaults). In the model with imperfect competition, the presence of a dominant player prevents new competitors from taking over the market share left unattended by failing banks as the big bank increases its market share in bad times (preventing small bank entry). Entry happens with imperfect competition only after aggregate productivity improves enough. This results in an entry rate that is pro-cyclical in line with what we observe in the data.

On the contrary, in the model with perfect competition, new banks are created even during bad times since there is no dominant player to make up for failed banks. This results in a counterfactual slightly countercyclical entry rate which is enough to induce counterfactual procyclical interest margins and markups and countercylical bank lending and equity issuance. Thus, banks build an equity buffer in good times for the imperfect competition model while the opposite occurs in the model with perfect competition. This provides some untargeted evidence in favor of our baseline model with imperfect competition.

The countercyclicality of loan interest rates (and margins) in the model with imperfect competition provides an endogenous amplification mechanism. Specifically, we find that the 2.4% drop in aggregate productivity (from z_M to z_C) induces an endogenous decline in aggregate loans *of up to*

[49]See Appendix A-4 for a set of scatter plots with the simulated data and the corresponding correlations. In addition, Table A.11 in the Appendix provides other untargeted moments in the data and the model.

[50]In the model with imperfect competition, while equity is procylical as in the data, risk-weighted capital ratios are also procyclical unlike the data since equity is more procyclical than assets.

4% in our baseline model (in contrast to a smaller decline in the model with perfect competition). This drop arises not only from a drop in loan demand by entrepreneurs, but also from a drop in loan supply emanating from deposit inflows via $G_{z'}(\delta'_\theta, \delta_\theta)$. The end result is a drop in total output *of up to* 5.1%. During the financial crisis, real GDP dropped by 3.9% from peak to trough (2007.Q4 - 2009.Q2), which is roughly consistent with our model.

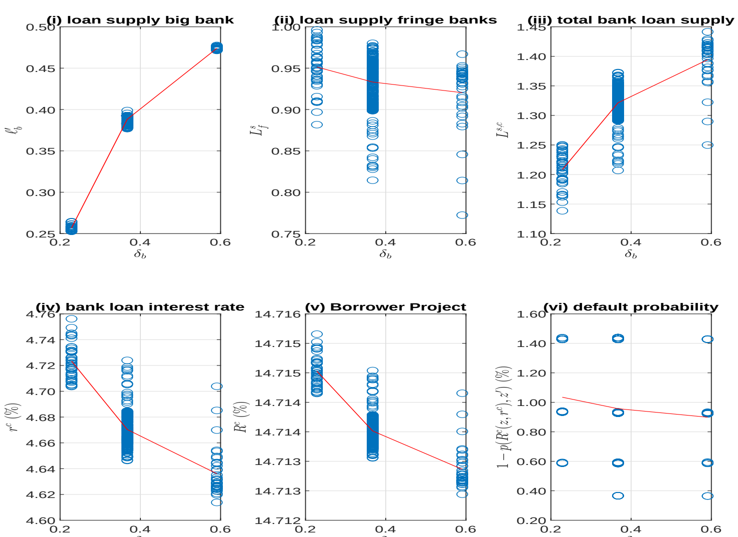
6 Model Properties

We focus on a small set of important properties of our benchmark model with imperfect competition in the loan market.

6.1 Systemic Spillovers

An imperfect competition model with non-atomistic banks implies that shocks to big banks can induce actions which spill over to the entire industry and affect financial stability. While a typical systemic spillover in banking models is via the interbank market (e.g. a default by one bank spills over to lower net worth of other banks in the network potentially inducing further defaults as in Acemoglu, Ozdaglar, Tahbaz-Salehi [2]), our paper provides a novel link via strategic interaction in the loan market. We describe this link by studying how idiosyncratic financing shocks to the big bank δ_b for a given value of aggregate productivity $z = z_M$ induce changes to big bank loan supply which spill over to the loan supply of fringe banks, the loan interest rate, borrower risk choices inducing endogenous entrepreneur default and aggregate loan losses. In this way, our framework exhibits granularity in the sense of Gabaix [34]; an idiosyncratic shock to the big bank has aggregate consequences.

Figure 4: Big Bank Shocks and Systemic Risk



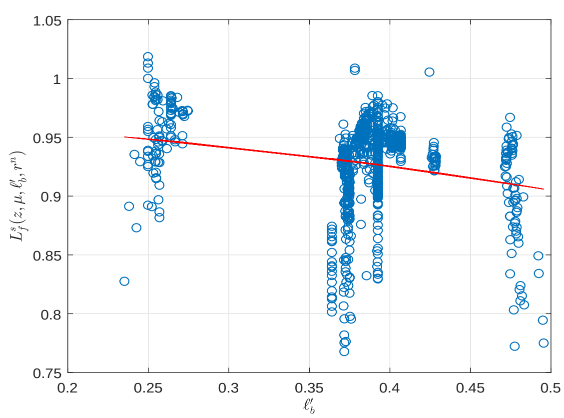
Note: This figure shows the loan supply of the big bank (panel i), the loan supply of fringe banks (panel ii), aggregate loan supply (panel iii), loan interest rates (panel iv), borrower risk choice (panel v), and default frequencies (panel v) at different values of financing shocks of the big bank δ_b conditional on aggregate productivity $z = z_M$ along the equilibrium path.

Figure 4 presents the observed values along the equilibrium path. For a given aggregate shock, a negative shock to big bank funding (i.e., a reduction in δ_b) leads to less big bank lending (panel

(i)). Panel (ii) makes clear that fringe banks respond by raising their lending, but this is insufficient to keep aggregate bank lending from falling in panel (iii). This quantitatively sizeable (13.26%) average drop in aggregate lending puts upward pressure on loan rates r^c in panel (iv); the average ensuing rise in loan interest rates is 10 basis points (which of course is mitigated by increased non-bank competition). The resulting higher interest rate induces borrowers to choose riskier projects evident in panel (v) which lead to higher default frequencies (panel (vi)) elevating the riskiness of the loan portfolio for the entire banking industry. Specifically, the rise in interest rates on loans induces an average rise in defaults of 14.4% with differences in realizations as large as 295% (from 0.36% to 1.44%). Importantly, the starting point for this spillover is an idiosyncratic drop in big bank finance which would be absent in a competitive model with atomistic banks. Thus, our model with imperfect competition can generate quantitatively large risk spillovers via strategic linkages in the loan market.

The strategic interaction between the big bank and the fringe is further illustrated in Figure 5, which effectively graphs an reaction function implicit in the loan market clearing condition (equation (30)). That is, $L^s_f(z, \mu, \ell'_b, r^n)$ can be thought of as a best response by the fringe banking sector to the loan supply choice ℓ'_b of the dominant bank. Figure 5 makes clear that big bank lending ℓ'_b and fringe bank aggregate lending L^s_f (which depends on both intensive and extensive margins) move in opposite directions. Since the variance of the aggregate loan supply can be written as $\text{Var}(L^{s,c}) = \text{Var}(\ell'_b) + \text{Var}(L^s_f) + 2\text{Cov}(\ell'_b, L^s_f)$, the negative covariance between ℓ'_b and L_f drives down the variance of the loan supply and with it all other variables in the imperfectly competitive environment. In Appendix A-4.1 we show that other key variables in the model with imperfect competition are less volatile than one with perfect competition.

Figure 5: Relationship between ℓ'_b and L^s_f with imperfect competition



Note: This figure shows the loan supply of fringe banks as a function of the loan supply of the big bank along the equilibrium path. The three clouds correspond to the three different values of δ_b and within each cloud variation depends on variation in μ induced by variation in z.

