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

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

December 27, 2019

Abstract

We quantify the impact of bank market power on the transmission of monetary policy through banks to borrowers. Using data from U.S. banks, we estimate a dynamic banking model in which monetary policy affects banks' funding costs. In the model, imperfectly competitive banks, facing capital and reserve regulations, optimally choose the degree to which they pass these cost shifts to borrowers and depositors. We find that in the last two decades, bank market power explains a significant portion of the transmission of monetary policy to borrowers. The quantitative effect is comparable in magnitude to the effect of bank capital regulation. We estimate that when the federal funds rate falls below 0.9%, market power interacts with bank capital regulation to produce a reversal of the effect of monetary policy.

JEL Codes: E51, E52, G21, G28

Keywords: Monetary policy transmission, banking competition, regulatory constraints.

1. Introduction

We examine the quantitative impact of bank market power for the transmission of monetary policy through the banking system. Interest in this question has sharpened after three decades of consolidation in the banking industry that has softened competitive pressure. Indeed, recent research offers qualitative evidence that bank market power affects the pass-through of monetary policy to the supply of loans (Scharfstein and Sunderam 2016; Drechsler, Savov, and Schnabl 2017). In contrast, traditional analysis of monetary policy transmission focuses on regulatory constraints, such as reserve or capital requirements, as the central frictions that influence monetary policy transmission (e.g., Bernanke and Blinder 1988; Kashyap and Stein 1995). Yet the qualitative nature of all of this evidence leaves open the question of the relative magnitude of traditional versus market-power channels for monetary policy transmission.

We answer this question by using data on U.S. banks to estimate a dynamic banking model with three frictions: regulatory constraints, financial frictions, and imperfect competition. The estimation allows our data to discipline the model parameters and thus expose the relative magnitude of these three frictions. We find that bank market power plays an important role in determining the degree of monetary policy transmission. In terms of magnitude, the effects of bank market power are comparable in magnitude to those of bank capital regulation, while the effects of bank reserve requirements are limited.

Our analysis produces two other key results. First, we show that banks face nontrivial costs when they access external financial markets. This friction plays a pivotal role in connecting banks' deposit market power to their lending decisions, as external financing costs serve to link banks' sources and uses of funds. We also show that these frictions help explain the differential sensitivity of lending to the policy rate for big and small banks.

Second, we show that bank market power interacts with capital regulation to reverse the effect of monetary policy when the federal funds rate is very low. Specifically, we estimate that, when the federal funds rate is below 0.9%, 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 a threshold predicted by the model.

To provide intuition for these results, we elaborate on the model. In an industry equilibrium, imperfectly competitive banks act as intermediaries between borrowers and depositors. Banks' lending decisions are dynamic for two reasons: a maturity mismatch between short-term deposits and long-term loans and financial frictions that induce precautionary capital accumulation. In this setting, monetary policy alters the federal funds rate. Because banks are not price takers in deposit or loan markets, they choose the extent to which they pass rate movements through to depositors and borrowers. The magnitude of this pass-through depends on the tightness of regulatory constraints, the severity of financial frictions, and the intensity of competition.

These frictions in our model map into four channels of monetary policy transmission. The first is the bank reserve channel, in which a high federal funds rate raises the opportunity cost of holding reserves, thus contracting deposit creation (Bernanke and Blinder 1988; Kashyap and Stein 1995). The second is the bank capital channel, in which a high federal funds rate reduces bank capital because of a maturity mismatch and thus constrains banks' capacity to lend (Bolton and Freixas 2000; Van den Heuvel 2002; Brunnermeier and Sannikov 2016). The third is the deposit market power channel, in which a high federal funds rate allows banks to charge higher markups on deposits, thus leading to a contraction in deposits and loans (Drechsler et al. 2017). The fourth is the loan market power channel, in which banks reduce markups to mitigate the effects of falling loan demand when the federal funds rate rises (Scharfstein and Sunderam 2016).

To gauge the quantitative importance of these transmission channels, we estimate our model using data on U.S. commercial banks from 1994 to 2017. 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). We begin by using demand estimation techniques to

obtain the elasticities of loans and deposits to interest rates. We then plug these estimates into our model and use simulated minimum distance to obtain estimates of parameters that quantify financial frictions and operating costs. The sequential use of these two techniques represents a methodological advance that enables us to consider a rich equilibrium model that would otherwise be intractable to estimate.

We use counterfactual experiments to measure the relative importance of each transmission channel. We start with a model with all frictions as estimated and then subtract each friction one at a time. We find that eliminating reserve requirements leaves the sensitivity of lending to the federal funds rate nearly unchanged. Eliminating either capital regulation or deposit market power reduces this sensitivity, while eliminating loan market power raises it.

These counterfactuals also show that rate cuts can be contractionary when rates are already low. Low rates depress bank profits by reducing bank deposit market power, as competition from cash intensifies. Lower profits then tighten the capital constraint and reduce lending. This result helps explain sluggish bank lending growth observed in the ultra-low interest rate environment after the 2008 financial crisis.

Our paper contributes to the literature on the role of banks in transmitting monetary policy (Bernanke and Blinder 1988; Kashyap and Stein 1995; Van den Heuvel 2002; Scharfstein and Sunderam 2016; Brunnermeier and Sannikov 2016; Drechsler et al. 2017). We are the first to structurally estimate a dynamic banking model to quantify various transmission channels. Prior to our work, little was known about the relative importance of these channels, as this type of quantitative exercise is difficult to undertake using reduced-form methods. Moreover, previous studies usually consider these transmission channels in isolation. For example, Scharfstein and Sunderam (2016) and Drechsler et al. (2017) study market power in the loan and deposit markets separately. However, little is known about the interactions between channels. Thus, an important contribution of this paper is to provide a unified framework within which to study these interactions. For example, we show that the relative importance of the deposit and loan 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.

Second, our paper is related to the theoretical literature on the effects of negative interest rate policies (Brunnermeier and Koby 2016; Eggertsson, Juelsrud, and Wold 2017; Wang 2019; Campos 2019). While these studies contain interesting insights, they typically treat the banking sector with a high level of abstraction. In contrast, we provide a model that is sufficiently realistic to be directly mapped onto microeconomic data. Our paper also contributes to the empirical literature on negative interest policy (Demiralp, Eisenschmidt, and Vlassopoulos 2017; Heider, Saidi, and Schepens 2018; Basten and Mariathasan 2018). These studies show that a negative interest rate policy can have perverse effects on bank lending. Our results suggest that such perverse effects can start to occur even before the policy rate turns negative because a near-zero policy rate results in a compression of banks' deposit spreads. Therefore, in countries such as the United States where the policy rate has never gone negative, the banking sector could nevertheless be hurt by an ultra-low rate monetary policy.

Third, our paper is related to the literature on external financial frictions. Romer and Romer (1990) argue that banks can easily replace deposits with external financing, so shocks to deposits are unlikely to affect bank lending. In contrast, Kashyap and Stein (1995) argue that external financing is costly for banks, so the quantity of deposits matters for bank lending. Our study contributes to this debate, as our structural estimation approach allows us to infer the degree of bank financing costs from the relative size of their deposit taking and external borrowing. We find that this cost is economically significant and that frictions related to bank balance sheets play an important role in the transmission of monetary policy.

Finally, our paper contributes to the structural industrial organization literature on the banking system (Egan, Hortacsu, and Matvos 2017a; Egan, Lewellen, and Sunderam 2017b; Buchak, Matvos, Piskorski, and Seru 2018; Xiao 2018). While this work usually features a static industry equilibrium, we introduce the dynamic adjustment of banks' balance sheets to study the role of maturity transformation and financial frictions. Our paper is also tan-

gentially related to recent work that uses dynamic banking models to study optimal capital regulation (Corbae and D'Erasmo 2013; Elenev, Landvoigt, and Van Nieuwerburgh 2016; Begenau 2018; Begenau and Landvoigt 2018). Our paper stands apart from this literature in that we emphasize the role of imperfect competition. Moreover, our approach is more empirical in nature, as we estimate, rather than calibrate, all of our model parameters.

2. Data and Stylized Facts

Our main dataset is the Consolidated Reports of Condition and Income (Call Reports). This dataset provides quarterly bank-level balance sheet information for U.S. commercial banks. It includes 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 on each bank since 1994 at an annual frequency. The sample period is 1994–2017.

We also use several other data sources. First, we retrieve publicly listed bank returns from CRSP and link these returns to bank concentration measures using the link table provided by the Federal Reserve Bank of New York. Banking industry stock returns are from Kenneth French's website. We collect Federal Open Market Committee meeting dates from the FOMC meeting calendar. Finally, we obtain the following time series from the Federal Reserve Economic Data (FRED) database: NBER recession dates, the effective federal funds rate, two-year and five-year Treasury yields, the aggregate amount of corporate bonds issued by U.S. firms, and the aggregate amounts of cash, Treasury bonds, and money-market mutual funds held by households. Details regarding the construction of our variables are in Table 1.

Table 2 provides summary statistics for our sample. Three patterns are of note. First, mean deposit and loan market shares for the U.S. national market lie near the 90$^{\text{th}}$ percentile, indicating a very skewed distribution of market shares in which a few large banks dominate the market. Second, we see little variation in the number of employees per branch, but we see both high variance and skewness in the number of branches per bank. This skewness is consistent with the skewness in market shares, as the number of branches is highly correlated

with bank size. Third, we find that average loan maturity is 3.429 years.

In Figure 1, we show the prices that banks charge for their deposits and loans. Panel A depicts a time-series kernel regression of the average quarterly U.S. bank deposit spread on the federal funds rate, where the deposit spread is the difference between the federal funds rate and the deposit rate. Because this spread measures the price that banks charge for their depository services, in the absence of market power, one would expect to observe constant deposit spreads that equal the marginal cost of providing deposits. However, we find a positive relation between deposit spreads and the federal funds rate, which steepens when the rate is close to zero. This relation implies that banks charge higher prices for their depository services as the federal funds rate rises. Intuitively, if banks have market power, a higher federal funds rate allows them to increase profits by raising markups above marginal costs because depositors find cash costly to hold (Drechsler et al. 2017).

Panel B of Figure 1 contains an analogous kernel regression of the average U.S. bank loan spread on the federal funds rate, where the loan spread is the difference between the loan rate, adjusted for loan loss provisions, and the 5-year Treasury yield. We find that the loan spread falls as the federal funds rate rises. This pattern is consistent with Scharfstein and Sunderam (2016), who show that banks lower markups on loans in the face of rising rates to mitigate the effects of falling demand. In sum, Figure 1 suggests that market power creates wedges between the federal funds rate and the rates at which banks borrow and lend. Furthermore, the sizes of these wedges depend on the level of interest rates.

