

Bank Market Power and Monetary Policy Transmission: Evidence from a Structural Estimation

Yifei Wang, Toni M. Whited, Yufeng Wu, and Kairong Xiao*

April 15, 2019

Abstract

We quantify the impact of bank market power on the pass-through of monetary policy to borrowers. To this end, we estimate a dynamic banking model in which monetary tightening increases banks' funding costs. Given their market power, banks optimally choose how much of a rate increase to pass on to borrowers. In the model, banks are subject to capital and reserve regulations, which also influence the degree of pass-through. Compared with the conventional regulation-based channels, we find that in the two most recent decades, bank market power explains a significant portion of monetary transmission. The quantitative effect is comparable in magnitude to the bank capital channel. In addition, the market power channel interacts with the bank capital channel, and this interaction can reverse the effect of monetary policy when the Federal Funds rate is low.

JEL Codes: E51, E52, G21, G28

Keywords: Monetary policy transmission, banking competition.

*Wang is from the University of Michigan; wangyf@umich.edu. Whited is from the University of Michigan and the NBER; twhited@umich.edu. Wu is from the University of Illinois; yufengwu@illinois.edu. Xiao is from Columbia University; kairong.xiao@gsb.columbia.edu.

1. Introduction

We examine the quantitative importance of bank market power as a transmission mechanism for monetary policy. This question rests against a backdrop of traditional theories and analysis of monetary policy transmission that largely focus on the impact of regulatory constraints on bank lending (e.g. [Bernanke and Blinder 1988](#); [Kashyap and Stein 1995](#)). Indeed, the banking industry is often assumed to be perfectly competitive, leaving only regulatory frictions, such as reserve or capital requirements, as a channel through which monetary policy affects loan supply. However, recent research suggests that the industrial organization of the banking sector may also play a role in the transmission of monetary policy ([Drechsler, Savov, and Schnabl 2017](#); [Scharfstein and Sunderam 2016](#)). While important, this newer literature has been qualitative in nature, leaving open the question of the relative importance of traditional versus market-power channels for the transmission of monetary policy.

Our paper tries to fill this void by constructing and estimating a dynamic banking model with three important frictions: regulatory constraints, financial frictions, and imperfect competition. The estimation allows our data on commercial banks to discipline the model parameters and thus expose the relative magnitude of these three frictions. We find that regulatory frictions restricting bank capital and bank market power both play an important role in monetary policy transmission, while reserve requirements are unimportant. In terms of magnitude, the effect of bank market power is comparable in magnitude to that of bank capital. We also find an interesting interaction between different monetary policy transmission channels. Specifically, we show that bank capital regulation interacts with market power and reverses the sign of monetary policy when the Federal Funds rate is very low. We estimate that when the Federal Funds rate is below 2.27%, further cuts in the policy rates can be contractionary. Moreover, we find external validation of this reversal rate by showing in a simple regression framework that the relation between bank capital and interest rates

switches sign at approximately this interest rate.

An understanding of the intuition behind these results requires a deeper description of the model. In the economy, banks act as intermediaries between borrowers and depositors. The lending decision is dynamic because deposits are short-term, while loans are long-term. Monetary tightening enters the picture by increasing banks' funding costs in the deposit market. Because they are not price takers in the deposit and loan market, banks choose how much of a rate increase to pass on to borrowers. The degree of pass-through is influenced by the tightness of regulatory constraints, the degree of financial frictions, and the intensity of competition.

These frictions in our model map into three monetary policy transmission channels emphasized in the literature. The first is the bank reserve channel in which a high Federal Funds rate raises the opportunity cost of holding reserves (Bernanke and Blinder 1988; Kashyap and Stein 1995) and thus contracts deposit creation. The second is the bank capital channel in which a rise in the Federal Funds rate exacerbates banks' natural maturity mismatch, thereby lowering bank capital and, consequently, the capacity to lend (Bolton and Freixas 2000; Van den Heuvel 2002; Brunnermeier and Sannikov 2016). The third is the market power channel emphasized by Drechsler, Savov, and Schnabl (2017) and Scharfstein and Sunderam (2016). Intuitively, after a rate increase, cash becomes less attractive to households relative to deposits. Monopolistically competitive banks can exploit this extra deposit demand by charging higher spreads on deposits. In equilibrium, total deposits fall because households substitute risk-free bonds for deposits. Because banks then need to fund marginal lending in the more expensive debt market, lending contracts.

To gauge the quantitative importance of these transmission channels, we estimate our model using a panel of U.S. commercial banks. Our estimation combines methods used in the industrial organization literature (Berry, Levinsohn, and Pakes 1995; Nevo 2001) with those used in the corporate finance literature (Hennessy and Whited 2005; Bazdresch, Kahn, and Whited 2018). As a first step, we use the demand estimation techniques from industrial

organization to obtain the elasticities of substitution in the deposit and loan markets. We then plug these estimates into our model, and use simulated method of moments to obtain estimates of parameters that quantify financial frictions and operating costs. The sequential use of these two techniques is a methodological advance that allows us to consider a rich equilibrium model that would otherwise be intractable to estimate.

To obtain our results on the relative importance of each transmission channel, we then use these parameter estimates to simulate counterfactual experiments in which we subtract each channel from the model one at a time. These counterfactuals also produce our interesting result that rate cuts can be contractionary when rates are already low. Low interest rates depress bank profits by reducing bank market power in the deposit market, as cash becomes less unattractive to households. Lower profits then tighten the capital constraint and result in less lending. This result sheds light on the sluggish bank lending growth post crisis, as an ultra-low rate policy can unintentionally reduce bank profitability, and consequently constrain banks' capacity to lend. Overall, our results suggest that Federal Reserve actions can have complicated effects on bank lending depending on the level of policy rates, the amount of bank capital, and the industrial organization of the banking sector.

Our paper contributes to the literature studying the role of banks in transmitting monetary policy. It is the first to estimate a structural dynamic banking model to quantify various transmission channels.¹ Prior to our work, little has been known about the relative importance of different transmission channels, as this type of quantitative exercise is difficult to undertake using reduced form methods.

Second, prior literature usually studies each transmission channel separately, but little is known about the interactions between different channels. Thus, an important contribution of this paper is to provide a unified framework to study these interactions. For example, [Drechsler, Savov, and Schnabl \(2017\)](#) and [Scharfstein and Sunderam \(2016\)](#) study the market power in the deposit market and loan market separately. We show that the relative

¹[Xiao \(2018\)](#) also uses a structural approach, but the main focus is on shadow banks.

importance of the two markets depends on the level of the Federal Funds rate. The deposit market is more important when the Federal Funds rate is low, while the loan market becomes more important when the Federal Funds rate is high. [Brunnermeier and Koby \(2016\)](#) put forth the theoretical possibility that monetary policy can switch sign at a certain threshold referred to as a “reversal rate.” Our paper provides an empirical estimation of this threshold.

Third, our paper is related to the literature on external financial frictions. Largely focused on industrial firms, this literature shows that financial frictions significantly affect corporate policies such as investment, cash holding, and dividend payout. We show that banks also face significant financial frictions despite their regular participation in the capital markets. This last result supports the arguments in [Kashyap and Stein \(1995\)](#) that banks’ high external financing costs can influence the quantity of bank lending.

2. Data

Our main data set is the Consolidated Reports of Condition and Income, generally referred to as the Call Reports. This data set provides quarterly bank-level balance sheet information for U.S. commercial banks, including deposit and loan amounts, interest income and expense, loan maturities, salary expenses, and fixed-asset related expenses. We merge the Call Reports with the FDIC Summary of Deposits, which provides branch-level information for each bank since 1994 at an annual frequency. We follow [Egan, Hortacsu, and Matvos \(2017\)](#) by excluding banks with fewer than ten domestic branches. This filter eliminates foreign banks with few U.S. branches and tiny local banks. The resulting sample period is 1994–2017.

Our analysis requires data from several further sources. First, we retrieve publicly listed bank returns from CRSP. We link the stock returns to bank concentration measures using the link table provided by the Federal Reserve Bank of New York. We obtain bank industry stock returns from Kenneth French’s website. We collect the Federal Open Market Committee meeting dates from the FOMC Meeting calendar. Finally, we obtain the following time

series from FRED (Federal Reserve Economic Data): NBER recession dates, the effective Federal Funds rate, the two-year and five-year Treasury yields, the aggregate amount of corporate bonds issued by U.S. firms, and the aggregate amount of cash, Treasury bonds, and money-market mutual funds held by households.

Table 1 provides summary statistics for this sample, which we use for our demand estimation. The first two lines report the mean, standard deviation, and several percentiles of bank market shares in both the deposit and loan markets. We define the entire United States as a unified market to compute these shares, with the total size of the deposit market defined as the sum of deposits, cash, and Treasury securities held by all U.S. households. The total size of the loan market is defined as the sum of bank loans and corporate bonds borrowed by U.S. firms. Interestingly, mean market shares in both the loan and deposit markets lie near the 90th percentile, indicating a very skewed distribution of market shares in which a few large banks dominate the market.

The next two lines report summary statistics for deposit and loan rates, which we impute by dividing total deposit interest expense by total deposits and total interest income by total loans. The next two lines report summary statistics for two non-rate bank characteristics used in the estimation: the number of branches and the number of employees per branch. While we see little variation in the number of employees per branch, we see both high variance and skewness in the number of branches per bank. The skewness is consistent with the skewness in market shares, as the number of branches is highly correlated with bank size. Lastly, we report summary statistics for the two supply shifters we use in our estimation: salaries and fixed asset expenses (Ho and Ishii 2011). Fixed asset expenses include all non-interest expenses stemming from use of premises, equipment, furniture, and fixtures. We scale both supply shifters by total assets.

3. Stylized Facts

To motivate our study of the relative importance of various monetary policy transmission channels, we first examine the relation between bank equity returns and monetary policy shocks. The conventional wisdom from the traditional transmission channels is that high interest rate monetary policy should have a negative impact on bank capital. Specifically, the bank capital channel predicts that because bank-loan maturity exceeds deposit maturity, an increase in interest rates reduces the value of assets more than the value of liabilities, so bank equity falls. The bank reserve channel predicts that a high Fed Funds rate increases the opportunity cost of holding reserves because reserves bear no interests. Therefore, a high interest rate leads to less deposit taking and less lending, which lowers banks' profits and consequently, bank capital.

We verify this prediction by regressing bank equity returns on the change in the two-year Treasury yield on Federal Open Market Committee (FOMC) meeting days. We focus on FOMC days following [Hanson and Stein \(2015\)](#), who use daily changes in the two-year Treasury yield surrounding FOMC meetings to measure monetary policy shifts. The advantage of examining the two-year Treasury yield instead of the Federal Funds rate is that the former captures the effects of “forward guidance” in the FOMC announcement, which has become increasingly important in recent years ([Hanson and Stein 2015](#)).² The identifying assumption is that unexpected changes in interest rates in a one-day window surrounding scheduled Federal Reserve announcements arise from news about monetary policy. While our sample period runs from 1994 to 2017, we exclude the burst of dot-com bubble (2000–2001) and the subprime financial crisis (2007–2009) because in these crisis times, monetary policy can be strongly influenced by the stock market rather than by inflation or economic growth ([Cieslak and Vissing-Jorgensen 2017](#)).

Table 2 reports the regression estimates. We find the conventional wisdom is generally

²We also use 1-year Treasury yield and the results are robust.

true when the startling level of the Fed Funds rate is above 2%. As shown in the regression in column 1, an increase in the short-term interest rates reduces bank equity value. However, we find that the negative relation between interest rates and bank equity reverse sign if the starting level of the Fed Funds rates are below 2%. As shown in column 2, an increase in the two-year Treasury yield is associated with positive significant returns for bank equity. In other words, the market expects that an increase in interest rates leads to an increase in bank capital. This low-interest-rate result stands in contrast to the conventional wisdom that monetary tightening reduces bank capital. This result is not driven by a steepening of the term structure, as we control for the change in the five-year term spread over two-year Treasury, where we pick the five-year spread to match the average maturity of bank loans in our sample, which is five years. As shown in Figure 1, the contrast between the results in Columns 1 and 2 can be seen in a simple scatter plot of bank industry excess returns against the change in the two-year Treasury yield. Moreover, in the Online Appendix, we examine returns for all 49 Fama-French industries. We find that the banking industry is the *only* industry exhibiting a switch from a negative interest sensitivity to positive interest sensitivity in the low interest environment. In summary, we find that monetary policy has a nonmonotonic effect on bank capital. When the Federal Funds rate is high, the relation between the short-term rates and bank capital seems to be negative, but when the Federal Funds rate is low, this relation becomes positive. As far as we know, this paper is the first to document this empirical relation.