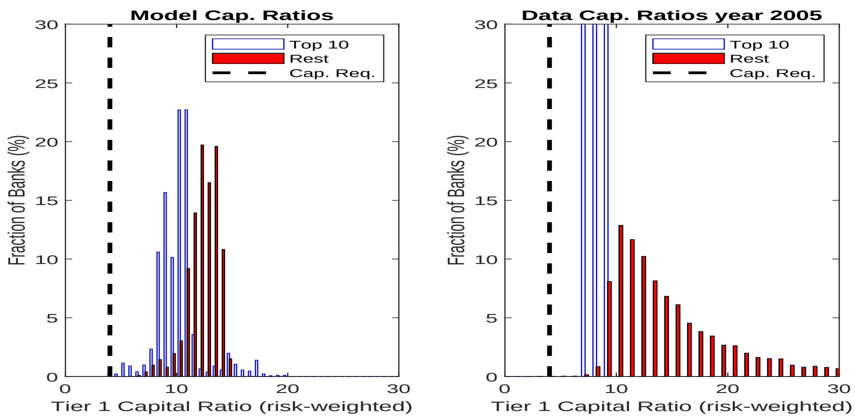
6.2 Capital Ratios

Figure 6 presents the distribution of risk-weighted bank capital ratios $(A'_\theta + \ell'_\theta + \pi'_\theta - d'_\theta)/\ell'_\theta$ by bank size in the model and its data counterpart for year 2005.[51] Higher volatility of external funding (i.e. $\sigma^d_{f,u} > \sigma^d_{b,u}$) and higher equity finance costs (i.e. $\zeta_{f,0} > \zeta_{b,0}$ and $\zeta_{f,1} > \zeta_{b,1}$) in Table

[51] To be precise, after simulating the model for T periods, we compute the average distribution of fringe banks $\bar{\mu}_f(n, \delta) = \sum_{t=1}^{T} \frac{\mu_{f,t}(n,\delta)}{T}$. We simulate the economy for 10,000 periods and drop the first 2,000 periods. Similarly, we compute the frequency of capital ratios that the big bank transits during the simulation.

5.b estimated from the data both contribute to induce fringe banks to hold higher risk-weighted capital ratios (as a buffer) than big banks on average. While we targeted this average (and hence the model distributions should be centered on their data counterparts), the other moments of the distributions were not targeted. Our model fails to generate the significant heterogeneity and long right tail of the distribution of risk-weighted capital of fringe banks. Some of this can be explained by the sparse finite state Markov approximation of the deposit shock process.

Figure 6: Capital Ratios by Bank Size



6.3 Competition-Stability Tradeoff

Many authors have estimated the relation between bank concentration (their proxy for competition) and banking system fragility using a reduced form approach. In this section, we follow this approach using simulated data from our model to show that the model is qualitatively consistent with findings from many of these empirical studies. As in Beck et al. [11], 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. [12], 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 and the exit rate are higher than two standard deviations from their mean; (ii) deposit insurance outlays as a fraction of GDP are higher than 2%; or (iii) large dominant banks are liquidated. The concentration index corresponds to the loan market share of the big bank. We use as extra regressors the growth rate of GDP and lagged growth rate of loan supply.[52] Table 8 displays the estimated coefficients and their standard errors.

Consistent with the empirical evidence in Beck, et al. [11], we find that banking system concentration is highly significant and negatively related to the probability of a banking crisis. Higher market power induces periods of higher profits that prevent bank exit (see also Corbae and Levine [23]). This is in line with the findings of Allen and Gale [5]. Consistent with the evidence in Berger et al. [12] we find that the relationship between concentration and loan portfolio risk is positive and significant, though not quantitatively large. This is in line with the view of Boyd and De Nicolo [17], who showed that higher concentration is associated with riskier loan portfolios.

[52]Beck et al. [11] also include other controls like “economic freedom” which are outside of our model.

Table 8: Competition and Stability

Model Dependent Variable	Logit Crisis_{t+1}	Linear Default Freq._{t+1}
Concentration_t	-27.224	0.057
	(2.018)***	(0.000)***
GDP growth in t	178.80	0.051
	(37.63)***	(0.029)*
Loan Supply Growth_t	1.992	-0.0163
	(5.106)	(0.008)**
R^2	0.69	0.47

Note: Standard Errors in parentheses. R^2 refers to Pseudo R^2 in the logit model. *** significant at 1%, ** at 5% and * at 10%.

6.4 The Bank Lending Channel

Since our focus is on the interaction of policy and market structure for commercial bank lending, it is important that the model be consistent with the bank lending channel. Kashyap and Stein ([42] and [43]) applied a corporate finance approach and argued that if the bank lending channel of monetary policy is correct, one should expect the loan portfolios of large and small banks to respond differently to a contraction in monetary policy. Kashyap and Stein asked of the data whether the impact of monetary policy on lending behavior is stronger for smaller banks who are more likely to have difficulty substituting into non-deposit sources of external finance. They found strong empirical evidence in favor of this bank lending channel. Their result is driven largely by the smaller banks (those in the bottom 95% of the size distribution). Here we ask the same question of our model.

We implement this policy experiment by analyzing how a permanent rise in the cost of external debt finance (in particular, an increase of 25 basis points in r^s from 0.688% to 0.938%) affects the balance sheet and lending behavior of banks of different sizes. We simulate the model and construct a pseudo-panel of banks under each value of r^s. We then follow a variant of Kashyap and Stein ([42]) and estimate the following panel regression by bank size:

$$\Delta\ell_{it} = a_0 + a_1\Delta r^s + a_2 X_{it} + u_{it}, \tag{42}$$

where $\Delta\ell_{it}$ denotes the growth rate of loans, Δr^s is the measure of monetary impulse and X_{it} are other bank and aggregate controls. To estimate this regression, we simulate many panels of banks pre- and post-policy change and use the period around the change to estimate the lending response of banks of different sizes.[53] The evaluation of the model response is done around the change in r^s.[54] Table 9 presents the estimated coefficients for banks of different sizes.[55]

The results (i.e., a negative and more sizable coefficient on the monetary impulse for small banks than big banks) in Table 9 are broadly consistent with Kashyap and Stein's findings. A contraction in monetary policy does indeed lead to a decline in lending in all size categories of small banks and importantly that the effect is larger the smaller the bank. This result stems from our estimation result that small banks find it harder to raise financing with instruments other than deposits (i.e. $\zeta_{f,0} > \zeta_{b,0}$ and $\zeta_{f,1} > \zeta_{b,1}$ in Table 5.b) consistent with a pecking order theory.[56]

[53] These panels differ in the realization of the stochastic processes of the model which in place lead to different dynamics of the endogenous variables.

[54] This experiment is analogous to a diff-in-diff regression where the control group is banks before the increase in r^s and the treatment group is banks after the change in r^s.

[55] See Table A.8 for other responses to the unanticipated change in monetary policy.

[56] In Appendix A-4.3 we further discuss the market structure and aggregate implications of an unanticipated

Table 9: Kashyap and Stein ('95) Regressions (Model Pseudo-Panel)

	Dependent Variable $\Delta \ell_{it}$
Specification	Coeff. on Monetary Impulse (Δr^s)
Small 98%	-0.180
s.e.	0.036***
Small 92%	-0.265
s.e.	0.036***
Small 68%	-0.356
s.e.	0.048***

Note: All specifications include one lag of the dependent variable, and growth rate of GDP. *** significant at 1% level, ** significant at 5% level, * significant at 10% level.