3. Model

While this evidence suggests interesting interactions between bank market power and monetary policy, it does not reveal any underlying mechanisms behind these patterns in the data. To understand this evidence further, we consider an infinite-horizon, bank industry equilibrium model with three sectors: households, firms, and banks. In the model, households and firms solve static discrete choice problems in which they choose from several saving and financing vehicles. Banks act as intermediaries between households and firms by taking short-term

deposits from households and providing long-term loans to firms.[1]

The richness of the model lies in the banking sector, as several frictions imply that monetary policy affects the extent of intermediation that banks provide. First, competition in the deposit and loan markets is imperfect, so banks strategically choose deposit and loan rates to maximize their profits. Second, banks are subject to regulation. Reserve regulation links the opportunity cost of taking deposits to the prevailing federal funds rate. Capital regulation incentivizes banks to optimize their lending intertemporally with an eye to preserving excess equity capital as a buffer against future capital inadequacy. Third, access to non-deposit external financing is more costly than taking deposits. This friction implies that shocks to the quantity of deposits are transmitted to the supply of loans, as banks cannot costlessly replace deposits with other borrowing. These frictions are important because, in their absence, banks are simply pass-through entities and bond market interest rates summarize monetary policy.

3.1. Households

In our infinite-horizon, equilibrium model, at each time t, the economy contains a mass W_t of 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 from among the following investment options for their endowments: cash, bonds, and deposits, where the deposits of each bank constitute a differentiated product. If we index each option by j, the households' choice set is given by $\mathcal{A}^d = \{0, 1, \ldots, J, J+1\}$, with option 0 representing cash, option $J+1$ representing short-term bonds, and options $1, \ldots, J$ 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 they have, and the probability of choosing each of the options can be interpreted as a portfolio weight.

Each option is characterized by a yield, r_j^d, and a vector of product characteristics, x_j^d,

---

[1]In reality, firms make deposits in banks, and households borrow from banks. Therefore, in our model, the households should be broadly interpreted as savers, and the firms should be broadly interpreted as borrowers.

which capture the convenience of each option. For instance, a household might value the number of branches and the number of employees per branch when choosing a bank. The yield on cash is 0, and the yield on bonds is the federal funds rate, f, where we abstract from differences between short-term Treasury yields and the federal funds rate.

The household chooses the best option to maximize its utility:

$$\max_{j \in \mathcal{A}^d} u_{i,j} = \alpha_i^d r_j^d + \beta^d x_j^d + \xi_j^d + \epsilon_{i,j}^d, \tag{1}$$

where households are indexed by $i \in 1, 2, \ldots, I$. The utility for household i from choosing option j is $u_{i,j}$, and α_i^d is the sensitivity to the yield r_j^d. We allow households to exhibit varying sensitivity to yields to capture the evidence that some depositors are less yield-sensitive than others, and that this heterogeneity impacts deposit rate-setting (Xiao 2018). We model the distribution of depositors' yield sensitivity as a uniform distribution with a mean, α^d, and a standard deviation, σ_α^d. The coefficients, β^d, are sensitivities to the non-rate product characteristics, x_j^d, and ξ_j is the unobservable product-level demand shock. $\epsilon_{i,j}^d$ is a relationship-specific shock to the choice of option j by household i. $\epsilon_{i,j}^d$ captures horizontal differentiation across banks. For instance, if household i lives close to bank j, then $\epsilon_{i,j}^d$ is large, so household i is more likely to choose bank j, holding other characteristics constant. The optimal choice for household i is given by an indicator function:

$$\mathbb{I}_{i,j}^d = \begin{cases} 1, & \text{if } u_{i,j} \geq u_{i,k}, \text{ for } k \in \mathcal{A}^d \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

We aggregate the optimal choices across all households to compute the deposit market share of each bank j. Adopting the standard assumption that $\epsilon_{i,j}^d$ follows a generalized extreme value distribution with a cumulative distribution function given by $F(\epsilon) = \exp(-\exp(-\epsilon))$, we can derive the standard logit market share:

$$s_j^d\left(r_j^d | f\right) \equiv \int \mathbb{I}_{i,j}^d dF(\epsilon) =$$

$$\sum_{i=1}^{I} \mu_i^d \frac{\exp\left(\alpha_i^d r_j^d + \beta^d x_j^d + \xi_j^d\right)}{\exp\left(\alpha_i^d f + \beta^d x_{J+1}^d + \xi_{J+1}^d\right) + \exp\left(\beta^d x_c^d + \xi_c^d\right) + \sum_{m=1}^{J} \exp\left(\alpha_i^d r_m^d + \beta^d x_m^d + \xi_m^d\right)}, \tag{3}$$

where μ_i^d is the fraction of total wealth, W, held by households of type i. The numerator represents the utility from depositing at bank j. Similarly, the first term in the denominator, $\exp\left(\alpha_i^d f + \beta^d x_{J+1}^d + \xi_{J+1}^d\right)$, represents the utility of holding Treasury bills, and the second term, $\exp\left(\beta^d x_c^d + \xi_c^d\right)$, is the utility of holding cash. The demand function for deposits of bank j is then given by the market share multiplied by total wealth,

$$D_j\left(r_j^d \,|\, f\right) = s_j^d\left(r_j^d \,|\, f\right) W. \tag{4}$$

3.2. Firms

There is a mass, K, of 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 long-term bank loans. We assume that each bank is a differentiated lender, given factors such as geographic location and industry expertise. Letting each option be indexed by j, the firms' choice set is given by $\mathcal{A}^l = \{0, 1, \ldots, J, J+1\}$, where option $J+1$ represents bonds, options $1, \ldots, J$ represent loans from each bank, and option 0 is the option to not borrow at all.

For tractability, we assume that both bonds and bank loans have the following repayment schedule. In each period, the firm has to pay back a fraction, η, of its outstanding principal plus interest. Thus, if the firm borrows one dollar at a fixed interest rate r, the repayment stream, starting in the next period, is $(1-\eta) \times (\eta + r)^t$, $t = 0, \ldots, \infty$. Accordingly, all firm debt has an average maturity of $1/\eta$ periods.

The long-term bond interest rate is given by an expected default cost, $\bar{\delta}$, plus the expected weighted average of future federal funds rates, $\bar{f}_t$, which is itself given by:

$$\bar{f}_t = \eta f_t + \mathbb{E}_t\left[\sum_{n=1}^{\infty} \eta(1-\eta)^n f_{t+n}\right]. \tag{5}$$

Each of the firm's financing options is characterized by a rate, r_j^l, and a vector of product characteristics, x_j^l, capturing the convenience of using each of the financing options.

The firm then chooses the best option to maximize its profits:

$$\max_{j \in \mathcal{A}^l} \pi_{i,j} = \alpha_i^l r_j^l + \beta^l x_j^l + \xi_j^l + \epsilon_{i,j}^l, \tag{6}$$

10

where $\pi_{i,j}$ is the profits of firm i from choosing option j, and α_i^l is the sensitivity to the interest rate r_j^l, which follows a uniform distribution with a mean, α^l, and a standard deviation, σ_α^l. The sensitivities to non-rate characteristics, x_j^l, are given by β^l, and ξ_j^l is the unobservable product-level demand shock. $\epsilon_{i,j}^l$ is an idiosyncratic shock when firm i borrows from bank j. The optimal choice of firm i is given by an indicator function:

$$\mathbb{I}_{i,j}^l = \begin{cases} 1, & \text{if } \pi_{i,j} \geq \pi_{i,k}, \text{ for } k \in \mathcal{A}^l \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

We aggregate the optimal choices across all the firms to compute the loan market share of each bank j. Assuming that $\epsilon_{i,j}^l$ follows a generalized extreme value distribution with a cumulative distribution function given by $F(\epsilon) = \exp(-\exp(-\epsilon))$, we can derive the standard logit market share:

$$s_j^l\left(r_j^l|f\right) \equiv \int \mathbb{I}_{i,j}^l dF(\epsilon) =$$

$$\sum_{i=1}^{I} \mu_i^l \frac{\exp\left(\alpha_i^l r_j^l + \beta^l x_j^l + \xi_j^l\right)}{\exp\left(\alpha_i^l(\overline{f} + \overline{\delta}) + \beta^l x_{J+1}^l + \xi_{J+1}^l\right) + \exp\left(\beta^l x_n^l + \xi_n^l\right) + \sum_{m=1}^{J} \exp\left(\alpha_i^l r_m^l + \beta^l x_m^l + \xi_m^l\right)}, \tag{8}$$

where μ_i^l is the fraction of type i firms, and $\overline{f}$ is the long-term bond interest rate. The numerator represents the utility from borrowing from bank j. Similarly, the first term in the denominator, $\exp\left(\alpha_i^l\left(\overline{f} + \overline{\delta}\right) + \beta^l x_{J+1}^l + \xi_{J+1}^l\right)$, represents the utility from issuing bonds, and the second term, $\exp\left(\beta^l x_n^l + \xi_n^l\right)$, is the utility of not borrowing. The demand function for loans is then given by the market share multiplied by the total loan market size:

$$B_j\left(r_j^l|f\right) = s_j^l\left(r_j^l|f\right)K. \tag{9}$$

3.3. The Banking Sector

Each bank simultaneously sets its deposit rate, $r_{j,t}^d$, and its loan rate, $r_{j,t}^l$, as a spread below or above the federal funds rate, f_t, which we assume is an exogenous state variable. These rate-setting decisions implicitly determine 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 the utility maximization problem as described in equation (1), which yields the quantity of

11

deposits supplied to bank j, $D_j(r_{j,t}^d)$, which is given by equation (4). Because households can hold cash, which has a zero nominal return, banks face a zero lower bound for deposit rates:

$$r_{j,t}^d \geq 0. \tag{10}$$

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, $B_j(r_{j,t}^l)$, given by equation (9). To simplify notation, we suppress the dependence of loans and deposits on the relevant interest rates, denoting them simply as $D_{j,t}$ and $B_{j,t}$.

Lending involves a maturity transformation between assets and liabilities. On the asset side, let $L_{j,t}$ denote the amount of loans the bank holds. As in the case of bonds, in each period, a fraction, η, of a bank's outstanding loans matures. 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 $1/\eta$. As noted above, banks can also issue new loans at an annualized interest rate of $r_{j,t}^l$. The new loans, once issued, have the same maturity structure as the existing ones, and the interest rate is fixed over the life of the new loans. From the bank's perspective, the present value of interest income is:

$$I_{j,t} = \sum_{n=0}^{\infty} \frac{(1-\eta)^n B_{j,t} r_{j,t}^l}{(1+\gamma)^n}, \tag{11}$$

where γ is the bank's discount factor, and a bank's outstanding loans evolve according to:

$$L_{j,t+1} = (1-\eta)\left(L_{j,t} + B_{j,t}\right). \tag{12}$$

We assume that in each period a random fraction of loans, $\delta_t \in [0, \eta]$, becomes delinquent. The bank takes δ as an exogenous state variable in its decision-making problem. Although we assume that the bank writes off delinquent payments, with charge-offs equal to $L_t \times \delta_t$, defaulting on a payment in one period does not exonerate the borrower from future payments. Therefore, delinquency does not affect the evolution of loans in equation (12).