To delve into the mechanism behind this basic result, we estimate a non-parametric relationship between the Fed Funds rates and the average deposit spread of the U.S. banks. The deposit spread is defined the difference between the Fed Funds rates and the deposit rates, which measures the price that banks charge for their depository services. One would expect a constant deposit spread that equals marginal costs of providing depository services if there is no market power or other frictions. However, we find a positive relation between the deposit spread and the Fed Funds rate, which becomes even stronger when the Fed

Funds rate is close to zero. In other words, banks are able to charge higher prices for their depository services when the Fed raises interest rates. This pattern is consistent with the idea that banks have market power as suggested by [Drechsler, Savov, and Schnabl \(2017\)](#). In presence of market power, high Fed Funds rate allows banks to raise the markups above marginal costs because an important outside option for depositors, cash, becomes more costly to hold. Therefore, banks make more profits in the deposit market when the Fed raises short-term rates. This effect is particular strong when the Fed Funds rate is close to zero when banks face intense competition from cash.

To verify this explanation, in Columns 3 and 4 of Table 2, we interact the change in the interest rate with a measure of bank market power, the Herfindahl-Hirschman index of the local deposit market in which the bank operates, where we define a local deposit market as a Metropolitan Statistical Area (MSAs). If a bank operates in several MSAs, the bank-level HHI is the weighted average of local HHIs, weighted by the deposits of the bank in the local market. We find that when the Federal Funds rate is below 2%, banks with greater market power in the deposit market experienced larger positive returns.

This evidence provides a possible explanation on the sluggish lending growth in the recent economy recovery. Figure 3 plots average U.S. bank loan-to-deposit ratios following the five recessions from 1973 to 2017. The loan-to-deposit ratio usually falls when a recession starts, as both loan demand and supply fall. In the first three recessions during this period, the loan-to-deposit ratio recovered three to four years after the start of the recession. However, the recoveries after the 2001 and 2008 recessions are notably slower. The loan-to-deposit ratio took six to seven years to recover after the 2001 recession, and this ratio still has not yet recovered 10 years after the 2008 recession. This sluggishness is surprising given that both recessions ended within two years.³

Although many factors such as regulatory changes could be driving the lengthy recovery

³One concern with the results in Figure 3 is that they may be driven largely by an increase in the denominator—deposits. To address this concern, in the Online Appendix, we plot the amount of loans following the onset of each recession and the result still holds.

in bank lending, one feature that these two episodes share is protracted periods in which the Federal Reserve lowered short term interest rates to near zero. As such, loose monetary policy is a possible culprit. As argued above, extreme low short-term rates may depress banks' profitability and slow down the capital accumulation. Since banks want to maintain certain level of capital ratio, a lower capital accumulation will constrain the lending growth.⁴

4. Model

While the facts documented in the previous section point to interesting interactions between bank capital, bank, competition and interest rates, this time-series evidence cannot help us understand the various mechanisms that drive these patterns in the data. To move in this direction, we next consider an infinite-horizon, equilibrium model with three sectors: households, firms, and banks. Banks act as intermediaries between households and firms by taking short-term deposits from households and providing long-term loans to firms. We model the households and firms as solving straightforward static discrete choice problems in which they choose from a variety of saving and financing vehicles. The richness of the model lies in the banking sector, as a variety of frictions imply that monetary policy affects the amount of intermediation that banks provide. These frictions are important because in a frictionless world where bank loans and bonds are perfect substitutes, bond market interest rates summarize monetary policy and banks are simply pass-through entities. However, if bank loans and bonds are imperfect substitutes ([Bernanke and Blinder 1988](#)), the supply of bank loans matters for monetary policy in its own right.

Since [Bernanke and Blinder \(1988\)](#), researchers have identified many frictions that affect monetary transmission through banks. Our model incorporates the following prominent channels featured in the literature. First, access to non-deposit external financing is more

⁴In the Online Appendix further separate banks into two groups according to their capital ratios. We find that banks with below-median capital ratios experience much lower loan growth following the recession. The gap between poorly-capitalized and well-capitalized banks is particularly large in 2001 and 2008 recessions.

costly than taking deposits. This friction implies that shocks to the quantity of deposits are transmitted to the supply of bank loans, as banks cannot costlessly replace deposits with non-reservable borrowing. Second, competition in the deposit and loan markets is imperfect. With market power, banks strategically choose deposit and loan rates to maximize their profits. This profit maximizing behavior, in turn, determines how monetary policy transmits through the banking system. Third, banks are subject to reserve regulation and capital regulation. Reserve regulation links the opportunity cost of taking deposits to the prevailing Federal Funds. Capital regulation incentivizes banks to optimize their loan supply intertemporally with an eye to preserving excess equity capital as a buffer against future capital inadequacy.

4.1 Households

At each time t , the economy contains W_t households, each of which is endowed with one dollar. Hereafter, we drop the time subscript for convenience, so aggregate household wealth is then W . Households choose among the following investment options for their endowments: cash, corporate bonds, and bank deposits, where the deposits of each individual bank constitute a differentiated product. If we index each option by j , the households' choice set is given by $\mathcal{A}^d = \{0, 1, \dots, J, J + 1\}$, with option 0 representing cash, option $J + 1$ representing short-term bonds, and options $1, \dots, J$ representing deposits in each bank. We further assume that each depositor can choose only one option. This one-dollar, one-option assumption is without loss of generality. For example, we can interpret this setting as if households make multiple discrete choices for each dollar that they have, and the probability of choosing each of the options can be interpreted as portfolio weights.

Each option is characterized by a yield, r_j^d , and a quality value, q_j^d . The yield on cash is 0, and the yield on bonds is the Federal Funds rate, f . q_j^d captures the convenience of option j to a depositor. For a bank, q_j^d can reflect the bank's number of branches or the number of employees per branch. For the other options, q_j^d can capture the ease with which

the household can use the option as a medium of transaction. We assume q_j^d varies only with banks and not with households, so different households cannot ascribe a different quality value to the same option. We allow households to belong to a finite set of types, indexed by $i \in 1, 2, \dots, I$. A household's utility from choosing option j is given by $u_{i,j} = \alpha_i^d r_j^d + q_j^d + \epsilon_{i,j}^d$, where α_i^d is the yield sensitivity and $\epsilon_{i,j}^d$ is a relationship specific shock for the choice of option j by household i . We assume $\epsilon_{i,j}^d$ follows a generalized extreme value distribution with a cumulative distribution function given by $F(\epsilon) = \exp(-\exp(-\epsilon))$. This distributional assumption is standard in the structural industrial organization literature and allows for a closed-form solution for the consumer's probability of making each choice.

The decision of each household is then to choose the best option to maximize its utility:

$$\max_{j \in \mathcal{A}^d} u_{i,j} = \alpha_i^d r_j^d + q_j^d + \epsilon_{i,j}^d. \quad (1)$$

The solution to (1) implies that the demand for the deposits of bank j is given by the following formula:

$$s_j^d(r_j^d) = \sum_{i=1}^I \mu_i^d \frac{\exp(\alpha_i^d r_j^d + q_j^d)}{\sum_{m \in \mathcal{A}^d} \exp(\alpha_i^d r_m^d + q_m^d)}, \quad (2)$$

where μ_i^d is the fraction of total wealth (W) held by households of type i . The quantity of deposits is then given by the market share multiplied by total wealth, $D_j = s_j^d W$.

Firms

There are K firms, each of which wants to borrow one dollar, so aggregate borrowing demand is K . Firms can borrow by issuing long-term bonds or taking out bank loans. We assume that each individual bank is a differentiated lender, where this assumption is motivated by such factors as geographic location or industry expertise. Letting each option be indexed by j , the firms' choice set is given by $\mathcal{A}^l = \{0, 1, \dots, J, J+1\}$, where option 0 represents bonds, option $1, \dots, J$ represents loans from each bank, and option $J+1$ is the option not

to borrow at all.⁵ Each option is characterized by a lending rate, r_j^l , and a vector of product characteristics, x_j^l . As above, these characteristics can include the number of branches or employees per branch.

For tractability, we assume that both bonds and bank loans have the following repayment schedule. Each period the firm has to pay back a fraction, μ , of its outstanding debt, where μ includes both interest accrued over the last period, as well as some amortized principal. For instance, if the firm borrow a nominal value of one dollar, the repayment stream, starting in the next period, is $\mu, (1 - \mu)\mu, (1 - \mu)^2\mu, \dots$. Accordingly, all firm debt has an average maturity of $\frac{1}{\mu}$ periods. This construction saves us from having to track the entire age distribution of a bank's loan portfolio. Thus, a sufficient statistic for the future repayment schedule is just the total amount outstanding.

We assume that the interest rate on the long-term bond (\bar{f}_t) is set according to:

$$\sum_{n=0}^{\infty} \frac{\mu(1 - \mu)^n}{(1 + \bar{f}_t)^{n+1}} = \frac{\mu}{1 + \bar{f}_t} + \mathbb{E}_t \left[\sum_{n=1}^{\infty} \frac{\mu(1 - \mu)^n}{\prod_{m=0}^n (1 + f_{t+m})} \right] \quad (3)$$

Put differently, \bar{f}_t is the fixed rate that produces the bond's fair present value when used to discount the bond's cash flows. Fair present value, in turn, is simply this expected value of this cash flow stream discounted at the (non-constant) Federal Funds rate, f_t . Each of the firm's financing option's is characterized by a rate, r_j^l , and a quality value, q_j^l , which reflects the effort a firm must exert to borrow via option j . In case of a bank loan, q_j^l can include the number of branches or the number of employees per branch. In the case of the corporate bond, q_j^l can capture the cost of hiring an underwriter. We assume q_j^l varies only with banks and not with firms, where, as in the case of the households, we allow firms to belong to a finite set of types, indexed by $i \in 1, 2, \dots, I$.

The profit for firm i from choosing option j is given by $\pi_{i,j} = \alpha_i^l r_j^l + q_j^l + \epsilon_{i,j}^l$, where α_i^l is the yield sensitivity, and $\epsilon_{i,j}^l$ is an idiosyncratic relationship shock when a firm i borrows from

⁵Our index choices recycle the notation i to index firms and j to index the firm's borrowing options.

bank j . We assume $\epsilon_{i,j}^l$ follows a generalized extreme value distribution with a cumulative distribution function $F(\epsilon) = \exp(-\exp(-\epsilon))$. Each firm's decision is to choose the best option to maximize its profit, as follows:

$$\max_{j \in \mathcal{A}^l} \pi_{i,j} = \alpha_i^l r_j^l + q_j^l + \epsilon_{i,j}^l. \quad (4)$$

The solution to (4) implies that the demand for the loans of bank j is given by:

$$s_j^l(r_j^l) = \sum_{i=1}^I \mu_i^l \frac{\exp(\alpha_i^l r_j^l + q_j^l)}{\sum_{m \in \mathcal{A}^l} \exp(\alpha_i^l r_m^l + q_m^l)}, \quad (5)$$

where μ_i^l is the fraction of type i firms. Loan quantity is given by the market share multiplied by the total market size, $B_j = s_j^l K$.

4.2 The Banking Sector

Given the Federal Funds rate, f_t , each bank simultaneously sets its deposit rate, $r_{j,t}^d$, and its loan rate $r_{j,t}^l$, thereby implicitly choosing the quantities of deposits to take from households and credit to extend to firms. For example, given each bank j 's choice of $r_{j,t}^d$, households solve their utility maximization problem described above, which yields the quantity of deposits supplied to bank j , $D_j(r_{j,t}^d)$. Similarly, given each bank j 's choice of $r_{j,t}^l$, firms solve their profit maximization problem, which yields the quantity of loans borrowed from bank j , $B_j(r_{j,t}^l)$. To simplify notation, in what follows, we suppress the dependence of loans and deposits on the relevant interest rates, denoting them simply as D_t and B_t .

This lending activity involves a maturity transformation. Let $L_{i,t}$ denote the amount of loans that the bank currently holds. In each period, as in the case of bonds, a fraction, μ , of a bank's outstanding loans matures, with principle plus interest payments equal to $\mu \times L_{i,t}$. This assumption about long-term loans captures a traditional maturity transformation role for banks, in which they convert one-period deposits into long-term bank loans with maturity

$\frac{1}{\mu}$. As noted above, banks can also issue new loans with an annualized interest rate of $r_{j,t}^l$. The new loans, once issued, have the same maturity structure as the existing ones, with a fraction $\mu B(r_{j,t}^l)$ becoming due each year. As such, a bank's outstanding loans evolve according to:

$$L_{j,t+1} = (1 - \mu)(L_{j,t} + B_{j,t}), \quad (6)$$

We assume that a random fraction of loans, $\delta_t \in [0, 1]$, falls delinquent in each period, with delinquency occurring when a loan comes due but the borrower fails to pay. Although we assume that delinquent payments are written off by the bank, with charge-offs equal to $\mu L_t \times \delta_t$, defaulting on a payment in one period does not exonerate the borrower from payments in future periods. Therefore, delinquency does not affect the evolution of loans in (6).