7 Policy Counterfactuals

After the financial crisis of 2008, efforts to increase financial stability via capital and liquidity requirements have been coordinated by the Basel Committee on Banking Supervision at the Bank for International Settlements (BIS) in what is referred to as Basel III rules.[57] In the U.S., Basel III has largely been implemented through the Dodd–Frank Act. According to the new rules, minimum tier 1 to risk-weighted assets equals 6 percent. In addition, all banks need to hold a capital conservation buffer of 2.5 percent of risk-weighted assets for a total of 8.5 percent tier 1 risk-weighted capital. That implies that $\varphi_{\theta,z} = 0.085$ for $\theta = \{b, f\}$ in the main experiment. Under the Dodd–Frank Act in the U.S., several policies are size dependent inducing differences in required capital ratios and liquidity ratios for the largest financial institutions and other banks. In particular, there is a capital surcharge for U.S. Global Systemically Important Banks (G-SIBs). Surcharges of U.S. G-SIBs currently range between 1.0 percent and 3.5 percent of risk-weighted assets. We identify G-SIBs in our model with the top 10 banks and set the surcharge to 2.5% in our size-dependent experiment, which increases $\varphi_{b,z} = 0.11$. There is also an option to incorporate countercyclical capital buffers (an increase in the required capital in good times) of up to 2.5 percent of risk-weighted assets also for large and internationally active banking organizations (those with more than $250 billion in assets, which corresponds to 13 institutions in the U.S.). While the countercyclical capital buffer has been kept at 0% to date in the U.S., we run a counterfactual where we allow $\varphi_{b,z} = \{0.1100, 0.1183, 0.1266, 0.1350\}$ (i.e., the minimum plus the GSIB surcharge plus the corresponding countercyclical capital buffer). Finally, liquidity regulation in the U.S. (Liquidity Coverage Ratio rule) requires large banks (also those with assets more than $250 billion) to maintain a minimum of liquid assets to withstand cash outflows over a 30-day horizon. In our model calibrated to a yearly frequency, that translates to $\varrho_b = 0.08$. The policy experiments analyze the response of the model in the short run and the long run. Appendix A-3 presents a description

change to a more contractionary monetary policy. We find that the model exhibits incomplete pass-through, which is consistent with models of imperfect competition such as Dreschler, Savov, and Schnabl [29], who find that interest spread betas (in their case the spread between the fed funds rate and the deposit rate with respect to changes in the fed funds rate) are less than one, and with Wang, Whited, Wu, and Xiao [65], who find that lending spreads decline as the federal funds rate rises. We find that for a 25 basis point increase in r^s, the loan interest rate r^L increases by 16 basis points in the short run, resulting in a 2.21% increase in the interest margin.

[57] It was understood that central banks and regulatory authorities would write the specific rules and timetables for implementation in their countries. The Bank for International Settlements discusses the evolution of global banking regulations at http://www.bis.org/about/chronology.htm.

of how policy counterfactuals are computed.[58]

7.1 Higher Capital Requirements

Since our focus is on the interaction between regulatory policy aimed towards stability and market structure, here we ask the question, how much does an increase from 4% to 8.5% in capital requirements affect bank exit, market shares of big and small banks, and other outcomes? Given our focus on market structure we consider the policy change in both our baseline model of imperfect competition and a model of perfect competition. Table 10 presents short and long run results for both cases.[59]

We start by describing the long-run effects of the policy change in the model with imperfect competition. As columns (2) and (5) in Table 10 show, we find that an increase in capital requirements from 4% to 8.5% leads to a 20% decline in long run exit (an intended consequence) and entry rates of small banks and decline in small bank loan market share of 6% leading to a more concentrated industry (an unintended consequence). The sizable change in big bank market share is explained by an extensive margin drop of 15% in the number (measure) of fringe banks as well as the larger increase in the intensive margin by big banks than that of surviving fringe banks (4% versus 0.7%). The net effect of higher big bank lending and lower small bank lending is a decrease in total bank lending of 9%. This leads to a 6 basis points increase in interest rates and a 1.4% increase in markups on bank loans (which benefits not only big banks with market power but also the fringe banks). The meager increase in bank interest rates is due to the shift towards non-bank lending so that total credit declines by only 0.3% and output decreases by only 0.2%. Due to selection effects, the decline in exit rates together with the fact that exiting banks are better capitalized and have a smaller share of assets invested in loans results in a nearly 70% decline in the taxes-to-output ratio (which is relevant for welfare effects we present later).[60]

The most significant difference between the short run impact and the long run effect of the policy (i.e. columns 4 and 5 in Table 10) is the increase in exit rates (37%) and the collapse of bank entry (a decline of 96%). The decline in bank entry is consistent with the data as there has been no denovo entry (i.e., newly created banks) between 2012 and 2016.[61] The resulting increase in loan market concentration (a 5% reduction in fringe market share in the short-run) leads to a significant increase in markups (a competitive effect together with a composition effect) in the short run. Specifically, the big bank increases its market share in response to the market share left by failing banks (a selection effect) and the reduction in lending by those fringe banks that remain

[58] We assume that the unanticipated policy change is announced and put into effect immediately. We approximate the response of the economy during the transition between the two long-run equilibria by computing decision rules under the new policy parameters assuming that beliefs about future moments switch from the initial approximating functions to the new long-run approximating functions at a random point in time chosen to minimize one-step ahead forecast errors. The switching probability is meant to approximate a learning process where beliefs can adjust slowly along the equilibrium path.

[59] See Table A.14 in the Appendix for further moments.

[60] From 2008 through 2013 almost 500 banks failed at a cost of approximately $73 billion to the Deposit Insurance Fund (DIF), which represents 0.51% of 2008 US GDP. Applying the 3.1% crisis probability that we observed in the data implies that the cost to the Deposit Insurance Fund equals 0.015% on average. The model counterpart for the baseline calibration is 0.011% slightly below but in line with the data. See https://www.fdic.gov/bank/historical/crisis/overview.pdf for an overview of the facts according to the FDIC.

[61] In addition, as evident in Figure 1, while the number of banks has been trending down during the last 30 years, there is a change in the rate of decline after the implementation of the Dodd-Frank Act. Specifically, the average yearly growth rate in the number of banks equals -2.67% for 1984-2007 and equals -3.43% for 2011-2016. The model's short-run response in the mass of fringe banks is consistent with the data as it shows an average decline of 0.77% per year (resulting in a 5 year model growth rate of -3.9%, surprisingly near the data growth rate of -3.4%).

active. This distributional change contributes to a reduction in average costs as well as an increase in price, leading to the increased markup.

Table 10: Benchmark Model vs Perfectly Competitive Model

	Baseline		High Capital Requirements			
(φ_f, φ_b)	$(0.04, 0.04)$		$(0.085, 0.085)$		$(\cdot, 0.085)$	
	Imperfect Comp.	Perfect Comp.	Imperfect Comp.		Perfect Comp.	
Moment	(%)	(%)	Short Run Δ (%)	Long Run Δ (%)	Short Run Δ (%)	Long Run Δ (%)
Capital Ratio (b, f)	$(10.24, 10.89)$	$(\cdot, 11.56)$	$(37.56, 35.80)$	$(44.24, 53.65)$	$(\cdot, 33.81)$	$(\cdot, 36.66)$
Exit Rate	0.87	2.24	37.09	-20.20	-76.08	-0.34
Entry Rate	0.90	2.27	-96.32	-20.32	-100.00	1.19
Prob. of Crisis	0.14	1.10	-	-20.00	-	12.73
Loan mkt sh. f	70.81	100.00	-4.85	-6.07	0.00	0.00
$L^{s,c}$/Total Credit	53.40	52.67	-7.11	-8.66	-3.74	-10.27
Loan Int. Rate r^c	4.67	4.58	0.95	1.17	0.50	1.41
Avg. Markup	95.71	84.92	12.59	1.37	0.49	0.00
Additional Moments						
Measure f Banks			-3.92	-15.19	-4.34	-12.18
Loans (b, f)			$(3.65, -7.98)$	$(4.45, 0.73)$	$(\cdot, 0.85)$	$(\cdot, 1.73)$
Bank Loan Supply			-7.19	-8.90	-3.77	-10.52
Total Credit			-0.08	-0.27	-0.05	-0.29
Output			0.00	-0.19	0.12	-0.22
Taxes/Output			-30.02	-68.50	-92.64	-32.74
Column	(2)	(3)	(4)	(5)	(6)	(7)