The rest of the asset side of each bank's balance sheet consists of reserves, R_t, and holdings of government securities, G_t, which the bank can accumulate if the supply of funds exceeds

12

demand from the lending market. These securities earn the federal funds rate, f_t.

Next, we describe the liabilities side of the balance sheet. In each period, the bank can obtain outside financing via deposits or via non-reservable borrowing, 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. This consideration implies that the marginal cost of non-reservable borrowing is likely an increasing function of the amount raised. Thus, we assume that non-reservable borrowing incurs a quadratic financing cost beyond the prevailing federal funds rate, as follows:

$$\Phi^N(N_t) = \left(f_t + \frac{\phi^N}{2} \cdot \frac{N_t}{D_t} \right) N_t. \tag{13}$$

The cost in equation (13) represents an important friction because, in its absence, banks could always raise non-reservable funding to compensate for any deposit shortfalls. This funding availability would disconnect banks' deposit-taking and lending decisions, so changes in bank deposits induced by federal funds rate shocks would have no impact on lending.

Banks also incur costs for serving depositors, such as hiring employees. We assume that costs are linear in the amount of deposits:

$$\Phi^d(D_t) = \phi^d D_t. \tag{14}$$

Similarly, we assume that lending incurs costs, such as paying loan officers to screen loans or maintain client relationships. Again, we assume a linear functional form:

$$\Phi^l(B_t) = \phi^l B_t. \tag{15}$$

Similarly, we model fixed operating costs and noninterest income, both of which we assume to be independent of the deposit and lending rate decisions. Specifically, we let ψ represent the difference between fixed operating expenses and noninterest income per unit of steady state equity capital, denoted by $\bar{E}$. Therefore, the net fixed operating cost is $\psi \bar{E}$.

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:

$$L_t + B_t + R_t + G_t = D_t + N_t + E_t, \tag{16}$$

where E_t is the bank's beginning-of-period book equity. E_t itself evolves according to:

$$E_{t+1} = E_t + \Pi_t \times (1 - \tau) - C_{t+1}, \tag{17}$$

where Π_t represents the bank's total operating profits from its deposit-taking, security investments, and lending decisions, τ denotes the linear tax rate on these profits, and C_{t+1} is the cash dividends distributed to the bank's shareholders. 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 equation (17) are in turn given by:

$$\Pi_t = I_t - (L_t + B_t) \times (\delta_t + \phi^l) + G_t \times f_t - \left(r_t^d + \phi^d\right) D_t - \left(f_t + \frac{\phi^N}{2} \cdot \frac{N_t}{D_t}\right) N_t - \psi \bar{E}. \tag{18}$$

A further central friction in the model lies in our assumption that a bank can increase its inside equity only via retained earnings. Thus, there is no new equity issuance, which implies:

$$C_{t+1} \geq 0 \quad \forall t. \tag{19}$$

This constraint reflects a bank's limited liability and thus implies that banks cannot raise equity capital to replace deposits or non-reservable borrowing. In Section 6, we replace this assumption with costly equity issuance, finding only a limited impact on our results, as banks' equity issuances are both tiny and rare, both in the extended model and the data.

The next important ingredient in our model is regulation, namely, a capital requirement and a reserve requirement:

$$E_{t+1} \geq \kappa \times (L_t + B_t), \tag{20}$$

$$R_t \geq \theta \times D_t. \tag{21}$$

Equation (20) implies that the bank's book equity at the beginning of the next period has

to be no smaller than a fraction, κ, of total loans outstanding. Equation (21) is the bank's reserve requirement, which states that the bank has to keep a fraction, 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 reserves, so equation (21) holds with equality. While the Federal Reserve has paid interest on reserves since 2008, we show in Section 5 that modeling this newer policy has a limited impact on our model solution.

3.4. Monetary Policy

We model monetary policy as a process for the real federal funds rate. This assumption is motivated by the existence of price stickiness in the final goods sector, so the central bank can alter the real rate. In addition, we allow the federal funds rate to be correlated with loan charge-offs, so their joint law of motion is given by:

$$\left[\begin{array}{c} \ln \delta_{t+1} - \mathbb{E}(\ln \delta) \\ \ln f_{t+1} - \mathbb{E}(\ln f) \end{array} \right] = \left[\begin{array}{cc} \rho_\delta & \rho_{\delta f} \\ 0 & \rho_f \end{array} \right] \cdot \left[\begin{array}{c} \ln \delta_t - \mathbb{E}(\ln \delta) \\ \ln f_t - \mathbb{E}(\ln f) \end{array} \right] + \left[\begin{array}{cc} \sigma_\delta & 0 \\ 0 & \sigma_f \end{array} \right] \varepsilon_{t+1}, \qquad (22)$$

where ε_{t+1} has a standard bivariant normal distribution.

Monetary policy affects banks in two ways. First, from equation (13), the federal funds rate affects the marginal funding costs that banks pay in the non-reservable funding market. Second, the short-term federal funds rate affects long rates through expectations. Thus, both long and short rate movements affect banks' market power in the deposit and loan markets.

3.5. Bank's Problem and Equilibrium

Figure 2 shows the sequence of events in each time period. The bank enters the period and observes the federal funds rate, f_t, and the realization of the fraction of defaults, δ_t. Next, banks interact with households and firms by setting the loan and deposit spreads. The loan and deposit demand functions then dictate the amount of deposits from households and 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, banks collect profits at the end of the period and distribute dividends to shareholders.

15

As noted above, loan and deposit demand depend on the rates offered 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 non-reservable borrowing, N_t, and investment in government securities, G_t, it rationally considers 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:

$$\Gamma_{t+1} = P^\Gamma(\Gamma_t). \tag{23}$$

In every period, after observing the federal funds rate, f_t, and the random fraction of defaulted loans, δ_t, banks choose the optimal policy to maximize the expected discounted cash dividends to shareholders:

$$V(f_t, \delta_t, L_t, E_t | \Gamma_t) = \max_{\{r_t^l, r_t^d, G_t, N_t, R_t, C_{t+1}\}} \frac{1}{1+\gamma} \left\{ C_{t+1} + \mathbb{E} V(f_{t+1}, \delta_{t+1}, L_{t+1}, E_{t+1} | \Gamma_{t+1}) \right\}, \tag{24}$$

$$s.t. \qquad (10), \ (13), \ (14), \ (15), \ (16), \ (17), \ (23).$$

We define equilibrium in this economy as follows.

Definition 1 *A stationary equilibrium occurs when:*

1. *All banks solve the problem given by* (24)*, 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. *In 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 the banks' problem (Γ_t) is an object whose dimension depends on the number of banks in the economy. This dimensionality poses a challenge to solving the banks' problem numerically. 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 that is 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 $\bar{r}^{l,i}(f_t)$ and $\bar{r}^{d,i}(f_t)$, respectively, as:

$$\exp(\alpha_i^d \bar{r}^{d,i}(f_t) + \beta^d x^d + \xi^d) \equiv \mathbb{E}\Big[\exp(\alpha_i^d r_{j,t}^d + \beta^d x^d + \xi^d)|f_t \Big], \tag{25}$$

$$\exp(\alpha_i^l \bar{r}^{l,i}(f_t) + \beta^l x^l + \xi^l) \equiv \mathbb{E}\Big[\exp(\alpha_i^l r_{j,t}^l + \beta^l x^l + \xi^l)|f_t \Big], \tag{26}$$

where $\bar{r}^{l,i}(f_t)$ and $\bar{r}^{d,i}(f_t)$ summarize the choices of other banks, thereby allowing each bank to derive its choices of deposit and loan rates.

In solving the model, which we describe in detail in Appendix A, we ensure that $\bar{r}^{l,i}(f_t)$ and $\bar{r}^{d,i}(f_t)$ are consistent with equilibrium bank choices by iterating over their values until we reach convergence. To check the accuracy of our solution, we regress the simulated evolution of the aggregate deposit and loan rates on the perceived law of motion based on the banks' beliefs. The R-squareds for these approximations are over 95% for the deposit market and 99% for the loan market. Thus, although the banks do not consider the full distribution, Γ_t, when making their decisions, their forecasting errors are small. This accuracy stems from two mechanisms in the model. First, without any financial or regulatory constraints, banks have static optimal deposit and loan rates. For example, in the lending market, the optimal level of loans is set to equalize expected marginal interest income and funding costs, which is a function of the current federal funds rate only. Therefore, the static optimal rate depends only on the federal funds rate, but not other aggregate moments. Second, taking the loan market as an example, although banks can deviate from the static optimum by charging higher loan spreads, this behavior is limited by competition from other banks. Thus, banks that deviate from the static optimum introduce only modest distortions into the other banks' rate forecasts.

3.6. Monetary Policy Transmission in a Static Model

In this subsection, we use a simplified version of the model to flesh out the intuition behind the effects of frictions on monetary policy transmission.

3.6.1. Frictionless benchmark

First, we examine how the economy behaves in a frictionless benchmark model. By frictionless, we mean a version of our model with the following six 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, so $\phi^N = 0$; (3) the bank faces no capital requirement, so $\kappa = 0$; (4) there is no reserve requirement, so $\theta = 0$; (5) there is no maturity transformation; and (6) the operating costs, ϕ^d and ϕ^l, are zero. These features imply that the banks' problem can be viewed as static in nature.

In this case, banks choose deposit and loan rates, r^d and r^l, to maximize one-period profits:

$$\Pi = \max_{\{r^l, r^d\}} \left\{ r^l B - r^d D - (B - D) f \right\}. \tag{27}$$

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

The optimal lending rates are given by the federal funds rate plus markups:

$$r^l = f + \left(-\frac{B'}{B} \right)^{-1}, \tag{28}$$

and the optimal deposit rates are given by the federal funds rate minus markups:

$$r^d = f - \left(\frac{D'}{D} \right)^{-1}. \tag{29}$$

Under perfect competition, the elasticities, $-B'/B$ and D'/D, approach infinity, the markups in equations (28) and (29) go to zero, and r^l and r^d approach the federal funds rate, f.

3.6.2. Imperfect competition

When competition is imperfect, market power creates wedges between the federal funds rate and the rates at which banks borrow and lend. As shown in Figure 3, the optimal

lending rate is higher than the federal funds rate and the optimal deposit rate is lower than the federal funds rate because banks charge markups on loans and deposits. Furthermore, the size of these wedges is also a function of the federal funds rate because its level influences the attractiveness of bank deposits and loans relative to other outside options available to households or firms. Specifically, in the deposit market, an increase in the federal funds rate increases the opportunity cost of holding cash, which allows banks to charge larger markups on deposits (e.g., Drechsler et al. 2017). 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, lending shrinks and banks optimally lower their markups on loans to mitigate the effects of lower loan demand.[2] Overall, if banks have market power, they adjust markups charged on deposits and loans in response to policy rate movements. Therefore, the pass-through of the policy rate to deposit and lending rates deviates from one-to-one.