We denote by $P(B_{j,t}, r_{j,t}^l)$ the amount bank j extends to firms. Fair pricing implies that this quantity is simply the discounted the payment stream at the constant loan rate $r_{j,t}^l$:

$$P(B_{j,t}, r_{j,t}^l) = \sum_{m=0}^{\infty} \frac{(1 - \mu)^m \mu B_{j,t}}{(1 + r_{j,t}^l)^{m+1}} = \frac{\mu}{r_{j,t}^l + \mu} B_{j,t} \quad (7)$$

Notice that if the loan rate $r_{j,t}$ is set to 0, then equation (7) simplifies to $P(B_{i,t}, r_{i,t}^l) = B_{i,t}$, showing that the amount the bank gives to firms at the present period simply equals the sum of all future payments. If μ is set to one, then debt has a maturity of one year, and the interest income from lending is $B_{j,t} - P_{j,t} = \frac{r_{j,t}^l}{1+r_{j,t}^l} B_{j,t} \approx r_{j,t}^l B_{j,t}$.

We summarize the rest of the bank's activities in the balance sheet given in Table 3, where we suppress the subscript for bank identity, j , for convenience. Here, we see that the banks assets consist of existing plus new loans, reserves, and holdings of government securities. Its liabilities consist of deposits, borrowing not subject to reserve requirements. The difference is then bank equity. We now go through these items in detail.

In each period, the bank can rely on deposits or internal retained earnings to finance its new loans, B_t . When the supply of funds falls short of loan demand, the bank can also

borrow via non-reservable securities, N_t . A typical example of non-reservable borrowing is large denomination CDs. As argued by [Kashyap and Stein \(1995\)](#), because non-reservable borrowing is not insured by FDIC deposit insurance, purchasers of this debt must concern themselves with the default risk of the issuing bank. These considerations make the marginal cost of non-reservable borrowing an increasing function of the amount raised, and motivate our next assumption, which is that non-reservable borrowing incurs a quadratic financing cost beyond the prevailing Federal Funds rate, as follows:

$$\Phi^N(N_t) = \phi^N N_t^2. \quad (8)$$

This assumption embodies two important frictions, the first of which is deposit insurance. Because a large fraction of deposits are insured, depositors are willing to accept a rate of interest that does not reflect default risk. The second friction is deadweight default losses, which affect the rate of return on non-reservable securities, which are subject to bank default risk.

Banks also incur operating costs, such as rents and wages. We assume that costs are linear in the amount of deposits:

$$\Phi^d(D_t) = \phi^d D_t. \quad (9)$$

Similarly, we also assume that lending activity itself incurs separate costs, such as the labor input necessary to screen loans or maintain client relationships. Again, we assume a linear functional form as follows:

$$\Phi^l(B_t) = \phi^l B_t. \quad (10)$$

If the total supply of funds exceeds the demand from the lending market, the bank can invest in government securities, G_t , where the return is the Federal Funds rate, f_t . The bank's holdings of loans, government securities, deposits, reserves, and non-reservable

borrowing must satisfy the standard condition that assets equal liabilities plus equity:

$$L_t + P(B_t, r_t) + R_t + G_t = D_t + N_t + E_t, \quad (11)$$

where R_t denotes bank reserves and E_t is the bank's begin-of-period book equity. E_t itself evolves according to:

$$E_{t+1} = E_t + \Pi_t \times (1 - \tau) - C_t \quad (12)$$

where τ denote the linear tax rate, and Π_t is the bank's total operating profit from its deposit taking, security investments, and lending decisions. This identity ends up being a central ingredient in the model, as it links bank competition, which is reflected in profits, with bank capital regulation.

The profits in (12) are in turn given by:

$$\Pi_t = B_t - P(B_t, r_t^l) - D_t \times r_t^d + G_t \times f_t - \Phi^l(B_t) - \Phi^d(D_t) - \Phi^N(N_t) - \mu L_t \times \delta_t. \quad (13)$$

Finally, C_t in equation (12) represents the cash dividends distributed to the bank's shareholders. We assume that a bank can only increase its inside equity via retained earnings, that is, there is no new equity issuance, so:

$$C_t \geq 0 \quad \forall t. \quad (14)$$

This constraint reflects a bank's limited liability, which prevents it from obtaining any external equity financing from shareholders. This constraint represents an important friction because in its absence, banks could always raise equity capital to fund any shortfall in the loanable funds market. This activity would disconnect banks' deposit taking decisions from their lending decisions, so changes in the Federal Funds rate would not have an impact on lending. Equation (14) implies that model cannot capture the equity issuances we see in the data. However, given that banks' equity issuances are both tiny and rare, we view this

drawback of our model as minor.

We now introduce the capital requirement and the reserve requirement:

$$E_{t+1} \geq \kappa \times L_{t+1} \quad (15)$$

$$R_t \geq \theta \times D \quad (16)$$

Equation (15) implies that the bank's book equity at the beginning of the next period has to be no smaller than a fraction, κ , of the loans outstanding. Equation (16) is the bank's reserve requirement, which says that the bank has to keep θ of its deposits in a non-interest bearing account with the central bank. Zero interest on reserves implies that the bank has no incentive to hold excess reserve, so equation (16) holds with equality.

To close the model, we assume that the law of motion for the bank loan charge-offs and the Federal Funds rate is given by:

$$\begin{bmatrix} \ln \delta_{t+1} \\ \ln f_{t+1} \end{bmatrix} = \begin{bmatrix} \rho_\delta & \rho_{\delta f} \\ 0 & \rho_f \end{bmatrix} \cdot \begin{bmatrix} \ln \delta_t \\ \ln f_t \end{bmatrix} + \begin{bmatrix} \sigma_\delta & 0 \\ 0 & \sigma_f \end{bmatrix} \cdot \mathcal{N}_2, \quad (17)$$

where \mathcal{N}_2 stands for the density function of a standard bi-variant normal distribution.

4.3 Bank's problem and equilibrium

Figure 7 summarizes the sequence of events in a typical time period. The bank enters the period and observes the Federal Funds rate, f_t , and the realization of the default fraction, δ_t . At that point, it takes the corresponding charge-offs. Next, banks interact with households and firms by setting the loan and deposit spreads, receiving the corresponding amount of deposits from households, and extending the corresponding amount of loans to firms. Depending on the extent of these activities, the banks adjust their reserves, holdings of government securities, and non-reservable borrowing. Finally, at the end of the period, a fraction μ of the loans matures, and banks collect profits and distribute dividends to shareholders.

As discussed above, loan and deposit demand depend on the rates put forth by all banks in the economy. Accordingly, when each bank chooses its own deposit and loan rates (r_t^d and r_t^l), as well as its non-reservable borrowings (N_t), and investment in government securities (G_t), it rationally takes into account the choices made by other banks in both the current and future periods. As such, all of a bank's optimal choices depend on the composition of the banking sector, that is, the cross-sectional distribution of bank states, which we denote by Γ_t . Letting P^Γ denote the probability law governing the evolution of Γ_t , we can express the evolution of Γ_t as:

$$\Gamma_{t+1} = P^\Gamma(\Gamma_t). \quad (18)$$

Every period, after observing the Federal Funds rate (f_t) and the random fraction of defaulted loans (δ_t), the banks choose the optimal policy to maximize its discounted cash dividends to shareholders:

$$V(f_t, \delta_t, L_t, E_t | \Gamma_t) = \max_{\{r_t^l, r_t^d, G_t, N_t, R_t\}} \left\{ C_t + \frac{1}{1 + \gamma} \mathbb{E}V(f_{t+1}, \delta_{t+1}, L_{t+1}, E_{t+1} | \Gamma_{t+1}) \right\} \quad (19)$$

s.t. (8), (9), (10), (11), (12), (18),

where γ is the bank's discount factor.

We define equilibrium in this economy as follows.

Definition 1 *A stationary equilibrium occurs when:*

1. *All banks solve the problem given by (19), taking as given the other banks' choices of loan and deposit rates.*
2. *All households and firms maximize their utilities given the list of rates put forth by banks.*
3. *Each period, the deposit and loan markets clear.*
4. *The probability law governing the evolution of the industry, P^Γ , is consistent with banks' optimal choices.*

One of the state variables for bank's problem (Γ_t) is an object whose dimension depends

on the number of banks in the economy. This dimensionality poses a challenge for numerically solving the banks' problem. To simplify the model solution, we follow [Krusell and Smith \(1998\)](#) by considering a low-dimensional approximation of Γ_t . Specifically, we postulate that all information about Γ_t relevant to banks' optimization can be summarized by the contemporaneous Federal Funds rate (f_t). Accordingly, we define the equilibrium "average" loan and deposit rates $\bar{r}_t^l(f_t)$ and $\bar{r}_t^d(f_t)$, respectively) as,

$$\exp(\alpha_i^d \bar{r}_t^d + q_i^d) \equiv \mathbb{E} \left[\exp(\alpha_i^d r^d + q_i^d) \right], \quad (20)$$

and

$$\exp(\alpha_i^l \bar{r}_t^l + q_i^l) \equiv \mathbb{E} \left[\exp(\alpha_i^l r^l + q_i^l) \right]. \quad (21)$$

With a reasonably large number of banks, $\bar{r}_t^l(f_t)$ and $\bar{r}_t^d(f_t)$ approximately summarizes the choices of other banks, thereby allowing each bank to derive its deposit and loan demand functions. In solving the model, we ensure that $\bar{r}_t^l(f_t)$ and $\bar{r}_t^d(f_t)$ are consistent with equilibrium bank choices by iterating over their values until convergence. As $\bar{r}_t^l(f_t)$ and $\bar{r}_t^d(f_t)$ are functions of f_t only, we drop Γ_t from banks' value functions in (19).

4.4 Monetary policy transmission

In this section, we use a simplified version of the model to illustrate the key transmission mechanisms.

4.4.1. Frictionless benchmark

First, we examine how the economy behaves in a frictionless benchmark model. By frictionless, we mean a simplified version of our model with the following five features: (1) the bank has no market power in either the deposit or the loan market, i.e., the deposit and loan demand elasticities are infinite; (2) there are no frictions related to non-reservable borrowing; (3) the bank faces no capital requirement, $\kappa = 0$; (4) there is no reserve requirement, $\theta = 0$;

and (5) there is no maturity transformation. These features imply that the bank's problem can be viewed as a static problem.

In this static frictionless model, banks choose deposit rates, r^d , and loan rates, r , to maximize one-period profits

$$\Pi = \max_{\{r^l, r^d\}} r^l B(r^l) - r^d D(r^d) - \phi^l B(r^l) - \phi^d D(r^d) - f(B(r^l) - D(r^d)). \quad (22)$$

When deposits fall short of loans, the bank can make up any funding shortfall, $B - D$, with non-reservable borrowing at a cost equal to the Federal Funds rate, f . There are no additional financing costs associated with non-reservable borrowing. When there are excess deposits, the bank can invest any this surplus, $D - B$, in government securities and earn the Federal Funds rate, f .⁶ In the absence of a balance sheet friction, the bank can optimize its choices for deposit and loan amounts (B and D) separately.

The optimal lending rates are given by the Federal Funds rate plus the marginal cost and the markup:

$$r^l = f + \phi^l + \left(-\frac{B'}{B}\right)^{-1},$$

and the optimal deposit rates are given by the Federal Funds rate minus the marginal cost and the markup:

$$r^d = f - \phi^d - \left(\frac{D'}{D}\right)^{-1}.$$

When there is perfect competition among banks, the demand elasticities, $-\frac{B'}{B}$ and $\frac{D'}{D}$, become infinite and the markups converge to zero. Deposit and lending rates converge to the Federal Funds rate minus or plus the marginal cost, as follows:

$$r^d \rightarrow f - \phi^d, \quad r^l \rightarrow f + \phi^l \quad (23)$$

⁶In reality, the Federal Funds rate is slightly higher than the risk-free Treasury yield because of the default risk of the bank. However, we assume there is no default by the bank, so these two rates are equal.

Under the frictionless benchmark, banks function as the bond market, passing through the interest rate changes to the exact same degree.