Note: Δ(%) refers to the percentage change relative to the baseline model with capital requirements at $\varphi_\theta = 0.04$ (columns (2) and (3)) for the corresponding model with imperfect competition or with perfect competition. The baseline columns in this table differ slightly from those of Table 6 due to the inclusion of the crisis state. The long-run moments are computed from 10,000 period simulations that discard the initial 2,000 periods. The initial baseline moments in Table 10 differ from those presented in Table 6 since we do not filter periods with $z = z_C$. Initial conditions for short-run simulations are consistent with the long-run average distribution of banks (big and fringe) as well as aggregate productivity. The moments reported as the "short-run" effects correspond to the moments that arise five periods after the policy change.

Moving to the role that market structure plays in policy outcomes (a comparison between columns 6 and 7 for the perfectly competitive outcomes and those outcomes from our baseline imperfect competition model in columns 4 and 5 in Table 10), we find that the perfectly competitive model generates a significantly different response in banking industry dynamics (entry and exit rate responses). In the long-run of the perfectly competitive model, the exit rate declines by 0.34%, which compares to the 20.2% decline in the model with imperfect competition. As in the model with imperfect competition, the entry rate plunges in the short-run. However, as there is no big bank competing for the market share of failing banks or preventing the entry of new banks, the entry rate eventually returns nearly to levels observed prior to the implementation of higher capital requirements. The short-run increases in loan interest rates and markups are 2 times and 25 times larger, respectively, in the model with imperfect competition than in the perfectly competitive model. At the center of this change is the systemic link between big bank lending and aggregates that is present in the model with imperfect competition (as shown in Figures 4 and 5) but absent in the model with perfect competition. These differences lead to intensive margin responses that move in opposite directions (negative in the model with imperfect competition and positive in the model with perfect competition).[62]

[62] It is instructive to present a back of the envelope partial equilibrium calculation on the effect on lending rates of a rise in capital requirements from 0.04 to 0.085 (i.e., a difference of 0.045). A full pass-through would equal 0.045 times the cost of bank equity funding relative to deposits, which in the long run is $((\beta\gamma)^{-1} - (1+\overline{r})) = 1.902\%$. That

Of particular importance for financial stability is the change in the probability of a crisis as a result of the policy.[63] Before turning to the effect of policy, we first note from columns (2) and (3) of Table 10 that a model with perfect competition has nearly a 10 times higher probability of a crisis. This is consistent with the competition-stability results we presented in Section 6.3; more competition leads to more fragility. Next, comparing columns (5) and (7), the policy change in the model with imperfect competition results in a decline of 20% in the probability of a crisis but implies a 13% increase in the model with perfect competition. In the model with imperfect competition, as the big bank gains market share, the reduction in the exit rate is accompanied by a decline in its volatility. The opposite is true in the model with perfect competition. These differences lead to a reduction in the likelihood of observing an exit rate above 2 standard deviations from its mean in the imperfectly competitive model (and hence a reduction in crisis probability) and an increase in the perfectly competitive model. In summary, bank regulation can have a significantly different response in the model with imperfect competition than in the model with perfect competition.

A natural question to ask after observing the effects of increasing capital requirements is what role do capital requirements play at all? That is, if banks hold a buffer of capital to insure their charter value, isn't that enough? To answer this question, we perform an experiment where we set capital requirements for all banks to zero and compare the short-run and long-run response for both models. Table A.18 in the Appendix presents the results. As expected, both big and small banks lower their capital ratios when the capital requirement is lifted. However, in keeping with the charter value hypothesis they do not set them to zero for fear of exit. Since capital requirements act like a tax on profitable lending opportunities, it is not surprising that lowering that tax to zero induces entry, much more so in the case of imperfect competition.[64] This experiment provides another example where regulation appears to be more important for an imperfectly competitive industry than in a perfectly competitive industry.

7.2 Size Dependent Capital and Liquidity Requirements

We begin with analyzing model predictions from an increase in size dependent policies for capital requirements and liquidity requirements.[65] In particular, liquidity requirements (ϱ_b) rise from zero in the benchmark to 8% (Liquidity Coverage Ratio) coupled with an additional increase in capital

is, the full partial equilibrium effect would equal 0.0856%. The observed changes in loan interest rates are 0.0549% and 0.0645% in the model with imperfect competition and in the model with perfect competition, respectively. Thus, in line with intuition, the pass-through is larger in the case of perfect competition than in the model with imperfect competition.

[63] We estimate this probability as we did in Section 6.3: the indicators are a tax-to-output ratio above 2%, exit and default rates above 2 standard deviations from their mean, and big bank failure. We find crisis events are almost always trigged by loan defaults and exit rates. In Table A.16 in the Appendix we provide additional measures of volatility and a decomposition of the calculated probability of a crisis across policy counterfactuals for our baseline model.

[64] Entry is more pronounced in the model with imperfect competition since in the face of massive entry the big bank strategically lowers its loan supply to try to counteract falling interest rates on loans (which itself induces entry). A further result of the decrease in "capital taxes" is a selection effect whereby less efficient banks enter. In the long-run, this increase in the number of banks leads to a more pronounced decline in markups and cash-flows (which leads to higher exit rates). All these effects combine to yield a nearly ten-fold increase in the probability of a crisis. While the model with perfect competition also predicts a rise in the probability of a crisis and taxes-to-output, the rise is more modest than in the baseline model owing to less entry by inefficient banks.

[65] It is difficult to apply size dependent policies (which include countercyclical capital buffers and liquidity requirements) in an environment with measure zero banks (i.e. perfect competition) since by construction, these banks lack the systemic relevance that we described in Section 6 that size dependent policies are attempting to address. In addition, in our calibration, the set of banks with the largest δ_f represents approximately 8%, while the target of the policy represents only 0.2% (13/5000). For this reason, we perform the following set of experiments only in our baseline model with imperfect competition.

requirements for big banks $\varphi_b = 0.11$ (a 2.5% G-SIB surcharge). Qualitatively, most of the results of this experiment (presented in columns (3) and (4) of Table A.15 in the Appendix) are in line with the results presented in columns (4) and (5) of Table 10. One notable difference, however, is the decline in lending by about 3% for big banks (as opposed to an increase in lending by about 4% without size dependent requirements). The shift in the balance sheet composition of large banks (a larger increase in the securities-to-asset ratio than in the previous experiment) is induced primarily by the new liquidity requirement which can be met by holding liquid securities. Not surprisingly, the size dependent policy results in a smaller rise in loan market concentration than in the experiments where capital requirements increased for all banks in Section 7.1.