3.6.3. Balance sheet frictions

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 is:

$$\Pi = \max_{\{r^l, r^d\}} \left\{ r^l B - r^d D - \Phi(N) \right\}, \tag{30}$$

where $\Phi(N)$ is the cost of non-reservable borrowing and $N = B - D$ is the funding imbalance. The presence of $\Phi(N)$ implies that the bank cannot costlessly replace lost deposits with non-reservable borrowing, so shocks to deposits are transmitted to loans.

3.6.4. Reserve requirement

Now consider the case in which banks face a reserve regulation requiring that, for every dollar of deposits, the bank needs to keep a fraction, θ, of these deposits as reserves. Assuming

---

[2]The transmission of the federal funds rate to the lending rate differs from a standard cost pass-through problem as in (Bulow and Pfleiderer 1983), which concerns only the pass-through of marginal costs to prices, holding the demand side constant. In contrast, in our setting, the federal funds rate affects both banks' marginal funding cost and the attractiveness of borrowers' outside option.

that the interest on reserves is zero, banks' optimization becomes:

$$\Pi = \max_{\{r^l, r^d\}} \left\{ r^l B - r^d D - (B + R - D) f \right\}, s.t.\ R \geq \theta D. \tag{31}$$

Because the interest rate on reserves is zero, the reserve constraint is binding, so we can solve for the deposit spread, $f - r^d$, as θf. Intuitively, the opportunity cost of holding reserves, θf, rises with the federal funds rate, widening the deposit spread, $f - r^d$, as f rises. Deposit intake then falls. If banks face balance sheet frictions so that they cannot perfectly replace deposits with non-reservable funding, then the supply of loans falls together with deposits.

3.6.5. Capital regulation

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

$$\Pi = \max_{\{r^l, r^d\}} r^l B - r^d D - (B - D) f + \lambda (E_0 + \Pi - \kappa B). \tag{32}$$

where E_0 is initial capital, Π is bank profit, κ is the minimum capital requirement, and λ is the Lagrange multiplier on the capital constraint. Using the first order-condition for profit maximization, we can solve for the lending spread, $r^l - f$ as $\lambda \kappa$.

In the presence of capital regulation, equation (32) shows that shocks to bank capital (E_0) affect lending capacity. Although outside the scope of this static environment, one important source of such shocks is maturity mismatch. Because deposits are short-term, deposit rates rise when the federal funds rate increases. However, loans are long-term, so only a fraction of loans matures, with the remaining loans carrying the same pre-fixed rate. Hence, an increase in the federal funds rate temporarily reduces bank capital and tightens the bank capital constraint.

Another way that monetary policy affects bank capital is through market power. When the federal funds rate approaches zero, the markup that banks can charge on their deposits shrinks because of increasing competition from cash. A lower deposit markup depresses banks' profits, so the capital constraint binds more tightly. As a result, an ultra-low-interest policy may have a contractionary effect on bank lending.

4. Estimation

In this section, we describe our estimation method, present results, and conduct counter-factuals to assess the relative importance of various frictions for monetary policy transmission.

4.1. Estimation procedure

We estimate the model in two stages. First, we estimate the loan and deposit demand functions. Second, we plug these estimates into the model and use simulated minimum distance (SMD) to estimate the remaining parameters that describe banks' balance sheet frictions.

We estimate deposit and loan demand, given by equations (4) and (9), using the methods in Berry et al. (1995) and Nevo (2001), where the set of bank characteristics used in the demand estimation includes the number of branches, the number of employees per branch, bank fixed effects, and time fixed effects. We provide a brief outline of our implementation of this procedure, with a detailed explanation in Appendix B.

We start with the definition of a market. For both our deposit and loan demand estimation, we use the U.S. national market as the market definition, where each year constitutes a separate market. This choice is necessary because data on many types of loans (e.g. commercial and industrial loans) are not available at subnational levels. However, as shown in Appendix C, our results are robust to estimating deposit demand at the local level.

A key challenge in identifying demand elasticities is the natural correlation between either deposit or loan rates and any unobservable demand shocks, ξ_j^d, that move the error terms in the estimating equations. For example, a positive deposit demand shock can induce banks to lower deposit rates. Therefore, we use a set of supply shifters as instrumental variables. In particular, following Dick (2007) and Ho and Ishii (2011), we use salaries and non-interest expenses related to the use of fixed assets. Our identifying assumption is that customers do not care about these costs, holding product characteristics constant. In this case, 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. Note that

our identification of the demand curve does not use variation in monetary policy. In fact, any aggregate shocks are absorbed by time fixed effects, so our identification strategy avoids a common challenge that studies of monetary transmission face, namely, the endogeneity of monetary policy and aggregate bank credit supply.

From this demand estimation, we obtain fitted values of the right-hand sides of equations (3) and (8) as follows:

$$D_j\left(r_j^d\,|f\right) = \sum_{i=1}^{I}\mu_i^d\frac{\exp\left(\hat{\alpha}_i^d r_j^d + q_j^d\right)}{\exp\left(\hat{\alpha}_i^d f\right) + \exp\left(q_c^d\right) + \sum_{m=1}^{\hat{J}}\exp\left(\hat{\alpha}_i^d r_m^d + q_m^d\right)}W \tag{33}$$

$$B_j\left(r_j^l\,|f\right) = \frac{\exp\left(\hat{\alpha}^l r_j^l + q_j^l\right)}{\exp\left(\hat{\alpha}^l\left(\overline{f}+\overline{\delta}\right)\right) + \exp\left(q_n^l\right) + \sum_{m=1}^{\hat{J}}\exp\left(\hat{\alpha}^l r_m^l + q_m^l\right)}K, \tag{34}$$

where $\hat{J}$ is the number of banks used in the second-stage estimation, and q generically stands for an option's quality value, which is the utility derived from non-rate product characteristics. As shown in Appendix B, q equals the fitted value of $\beta x + \xi$. We normalize to zero the quality values of saving via Treasury bills and of borrowing from the bond market. Two features of equations (33) and (34) are of note. First, we assume homogeneous sensitivity to loan rates, as allowing for heterogeneous sensitivities slows down the estimation but minimally affects banks' rate-setting decisions. Second, we cannot estimate the quality value of not borrowing, q_n^l, from the demand estimation because we do not observe its share, so we estimate it in our second-stage estimation.

The final plug-in problem consists of inserting the estimated demand functions described in equations (33) and (34) into the banks' dynamic problem (24). This plug-in problem operationalizes the notion that banks set deposit and loan rates facing the above-specified demand curves for deposits and loans.

In the second stage, we estimate seven additional parameters using SMD, which chooses parameter values that minimize the distance between the moments generated by the model and their analogs in the data. We use ten moments to identify the remaining seven model parameters. Parameter identification in SMD 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 deposits to identify the cost of holding non-reservables, ϕ^N. Intuitively, higher 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. Higher marginal costs lead banks to charge higher spreads in both deposit and loan markets. Next, we use two moments to identify the net fixed operating cost, ψ. The first is average net non-interest expenses, scaled by assets. This moment measures the costs that banks pay outside of their routine deposit-intaking and loan-servicing business. The second moment is banks' average leverage ratio, which indirectly reflects fixed operating costs, as higher fixed costs induce banks to operate with lower leverage. Next, we use banks' average dividend yield to identify the discount rate, γ. Intuitively, a high discount rate makes banks impatient, so they pay out more of their profits to shareholders instead of retaining the funds to finance future business. Next, to identify the relative size of the deposit market, W/K, as well as the value of firms' outside option of not borrowing, q_n^l, we include banks' average deposit-to-asset ratio and the sensitivity of total borrowing to the federal funds rate, which we estimate using a vector autoregression (VAR), the details of which are in Appendix D. These two moments suit this purpose because holding banks' market shares constant, when W/K increases, the value of deposits rises relative to loan value, leading to a higher deposit-to-asset ratio. Next, when the outside option becomes less valuable, its market share remains low, regardless of the federal funds rate. Thus, the sensitivity of aggregate corporate borrowing to the federal funds rate falls as q_n^l falls. In addition, the high loan-to-deposit ratio is inversely related to q_n^l because loan demand is weaker when firms value the option to not borrow. Finally, we also include banks' average market-to-book ratio to ensure that our model predicts the right valuation for banks.

4.2. Baseline Estimation Results

Table 3 presents the point estimates for the model parameters. In Panel A, we start with the statutory parameters. We set the corporate tax rate to its statutory rate of 35% and the capital requirement to 6% according to Basel accord. According to the Federal Reserve Board's Regulation D, the reserve ratios are 10% for transaction deposits, 1% for saving deposits, and 0% otherwise. In our model, we include only one type of deposit, so our estimate of the reserve 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 calibrate the number of representative banks to be 6, which matches the average county-level banking concentration in the data.

Panel B presents the parameters that we can directly quantify in the data. Specifically, we obtain the means, standard deviations, and autocorrelations of the federal funds rate and the bank-level loan default rate by direct estimation of the equation (22). Next, average loan maturity, defined in Table 1, is approximately 3.5 years.

Panel C in Table 3 provides the demand parameters from the first-stage BLP estimation, with details of the estimation results presented in Table E.1 in Appendix E. 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. A one percentage point increase in the deposit rate increases a bank's market share by 0.968%, while a one percentage point increase in the loan rates decreases its market share by 1.424%. We also find that depositors exhibit significant dispersion in their rate sensitivity. Finally, we estimate depositors' and borrowers' sensitivities to non-rate bank characteristics. The estimates are also both statistically and economically significant. A one percentage point increase in the number of branches increases a bank's market share by 0.804% in the deposit market and 0.944% in the loan market. In comparison, the sensitivity to the number of employees per branch is smaller. A one percentage point increase in the number of employees per branch increases a bank's market share by 0.714% in the deposit market and 0.630% in the lending market.

Panel D of Table 3 presents the parameters from our second-stage SMD estimation. We find that banks have a subjective discount rate of 4.5%, which is higher than the average federal funds rate observed in the data. Given the discount rate, banks pay out 3% of their equity value as dividends. Next, the cost of non-reservable borrowing is both statistically and economically significant. At the average level of non-reservable borrowing (30% of total deposits), a marginal dollar of non-reservable borrowing costs a bank 30 basis points above the cost implied by the prevailing federal funds rate, where we calculate this cost as $\partial\Phi^N/\partial N = \phi^N N = 0.010 \times 0.3 = 0.003$. Note that the mean deposit spread is 1.3%. Because banks equate the marginal costs of their funding sources, these numbers imply that the marginal cost of expanding deposits averages 1.6%. 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 0.9% cost of maintaining deposits and a slightly lower 0.7% cost of servicing their outstanding loans.