4.4.2. Imperfect competition

When competition is imperfect, market power creates a wedge between the Federal Funds rate and the rates at which banks borrow and lend. Monetary policy can affect the market power of banks by influencing the attractiveness of bank deposits or loans relative to other outside options available to households or firms.

In the model, bank deposits face competition from both bonds and cash. Investors get higher returns from investing in bonds but endure low liquidity. On the contrary, cash holdings offer high liquidity but zero return. Bank deposits are somewhere in between, offering investors both liquidity and some non-zero return. When the interest rate is high, the opportunity cost of holding cash increases and investors move out of cash and into bank deposits and bonds. Consequently, banks enjoy an outward shift in their deposit demand function and they can charge larger markup on deposits (e.g., [Drechsler, Savov, and Schnabl 2017](#)), so:

$$\frac{\partial \left(\frac{D'}{D}\right)^{-1}}{\partial f} > 0. \quad (24)$$

In the lending market, an increase in the Federal Funds rate makes bank loans less attractive to firms relative to the outside option of not borrowing. Therefore, total lending shrinks and banks optimally lower the markups they on loans to mitigate the effect of lower lending demand.

$$\frac{\partial \left(-\frac{B'}{B}\right)^{-1}}{\partial f} < 0 \quad (25)$$

4.4.3. Balance sheet frictions

In the frictionless benchmark, the deposit market and loan market are entirely separable because the bank can costlessly use non-reservable borrowing and government securities as

buffers. Now consider the case in which banks face balance sheet frictions, so they incur additional costs when using non-reservable borrowing. In this case, the banks' optimization problem becomes:

$$\Pi = \max_{\{r^l, r^d\}} r^l B(r^l) - r^d D(r^d) - \phi^l B(r^l) - \phi^d D(r^d) - f(B(r^l) - D(r^d)) - \Phi(N),$$

where $\Phi(N)$ is the cost of non-reservable borrowing or investing excess funding and $N = B - D$ is the funding imbalance. In the presence of balance sheet frictions, the bank cannot costlessly replace any lost deposits with wholesale borrowing. Therefore, shocks to deposits will be transmitted to loans.

4.4.4. Reserve requirement

Now consider the case in which banks face reserve regulation that requires that for every dollar of deposits, the bank needs to keep a fraction, θ , of these deposits as reserves. Assuming that the interest on reserves is zero, banks' optimization becomes

$$\Pi = \max_{\{r^l, r^d\}} r^l B(r^l) - r^d D(r^d) - \phi^l B(r^l) - \phi^d D(r^d) - f(B(r^l) + R - D(r^d))$$

$$s.t. R \geq \theta D(r^d)$$

Because the interest rate on reserves is zero, the reserve constraint is binding. We can solve for the optimal deposit rate as

$$r^d = f - \phi^d - \left(\frac{D'}{D}\right)^{-1} - \theta f. \quad (26)$$

Here we see that higher Federal Funds rates increase the opportunity cost of holding reserves, θf , which lowers deposit rates. Lower deposit rates, in turn, reduce the quantity of deposits and optimal supply of loans.

4.4.5. Capital regulation

Now consider the case in which banks face capital regulation that requires bank capital to exceed a certain fraction of bank assets. In this case, the banks' optimization problem becomes:

$$\Pi = \max_{\{r^l, r^d\}} r^l B(r^l) - r^d D(r^d) - \phi^l B(r^l) - \phi^d D(r^d) - f(B(r^l) - D(r^d)),$$

$$s.t. E_0 + (1 - \tau_c)\Pi \geq \kappa B(r^l).$$

In the presence of capital regulation, shocks to bank capital affect lending capacity. One way that monetary policy affects bank capital is through maturity mismatch. Because deposits are short-term, an increase in the Federal Funds rate raises the rate that the bank has to pay on all deposits. However, loans are long-term, so only a fraction of loans matures, with the remaining outstanding loans commanding a lower rate. Hence, an increase in the Federal Funds rate temporarily reduces bank capital and tightens the bank capital constraint in equation (15).

Another way that monetary policy affects bank capital is through market power. When the Federal Funds rate increases, bank profits from the deposit market increase, as the competition from cash lessens, but bank profits from the loan market decrease as borrowers are less willing to borrow at higher rates. The effect from the deposit market is likely to dominate the loan market effect when rates are very low, and especially, when rates are near the zero lower bound, leading to a tighter capital constraint in this region.

5. Estimation

5.1 Estimation procedure

We divide the estimation procedure into two stages. In the first stage, we estimate the demand elasticities and liquidity values for deposits and loans following [Berry, Levinsohn, and Pakes \(1995\)](#). In the second stage, we estimate the remaining parameters describing banks' balance sheet frictions using simulated method of moments (SMM).

We first parameterize the quality of each option as a function of characteristics, $q_j^k = \beta^k x_j + \xi_j^k$, $k = d, l$, where x_j is a vector of bank characteristics that includes the number of branches, the number of employees per branch, banks fixed effects, and time fixed effects. ξ_j^k is an unobservable demand shock associated with product j . Next, we allow for heterogeneity in rate sensitivity in the deposit market, so each individual depositor's rate sensitivity can be written as a mean rate sensitivity and a deviation from the mean, $\alpha_i = \alpha + \sigma_\alpha v_i$, where v_i follows a uniform distribution. We assume the sensitivity to non-rate characteristics is homogeneous across both borrowers and depositors. Taken together, the demand functions for bank deposits and loans are characterized by the preference parameters, $(\alpha^k, \beta^k, \sigma_\alpha^k)$, where $k = d, l$. For simplicity, we drop the superscript indicating the deposit or loan market in the following discussion.

We use the methods from [Berry, Levinsohn, and Pakes \(1995\)](#) to estimate the demand parameters, $\Theta = (\alpha, \beta, \sigma_\alpha)$. First, we divide the parameters into linear parameters, (α, β) , and non-linear parameters, σ_α . Second, for a given value of σ_α , there is a relation between mean utility, $E[u_{ij}] = u_j = \alpha r_j + \beta x_j + \xi_j$, and the observed market share, s_0 , given by $s(u|\sigma_\alpha) = s_0$, where $s(\cdot)$ is the market share, which is defined by equations (2) and (5). Third, we solve for the implicit function, $s^{-1}(\cdot)$, using the nested fixed-point algorithm described in [Nevo \(2001\)](#). Fourth, using this implicit function, we can solve for the unobservable demand shocks as functions of the observable market share and the demand parameters, as

follows:

$$\xi(\Theta) = s^{-1}(s_0|\sigma_\alpha) - (\alpha r_j + \beta x_j)$$

A key challenge in identifying the demand parameters is the natural correlation between deposit rates and unobservable demand shocks, ξ_j . Following the industrial organization literature (Nevo 2001), we use a set of supply shifters, c_j , as instrumental variables. Our particular instruments are bank salaries and non-interest expenses related to the use of fixed assets. Our identifying assumption is that these supply shifters are orthogonal to unobservable demand shocks and thus shift the supply curve along the demand curve, allowing us to trace out the slope of the demand curve.

Formally, define $Z = [x, c]$, where x is a vector of bank characteristics and c is a vector of supply shifters. The moment condition for this estimation is the orthogonality condition between the unobservable demand shocks, ξ_j and the exogenous variables, z_j , as follows:

$$\mathbb{E}[\xi_j z_j] = 0.$$

Define W as a consistent estimate of $\mathbb{E}[Z'\xi\xi'Z]$. The GMM estimator of the demand parameters is then given by:

$$\hat{\Theta} = \arg \min_{\Theta} \xi(\Theta)' Z' W^{-1} Z \xi(\Theta).$$

The data used for the deposit demand estimation include each bank's deposit market share, the proportions of cash and bonds in the household portfolio, non-rate bank characteristics such as the number of branches and the number of employees per branch, and each bank's deposit rate, where we use a weighted average of deposit rates for different types of deposits, where the weights are the relative quantities of each deposit type. We also include bank and time fixed effects. The data used for the loan demand estimation include loan market shares for each bank, the corporate bonds market share, each bank's lending rate,

and any non-rate characteristics.

We then plug the first-stage estimates of the deposit and loan demand functions into our model of the banking sector in Section 4.2 for the second stage SMM estimation. For this stage, we have to make one further simplification. Our data contain a large number of very small banks. This feature of the data poses a challenge because solving a model with a number of banks equal to the number of banks in our data would be intractable. Therefore, we solve for an equilibrium with \hat{J} ex ante symmetric representative banks, where the number of representative banks \hat{J} is calibrated to match the HHI in the data. Because the size distribution has a heavy left tail, this approach substantially reduces the number of banks in the model while keeping the market concentration similar to that in the data. An alternative approach is to limit the sample to the largest banks. Our results are robust to the alternative approach.

Next, because we have a large number of non-rate bank characteristics that enter linearly in the utility function, we summarize these non-rate characteristics with a composite index that we interpret as “quality.” For simplicity, we assume quality is the same for all the representative banks. The quality value can be calculated as follows:⁷

$$\hat{q} = \log \left(\frac{1}{\hat{J}} \sum_{j=1}^{\hat{J}} \exp(x_j \hat{\beta}) \right). \quad (27)$$

With the quality value in hand, we parameterize the deposit and loan demand functions as:

$$D_j(r_j^d | f) = \sum_{i=1}^I \mu_i^d \frac{\exp((\hat{\alpha}^d + \hat{\sigma}_{\alpha}^d v_i) r_j^d + \hat{q}_j^d)}{\sum_{m \in \mathcal{A}} \exp((\hat{\alpha}^d + \hat{\sigma}_{\alpha}^d v_i) r_m^d + \hat{q}_m^d)} W \quad (28)$$

$$B_j(r_j^l | f) = \sum_{i=1}^I \mu_i^l \frac{\exp((\hat{\alpha}^l + \hat{\sigma}_{\alpha}^l v_i) r_j^l + \hat{q}_j^l)}{\sum_{m \in \mathcal{A}} \exp((\hat{\alpha}^l + \hat{\sigma}_{\alpha}^l v_i) r_m^l + \hat{q}_m^l)} K, \quad (29)$$

⁷Alternatively, we can directly use the characteristics of the N largest banks in the data and ignore the smaller banks. This alternative approach biases the HHI in the model upwards, but the magnitude of the bias becomes quite small if N is larger than 10. The small bias occurs because the 10 largest banks in the United States account for a disproportionately large market share, so ignoring the small banks has a limited impact on the market equilibrium in the national market.

in which $\hat{\alpha}$ and $\hat{\sigma}_\alpha$ are the mean and standard deviation of the deposit or loan rate sensitivities, which are estimated from the first stage. Correspondingly, v_i and μ_i , $i = 1, 2, \dots, I$ are a discrete approximation of a uniform distribution, and q_j is the quality value associated with option $j \in 0, 1, \dots, J + 1$. In the deposit market, we normalize the quality value of cash to zero, and we denote by q_d^d and q_b^d the quality values of bank deposits and short-term bonds, respectively. In the loan market, we normalize the quality of bonds to zero, and we denote by q_l^l and q_n^l the quality values of loans and not borrowing, respectively. Note that the quality value of not borrowing cannot be estimated from the demand estimation because we do not observe its share. Therefore, we relegate this parameter to SMM.

The final plug-in problem consists of inserting (28) and (29) into the definition of bank profits given by (13) before solving and simulating the bank model for the SMM estimation. This plug-in problem operationalizes the notion that banks set deposit and loan rates facing the demand for deposits and loans, banks set deposit and loans rates. It is important to note that, the pricing decision is dynamic because deposits are short-term, while loans are long-term.

In the second stage, we estimate five additional parameters using simulated method of moments (SMM), which chooses parameter values that minimize the distance between the moments generated by the model and their analogs in the data. We use eight moments to identify the remaining five model parameters. Parameter identification in SMM requires choosing moments whose predicted values are sensitive to the model's underlying parameters. Our identification strategy ensures that there is a unique parameter vector that makes the model fit the data as closely as possible.