An unintended consequence of the size dependent policy is a slight increase in the probability of a crisis (from the baseline of 0.14% to 0.17%). We find this result arises primarily from the introduction of size dependent liquidity requirements.[66] Two forces are at play. As banks shifts their portfolio towards safe assets the probability of a crisis declines. However, as big bank lending becomes more sensitive to external financing shocks δ_b, size dependent liquidity requirements lead to a higher and more volatile default frequency which in place lead to a rise in the crisis probability.[67]

While countercyclical capital buffers (i.e. higher capital requirements on banks with assets over $250 billion in good times) are part of the Dodd-Frank plan's implementation of Basel III, the Federal Reserve Board has voted not to raise the requirement in good times. Here we consider the counterfactual consequences of adding up to 0.025 to the capital buffer of 0.11 contingent on the state of the economy (i.e. $\varphi_{b,z} \in [0.11, 0.135]$) along with the higher liquidity.[68] The results of this experiment are shown in columns (5) and (6) in Table A.15 in the Appendix. Qualitatively, the addition of countercyclical capital requirements makes little difference relative to the other size dependent policy. One notable difference, however, is that the augmented capital buffer allows the big bank to slightly increase its loan supply in the long run since the capital accumulated during good times reduces the volatility of its loan supply. A second notable unintended consequence is a further rise in the probability of a crisis (a 30% rise from the baseline of 14% to 18%). We find this result again primarily arises from the size dependent liquidity requirements. In fact, if only the countercyclical capital requirements were imposed, the probability of a crisis would fall, rather than rise.[69] This is driven by the fact that the big bank does more lending, which dampens the risk-shifting effect on the default frequency.

As discussed above, changes to capital requirements interact with changes to liquidity requirements. In order to understand the contribution of each of the capital and liquidity regulations implemented by the Dodd-Frank Act, Table A.17 presents a partial decomposition of the size dependent experiments. This table shows that most of the effect on exit and entry rates comes from the increase in capital requirements (size dependent or countercyclical buffer) since the increase in liquidity regulation reduces exit and entry rates by roughly one half of the combined effect. The negative aggregate effect on lending is also explained almost completely by the increase in capital requirements leading to a decline in the measure of small banks and a loss of market share (nearly -7% for both). Alternatively, the imposition of liquidity requirements alone on the big bank leads

[66] This can be seen from a decomposition of the two parts of the policy in Table A.17 in the Appendix. Specifically, the size dependent capital requirement policy only (column (4)) results in a rise from 0.14% to 0.15% while the size dependent liquidity requirement policy only (column (8)) results in a rise from 0.14% to 0.19%.

[67] Table A.16 in the Appendix shows that, while the volatility of default frequency falls on average, it exceeds the threshold right when default frequency rises (during crisis states) pushing it over the threshold thereby raising the probability of a crisis.

[68] Specifically $\{\varphi_{b,z_C} = 0.1100, \varphi_{b,z_B} = 0.1183, \varphi_{b,z_M} = 0.1266, \varphi_{b,z_G} = 0.1350\}$.

[69] This can be seen from a decomposition of the two parts of the policy in Table A.17 in the Appendix. Specifically, the countercyclical capital requirement policy only (column (6)) results in a fall from 0.14% to 0.13% while the size dependent liquidity requirement policy only (column (8)) results in a rise from 0.14% to 0.19%.

to a rise in the measure of fringe banks and fringe bank loan market share (+3.7%). As discussed above, liquidity requirements alone have a major unintended consequence of raising the probability of default.

7.3 Policy Implications for Allocative Efficiency

An important aspect of the policy reforms we study is that they may change the level of allocative efficiency in the economy by shifting lending between heterogeneous banks. In order to explain this change and to provide a measure that captures how efficiently lending is allocated in the economy, we use the following decomposition of weighted average bank-level marginal cost (proposed originally by Olley and Pakes [52] to measure productivity):

$$\widehat{c}(z) \equiv \sum_{\theta} \int c_{\theta}(\ell'_{\theta}, z)\omega(\ell'_{\theta})d\mu_{\theta}(k_{\theta}, \delta_{\theta}) = \overline{c}(z) + cov(c_{\theta}(\ell'_{\theta}, z), \omega(\ell'_{\theta})), \tag{43}$$

where $\widehat{c}(z)$ is the loan-weighted average of bank-level cost in a period when aggregate productivity is z, $c_{\theta}(\ell'_{\theta}, z)$ is the net cost of extending ℓ'_{θ} (as defined in equation (41)), $\omega(\ell'_{\theta})$ is the loan share, and $\overline{c}(z)$ is the un-weighted mean cost (i.e., $\sum_{\theta} \int c_{\theta}(\ell'_{\theta}, z)d\mu_{\theta}$).[70] That is, loan weighted average cost can be decomposed into two terms: the un-weighted average of bank-level cost and a covariance term between loan shares and cost. A smaller value for the covariance term captures an improvement in allocative efficiency (since the distribution of loans shifts towards banks with lower costs). Since we documented differences in costs $c_{\theta}(\cdot)$ between big and fringe banks in Table 3, allocative efficiency can improve as loans are reallocated from high cost fringe banks to big banks with declining average costs (increasing returns over the region we have estimated).

Since access to cheap external finance is capacity constrained by $\delta_{\theta}(z)$, we also present a decomposition of average (loan-weighted) deposits $\widehat{\delta}(z) = \overline{\delta}(z) + cov(\delta_{\theta}(z), \omega(\ell'_{\theta}))$. This decomposition shows how the distribution of lending and access to external financing (which affects the cost of financing loans) change with regulatory policy. For a given level of aggregate lending, as banks with a higher δ_{θ} take over a larger portion of the market, access to cheap external financing increases. For this reason, a larger value for the covariance term $cov(\delta_{\theta}(z), \omega(\ell'_{\theta}))$ captures an improvement in allocative efficiency (since the distribution of loans shifts towards banks with higher δ_{θ}). We view allocative efficiency as measured by $cov(c, \omega)$ as capturing costs on the asset side and $cov(\delta, \omega)$ as capturing costs on the liabilities side of the bank balance sheet.

Table 11 shows the values for the two decompositions in the baseline model with imperfect competition and the model with perfect competition. Starting with a comparison of the average cost results in the baseline parameterization, we observe that average (unweighted) costs are almost identical between the model with imperfect competition (column 2) and perfect competition (column 3) but allocative efficiency as measured by $cov(c, \omega)$ is significantly higher (i.e., the covariance is more negative) in the model with imperfect competition. This better allocation of lending activity is reflected in a significantly lower measured loan weighted average cost $\widehat{c}$ in the model with imperfect competition than in the model with perfect competition as a significant fraction of loans are extended by big banks, which have lower average costs than fringe banks. Allocative efficiency as measured by $cov(\delta, \omega)$ is also better (more positive) in the model with imperfect competition than the perfectly competitive model, again due to distributional differences.

Turning to the effect of higher capital requirements, they result in an increase in allocative efficiency (as measured by the average (loan-weighted) costs) but the increase is quantitatively

[70] Like the long-run moments presented in the previous section, we compute the average (loan-weighted) cost in equation (43) and all its components every period during the simulation of the model. Table 11 presents the corresponding time series average.

more significant in the model with imperfect competition (column 2 versus column 4) as opposed to the model with perfect competition (column 3 versus column 5). The differences arise as the change in average costs comes not only from changes within the distribution of fringe banks but also from changes in loan market concentration. In addition, we observe that the differences in bank industry dynamics described in the previous section also have important implications for the distribution of δ. In the model with imperfect competition, the increase in concentration and the reduction in bank failure after the increase in capital requirements are large enough to induce an increase in average (loan-weighted) deposits $\widehat{\delta}(z)$ that is explained completely by an increase in allocative efficiency as measured by $cov(\delta, \omega)$ (as in fact average deposits $\overline{\delta}(z)$ decline). On the other hand, the significantly smaller decline in bank exit (and the absence of changes in big bank market share) in the perfectly competitive case (column 3 versus column 5) leads to a slight decline in $\widehat{\delta}(z)$ that is explained by a decline in allocative efficiency as $cov(\delta, \omega)$ decreases (as $\overline{\delta}$ increases).[71]

Table 11: Allocative Efficiency of Capital and Liquidity Requirements

	Baseline		High Cap. Req.	
(φ_f, φ_b)	$(0.04, 0.04)$		$(0.085, 0.085)$	
(ϱ_f, ϱ_b)	$(0.0, 0.0)$		$(0.0, 0.0)$	
	Imperfect	Perfect	Imperfect	Perfect
Moment (%)	Comp.	Comp.	Comp.	Comp.
Avg. (loan-weighted) cost $\hat{c}$	1.718	1.802	1.719	1.824
Avg. cost $\overline{c}$	1.759	1.759	1.807	1.791
$Cov(c, \omega)$	-0.041	0.043	-0.089	0.033
Avg. (loan-weighted) deposit $\hat{\delta}$	0.229	0.271	0.241	0.271
Avg. deposit $\overline{\delta}$	0.202	0.261	0.198	0.264
$Cov(\delta, \omega)$	0.027	0.010	0.043	0.007
Fringe Loan Market Share	70.810	100.000	66.506	100.000
Column	(2)	(3)	(4)	(5)

Note: Moments presented here correspond to time series averages over z.