In Table 4, we compare the empirical and model-implied moments. The model is able to match closely the banks' balance sheet quantities, the spreads they charge, and their valuations. In both the data and the model, banks borrow non-reservable securities, which amount to 30% of the deposit intake. In both the model and the data, the spreads that banks charge in the deposit market are significantly smaller than those in the loan market. When the federal funds rate is low, as it is in much of our sample, banks face stiffer competition from cash, so they shrink deposit spreads.

As an external model validation, in Figure 4 we plot the relation between banks' deposit and loan rates and the federal funds rate, as implied by our model and as calculated from our data, where we accompany the data calculations with a quarterly scatter plot. Note that we do not use these relations in our moment-matching exercise, as we target only the average levels of these rates. Figure 4 shows that the pass-through of the federal funds rate in both the deposit and loan markets is less than one to one, as indicated by the less than unitary slope of the plots. This result is consistent with the message in Table 3 and Table E.1 in Appendix

E, as it suggests that banks have significant market power. In addition, our model-predicted deposit and loan rates track the pattern that we see in the actual data, indicating that our model can quantitatively capture banks' pricing of their products in both the deposit and loan markets.

In a similar spirit, we examine how the market value of bank equity reacts to an unexpected federal funds rate shock. In the data, a one percentage point increase in the federal funds rate leads to a 1.93% drop in bank equity value. Although our model is not geared to match asset-pricing moments, the same magnitude shock generates a 2.84% drop in equity value. This result is important because models without market power can over-predict this response. For example, if we model a bank as a replicating portfolio with a long position in 3.5-year Treasuries and a short position in the federal funds rate with a leverage ratio of 10, then we find an 18% drop in equity value.[3] On an intuitive level, bank market power matters for this result because it allows banks to borrow at deposits rates that are relatively insensitive to the policy rate, thereby dampening the impact of rate hikes on their net interest margin (Drechsler et al. 2017). This result highlights the importance of accounting for bank market power when analyzing banks' reactions to monetary policy shocks.

5. Counterfactuals

5.1. Decomposing Monetary Policy Transmission

We now examine the quantitative forces that shape the relation between aggregate bank lending and monetary policy, as embodied in the federal funds rate. Table 5 depicts, for different versions of our model, the percentage change in aggregate bank lending in response to a one percentage point change in the federal funds rate. In row (1), for our model exactly as specified in Section 3, with all frictions as estimated in Table 3, find a response of 1.55%. As shown in Appendix D, in our data, the impulse response of aggregate lending to federal funds rate shocks is 1.6% over a three-year horizon. Although we do not target this sensitivity, it is

---

[3]A maturity mismatch of 3.5 years implies a drop in asset value of 3.5%. A book leverage ratio of 10 and a market-to-book ratio of approximately two implies that equity value drops by $18\% = 3.5\% \times 10/2$.

remarkably close to the figure produced by the model.

We proceed by eliminating from the model the regulatory constraints and banks' market power one at a time. As such, we analyze how the absence of each of these frictions affects the transmission of monetary policy. Row (2) in Table 5 presents the results from a version of the model without the reserve requirement. We find that the sensitivity of bank lending to the federal funds rate decreases by 8%. This modest magnitude reflects the small amount of non-interest-bearing reserves held by banks in our sample period.[4] As a result, monetary policy has a limited effect on banks' marginal costs of lending through the reserve requirement.

This result provides insights into a recent policy debate over interest on excess reserves. In October 2008, the Federal Reserve started paying interest on reserves. This move spawned worry over the power of monetary policy to affect bank lending. For instance, a January 1, 2019 *Wall Street Journal* article titled "The Fed's Obama-Era Hangover" argues, "by paying banks not to lend, the central bank diminished its ability to control interest rates" (Gramm and Saving 2019). Another example is in an article in the *American Banker* on June 22, 2019 titled "Fed must stop rewarding banks for not lending," which says, "It was thanks to interest on excess reserves that the Fed ended up stimulating so little in the economy, despite its efforts to ease so much" (Michel and Selgin 2019). However, we find that this concern is unwarranted, as the bank reserve transmission channel is not important during our sample period. This result is also consistent with Xiao (2018), who shows that the reserve requirement is not a quantitatively important feature that distinguishes commercial banks from shadow banks.

Row (3) in Table 5 presents the results from a version of the model that excludes the capital requirement. We find that the presence of the capital requirement enhances monetary policy transmission by 28% ($1.548/1.120 - 1$). This result connects two long-standing sets of reduced-form evidence on the bank capital channel. The first set establishes that monetary policy shocks can trigger movements in bank capital because of the maturity mismatch on banks' balance sheets (Flannery and James 1984; English, Van den Heuvel, and Zakrajšek

---

[4]Since 2008, bank reserves have increased substantially. However, in this period, reserves started bearing interest, which effectively eliminated the reserve channel.

2018). The second set shows that bank capital has an economically significant impact on bank lending (Peek and Rosengren 2000; Mora and Logan 2012). Because of the endogeneity of bank capital, this second literature often exploits exogenous shocks to bank capital instead of directly focusing on the role of monetary policy. Our paper bridges the two bodies of empirical evidence by connecting monetary policy to bank lending through the bank capital requirement. Moreover, we measure quantitative magnitude of this long-established channel.

Row (4) of Table 5 shows the results from removing banks' deposit market power from the model. In this case, banks receive fixed lump-sum profits equal to their oligopolistic profits in the baseline case. They also use marginal cost pricing for deposit-intake decisions, setting the deposit rate equal to the federal funds rate minus the bank's marginal cost of servicing deposits. They then take as many deposits as depositors offer, given the deposit rate. We find that once we eliminate market power in the deposit market, bank lending becomes less sensitive to changes in the federal funds rate. A 1% increase in the federal funds rate causes an almost one-to-one decrease in aggregate lending. This sensitivity is 36% ($1.548/0.992 - 1$) smaller than the 1.55% sensitivity observed in the baseline case. Moreover, the change in sensitivity is larger than that observed when we eliminate the capital requirement.

Intuitively, if deposit market power is in place, when the federal funds rate increases, the households' opportunity cost of holding cash rises, making cash less attractive relative to bank deposits. Banks react by charging higher deposit spreads, thus lowering the amount of deposits. Banks' lending decisions partially echo this decline in deposits because they need to use expensive non-reservable borrowing to finance their loans when the amount of loans exceeds deposits. Thus, bank market power, combined with the non-reservable borrowing cost, contributes to a negative relationship between bank lending and the federal funds rate. Finally, this result is important because it highlights the interconnectedness of banks' deposit-taking and lending businesses. Banks' market power in the deposit market is passed on to the loan market and contributes to the sensitivity of bank lending to the federal funds rate.

Row (5) of Table 5 shows the results from removing banks' lending market power from the

model. To isolate the effect of loan market power, we allow banks to retain their oligopolistic market power in the deposit market, but in the loan market, we assume banks act as price takers. They adopt marginal cost pricing and set their loan rates equal to the funding cost. We find that the presence of banks' loan market power makes the aggregate quantity of loans less sensitive to the federal funds rate, with the sensitivity changing by -24% $(1.548/1.910 - 1)$. This result quantifies the intuition in Scharfstein and Sunderam (2016) and Corbae and Levine (2019), who argue and show that loan market power allows banks to cushion the effects of monetary tightening on lending by reducing markups on loans to mitigate reductions in loan demand induced by higher rates.

Overall, we find that monetary transmission channels based on market power have comparable, if not larger, effects than channels channels based on regulation. Thus, our findings highlight the importance of accounting for the banking system market structure in assessing monetary transmission mechanisms.

5.2. Reversal Rate

Does the sensitivity of lending to the federal funds rate depends on its level? In Panel A of Figure 5, we show the amount of bank lending that corresponds to different levels of the federal funds rate, where we normalize steady-state lending to one. We find that aggregate bank lending in the economy is hump-shaped. When the federal funds rate rises above a certain threshold, a further increase has the usual effect of tightening lending. However, when the rate is below the threshold, a rate increase actually expands lending.

In Panel B, we plot the impulse response of aggregate bank lending to a monetary policy shock, conditional on the current federal funds rate. A point (x, y) in the plot means that if the current federal funds rate is x, aggregate lending changes by y percent after a one percentage point federal funds rate shock. As expected, we find a negative overall relation between the federal funds rate and lending. This negative relation is reversed, however, in a region where this rate is below 0.9%. We call this region a reversal-rate environment.

To understand the mechanism behind the reversal rate, in Panel C, we plot the amount of

desired bank lending in a world with no capital requirements, and in Panel E, we plot the level of bank capital. The corresponding impulse responses are in Panels D and F. First, we see that desired lending always falls with the federal funds rate. However, the relation between bank equity capital and the federal funds rate is also hump shaped, with a 2% turning point at which the relation between bank equity and the federal funds rate flips sign.

Two patterns in Panels C and E underlie the reversal rate. The intuition behind Panel C is straightforward, as high funding costs deter firms from borrowing in equilibrium. The intuition behind Panel E is more intricate and depends on the relative profitability of lending and deposit taking. First, as the federal funds rate rises, depositors increasingly find holding cash to be costly, so banks face weaker competition from cash in the deposit market. Hence, bank profits from the deposit market rise with the federal funds rate. Second, bank profits from lending decrease with the federal funds rate, as higher funding costs make firms' outside option of not borrowing more appealing. Our parameter estimates imply that the deposit market exerts more pressure on profits than the lending market when the federal funds rate is low. Thus, an increase in the rate leads to higher bank profits, which increases the equity capital base. Banks accumulate equity capital instead of paying out their profits to shareholders to avoid being capital constrained in the future. In contrast, in a region of high federal funds rates, further rate increases erode bank capital via a standard maturity mismatch argument.

The reversal rate in Panel A arises because optimal lending is the smaller of two quantities: desired and feasible lending. The former is the optimal amount of lending in the absence of a capital requirement, and the latter is the maximal lending permitted by a bank's equity capital. When the federal funds rate is low, given firms' equilibrium heavy demand for loans in a low-interest-rate environment, desired lending exceeds the amount allowed by the bank's equity. Thus, the capital requirement binds, and actual lending tracks the bank's equity capital, which increases with the federal funds rate. When this rate is high, the capital requirement is slack, and the actual quantity of lending is the desired amount.

To understand more fully the dynamic response of bank lending to monetary policy shocks,

in Figure 6, we report the simulated response of bank lending to federal funds rate shocks. The economy starts at time zero in an initial steady state with the federal funds rate equal to the inflection point of 0.9%. At time 1, the federal funds rate either increases to 2% or decreases to 0.1%, 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 initial steady state.