First, we use banks' average non-reservable borrowing as a fraction of their deposits to identify the cost of holding non-reservables (ϕ^N). Intuitively, larger financing costs induce banks to finance loans mainly through deposits, and less via borrowing. Next, we use the average deposit and loan spreads to identify banks' marginal costs of generating deposits (ϕ^d) and servicing loans (ϕ^l). Deposit spreads are defined as the difference between the Federal

Funds rate and deposit rates, while loan spreads are the difference between loan rates and Treasury yields matched by maturity. In our model, banks with market power optimally choose to pass a fraction of their operating costs onto the depositors and borrowers. Hence, higher operating costs lead to monotonically higher spreads that banks charge in the deposit and lending markets. In addition, banks' market power also determines the fraction of banks' marginal costs that get passed onto customers. Market power depends critically the Federal Funds rate, as the attractiveness to households of alternative investments, as such cash and long-term bonds, changes with the Federal Funds rate. Therefore, we also include the correlation between the Federal Funds rate and both loan and deposit spreads to ensure that our model captures this important mechanism. Next, we use banks' average dividend yield to identify the discount rate, γ . Intuitively, a high discount rate makes the banks impatient, so they pay out a larger fraction of their profit to shareholders instead of retaining it to finance future business. Finally, to identify the value of firms' outside option of not borrowing, q_n^l , we include banks' average loan to deposit ratio and the sensitivity of total corporate borrowing to the Federal Funds rate. These two moments suit this purpose because when the outside option becomes less valuable, its market share remains low regardless of the current Federal Funds rate. Thus, the sensitivity of the aggregate corporate borrowing to the Federal Funds rate should fall as q_n^l falls. In addition, a high loan-to-deposit ratio should be inversely related to q_n^l because when aggregate borrowing from the corporate sector is high, bank loans face proportionally higher demands.

5.2 Estimation Results

Table 4 presents the point estimates for the 22 model parameters. In Panel A, we start with the parameters that we can directly quantify in the data. Specifically, we set the corporate tax rate to its statutory rate of 35%. Capital regulation stipulates that banks keep no less than 6% of their loans as book equity. Reserve regulation requires a 10% reserve ratio for transaction deposits, 1% for saving deposits, and 0% otherwise. In our model, we only

have one type of deposit, so our estimate of the deposit ratio is a weighted average of these three requirements, where the weights are the shares of a particular type of deposit in total deposits. We model the Federal Funds rate and the bank-level loan default rate as log AR(1) processes, and we directly calculate their means, standard deviations, and autocorrelations from the data. Finally, we set the maturity of loans in our model to average loan maturity in the data, which is approximately five years.

Panel B in Table 4 presents the demand parameters from the first stage BLP estimation.⁸ Not surprisingly, we find that depositors react favorably to high deposit rates while borrowers react negatively to high loan rates. Both yield sensitivities are precisely estimated, and the economic magnitudes are significant as well. A 1% increase in the deposit rate increases the market share of a bank by approximately 0.8%, while a 1% increase in the loan rates decreases bank market share by 0.9%. We also find that depositors exhibit significant dispersion in their rate sensitivity. Finally, we estimate depositors' and borrowers' sensitivities to non-rate bank characteristics such as the number of branches and the number of employees per branch. The estimates are also both statistically and economically significant. A 1% increase in the number of branches increases bank market share by 0.868% in the deposit market and 1.117% in the loan market. In comparison, the sensitivity to the number of employees per branch is smaller. A 1% increase in the number of employees per branch increases bank market share by 0.587% in the deposit market and 0.694% in the lending market. Using these estimates combined with bank and time fixed effects, we can calculate the implied quality parameter using (27), which we plug into the second-stage SMM estimation.

Panel C in Table 4 presents the balance sheet parameters from our second stage SMM estimation. We find that banks have a subjective discount rate of 5%, which is only slightly higher than the average Federal Funds rate observed in the data. Given the discount rate, banks pay out 2.6% of their equity value as dividends. We also find the cost of non-reservable borrowing both statistically and economically significant. At the average amount of non-

⁸Detailed estimation results are presented in Table 6.

reservable borrowing (30%), a marginal dollar of non-reservable borrowing costs the bank 0.6 cents above the cost implied by the prevailing Federal Funds rate. Note that the average deposit spread is 0.8%. Because banks equate the marginal costs of their funding sources, these numbers imply the marginal cost of expanding deposit averages 1.4%, suggesting a large role for deposit market power.⁹ This result implies that banks cannot easily replace deposits with other funding sources. Therefore, shocks to bank deposits are likely to be transmitted to bank lending. Finally, we find that banks incur a 0.7% cost of maintaining deposits and a 0.9% cost of servicing their outstanding loans.

Table 5 compares the empirical and model-implied moments. The model is able to match closely the banks' market shares, average spreads, and the sensitivity of the Federal Funds rate to the deposit spread. In both the data and the model, banks borrow non-reservable securities, which amount to 30% of their total deposit intake. The spread that banks charge in the deposit market is significantly lower than the spread they receive in the loan market. This result arises because as the Federal Funds rate approaches zero, bank deposits face increasing competition from cash. Thus, banks market power falls and deposit spreads become compressed. A 1% decrease in the Federal Funds rate decreases the deposit spread by 30 basis points. However, as the Federal Funds rate falls lower, loan demand rises. A 1% decrease in the Federal Funds rate increases loan spreads by 25 basis points. These results are consistent with our intuition and the data.

6. Counterfactuals

6.1 Decomposing Monetary Policy Transmission

Now we examine the quantitative forces that shape the relation between monetary policy, as embodied in changes in the Federal Funds rate, and aggregate bank lending. To this end, we start with the baseline model in row (1) of Table 7, where we see that on average in our

⁹The marginal cost of non-reservable borrowing is $\frac{\partial \Phi}{\partial N} = 2\phi^N N = 2 \times 0.05 \times 0.3 = 0.03$.

model, a one percent change in the Federal Funds rate translates into a 3.88% decrease in aggregate lending. In the column labeled “Aggregate Lending,” we have normalized lending to 100% for this baseline case.

We proceed by eliminating the banks’ local market power and the regulatory constraints one by one to examine the cumulative effect of removing these frictions. As such, we analyze how the absence of each model ingredient influences aggregate lending in the economy and the transmission of the Federal Reserve’s monetary policy.

Row (2) presents the results from a version of the model without a reserve requirement. The nearly identical results imply that the reserve regulation has a minimal effect on banks’ lending decisions.

In row (3), we remove from the model banks’ market power in the deposit market. In this case, banks receive a fixed lump-sum profit equal to their oligopolist profit in the baseline case, and they use marginal cost pricing for deposit-intake decisions. Namely, they set the deposit rate equal to the Federal Funds rate minus the bank’s marginal cost of servicing deposits, and they take as many deposits as depositors offer, given that deposit rate. If deposit market power is in place, when the Federal Funds rate increases, banks enjoy higher market power in the deposit market because consumers place less value on cash as an investment option. Put differently, the competitiveness of cash relative to deposits falls. Banks react by charging higher deposit spreads and consequently lowering the amount of deposit intake. Banks’ lending decisions partially echo this decline in deposit in-take because when the amount of lending exceeds deposits, banks need to use expensive non-reservable borrowing to finance their loans. Thus, the bank’s market power, combined with the non-reservable borrowing cost, contribute to a negative relation between banks’ lending and the Federal Funds rate. Our results confirm this intuition. Once we eliminate market power in the deposit market, bank lending becomes less sensitive to change in the Federal Funds rate. A 1% increase in the Federal Funds rate causes a 2.57% decrease in aggregate lending. This sensitivity is 30% smaller than the 3.81% sensitivity observed in the baseline case. Finally,

this result is important in that it highlights the interconnectedness of banks' deposit taking and lending businesses. Banks' market power in the deposit market gets passed on to the loan market and contributes to the sensitivity of bank lending to the Federal Funds rate.

Does the above result imply that bank market power plays a quantitative important monetary policy transmission, relative to other candidate transmission mechanisms? To answer this question, we compare the magnitudes in line (3) to the effects of relaxing the capital requirement, the results of which are in line (4). We find that without a capital requirement, the sensitivity of bank loans to the Federal Funds rate drops sharply, with a change in the Federal Funds rate translating almost one-to-one into a change in bank loans. A comparison of the results in lines (3) and (4) shows that the capital requirement enhances monetary policy transmission by 1.62% ($2.57\% - 0.95\%$). This effect is roughly 25% larger than the effect of deposit market power. Of course, this 1.62% magnitude captures both the effects of the capital requirement itself and also of its interaction with bank market power. More specifically, when the Federal Funds rate goes up, bank capital takes a hit because of the maturity mismatch. The profit resulting from bank's market power also changes, thus interacting with the capital requirement.

Finally, we turn to banks' market power in the lending market. Our results show that having market power in the lending market has a large impact on the quantity of aggregate lending. The magnitude is consistent with the sizable spread that banks charge. In both the data and our baseline model, banks charge an average loan spread of 2.7%, which is significantly higher than the expected default cost plus the marginal cost of servicing loans. Once the banks switch to marginal cost pricing, the aggregate amount of lending in the economy goes up by 34%, from 128% of the baseline to 193% of the baseline. At the same time, removing banks' loan market power also makes the aggregate quantity of loans more sensitive to the Federal Funds rate. This sensitivity goes up from -0.95% in the case in which loan market power is present to -1.38% in the case in which banks use marginal cost pricing in both the deposit and loan markets.

6.2 Reversal Rate

In our previous analysis, we did not break down the effects of changes in the Federal Funds rate as a function of the rate level. We now turn to this question in Figure 5, which shows the amount of bank lending for different levels of the Federal Funds rate. Overall, and as expected, there is a negative relationship between the Federal Funds rate and the amount of lending. However, the negative relation is reversed in the region where the Federal Funds rate is close to zero.

To understand this pattern, we also plot in Figure 5 the level of bank capital and the optimal amount of bank lending in a world with no capital requirements. We find that aggregate bank capital in the economy is inversely U-shaped. When the Federal Funds rate is above the 2.27% threshold, an increase in the Federal Funds rate has the usual effect of tightening lending. However, when the Federal Funds rate is below the 2.27% threshold, an increase in the Federal Funds rate actually has the opposite effect of expanding lending. We call the region in which the Federal Funds rate is below 2.27% a “reversal rate” environment.

The pattern in bank capital is central to answering the question of why monetary policy, as embodied in changes in the Federal Funds rate, has opposite signs on the two sides of the threshold. To understand this connection, note that optimal lending is the smaller of two quantities: desired lending and feasible lending. The former is the optimal amount of lending in the absence of a capital requirement, and the latter is maximal lending permitted by bank’s equity capital. In equilibrium, desired lending is always decreasing in the Federal Funds rate, as high funding costs deter firms from borrowing. Hence, in high-rate regions, the capital requirement is slack, and the actual quantity of lending is the desired amount. On the other hand, when the Federal Funds rate is low, desired lending exceeds that allowed by the bank’s equity. Thus, the capital requirement binds, and the actual lending tracks the bank’s equity capital, which increases in the Federal Funds rate. Figure 5 confirms this intuition. When the Federal Funds rate is above 2.27%, actual lending and desired lending

closely track each other, whereas when the Federal Funds rate is below this threshold, actual lending falls short of the desired quantity.

The excess of desired over capital-constrained lending makes sense given firms' equilibrium high demand for loans in a low interest rate environment. However, the question remains of the forces behind the positive relation between bank equity and the Federal Funds rate when the latter is low. This result stems from the relative magnitudes of profits from lending and deposit taking. First, changes in the Federal Funds rate have opposite effects on bank profits in the deposit and lending markets. When the Federal Funds rate is high, holding cash is highly unattractive from the depositors' point of view, and banks face consistent weak competition from cash in the deposit market. Hence, bank profit from the deposit market increases in the Federal Funds rate. In contrast, bank profit from lending decreases in the Federal Funds rate, as higher funding costs makes the firms' outside option of not investing more appealing. Our parameter estimates imply that when the Federal Funds rate is low, the effect on profits from the deposit market dominate the effects from the lending market. Thus, an increase in the Federal Funds rate leads to higher bank profits, which in turn feed into the equity capital base, as banks find it optimal to shore up capital to deflect future financing costs by not paying out all of their profits to shareholders.

To understand more fully the dynamic response of bank lending to monetary policy shocks, in Figure 6, we simulate the response of bank lending to a shock to the Federal Funds rate. The economy starts at time zero in an initial steady state with the Federal Funds rate equal to the inflection point of 2.27%. At time one, the Federal Funds rate either increases or decreases by two standard deviations, and it stays at that level afterwards until the economy reaches a new steady state. Each variable in the graph is scaled by its level in the old steady state, that is, when the Federal Funds rate is 2.27%.

The top panel depicts the response to an increase in the Federal Funds rate. In this case, banks faces less competition from households' demand for cash in the deposit market. Thus, they can behave more like monopolists by charging higher spreads and cutting the amount of

deposits. Lower deposit intake increases the banks' marginal cost of lending because when their lending exceeds their capital plus deposit intake, they must turn to the market for non-reservable borrowing, in which they face increasing marginal external financing costs. A positive shock to the Federal Funds rate also increases the cost of capital in corporate sector, making firms more likely to switch to the outside option of not borrowing. Both effects shrink the amount of lending. Because deposits have shorter duration than loans, deposits drop sharply and converge almost instantaneously to the new steady state. Non-reservable borrowing increases to fill the gap between deposits and loans. In contrast, loan quantity first overshoots and then reverts slowly back to the new steady state. Intuitively, the bank's equity takes a hit with a positive shock to the Federal Funds rate because of the maturity mismatch on its balance sheet. It takes time for the bank to restore its capital stock by retaining profits. Loan quantity converges even more slowly as the bank only replaces a fraction, μ , of its long-term loans each period.