In sum, after all policy changes, the distribution of loans shifts towards banks with lower costs both in monitoring loans and in obtaining cheap external finance, which drives the improvement in allocative efficiency in our baseline model of imperfect competition (unlike the perfectly competitive model). This relationship between banking regulation and allocative efficiency is consistent with the findings of Berger and Hannon [14] that present evidence in favor of efficiency gains from an increase in bank concentration.

7.4 Welfare Implications of Capital and Liquidity Requirements

To assess the welfare consequences of changes in banking regulation that we presented in the previous subsections, we ask the question, "What would households and entrepreneurs be willing to pay (or be paid) to increase capital and liquidity requirements?"

To answer this question, we calculate consumption equivalents for each type of representative agent ($h \in \{H, E\}$). Specifically, let $C_{h,t}^{pre}$ denote type h equilibrium aggregate consumption in the baseline (pre-reform) economy and $C_{h,t}^{post}$ denote type h equilibrium aggregate consumption post-reform. Then the consumption equivalent α_h^T for a representative type h agent over a period T as a result of the reforms is defined as the constant percentage increase in consumption in the pre-reform

[71] Table A.19 in Appendix A-5.6 shows that allocative efficiency increases monotonically (as measured by a decline in $cov(c, \omega)$ and an increase in $cov(\delta, \omega)$) with the market share of the big bank after the introduction of size dependent capital and liquidity regulations.

case that gives the agent the same expected utility in the post-reform economy. Specifically, α_h^T solves:

$$E_0\left[\sum_{t=0}^{T}\beta_h^t C_{h,t}^{post}\right] = E_0\left[\sum_{t=0}^{T}\beta_h^t(1+\alpha_h^T)C_{h,t}^{pre}\right]. \quad (44)$$

We compute both short-term ($T = 5$) and long-term ($T = \infty$) gains or losses as in our policy experiments.[72] We let the ex-ante value (i.e., before being born as $h \in \{H, E\}$) for an agent in the pre-reform economy be[73]

$$W_{pre}^T = \frac{\mathcal{H}}{1+\mathcal{H}}E_0\left[\sum_{t=0}^{T}\beta_H^t C_{H,t}^{pre}\right] + \frac{1}{1+\mathcal{H}}E_0\left[\sum_{t=0}^{T}\beta_E^t C_{E,t}^{pre}\right]. \quad (45)$$

We define the ex-ante value post-reform similarly. Then, we can write the ex-ante value of the reform in consumption equivalent terms $\overline{\alpha}^T$ as $W_{post}^T = (1+\overline{\alpha}^T)W_{pre}^T$.

Equation (44) makes clear that, since preferences are linear, computing ex-ante consumption equivalents in the long-run amounts to computing the difference between expected consumption pre- and post-reform. Since we work with linear preferences to avoid complications with stochastic discount factors, our welfare measure does not capture the effects of changes in aggregate volatility.[74] Hence, our measure of welfare gains should be taken as a lower bound in cases in which consumption volatility declines and as an upper bound in cases in which volatility increases.

Table 12 presents the average welfare gains (or losses if negative) as well as the change in the coefficient of variation of consumption for each type of agent in the long run (ΔCV_{C_H} and ΔCV_{C_E} for consumers and entrepreneurs, respectively) and for the aggregate (i.e, population weighted average, ΔCV_C) for each policy experiment we performed. In the model with imperfect competition (column (2)), increasing capital requirements to 8.5% for all banks induces a welfare loss in the short run and a welfare gain in the long run (nearly 1 tenth of 1%).[75] Long-run household gains (1 tenth of 1%) arise from lower taxes due to the decline in bank failures and better capitalized failing banks (which reduce deadweight losses from liquidation costs ξ that represent up to 20 percent of the loan portfolio), higher dividends resulting from higher markups (at the expense of entrepreneurs), and a relatively small drop in output due to the increase in the share of non-bank lending. Long-run entrepreneur losses (nearly 3 tenths of 1%) arise from the increase in bank loan interest rates (which due to the risk-shifting effect induce higher deadweight losses associated with project failure λ that represent up to 40 percent of the value of the project) and the decline in projects being operated due to a drop in total credit. Since entrepreneurs represent a relatively small fraction of the population the ex-ante welfare change is positive in the long run.[76] The long run gains contrast with short run losses that arise from the increase in small bank failure and the associated increase in bank loan interest rates and markups.

[72] As we discussed in Section 7, all experiments are run with initial conditions that are consistent with the long-run equilibrium in the pre-reform economy. To compute expectations we take the average across a large number of simulations of the model from this given set of initial conditions.

[73] Since there is a mass $\mathcal{H}$ of households and a unit mass of entrepreneurs the probability of being born as a household is $\frac{\mathcal{H}}{1+\mathcal{H}}$.

[74] Moreover, in this paper, we focus on an equilibrium that takes as given (i.e., as determined uniquely by government policy) the return on safe securities r^A. This price would be subject to change in a full general equilibrium model (i.e., a model that captures how changes in demand for securities affect its price).

[75] The long-run welfare gains in our model are somewhere between those (0.035%) in Begenau [9] and those (1.8%) in Clerc et. al [19] at the optimum.

[76] We have calibrated $\mathcal{H}$ (the mass of households) so that entrepreneurs in the model represent 17% of the population (consistent with the evidence presented in Quadrini [56]). A different target will affect ex-ante welfare but not other moments in the model due to our normalization that $\mathcal{H}y = 1$. Simple algebra shows that economy-wide long-run ex-ante welfare would decline if $\mathcal{H}$ is calibrated using a fraction of entrepreneurs larger than 30%.

Table 12 (column (5)) shows that the model with perfect competition generates a much smaller long run welfare gain from the rise in capital requirements due to its near absent effect on exit rates and the larger decline in bank loan supply. In the short run (column (4)), however, the policy change induces a much larger decline in exit rates than the imperfectly competitive model, which explains the qualitative difference (i.e. positive in column (4) and negative in column (3)) in short run welfare between the two models.[77]

Of particular importance is the change in consumption volatility and the difference between the two models. While with linear preferences this change in volatility does not have direct implications for our measure of welfare gains or losses , the observed changes in consumption volatility help to set a bound on welfare values. In particular, the decline of consumption volatility of -1.6% observed in the model with imperfect competition (column (2)) suggests that the reported average welfare gain can be taken to be a lower bound of the benefits of this policy change. On the contrary, the increase in consumption volatility of 5.8% observed for the model with perfect competition (column (3)) implies that the small welfare gain should be taken to be an upper bound of the effects of this policy on an environment with only perfectly competitive banks.