Panel A of Figure 6 depicts the response to an increase in the federal funds rate. In this case, banks faces less competition from household demand for cash in the deposit market. Thus, they behave more like monopolists by charging higher spreads, which in turn lower household deposit demand. Lower deposit intake increases the need for banks to fund their lending by turning to the market for non-reservable borrowing, which carries increasing marginal costs. A positive federal funds rate shock also increases the cost of capital in the corporate sector, making firms more likely to switch to the outside option of not borrowing. Both effects shrink lending. Because deposits have shorter durations than loans, deposits drop sharply and converge almost instantaneously to the new steady state. In contrast, loan quantity converges slowly as the bank replaces only a fraction, η, of its long-term loans in each period. Non-reservable borrowing increases to fill the gap between deposits and loans.

Panel B of Figure 6 depicts the responses to a decrease in the federal funds rate. On the one hand, when this rate decreases, banks profit from having a maturity mismatch on their balance sheets as the rates they pay on short-term liabilities decrease instantly, while most of their long-term assets keep generating higher rates of return. This effect diminishes gradually over time as existing loans mature and are repriced. On the other hand, a lower federal funds rate leads to increasingly intense competition from cash in the deposit market. The effect is especially strong as the federal funds rate approaches the zero lower bound, in which case, the spreads that banks can charge in the deposit market are squeezed, leading to a sharp drop in their profits. Given the persistence of the federal funds rate, lower profits translate into slower retained earnings accumulation and lead to lower bank capital. In the new steady state, banks take more deposits, which can support increased lending. Indeed, lending increases in the first

year. However, banks cannot sustain this higher level of lending, as their capital requirements tighten in this extremely low federal funds rate environment. Because total lending decreases, banks need less external financing and thus less non-reservable borrowing.

Note that in Figure 6, the lending falls when the federal funds rate changes in either direction. Although the lending moves in the same direction, the driving force differs 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.

We obtain external validation of the reversal-rate prediction from a reduced-form regression in which we examine the relation between bank equity returns and monetary policy news on Federal Open Market Committee (FOMC) meeting days. We measure monetary policy news released during FOMC meetings as changes in the two-year Treasury yield on FOMC meeting days, following Hanson and Stein (2015). 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 FOMC announcements, which has become increasingly important in recent years (Hanson and Stein 2015).[5] The identifying assumption is that unexpected changes in interest rates in a one-day window surrounding scheduled Federal Reserve announcements arise largely from news about monetary policy because macroeconomic fundamentals would not change discretely within such a short window. While our sample period runs from 1994 to 2017, we exclude the dot-com bubble collapse (2000–2001) and the financial crisis (2007–2009) because, in these crisis times, information other than conventional monetary policy news could also be released in FOMC meetings.[6]

---

[5]As shown in Appendix F, our results are robust to using use one-year Treasury yields.

[6]For instance, on January 28, 2009, the FOMC expressed its intent to purchase "large quantities of agency debt and MBS ... and stands ready to expand the quantity of such purchases and the duration of the purchase program as conditions warrant." This unusual action signaled to the market that the Fed was willing to support the banking system, leading to 12% one-day banking stock returns. Thus, in this event, the Fed affected stock

Our model suggests that the relation between bank capital and the federal funds rate changes sign around 2%, as shown in Panel E of Figure 5. Therefore, we split our data sample using a 2% cutoff for the federal funds rate. In Table 6, we report the regression estimates. As shown in column (1), when rates are high, interest rates and returns are negatively related. However, this conventional negative relation reverses sign in a low-rate environment. As shown in column (2), a rate increase is associated with positive significant bank equity returns. In other words, the market expects that an increase in rates will lead to an increase in bank capital. This result is not driven by a steepening of the term structure, as we control for changes in term spreads. As shown in Figure 7, the contrast between the results in columns (1) and (2) can be seen in a simple scatter plot of bank industry excess returns against monetary policy shocks on FOMC days.

It is possible that the positive relation between monetary policy shocks and bank stock returns during low-interest-rate environments could be driven by the central bank's economic outlook (Nakamura and Steinsson 2018). To address this concern, in Appendix F, 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 to a positive interest sensitivity in the low-interest environment. Furthermore, we use an alternative measure of monetary policy shocks constructed by Jarocinski and Karadi (2018), who disentangle the potential information shocks from monetary policy shocks. Our results are robust to this alternative measure.

In columns (3) and (4) of Table 6, we report the results of interacting changes in the federal funds rate with a reduced-form measure of deposit market power, the Herfindahl-Hirschman Index (HHI) of the local deposit market in which a bank operates, where we define a local deposit market as a county. If a bank operates in several counties, the bank-level HHI is the weighted average of local HHIs, weighted by the deposits of the bank in each county. We find that in a low-interest-rate environment, banks with greater deposit market power experience higher positive returns. 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

prices through the Fed put channel (Cieslak and Vissing-Jorgensen 2017), which is outside our model.

rates and bank capital is negative, but when this rate is low, the relation is positive.

Our finding that ultra-low interest rate policy impedes lending growth is related to Heider et al. (2018), who show that the introduction of negative policy rates by the European Central Bank in mid-2014 reduced lending by eurozone banks with greater reliance on deposit funding. Our results show that such perverse effects can occur even when the rates are positive because banks' deposit spreads are compressed by the low-rate environment.

The reversal-rate result is informative about the sluggish recovery of bank lending in the United States since the 2008 financial crisis. By the end of 2018, bank lending had increased only about 25% cumulatively from its low in August 2009. In contrast, from trough to peak, in all recessions since 1974, bank lending grew by 60%–120%. Although many factors such as banking regulation may have contributed to this slow recovery, the ultra-low rate policy could be an important factor.

6. Extensions and Robustness

6.1. Heterogeneous Transmission through Large and Small Banks

We extend our analysis by examining whether monetary transmission depends empirically on bank size, motivated by the finding in Kashyap and Stein (1995) that monetary policy has a particularly strong effect on lending by small banks, as they cannot replace deposits with frictionless access to non-reservable funding. However, one limitation of the data they use is the absence of a measure of the actual cost of external finance. While they use bank size as a proxy for this cost, size is correlated with many other bank attributes, whose presence compromises the interpretation of their results. For instance, small banks tend to lend to small firms, whose credit demand is more cyclical. Therefore, the higher sensitivity of smaller bank lending might be driven by demand rather than by financing frictions on the supply side.

To explore this issue, we split our sample at the 10th percentile of the bank size distribution and estimating a subset of the model parameters separately for the large and small banks. For these estimations, we hold constant across subsamples parameters describing household

preferences and macroeconomic conditions. We also fix the banks' discount rate because we cannot identify it in the small-bank sample. Identification using dividend yields is infeasible because many small banks do not report dividends. Also, many small banks are private, so we cannot calculate their market-to-book ratios. We re-estimate all the remaining parameters.

The results are in Table 7. We find that both external financing costs, ϕ^N, and fixed operating costs, ψ, are larger for small banks. This result is striking because we do not use bank size as an identifying moment. Instead, ϕ^N is identified from the fraction of assets financed by non-reservables, and ψ is identified from the banks' net non-interest expense and leverage ratios. These parameter differences across samples imply that small-bank loans are 37% more sensitive to the federal funds rate, with a -1.758 sensitivity for small banks and a -1.103 sensitivity for large banks.

Next, we examine how the sensitivity of lending to the federal funds rate depends on the frictions embodied in ϕ^N and ψ. If we reduce the small-bank financing cost, ϕ^N, to the level estimated for the large banks (0.006), the small-bank sensitivity of loans to the federal funds rate falls to -1.364, a drop representing 61% of the difference in loan sensitivity between large and small banks. Because we allow only four parameters to differ across large and small banks, we effectively hold loan demand constant, thus isolating the effect of financial frictions. This result is consistent with the hypothesis in Kashyap and Stein (1995) that large and small banks have different external financing costs, which lead to differences in the transmission of monetary policy to credit supply. The fixed operating cost, ψ, explains most of the remaining variation in monetary transmission. Small banks have relatively few sources of non-interest income, so they have higher net operating costs, and they accumulate equity buffers more slowly. As a result, when rate hikes erode bank capital via the maturity mismatch, this effect hits small banks more strongly, leading to a sharper reduction in lending.

6.2. Changes in Transmission over Time

In this subsection, we examine how the impact of monetary policy on bank lending has changed over time. For example, in the past few decades, the average interest rate has declined

substantially, and the banking industry itself has experienced a large volume of mergers, leading to increased concentration.

We split our sample into two sub-periods—early (1994–2005) and late (2006–2017). We then reestimate all model parameters for the two sub-periods. The parameter estimates in Table 8 imply that the sensitivity of lending to the federal funds rate has declined from 1.727% to 0.987% over time. This result is consistent with evidence in the literature that monetary policy has had more muted effects on real activity and inflation in recent decades (Boivin, Kiley, and Mishkin 2010).

To understand the declining impact of monetary policy on bank lending, we categorize our parameters into three groups. The first group includes macroeconomic conditions, such as the federal funds rate and the loan charge-off processes, the regulatory constraints, and the loan and deposit market sizes. The second group consists of measures of bank operating efficiency and financial frictions, namely discount rates, operating costs, and external financing costs. The last group consists of parameters that govern banks' market power. It includes the number of competing banks in the local market, $\hat{J}$, as well as the rate sensitivities that banks face in the deposit and loan markets, α^d and α^l. As seen in Table 8, over time, banks' market concentration has risen, with the number of competing banks in the local market falling from 7 to 5. However, both depositors and borrowers have become more rate-sensitive. The adoption of new technology and a surge in Internet and mobile banking has lowered the cost of searching. Thus, deposits and borrowers are more reactive to banks' rate setting. Holding all else equal, this increased sensitivity decreases banks' market power.

To gauge the overall effect of bank market power on the observed change in the federal funds rate sensitivity, we eliminate the difference in bank market power parameters across the two subsamples by setting the late period values to those from early period. We find that the gap between the early and late sensitivities declines by 18%.

Table 8 also shows that fixed operating costs have fallen, likely because of bank mergers. The consequent rise in profitability allows banks to accumulate healthier capital buffers

that reduce their exposure to monetary policy. Furthermore, the cost of accessing the non-reservable funding market has declined, so banks can better cushion fluctuations in deposits. If we eliminate the difference in bank operating and financing costs, the gap between the early and late sensitivities declines by 16%.

The remaining 66% of the gap is attributable to changes in macro-economic conditions. In particular, we find that changes in the federal funds rate process play the most important role in explaining the declining trend in the sensitivity of lending to the federal funds rate. In particular, the average federal funds rate is much lower in our late period, so the economy spends more time around the reversal-rate region, where monetary policy has a weaker, or even opposite, effect on bank lending decisions.