The bottom panel depicts the response to a decrease in the Federal Funds rate. As it approaches the zero lower bound, banks face increasingly intense competition from cash in the deposit market. As a result, the spread that banks can charge in the deposit market is squeezed, leading to a sharp drop in banks' profits. Given the high persistence in the Federal Funds rate, this lower profit translates into slower retained earnings accumulation over time and leads to decreased bank capital. In the new steady state, banks take large deposits in the deposit market, which can support increased lending. However, banks cannot lend more because their capital requirements tighten in the extremely low Federal Funds rate environment. Because total lending decreases, the banks face less of a need to seek external financing, and they use less non-reservable borrowing. As before, loan quantity initially overshoots because the longer maturity of bank loans induces a temporary relaxation in the capital requirement.

It is interesting to note that in the two graphs in Figure 6, the loan amount decreases when the Federal Funds rate changes in either direction. Although the loan amount moves

in the same direction, the driving force is different in the two cases. When the Federal Funds rate increases, loans fall because higher spreads in the deposit market discourage households from making deposits. Banks turn to non-reservable borrowing to fund loans, and because of increasing costs in this market, the amount of lending is highly dependent on the quantity of deposits. Instead, when the Federal Funds rate decreases, the loan amount decreases because of the binding capital requirement, which in turn echoes changes in the banks' profit accumulation. The distinct driving forces underlying the above two plots are also reflected by the differential trends in banks' deposit quantity and non-reservable borrowing.

6.3 Heterogeneous Monetary Transmission across Regions

There are significant cross-region variations in banking market concentration in the U.S. In 2017, the 10th percentile MSA has 2 banks while the 90th MSA has 238 banks. Such variations in market concentration may rise to heterogeneous monetary transmission across regions.

The effect of market concentration on monetary transmission may be quite complicated because of two opposing forces. On one hand, [Drechsler, Savov, and Schnabl \(2017\)](#) argue that an increase in market concentration in the deposit market enhances monetary transmission. On the other hand, [Scharfstein and Sunderam \(2016\)](#) argue that an increase in market concentration in the loan market dampens monetary transmission. The net effect of these two forces, however, is unclear.

Our structural model is able to speak to this issue. In Table 8, we consider three scenarios with different level of market concentration. We first calculate the overall effect of monetary policy for each scenario and then decompose the overall effect by each channel of transmission. There are two takeaways from this table. First, when market concentration increases, monetary policy transmission becomes stronger because the enhancing effect of the deposit market dominates the dampening effect from the lending market. Second, when the market becomes more concentrated, the bank capital channel becomes weaker because

banks are able to charge higher spreads in both the deposit and the lending markets, which makes capital regulation less binding.

6.4 Heterogeneous Monetary Transmission across Banks

Monetary transmission may also be heterogeneous across banks. [Kashyap and Stein \(1995\)](#) find that the impact of monetary policy on lending is stronger for small banks. They suggest that small banks are not able to frictionlessly raise wholesale funding to replace deposits. Therefore, shocks to deposits from monetary tightening are more likely to be transmitted to the supply of bank loans. However, [Kashyap and Stein \(1995\)](#) face a major challenge which is they do not directly measure the cost of external financing. They only have the size of the bank as a proxy which may affect many other things at the same time. This opens up alternative interpretations of their results. For instance, some may argue that small banks lend to small firms, whose credit demand is more cyclical. Therefore, the larger sensitivity of smaller bank lending is driven by the demand side rather than the financing friction in the supply side.

Our structural model sheds light on this issue because we are able to directly estimate the cost of external financing for each type of banks. Specifically, we re-estimate the balance sheet parameters for large and small banks separately using the data moments for each subsample. The results are reported in Table 9. We find that the largest difference between large versus small banks lies in their external financing cost. Note that we do not use bank size to estimate this parameter. Instead, this parameter is identified off the fraction of assets financed by nonreservables. This additional data moment lends support to the [Kashyap and Stein \(1995\)](#) argument that large and small banks mainly differ in the dimension of external financing costs. In addition, in the simulation exercise, we are also able to hold the loan demand constant to isolate the effect of the financing friction, which is hard to do with reduced-form approaches. Overall, we find that lending in small banks is more sensitive to the FFR. This is consistent with the hypothesis proposed by [Kashyap and Stein \(1995\)](#)

that different external financing costs across large and small banks leads to heterogeneous transmission of monetary policy to credit supply.

7. Conclusion

The U.S. banking sector has experienced an enormous amount of consolidation. The market share of the top five banks has increased from less than 15% in the 1990s to over 45% as of 2017. This consolidation begs the question of whether bank market power has a quantitatively important effect on the transmission of monetary policy. We study this question by formulating and estimating a dynamic banking model with regulatory constraints, financial frictions, and imperfect competition. This unified framework is useful because it allows us to gauge the relative importance of different monetary policy transmission channels.

In our counterfactuals, we show that the channel related to reserve requirements has minor quantitative importance. In contrast, we find that channels related to bank capital requirements and to market power are very important. We also find an interesting interaction between the market power channel and the bank capital channel. If the Federal Funds rate is low, depressing it further can actually contract bank lending, as the drop in bank profits in the deposit market has a negative impact on bank capital. Lastly, we show that financial frictions on banks' balance sheet play an important role in transmitting shocks from bank deposits to bank loans.

Our work contributes to the historical debate between the “money view” and the “lending view” of banking monetary policy transmission ([Romer and Romer 1990](#); [Kashyap and Stein 1995](#)). The “money view” postulates that the quantity of deposits matters for economic activity as a medium of exchange and that monetary policy influences deposit quantity through bank reserves. While our results do not negate the premise that the quantity of deposits matters, we show that the channel through which monetary policy affects bank deposits is increasingly through the market power channel rather than the reserve channel,

at least after the 1990s.

An alternative view of monetary transmission is the “lending view”, which rests on the idea that monetary policy also has a separate effect on the supply of loans by influencing the quantity of deposits. [Romer and Romer \(1990\)](#) argue that the “lending view” is unlikely to be important because banks can always easily replace deposits with external financing. In contrast, [Kashyap and Stein \(1995\)](#) argues that the banks can face costly external financing. Our study sheds new light on this debate, as the structural estimation approach allows us to infer the degree of the bank financing costs from the relative size of their non-reservable borrowing and deposit taking. We find the magnitude of this cost is economically significant and frictions related to bank balance sheets play an important role in the transmission of monetary policy.

References

- Bazdresch, Santiago, R. Jay Kahn, and Toni M. Whited. 2018. “Estimating and Testing Dynamic Corporate Finance Models.” *Review of Financial Studies* 31 (1):322–361.
- Bernanke, Ben S. and Alan S. Blinder. 1988. “Credit, Money, and Aggregate Demand.” *American Economic Review* 78 (2):435–439.
- Berry, Steven, James Levinsohn, and Ariel Pakes. 1995. “Automobile Prices in Market Equilibrium.” *Econometrica* 63 (4):841–90.
- Bolton, Patrick and Xavier Freixas. 2000. “Equity, bonds, and bank debt: Capital structure and financial market equilibrium under asymmetric information.” *Journal of Political Economy* 108 (2):324–351.
- Brunnermeier, Markus K. and Yann Koby. 2016. “The reversal interest rate: An effective lower bound on monetary policy effective lower bound on monetary policy.” Manuscript, Princeton University.
- Brunnermeier, Markus K. and Yuliy Sannikov. 2016. “The I theory of money.” Manuscript, Princeton University.
- Cieslak, Anna and Annette Vissing-Jorgensen. 2017. “The economics of the Fed put.” Manuscript, Duke University.
- Drechsler, Itamar, Alexi Savov, and Philipp Schnabl. 2017. “The deposits channel of monetary policy.” *Quarterly Journal of Economics* 132 (4):1819–1876.
- Egan, Mark, Ali Hortacsu, and Gregor Matvos. 2017. “Deposit Competition and Financial Fragility: Evidence from the U.S. Banking Sector.” *American Economic Review* 107 (1):169–216.
- Hanson, Samuel G. and Jeremy C. Stein. 2015. “Monetary policy and long-term real rates.” *Journal of Financial Economics* 115 (3):429–448.
- Hennessy, Christopher A. and Toni M. Whited. 2005. “Debt dynamics.” *Journal of Finance* 60 (3):1129–1165.
- Ho, Katherine and Joy Ishii. 2011. “Location and competition in retail banking.” *International Journal of Industrial Organization* 29 (5):537–546.
- Kashyap, Anil K. and Jeremy C. Stein. 1995. “The impact of monetary policy on bank balance sheets.” In *Carnegie-Rochester Conference Series on Public Policy*, vol. 42. Elsevier, 151–195.
- Krusell, Per and Anthony A. Smith. 1998. “Income and Wealth Heterogeneity in the Macroeconomy.” *Journal of Political Economy* 106 (5):867–896.

- Nevo, Aviv. 2001. “Measuring Market Power in the Ready-to-Eat Cereal Industry.” *Econometrica* 69 (2):307–342.
- Romer, Christina D. and David H. Romer. 1990. “New evidence on the monetary transmission mechanism.” *Brookings Papers on Economic Activity* 1:149–213.
- Scharfstein, David and Adi Sunderam. 2016. “Market power in mortgage lending and the transmission of monetary policy.” *Working paper, Harvard Business School* .
- Van den Heuvel, Skander J. 2002. “The bank capital channel of monetary policy.” Manuscript, University of Pennsylvania.
- Xiao, Kairong. 2018. “Monetary Transmission through Shadow Banks.” Manuscript, Columbia University.

Table 1: Summary Statistics

	mean	sd	p10	p25	p50	p75	p90
Deposit Share	0.086	0.535	0.004	0.005	0.010	0.024	0.088
Loan Share	0.028	0.160	0.001	0.001	0.002	0.006	0.027
Deposit Rates	1.774	1.275	0.117	0.562	1.667	2.883	3.546
Loan Rates	6.521	1.629	4.416	5.269	6.383	7.813	8.682
No. of Branches	80.963	342.600	12.000	14.000	20.000	40.000	115.000
No. of Employees per Branch	17.443	16.131	9.063	11.015	13.905	18.571	26.533
Expenses of Fixed Assets	0.458	0.153	0.271	0.342	0.437	0.564	0.729
Salary	1.683	0.451	1.065	1.343	1.634	1.990	2.493

This table reports summary statistics of the sample for BLP estimation. The sample period is from 1994 to 2017. Deposit share and loan share are computed taking the entire United States as a unified market. The total size of the deposit market is defined as the sum of deposits, cash, and bonds held by all the U.S. households. The total size of the loan market is defined as the sum of bank loans and corporate bonds issued by the U.S. firms. Deposit and loan rates are imputed using the interest expense and income from Call report. Expense of fixed assets and salary are scaled by total assets. Deposit share, loan share, deposit rates, loan rates, expense of fixed assets and salary are reported in percentage. The data is from Call report and FDIC Summary of Deposits.

Table 2: Monetary Policy Shocks and Bank Equity Returns on FOMC Days

	(1) FFR > 2%	(2) FFR ≤ 2%	(3) FFR > 2%	(4) FFR ≤ 2%
Δ 2-year Yield	-1.292** [0.615]	2.202** [0.879]	-0.639 [0.653]	-1.393 [0.852]
HHI*Δ 2-year Yield			-0.134 [0.145]	0.562*** [0.153]
Δ Term Spread	-0.634 [1.265]	2.336* [1.350]	-0.667 [1.257]	1.827 [1.293]
Market Return	0.297*** [0.072]	0.730*** [0.070]	0.295*** [0.071]	0.733*** [0.070]
Observations	27,257	33,805	27,257	33,805
Adj, R-squared	0.015	0.123	0.016	0.125

This table reports the estimates of the relation between bank equity returns and monetary policy shocks on FOMC Days. Monetary shocks are measured by the daily change in the two-year Treasury yield on FOMC days. HHI is the the Herfindahl-Hirschman index of the local deposit market in which the bank operates. A local deposit market is defined as a Metropolitan Statistical Area (MSAs). If a bank operates in several MSAs, the bank-level HHI is the weighted average of local HHI, weighted by the deposits of the bank in the local market. The sample includes all publicly traded U.S. banks from 1994 to 2017. We exclude observations during the burst of dot-com bubble (2000-2001) and the subprime financial crisis (2007-2009). The standard errors are clustered by time.