Table 12: Welfare Consequences of Capital Requirements

	High Cap. Req.			
(φ_f, φ_b)	$(0.085, 0.085)$			
(ϱ_f, ϱ_b)	$(0.0, 0.0)$			
	Imperfect Comp.		Perfect Comp.	
	short-run	long-run	short-run	long-run
α_H	-0.004	0.114	0.030	0.024
ΔCV_{C_H}	-	-1.822	-	6.736
α_E	-0.280	-0.268	0.015	-0.325
ΔCV_{C_E}	-	-0.673	-	1.425
$\overline{\alpha}$	-0.034	0.100	0.028	0.005
Avg. ΔCV_C	-	-1.627	-	5.833
Column	(2)	(3)	(4)	(5)

Note: α_H and α_E are defined in equation (44). Positive values correspond to a welfare gain from the reform and a negative value corresponds to a welfare loss. ΔCV_{C_H} and ΔCV_{C_E} refer to the change in the coefficient of variation of long-run consumption for households and entrepreneurs, respectively. All values reported in percentage terms.

8 Directions for Future Research

The focus of our paper is the interaction of policy to promote stability and market structure. Our policy experiments document how policy affects market shares of big and small banks (as well as the loss of commercial bank market share to the non-bank sector due to regulatory arbitrage) both through intensive and extensive margins. We also show how market structure affects the efficacy of policy, providing a quantitative dynamic framework to consider possible nonlinear tradeoffs between competition and stability discussed in Martinez-Miera and Repullo [46] and Corbae and Levine [23].

While we model the U.S. banking sector using a simple dominant-fringe framework owing to the roughly 50-50 split between the top 10 big banks and the remaining 5000 smaller banks at the time of this writing, a direction for future research would consider more complicated strategic market structures. For instance, Corbae and D'Erasmo [21] assume strategic dominant firms of different sizes to capture spatial heterogeneity. Another important direction for future research is

[77] Table A.20 in Appendix A-5.7 presents the welfare effects of size dependent policies. Columns (6)-(9) show similar qualitative patterns; size dependent capital and liquidity requirements result in average welfare losses in the short run and welfare gains in the long run.

to incorporate mergers into this framework. Corbae and D'Erasmo [22] take a step in that direction by employing the Ericson and Pakes [32] framework to consider the growth of big banks. Finally, while our Call Report data does not admit borrower level data, as such data becomes available (as in the Spanish case of Jimenez et al. [41]), ex-ante heterogeneity in borrower type could be added to the model in future research.

References

[1] Acharya, V. and A. Thakor (2016) "The dark side of liquidity creation: Leverage and systemic risk," *Journal of Financial Intermediation*, vol. 28, 4-21. https://doi.org/10.1016/j.jfi.2016.08.004

[2] Acemoglu, D., A. Ozdaglar, and A. Tahbaz-Salehi (2015) "Systemic risk and stability in financial markets," *American Economic Review*, vol. 105, 564-608. https://doi.org/10.1257/aer.20130456

[3] Aguirregabiria, V., R. Clark, and H. Wang (2016) "Diversification of geographic risk in retail bank networks: Evidence from bank expansion after the Riegle-Neal Act," *Rand Journal of Economics*, vol. 47, 529-572. https://doi.org/10.1111/1756-2171.12136

[4] Aliaga-Diaz, R. and M., Olivero (2011) "Do bank capital requirements amplify business cycles? Bridging the gap between theory and empirics," *Macroeconomic Dynamics*, doi:10.1017/S1365100510000623, 1-38. https://doi.org/10.1017/S1365100510000623

[5] Allen, F. and D. Gale (2004) "Competition and financial stability," *Journal of Money, Credit, and Banking*, 36, p. 453-80. https://doi.org/10.1353/mcb.2004.0038

[6] Arellano, Manuel and Bond, Stephen, (1991) "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *Review of Economic Studies*, 58(2), p. 277-97. https://doi.org/10.2307/2297968

[7] Basel Committee on Banking Supervision (2010) "Basel III: International framework for liquidity risk measurement, standards, and monitoring", *Bank for International Settlements*.

[8] Bassett, W., M. Chosak, J. Driscoll, and E. Zakrajšek (2014) "Changes in bank lending standards and the macroeconomy," *Journal of Monetary Economics,* 62, 23-40. https://doi.org/10.1016/j.jmoneco.2013.12.005

[9] Begenau, Juliane (2020) "Capital requirements, risk choice, and liquidity provision in a business-cycle model," *Journal of Financial Economics* 136.2, 355-378. https://doi.org/10.2139/ssrn.2576277

[10] Bianchi, J. and S. Bigio (2017) "Banks, liquidity management and monetary policy", mimeo. https://doi.org/10.21034/sr.503

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

[12] Berger, A., L. Klapper and R. Turk-Ariss (2008) "Banking competition and financial stability," *Journal of Financial Services Research*, pp. 99-118. https://doi.org/10.1007/s10693-008-0050-7

[13] Bernanke, B. and M. Gertler (1989) “Agency costs, net worth, and business fluctuations,” *The American Economic Review,* 79(1), pp.14-31.

[14] Berger, A., and T. Hannan (1998) “The efficiency cost of market power in the banking industry: A test of the “quiet life” and related hypotheses,” *Review of Economics and Statistics,* 80.3 (1998): 454-465. https://doi.org/10.1162/003465398557555

[15] Berger, A. and L. Mester (1997) “Inside the black box: What explains differences in the efficiencies of financial institutions?”*Journal of Banking and Finance*, vol. 21, pp. 895-947. https://doi.org/10.1016/S0378-4266(97)00010-1

[16] BIS (2011) “Basel III: A global regulatory framework for more resilient banks and banking systems,” Basel Committee on Banking Supervision.

[17] Boyd, J. and G. De Nicolo (2005) “The theory of bank risk-taking and competition revisited”, *Journal of Finance*, 60, p. 1329-43. https://doi.org/10.1111/j.1540-6261.2005.00763.x

[18] Buchak, G., G. Matvos, T. Piskorski, and A. Seru (2018) “Fintech, regulatory arbitrage, and the rise of shadow banks”, *Journal of Financial Economics*, 130, p. 453-483. https://doi.org/10.1016/j.jfineco.2018.03.011

[19] Clerc, L, A. Derviz, C. Mendicino, S. Moyen, K. Nikolov, L. Stracca, J. Suarez, A. Vardoulakis (2015) “Capital regulation in a macroeconomic model with three layers of default”, *International Journal of Central Banking*, vol. 11, p. 9-63. https://doi.org/10.2139/ssrn.2544607

[20] Cooley, T. and V. Quadrini (2001) “Financial markets and firm dynamics”, *American Economic Review*, vol. 91, pp. 1286-1310. https://doi.org/10.1257/aer.91.5.1286

[21] Corbae, D. and P. D’Erasmo (2011) “A quantitative model of banking industry dynamics”, mimeo.

[22] Corbae, D. and P. D’Erasmo “Rising bank concentration,” forthcoming *Journal of Economic Dynamics and Control*.

[23] Corbae, D. and R. Levine (2018) “Competition, stability, and efficiency in the banking industry,” mimeo.

[24] den Haan, W., S. Sumner and G. Yamashiro (2007) “Bank loan portfolios and the monetary transmission mechanism,” *Journal of Monetary Economics*, 54(3), 904-924. https://doi.org/10.1016/j.jmoneco.2006.01.008

[25] De Loecker, J., J. Eeckhout, and G. Unger (2019) “The rise of market power and the macroeconomic implications,” forthcoming *Quarterly Journal of Economics*.