6.3. Model Robustness

In this subsection, we examine the implications of several ingredients that we have left out of the baseline model. First, instead of requiring dividends to be positive, we allow them to be negative, subject to a linear equity issuance cost, ϕ^e. We re-estimate our model, with the parameter ϕ^e being identified by an additional moment: the ratio of bank equity issuance to total assets. In the data, this moment is 2%. In our model, we find that an equity issuance cost of 10% yields a predicted equity issuance rate that matches the value in the data. The value of ϕ^e is comparable the estimates for industrial firms in Hennessy and Whited (2007).

Second, we introduce time-varying discount rates. Specifically, we assume that banks apply a discount rate of $f_t + \omega$. In our estimation, we identify ω from banks' dividend yields, the same moment that we use to identify the constant discount factor in our baseline model. We find that $\omega = 1.6\%$. As is the case with the baseline estimates, the spread between the federal funds rate and banks' discount rates indicates that banks face substantial frictions in their maturity transformation activity. Finally, our results concerning the decomposition of monetary transmission and reversal rate are robust to these alternative model specifications.

Another concern is that agents in our model are risk-neutral, while in reality loan spreads contain a risk premium. To make the model and data moments comparable, we adjust the

data moment by subtracting a risk premium, which we calibrate following Giesecke, Longstaff, Schaefer, and Strebulaev (2011). They show that the credit risk premium in the bond market roughly equals the expected default loss. After adjusting the data moment, we re-estimate the model. We find that the only notable difference in the new results lies in banks' estimated cost of serving loans, which becomes insignificantly different from zero. This result suggests that banks' perceived costs in the loan market are driven almost entirely by the default risk. Finally, we verify that our main predictions remain robust under this re-estimation.

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 limited quantitative importance. In contrast, we find that channels related to bank capital requirements and market power are very important. These quantitative findings are new to an empirical literature dominated by qualitative results (Kashyap and Stein 1995; Scharfstein and Sunderam 2016; Drechsler et al. 2017). 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 reduced profits in the deposit market impact bank capital negatively. Lastly, we show that accounting for bank market power is key to understanding cross-sectional variation in banks' responsiveness to monetary policy, while the interaction of bank market power with regulatory constraints explains most of the decline in monetary transmission effectiveness over time.

References

Basten, Christoph, and Mike Mariathasan, 2018, How banks respond to negative interest rates: Evidence from the Swiss exemption threshold, CESifo Working Paper Series 6901.

Bazdresch, Santiago, R. Jay Kahn, and Toni M. Whited, 2018, Estimating and testing dynamic corporate finance models, *Review of Financial Studies* 31, 322–361.

Begenau, Juliane, 2018, Capital requirements, risk choice, and liquidity provision in a business cycle model, *Journal of Financial Economics*, forthcoming.

Begenau, Juliane, and Tim Landvoigt, 2018, Financial regulation in a quantitative model of the modern banking system, Manuscript, Stanford University.

Bernanke, Ben S., and Alan S. Blinder, 1988, Credit, money, and aggregate demand, *American Economic Review* 78, 435–439.

Berry, Steven, James Levinsohn, and Ariel Pakes, 1995, Automobile prices in market equilibrium, *Econometrica* 63, 841–90.

Boivin, Jean, Michael T. Kiley, and Frederic S. Mishkin, 2010, How has the monetary transmission mechanism evolved over time?, in *Handbook of monetary economics*, volume 3, 369–422 (Elsevier).

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, 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.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru, 2018, Fintech, regulatory arbitrage, and the rise of shadow banks, *Journal of Financial Economics* 130, 453–483.

Bulow, Jeremy I., and Paul Pfleiderer, 1983, A note on the effect of cost changes on prices, *Journal of Political Economy* 91, 182–185.

Campos, Mauricio, 2019, Going negative at the zero lower bound: The effects of negative nominal interest rates, Working paper.

Cieslak, Anna, and Annette Vissing-Jorgensen, 2017, The economics of the Fed put, Manuscript, Duke University.

Corbae, Dean, and Pablo D'Erasmo, 2013, A quantitative model of banking industry dynamics, Manuscript, University of Wisconsin.

Corbae, Dean, and Ross Levine, 2019, Competition, stability, and efficiency in the banking industry, Manuscript, University of Wisconsin.

Demiralp, Selva, Jens Eisenschmidt, and Thomas Vlassopoulos, 2017, Negative interest rates, excess liquidity and bank business models: Banks' reaction to unconventional monetary policy in the Euro area, ECB Working Paper.

Dick, Astrid A, 2007, Market size, service quality, and competition in banking, *Journal of Money, Credit and Banking* 39, 49–81.

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2017, The deposits channel of monetary policy, *Quarterly Journal of Economics* 132, 1819–1876.

Egan, Mark, Ali Hortacsu, and Gregor Matvos, 2017a, Deposit competition and financial fragility: Evidence from the U.S. banking sector, *American Economic Review* 107, 169–216.

Egan, Mark, Stefan Lewellen, and Adi Sunderam, 2017b, The cross section of bank value, Manuscript, Harvard University.

Eggertsson, Gauti B., Ragnar E. Juelsrud, and Ella Getz Wold, 2017, Are negative nominal interest rates expansionary?, Manuscript, Brown University.

Elenev, Vadim, Tim Landvoigt, and Stijn Van Nieuwerburgh, 2016, Phasing out the GSEs, *Journal of Monetary Economics* 81, 111–132.

English, William B, Skander J. Van den Heuvel, and Egon Zakrajšek, 2018, Interest rate risk and bank equity valuations, *Journal of Monetary Economics* 98, 80–97.

Flannery, Mark J, and Christopher M James, 1984, The effect of interest rate changes on the common stock returns of financial institutions, *Journal of Finance* 39, 1141–1153.

Giesecke, Kay, Francis A. Longstaff, Stephen Schaefer, and Ilya Strebulaev, 2011, Corporate bond default risk: A 150-year perspective, *Journal of Financial Economics* 102, 233–250.

Gramm, Phil, and Thomas Saving, 2019, The Fed's Obama-era hangover, *The Wall Street Journal*.

Hanson, Samuel G., and Jeremy C. Stein, 2015, Monetary policy and long-term real rates, *Journal of Financial Economics* 115, 429–448.

Heider, Florian, Farzad Saidi, and Glenn Schepens, 2018, Life below zero: Bank lending under negative policy rates, *Review of Financial Studies*, forthcoming.

Hennessy, Christopher A., and Toni M. Whited, 2005, Debt dynamics, *Journal of Finance* 60, 1129–1165.

Hennessy, Christopher A., and Toni M. Whited, 2007, How costly is external financing? Evidence from a structural estimation, *Journal of Finance* 62, 1705–1745.

Ho, Katherine, and Joy Ishii, 2011, Location and competition in retail banking, *International Journal of Industrial Organization* 29, 537–546.

Jarocinski, Marek, and Peter Karadi, 2018, Deconstructing monetary policy surprises: The role of information shocks, *American Economic Journal: Macroeconomics*, forthcoming.

Kashyap, Anil K., and Jeremy C. Stein, 1995, The impact of monetary policy on bank balance sheets, *Carnegie-Rochester Conference Series on Public Policy* 42, 151–195.

Krusell, Per, and Anthony A. Smith, 1998, Income and wealth heterogeneity in the macroeconomy, *Journal of Political Economy* 106, 867–896.

Michel, Norbert, and George Selgin, 2019, Fed must stop rewarding banks for not lending, *American Banker*.

Mora, Nada, and Andrew Logan, 2012, Shocks to bank capital: Evidence from UK banks at home and away, *Applied Economics* 44, 1103–1119.

Nakamura, Emi, and Jón Steinsson, 2018, High-frequency identification of monetary non-neutrality: The information effect, *Quarterly Journal of Economics* 133, 1283–1330.

Nevo, Aviv, 2001, Measuring market power in the ready-to-eat cereal industry, *Econometrica* 69, 307–342.

Peek, Joe, and Eric S. Rosengren, 2000, Collateral damage: Effects of the Japanese bank crisis on real activity in the United States, *American Economic Review* 90, 30–45.

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, Manuscript, Harvard Business School.

Van den Heuvel, Skander J., 2002, The bank capital channel of monetary policy, Manuscript, University of Pennsylvania.

Wang, Olivier, 2019, Banks, low interest rates, and monetary policy transmission, Manuscript, University of Pennsylvania.

Xiao, Kairong, 2018, Monetary transmission through shadow banks, *Review of Financial Studies*, forthcoming.

Table 1: Variable Definitions

Variable	Details of construction
Deposit market share	Deposits of a bank divided by the sum of deposits, cash, and Treasury bills in the U.S. economy.
Loan market share	Loans of a bank divided by the sum of U.S. corporate and household debt.
Deposit rates	Deposit interest expense divided by deposits.
Loan rates	Loan interest income divided by loans outstanding.
No. of branches	Number of local branches.
No. of employees per branch	Number of employees divided by number of branches.
Expenses related to fixed assets	Non-interest expenses related to the use of fixed assets divided by total assets.
Salary	Total salary expense divided by total assets.
Reserve ratio	10% times the weight of transaction deposits plus 1% times the weight of saving deposits.
Average loan maturity	Estimated maturity of each type loan weighted by the portfolio weight. Non-mortgage loan maturity is the repricing maturity and average prepayment adjusted mortgage duration is from Elenev et al. (2016).
Non-reservable borrowing share	Non-reservable borrowing divided by total deposits.
Deposit spread	federal funds rate minus a deposit rate.
Loan spread	A loan rate minus the corresponding five-year Treasury yield.
Deposit-to-asset ratio	Deposits divided by total assets.
Net noninterest expense	Noninterest expense minus noninterest income, divided by total assets.
Leverage	Total assets divided by the book value of equity.
Market-to-book ratio	The market value of equity divided by the book value of equity.
Total credit-FFR sensitivity	Three-year impulse response coefficient of total credit to the federal funds rate estimated from a VAR.