Table 3: Bank Balance Sheet

Assets		Liabilities	
Existing loans	L_t	Deposits	D_t
New loans	$P(B_t, r_t^l)$	Non-reservable borrowings	N_t
Reserves	R_t	Equity	E_t
Government securities	G_t		
<i>Total Assets</i>	$L_t + P(B_t, r_t^l) + R_t + G_t$	<i>Total Liabilities and Equity</i>	$D_t + N_t + E_t$

This table illustrates the balance sheet of a typical bank at the beginning of the period.

Table 4: Parameter Estimates

Statutory Parameters			
τ_c	Corporate tax rate	0.35	
θ	The reserve ratio	0.022	
κ	The capital ratio	0.06	
Parameters Estimated Separately			
μ	Average loan maturity	5	
\bar{f}	Log Federal Funds rate mean	0.60	
σ_f	Log Federal Funds rate variance	0.65	
ρ_f	Log Federal Funds rate persistence	0.91	
$\bar{\delta}$	Log loan chargeoffs mean	-0.89	
σ_δ	Log loan chargeoffs variance	1.24	
ρ_δ	Log loan chargeoffs persistence	0.50	
\hat{J}	Number of representative banks	14	
Parameters Estimated via BLP			
α^d	Depositors' sensitivity to deposit rates	0.80	[0.16]
σ_{α^d}	The dispersion of depositors' sensitivity to deposit rates	1.58	[0.44]
α^l	Borrowers' sensitivity to loan rates	-0.90	[0.16]
q_d^d	Convenience of holding deposits	0.95	[0.19]
q_b^d	Convenience of holding bonds	-0.10	[0.16]
q_l^l	Convenience of borrowing through loans	-0.10	[0.59]
Parameters Estimated via SMM			
γ	Banks' discount rate	0.052	
ϕ^N	Cost function of non-reservable borrowing	0.010	
ϕ^d	Bank's cost of taking deposits	0.007	
ϕ^l	Bank's cost of servicing loans	0.009	
q_n^l	The value of firms' outside option	-6.493	

This table reports the model parameter estimates. Panel A presents results for parameters that represent statutory rates. Panel B presents results from parameters that can be calculated as simple averages or by simple regression methods. Panel C presents results from parameters estimated via BLP. Panel C presents results from parameters estimated via SMM.

Table 5: Moment Conditions

	Actual Moment	Simulated Moment
Dividend yield	2.53%	2.64%
Non-reservable borrowing share	30%	30.8%
Deposit spread	1.46%	1.66%
Loan spread	2.78%	2.77%
Loan/Deposit Ratio	0.96	0.978
Corporate Borrowing-FFR Sensitivity	-0.50	-0.562
Deposit spread - FFR sensitivity	0.30	0.285
Loan spread - FFR sensitivity	-0.25	-0.241

This table reports the moment conditions in the simulated method of moment (SMM) estimation.

Table 6: Demand Estimation

	Deposit	Loan
Yield sensitivity (α)	0.800*** [0.158]	-0.904*** [0.163]
Log number of branches (β_1)	0.868*** [0.009]	1.117*** [0.000]
Log number of employees (β_2)	0.587*** [0.016]	0.694*** [0.031]
Yield sensitivity dispersion (σ_α)	1.579*** [0.439]	
Sector F.E.	Y	Y
Time F.E.	Y	Y
Observations	15575	15575
Adj. Rsq	0.982	0.867

This table reports the estimated parameters of the deposit and loan demand. The first column reports parameters of deposit demand. The second column reports parameters of loan demand. Yield sensitivity (α) refers to the average sensitivity of the depositors (firms) to deposit rates (loan rates). Log No. of Branches (β_1) refers the sensitivity of the depositors (firms) to log number of branches that each bank has. Log No. of Employees (β_2) refers the sensitivity of the depositors (firms) to log number of employees per branch. Yield sensitivity dispersion (σ_α) refers to the dispersion in the sensitivity of the depositors to deposit rates (the dispersion is set to 0 for firms). The sample includes all the U.S. commercial banks from 1994 to 2017 with domestic branches higher than 10. The data is from the Call report and the Summary of Deposits.

Table 7: Determinants of Monetary Policy Transmission

		Sensitivity of Loans to FFR ($\frac{\Delta(l)}{\Delta f}$)	Aggregate Bank Loans
(1)	Baseline	-3.88%	100%
(2)	- Reserve Regulation	-3.81%	101%
(3)	- Deposit Market Power	-2.57%	118%
(4)	- Capital Constraint	-0.95%	128%
(5)	- Loan Market Power (quantities)	-1.38%	193%

This table depicts a series of counterfactual experiments in which we examine the cumulative effect of removing frictions from our model on two important quantities. The first is the sensitivity of loans to the Federal Funds rate (FFR), and the second is the aggregate amount of borrowing, which is normalized to 100% in the baseline case. Each line of the table presents the results from eliminating the corresponding friction.

Table 8: High versus Low Market Concentration

	Low Concentration	Median Concentration	High Concentration
	$N = 200$	$N = 14$	$N = 2$
(1) Baseline	-3.52%	-3.88%	-5.21%
(2) – Deposit Market Power	-3.18%	-2.57%	-3.27%
(3) – Capital Constraint	-0.96%	-0.95%	-0.93%
(4) – Loan Market Power	-1.35%	-1.38%	-1.59%

This table illustrates how the monetary policy transmission is influenced by different frictions under three economic regimes: high bank concentration, median bank concentration, and low bank concentration. Bank concentration is measured by the number of competing banks, N , in the local market. Monetary policy transmission is captured by the sensitivity of loans to the Federal Funds rate (FFR). Each line of the table presents the results from eliminating the corresponding friction.

Table 9: Large versus Small Banks

	Parameters Estimates					Sensitivity of Loans to FFR($\frac{\Delta(l)}{\Delta f}$)
	γ	q_n^l	ϕ^N	ϕ^d	ϕ^l	
Full Sample	0.052	-6.439	0.004	0.008	0.006	-2.48
Big Banks	0.052	-6.439	0.004	0.008	0.005	-2.31
Small Banks	0.052	-6.439	0.003	0.009	0.012	-3.24

This table presents the parameter estimates and the degree of monetary policy transmission in subsamples of large and small banks. ϕ^d and ϕ^l are banks' marginal costs of intaking deposits and servicing loans, respectively, and ϕ^N is the quadratic cost of borrowing non-reservables. The last column reports the sensitivity of loans to the Federal Funds rate (FFR), which is a measure of monetary policy transmission. Standard errors clustered at the firm level are in parentheses under the parameter estimates.

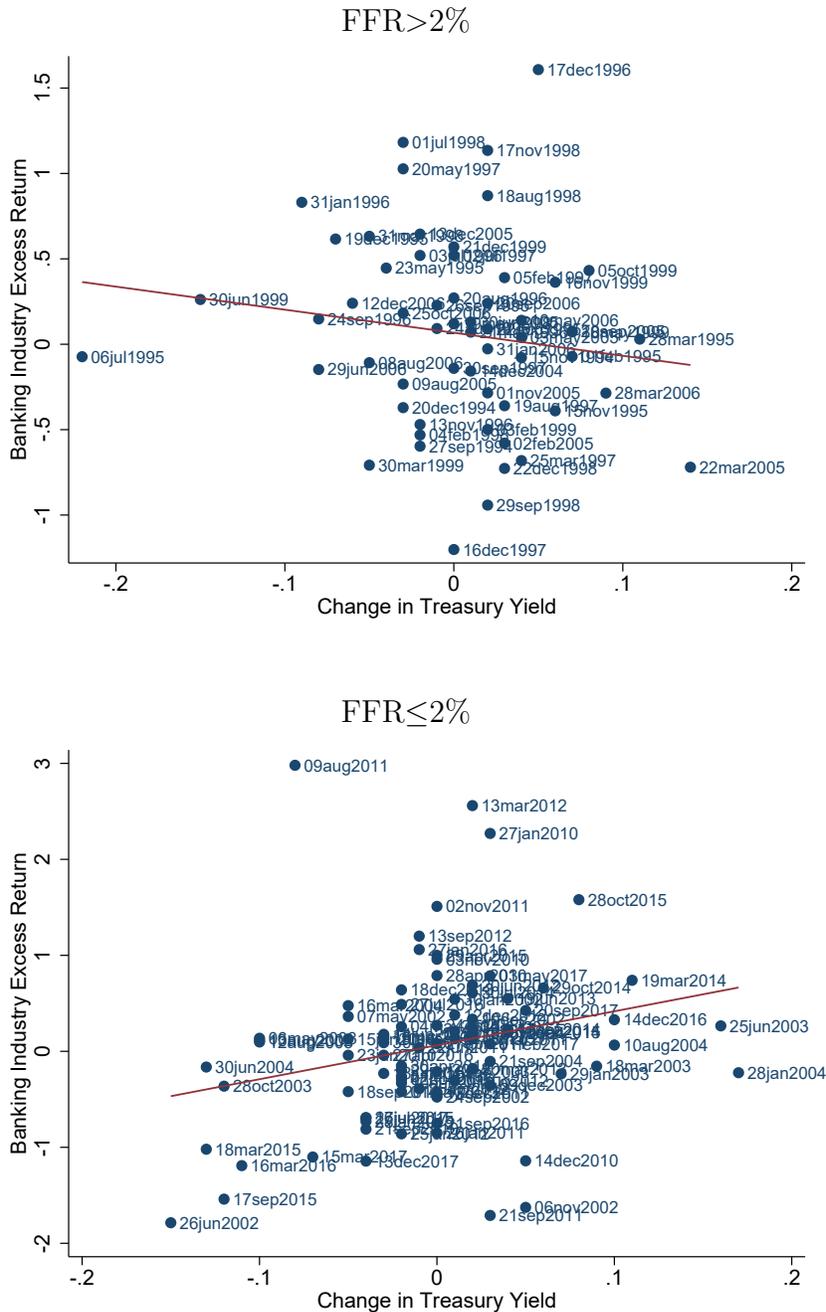


Figure 1: Monetary Policy Shocks and Bank Equity Returns

This figure provides the scatter plot of the bank industry excess returns against the daily change in two-year Treasury yield on FOMC days from 1994 to 2017. The excess return is defined as the different between bank industry index return and the market return. The sample of the upper panel is when the Federal Funds rate is above 2% and the sample of the lower panel is when the Federal Funds rate is below 2%. We exclude observations during the burst of dot-com bubble (2000-2001) and the subprime financial crisis (2007-2009) because the stock returns are extremely volatile. The bank industry stock returns are retrieved from Kenneth French's website and the two-year Treasury yield is retrieved from FRED database of the Federal Reserve Bank of St. Louis.

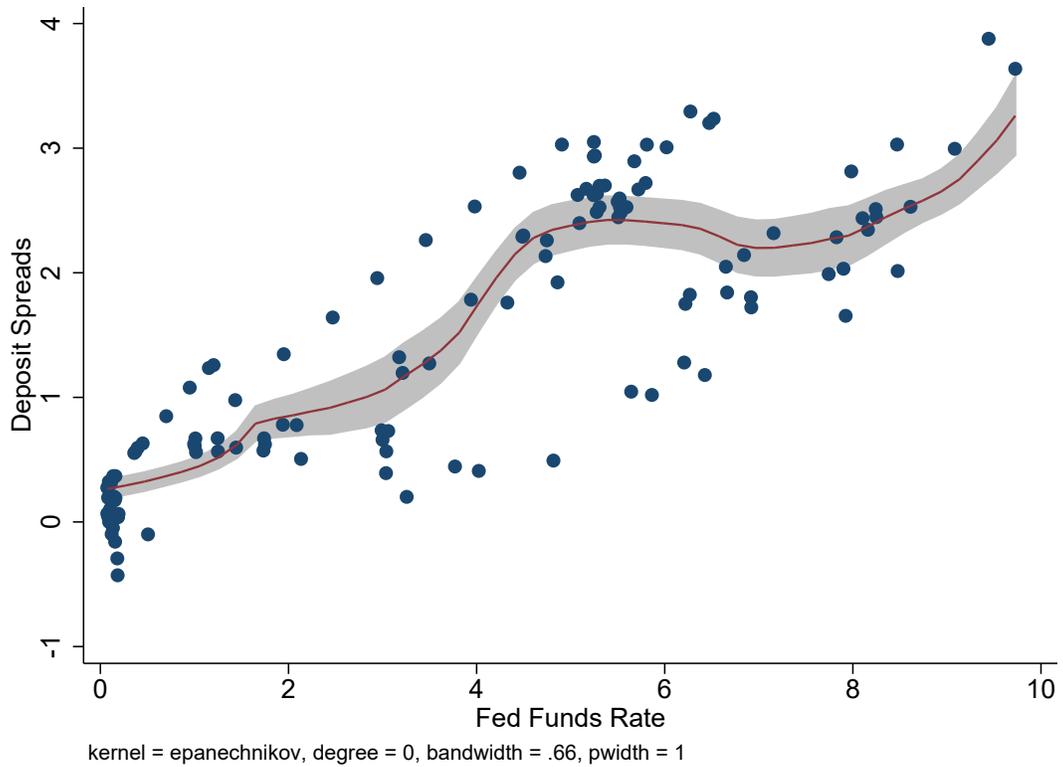


Figure 2: The Average Deposit Spread and the Fed Funds Rates

This figure plots the non-parametric relationship between the Fed Funds rate and the average deposit spread of U.S. banks. The sample is from 1985 to 2017. The data frequency is quarterly. The deposit spread is constructed using the Call report and the Fed Funds rate is retrieved from FRED database of the Federal Reserve Bank of St. Louis.