[26] De Nicolo, G., A. Gamba, M. Lucchetta (2014) “Microprudential regulation in a dynamic model of banking”, *Review of Financial Studies*, vol. 27, pp. 2097-2138. https://doi.org/10.1093/rfs/hhu022

[27] Diamond, D. (1984) “Financial intermediation and delegated monitoring”, *Review of Economic Studies*, 51, p. 393-414. https://doi.org/10.2307/2297430

[28] 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. https://doi.org/10.1111/j.1468-0297.2008.02208.x

[29] Drechsler, I., A Savov, and P. Schnabl (2017) "The deposits channel of monetary policy", *Quarterly Journal of Economics*, vol. 132, pp. 1819-1876. https://doi.org/10.1093/qje/qjx019

[30] Egan, M., A. Hortascu, and G. Matvos (2017) "Deposit competition and financial fragility: evidence from the US banking sector", *American Economic Review*, 107, p. 169-216. https://doi.org/10.1257/aer.20150342

[31] Elenev, V., T. Landvoight, S. Van Nieuwerburgh (2020) "A macroeconomic model with financially constrained producers and intermediaries," mimeo.

[32] Ericson, R. and A. Pakes (1995) "Markov-Perfect industry dynamics: A framework for empirical work", *Review of Economic Studies*, 62, p. 53-82. https://doi.org/10.2307/2297841

[33] Fernald, J. (2014) "A quarterly, utilization-adjusted series on total factor productivity," Federal Reserve Bank of San Francisco.

[34] Gabaix, X. (2011) "The Granular Origins of Aggregate Fluctuations", *Econometrica*, 79, p.733-772. https://doi.org/10.3982/ECTA8769

[35] Gallin, J. (2013) "Shadow banking and the funding of the nonfinancial sector," FEDS Working Paper No. 2013-50, pp. 89-123. https://doi.org/10.2139/ssrn.2340684

[36] Gomes, J. (2001) "Financing investment", *American Economic Review*, 91, p. 1263-1285. https://doi.org/10.1257/aer.91.5.1263

[37] 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. https://doi.org/10.2307/1593708

[38] Hennessy, C. and T. Whited (2007) "How costly is external financing? Evidence from a structural estimation," *Journal of Finance*, vol. 62, p. 1705-1745. https://doi.org/10.1111/j.1540-6261.2007.01255.x

[39] Hughes, J. , L. Mester, and C. Moon (2001) "Are scale economies in banking elusive or illusive?: Evidence obtained by incorporating capital structure and risk-taking into models of bank production," *Journal of Banking and Finance*, 25(12), 2169-2208. https://doi.org/10.1016/S0378-4266(01)00190-X

[40] Ifrach, B. and G. Weintraub (2017) "A framework for dynamic oligopoly in concentrated industries", *Review of Economic Studies*, 84, p. 1106-1150. https://doi.org/10.1093/restud/rdw047

[41] Jimenez, G., S. Ongena, J.L. Peydro, and J. Saurina (2014) "Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?", *Econometrica*, 82, p.463-505. https://doi.org/10.3982/ECTA10104

[42] Kashyap, A. and J. Stein (1995) "The impact of monetary policy on bank balance sheets" *Carnegie-Rochester Conference Series on Public Policy*, vol. 42, p. 151-195. https://doi.org/10.1016/0167-2231(95)00032-U

[43] 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. https://doi.org/10.1257/aer.90.3.407

[44] Krusell, P. and Anthony Smith (1998) "Income and wealth heterogeneity in the macroeconomy," *Journal of Political Economy*, 106(5), 867-896. https://doi.org/10.1086/250034

[45] Liang, N. and S. Rhoades (1988) "Geographic diversification and risk in banking" *Journal of Economics and Business*, vol. 40, p. 271-284. https://doi.org/10.1016/0148-6195(88)90026-4

[46] Martinez-Miera, D. and R. Repullo (2010) "Does competition reduce the risk of bank failure?,"*Review of Financial Studies*, 23, p. 3638-3664. https://doi.org/10.1093/rfs/hhq057

[47] McLean, R. and M. Zhao (2014) "The business cycle, investor sentiment, and costly external finance," *The Journal of Finance,* 69(3), 1377-1409. https://doi.org/10.1111/jofi.12047

[48] Menzio, G. and S. Shi (2010) "Block recursive equilibria for stochastic models of search on the job," *Journal of Economic Theory* 145(4), 1453-1494. https://doi.org/10.1016/j.jet.2009.10.016

[49] Meyer, A. (2018) "Market concentration and its impact on community banks," *The Regional Economist* 26.1, Federal Reserve Bank of St. Louis.

[50] Minnick, K. and L. Rosenthal (2014) "Stealth compensation: Do CEOs increase their pay by influencing dividend policy?" *Journal of Corporate Finance*, 25, p. 435-454. https://doi.org/10.1016/j.jcorpfin.2014.01.005

[51] Nguyen, T. (2013) "Bank capital requirements: A quantitative analysis,"mimeo. https://doi.org/10.2139/ssrn.2356043

[52] Olley, G. S. and A. Pakes (1996) "The dynamics of productivity in the telecommunications equipment industry,"*Econometrica*, vol. 64, p. 1263-1297. https://doi.org/10.2307/2171831

[53] 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. https://doi.org/10.2307/2555975

[54] Perotti, E. and J. Suarez (2002) "Last bank standing: What do I gain if you fail?," *European Economic Review*, 46, p. 1599-1622. https://doi.org/10.1016/S0014-2921(02)00241-6

[55] Pozsar, Z., T. Adrian, A. Ashcraft, and H. Boesky 2013 "Shadow banking," Federal Reserve Bank of New York *Economic Policy Review*, 19(2), 1.

[56] Quadrini, V. (2000) "Entrepreneurship, saving, and social mobility," *Review of Economic Dynamics*, 3(1), 1-40. https://doi.org/10.1006/redy.1999.0077

[57] Repullo, R. (1986) "Capital requirements, market power, and risk-taking in banking," *Journal of Financial Intermediation*, vol. 13, p. 156-182. https://doi.org/10.1016/j.jfi.2003.08.005

[58] Rios-Rull V., T. Takamura, and Y. Terajima (2018) "Banking dynamics and capital requirements," mimeo.

[59] Rossi-Hansberg, E., P. Sarte, and N. Trachter (2020) "Diverging trends in national and local concentration," forthcoming *NBER Macroeconomics Annual.*

[60] Shaffer, S. (2004) "Patterns of competition in banking", *Journal of Economics and Business*, Vol. 56, p. 287-313. https://doi.org/10.1016/j.jeconbus.2003.10.003

[61] Stein, J. (1988) “Takeover threats and managerial myopia,” *Journal of Political Economy*, vol. 96, p. 61-80. https://doi.org/10.1086/261524

[62] Stiglitz, J. and A. Weiss (1981) “Credit Rationing in Markets with Imperfect Information,” *Journal of Political Economy*, vol. 96, p. 61-80. https://doi.org/10.1086/261524

[63] Tauchen, G. (1986) “Finite state Markov-chain approximations to univariate and vector autoregressions,” *Economics Letters*, 20(2), p. 177-181. https://doi.org/10.1016/0165-1765(86)90168-0

[64] Van Den Heuvel, S. (2008) “The welfare cost of bank capital requirements,” *Journal of Monetary Economics*, 55, p. 298-320. https://doi.org/10.1016/j.jmoneco.2007.12.001

[65] Wang, Y., T. Whited, Y. Wu, and K. Xiao (2020) “Bank market power and monetary policy transition: Evidence from a Structural Estimation,” mimeo.

[66] Wheelock, D. and P. Wilson (2017) “The Evolution of Scale Economies in U.S. Banking,” Federal Reserve Bank of St. Louis Working paper 2015-021C.