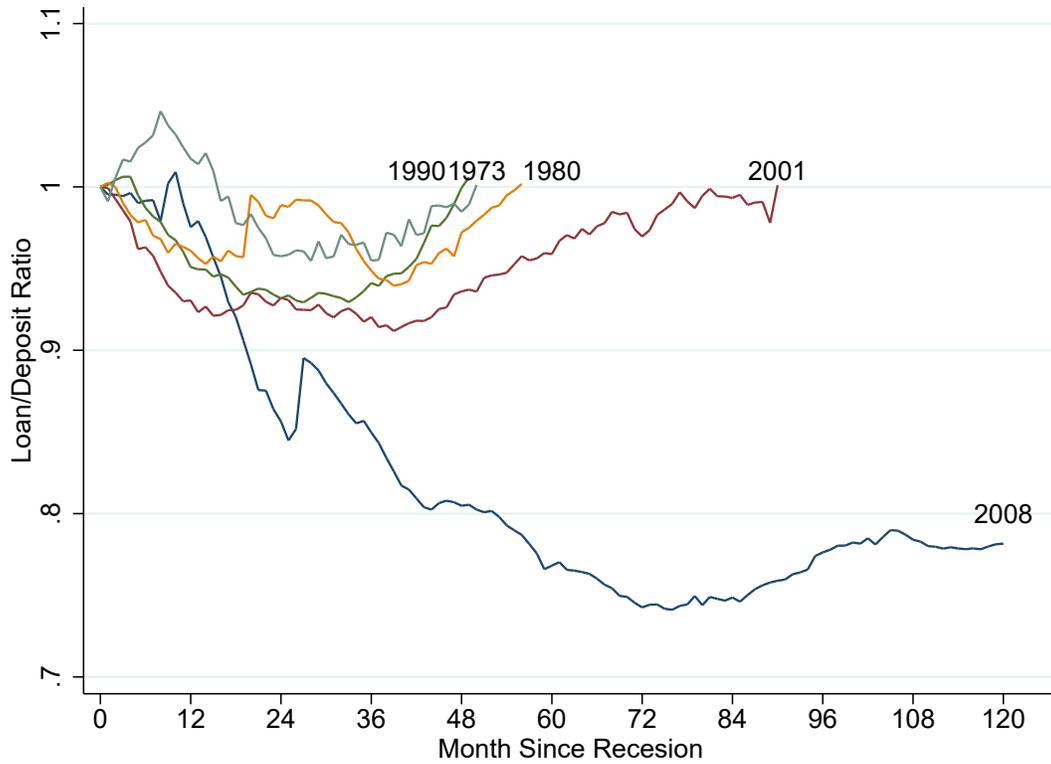


Figure 3: Loan-to-Deposit Ratios for U.S. Banks

This figure plots the loan-to-deposit ratio of U.S. banks after the start of five recessions from 1973 to 2017. The x-axis is the month since the start of recession and the y-axis the loan-to-deposit ratio. We normalize the ratio in month 0 to 1. We plot the path of the ratio until the ratio recovers to the pre-recession level. The data is retrieved from FRED database of the Federal Reserve Bank of St. Louis.

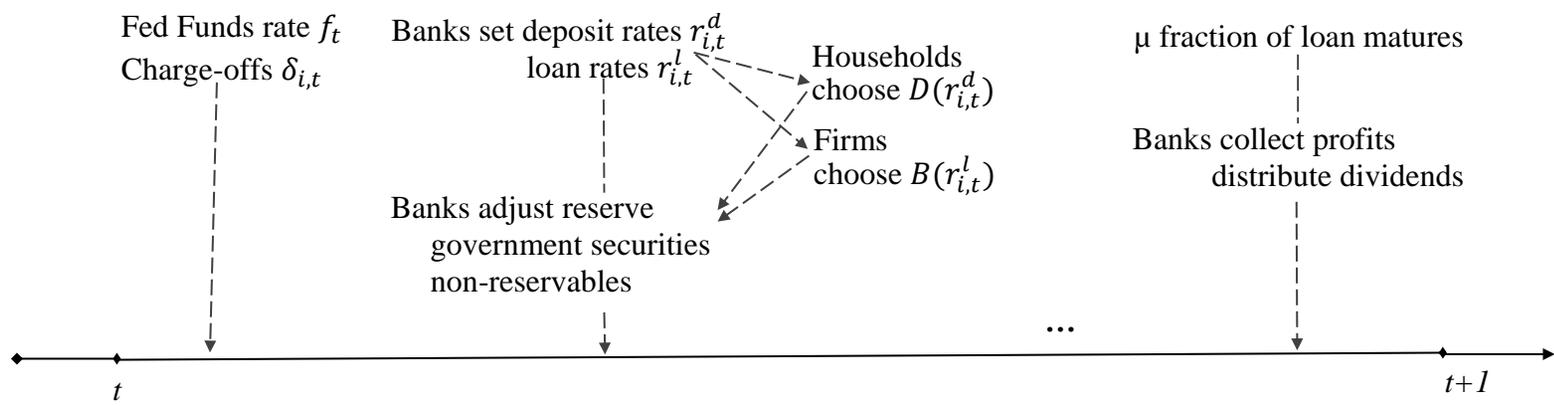


Figure 4: Timeline within a Period

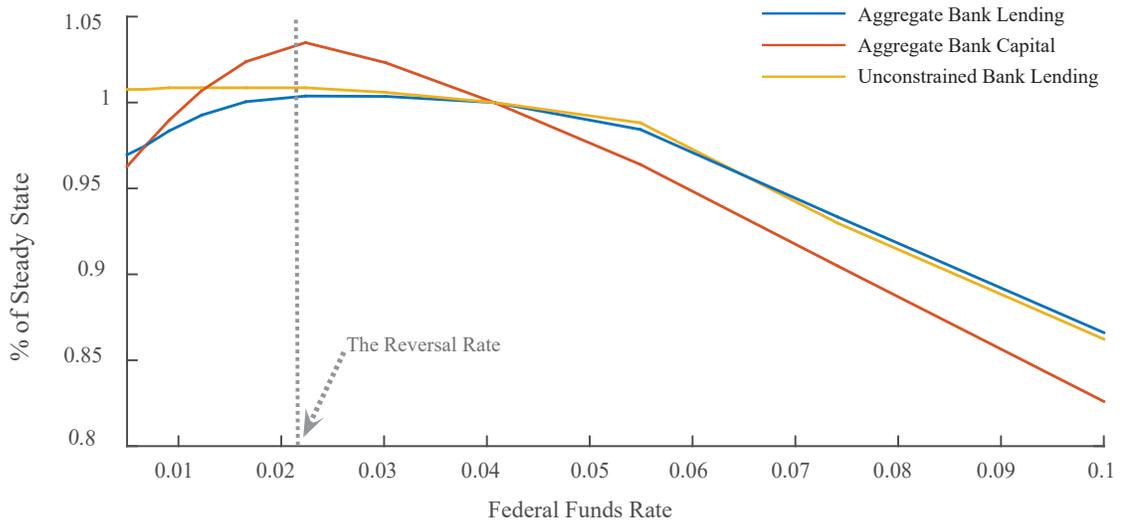
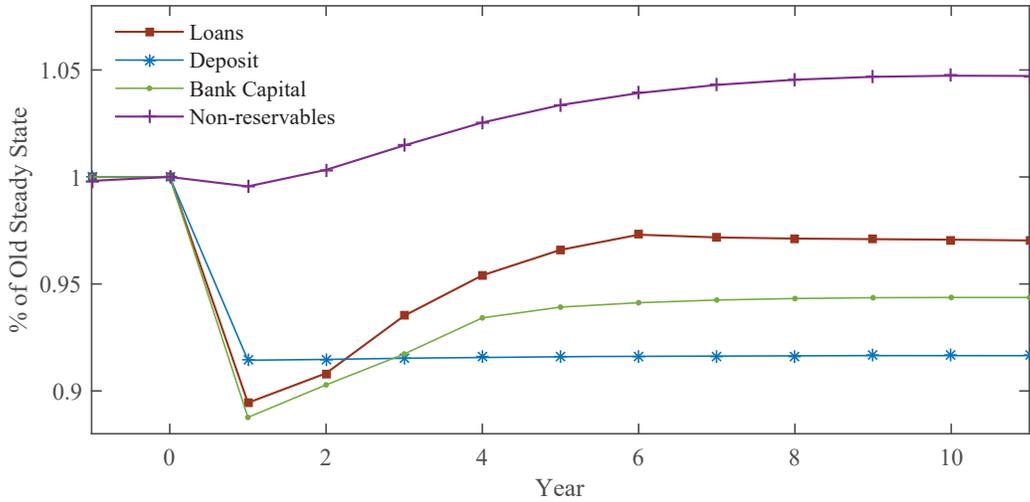


Figure 5: Fed Funds Rate and Bank Characteristics

This figure illustrates how bank capital and optimal lending vary with the Fed Funds rate. Banks' optimal lending is calculated under two alternative cases: the baseline line where banks are subject to the capital regulation and an alternative unconstrained case where the capital regulation is removed. The Fed Funds rate is on the x-axis; bank characteristics, scaled by their respective steady state values (when the Fed funds rate is 0.02), is on the y-axis.

Panel A: Two Std Increase in Fed Funds Rate



Panel B: Two Std Decrease in Fed Funds Rate

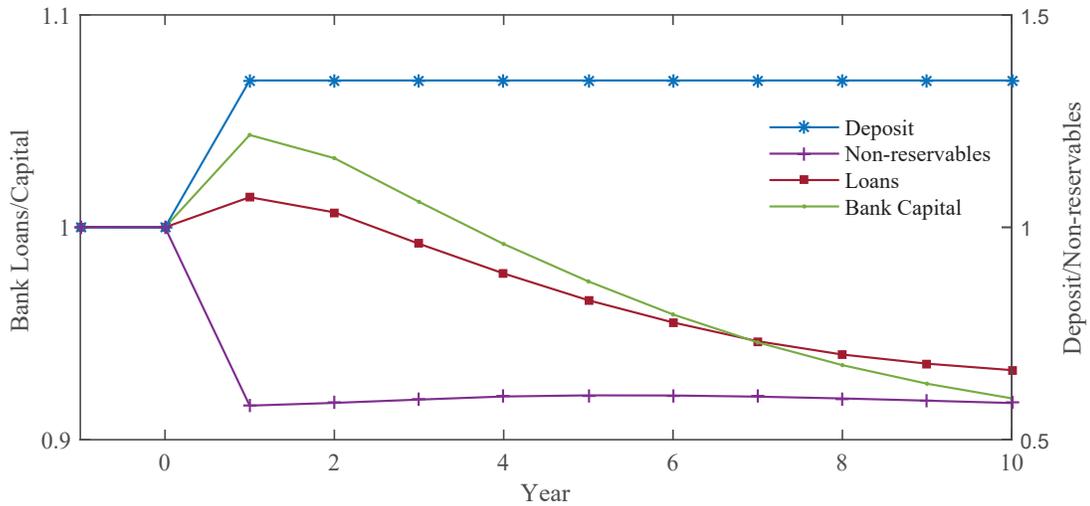


Figure 6: Impulse Response to Fed Funds Rate Shocks

This figure illustrates banks' impulse response to Fed fund rate shocks. The economy starts at Year 0 when it is in the old steady state with the FFR equal to 0.02; In Year 1, the FFR either increases or decreases by two standard deviations, and it stays at that level afterwards until the economy reaches the new steady state. Each variable in the graph is scaled by the level in the old steady state (when FFR = 0.02).