Table 2: Summary Statistics

In this table, we report summary statistics for our sample. The sample period is 1994–2017. The total size of the deposit market is defined as the sum of deposits, cash, and Treasury bills held by all U.S. households and non-financial corporations. The total size of the loan market is defined as the sum of U.S. corporate and household debt. Deposit and loan rates are calculated using interest expense and income. Expense related to fixed assets and salaries are scaled by total assets. Deposit shares, loan shares, deposit rates, loan rates, expenses related to fixed assets, salaries, and net noninterest expenses are reported in percentages. Asset maturity is reported in years. "(vw)" indicates an asset-weighted mean, and "(ew)" indicates an equal-weighted mean . The data sources are the Call Reports and the FDIC Summary of Deposits.

	mean(vw)	mean(ew)	sd	p10	p25	p50	p75	p90
Deposit market shares	3.519	0.079	0.523	0.003	0.005	0.009	0.021	0.077
Loan market shares	1.368	0.033	0.207	0.001	0.002	0.004	0.009	0.034
Deposit rates	1.706	2.032	1.292	0.166	0.873	2.085	3.150	3.714
Loan rates	5.935	6.921	1.725	4.540	5.599	6.959	8.286	9.061
No. of branches	1778	69.753	315.678	7.000	11.000	17.000	34.000	94.000
No. of employees per branch	53.736	18.338	17.433	9.109	11.188	14.306	19.556	28.500
Expenses of fixed assets	0.454	0.480	0.165	0.270	0.347	0.448	0.584	0.798
Salaries	1.590	1.725	0.486	1.061	1.348	1.650	2.036	2.646
Net noninterest expenses	1.230	2.778	0.830	1.904	2.246	2.653	3.142	3.743
Loan-to-deposit ratio	0.816	0.815	0.170	0.598	0.710	0.821	0.925	1.022
Borrowing-to-deposit ratio	0.699	0.136	0.138	0.013	0.041	0.096	0.181	0.308
Deposit-to-asset ratio	0.707	0.805	0.082	0.691	0.763	0.822	0.866	0.895
Book leverage	11.464	11.114	2.577	7.947	9.408	10.990	12.656	14.390
Asset maturity	3.429	3.772	1.402	2.163	2.764	3.560	4.604	5.698

Table 3: Parameter Estimates

In this table, we report the model parameter estimates. Panel A presents calibrated parameters. Panel B presents values for parameters that can be calculated as simple averages or by simple regression methods. Panel C presents results for parameters estimated via BLP. Panel D presents results for parameters estimated via SMD. Standard errors for the estimated parameters are clustered at the bank level and reported in brackets.

	Panel A: Calibrated Parameters		
τ_c	Corporate tax rate	0.350	
θ	The reserve ratio	0.024	
κ	The capital ratio	0.060	
$\hat{J}$	Number of representative banks	6	
	Panel B: Parameters Estimated Separately		
μ	Average loan maturity	3.429	[1.445]
$\bar{f}$	Log Federal Funds rate mean	0.297	[0.352]
σ_f	Std of Federal Funds rate innovation	0.554	[0.135]
ρ_f	Log Federal Funds rate persistence	0.902	[0.103]
$\bar{\delta}$	Log loan chargeoffs mean	-1.300	[0.004]
σ_δ	Std log loan chargeoffs innovation	0.961	[0.003]
ρ_δ	Log loan chargeoffs persistence	0.598	[0.002]
$\rho_{\delta f}$	Corr of Federal Funds rate innovation and log loan chargeoffs	-0.108	[0.003]
	Panel C: Parameters Estimated via BLP		
α^d	Depositors' sensitivity to deposit rates	0.968	[0.140]
σ_{α^d}	Dispersion of depositors' sensitivity to deposit rates	1.916	[0.403]
α^l	Borrowers' sensitivity to loan rates	-1.424	[0.292]
q_d^d	Convenience of holding deposits	3.440	[0.251]
q_c^d	Convenience of holding cash	1.985	[0.242]
q_l^l	Convenience of borrowing through loans	1.094	[1.065]
	Panel D: Parameters Estimated via SMD		
γ	Banks' discount rate	0.045	[0.006]
W/K	Relative size of the deposit market	0.217	[0.040]
q_n^l	Value of firms' outside option	-9.641	[0.414]
ϕ^N	Quadratic cost of non-reservable borrowing	0.010	[0.003]
ϕ^d	Bank's cost of taking deposits	0.009	[0.001]
ϕ^l	Bank's cost of servicing loans	0.007	[0.003]
ψ	Net fixed operating cost	0.015	[0.001]

Table 4: Moment Conditions

In this table, we report the simulated versus the actual moments in the SMD estimation, along with the t-statistics for the pairwise differences. The dividend yield is defined as dividends over bank equity value; the non-reservable borrowing share is defined as the ratio of non-reservable borrowing to total assets; the sensitivity of corporate borrowing to the federal funds rate (FFR) is estimated via a vector autoregression.

	Actual Moment	Simulated Moment	t-stat
Dividend yield	3.38%	3.00%	-0.687
Non-reservable borrowing share	29.9%	27.0%	-1.782
Std of non-reservable borrowing	12.6%	14.8%	1.337
Deposit spread	1.29%	1.28%	-0.113
Loan spread	2.81%	2.75%	-0.374
Deposit-to-asset ratio	0.699	0.728	1.022
Net noninterest expenses	1.23%	1.20%	-0.822
Leverage	11.20	11.69	2.337
Market-to-book ratio	2.061	1.814	-1.424
Credit-FFR sensitivity	-0.995	-0.975	0.083

Table 5: Determinants of Monetary Policy Transmission

This table presents the results of a series of counterfactual experiments in which we examine the effects of removing frictions from our model. The first column lists the frictions that are removed from the model. The second column presents the sensitivity of loans to the federal funds rate (FFR) when the corresponding frictions are removed. The third column presents the contributions of the corresponding frictions. All model solutions are under the same set of parameters reported in Table 3.

		Sensitivity of Loans to FFR $\left(\frac{\Delta \ln l}{\Delta f}\right)$	Contribution %
(1)	All frictions are present	-1.548%	/
(2)	$-$ Reserve Regulation	-1.426%	7.88%
(3)	$-$ Capital Regulation	-1.120%	27.65%
(4)	$-$ Deposit Market Power	-0.992%	35.91%
(5)	$-$ Loan Market Power	-1.910%	-23.39%

Table 6: Monetary Policy Shocks and Bank Equity Returns on FOMC Days

In this table, we report the estimates of the relation between bank equity returns and monetary policy shocks on FOMC Days. Monetary policy shocks are measured by one-day changes in the two-year Treasury yield on FOMC days. HHI is the Herfindahl-Hirschman Index for the local deposit market in which a bank operates. The sample includes all publicly traded U.S. banks from 1994 to 2017. The sample for columns (1) and (3) constitutes observations in which the starting level of the federal funds rate (FFR) is above 2%. The sample for columns (2) and (4) constitutes observations in which the starting level of the federal funds rate is below 2%. We exclude observations during the collapse of the dot-com bubble (2000–2001) and the subprime financial crisis (2007-2009). Standard errors are clustered by time.

	(1) High FFR	(2) Low FFR	(3) High FFR	(4) Low FFR
Policy shock	-1.292**	2.202**	-0.639	-1.393
	[0.615]	[0.879]	[0.653]	[0.852]
HHI*Policy shock			-0.134	0.562***
			[0.145]	[0.153]
Δ Term spread	-0.634	2.336*	-0.667	1.827
	[1.265]	[1.350]	[1.257]	[1.293]
Market return	0.297***	0.730***	0.295***	0.733***
	[0.072]	[0.070]	[0.071]	[0.070]
Observations	27,257	33,805	27,257	33,805
Adj. R^2	0.015	0.123	0.016	0.125

47

Table 7: Large Banks versus Small Banks

In this table, we report the SMD estimation results for subsamples of large and small banks. Panel A contains the simulated versus the actual moments, along with the t-statistics for the pairwise differences. In Panel B, we report the parameter estimates. ϕ^d and ϕ^l are banks' marginal costs of intaking deposits and servicing loans, respectively; ψ is the net fixed operating cost, and ϕ^N is the quadratic cost of borrowing non-reservables. The last column presents the sensitivity of loans to the federal funds rate (FFR). Standard errors clustered at the bank level are in brackets under the parameter estimates.

Panel A: Moment conditions

	Large Banks			Small Banks		
	Actual	Simulated	t-stat	Actual	Simulated	t-stat
Dividends	3.60%	2.95%	-1.929	/	/	/
Non-reservable borrowing share	35.5%	38.8%	2.130	15.3%	15.1%	-0.577
Std of non-reservable borrowing	19.0%	20.6%	0.726	8.70%	7.58%	-1.992
Deposit spread	1.32%	1.33%	0.133	1.25%	1.29%	0.915
Loan spread	2.67%	2.71%	0.280	2.98%	2.83%	-3.193
Deposit-to-asset ratio	0.666	0.672	0.201	0.784	0.793	0.788
Net noninterest expense	0.96%	1.01%	0.994	1.70%	1.58%	-3.918
Leverage	11.36	12.45	2.448	10.78	11.16	2.499
Market-to-book ratio	2.061	2.009	-0.894	/	/	/
Total credit-FFR sensitivity	-0.995	-0.898	0.177	-0.995	-0.960	0.064

Panel B: Parameter Estimates

	ϕ^N	ϕ^d	ϕ^l	ψ	Sensitivity of Loans to FFR $\left(\frac{\Delta \ln l}{\Delta f}\right)$
Large Banks	0.006	0.010	0.006	0.002	-1.103
	[0.0011]	[0.0003]	[0.0011]	[0.0001]	
Small Banks	0.016	0.009	0.007	0.029	-1.758
	[0.0009]	[0.0005]	[0.0009]	[0.0001]	

Table 8: Subsample Estimates: Early versus Late

In this table, we report the model parameter estimates for the early (1994–2005) and late (2006–2017) subsamples. Panel A presents calibrated parameters. Panel B presents values for parameters that can be calculated as simple averages or by simple regression methods. Panel C presents results for parameters estimated via BLP. Panel D presents results for parameters estimated via SMD. Standard errors for the estimated parameters are clustered at the bank level and reported in brackets.

		Early Subsample		Late Subsample	
Panel A: Calibrated Parameters					
τ_c	Corporate tax rate	0.350		0.350	
θ	The reserve ratio	0.028		0.022	
κ	The capital ratio	0.060		0.060	
$\hat{J}$	Number of representative banks	7		5	
Panel B: Parameters Estimated Separately					
μ	Average loan maturity	3.178	[1.402]	3.195	[1.448]
$\bar{f}$	Log federal funds rate mean	1.245	[0.205]	-0.920	[0.482]
σ_f	Std of log federal funds rate innovation	0.552	[0.115]	0.951	[0.282]
ρ_f	Log federal funds rate persistence	0.700	[0.252]	0.623	[0.213]
$\bar{\delta}$	Log loan chargeoffs mean	-1.647	[0.005]	-1.221	[0.007]
σ_δ	Std of log loan chargeoffs innovation	0.863	[0.004]	0.935	[0.005]
ρ_δ	Log loan chargeoffs persistence	0.611	[0.003]	0.722	[0.004]
$\rho_{\delta f}$	Corr of federal funds rate innovation and chargeoffs	-0.064	[0.011]	-0.128	[0.006]
Panel C: Parameters Estimated via BLP					
α^d	Depositors' sensitivity to deposit rates	0.743	[0.165]	0.925	[0.399]
σ_{α^d}	Dispersion of depositors' sensitivity to deposit rates	1.467	[0.498]	1.830	[0.965]
α^l	Borrowers' sensitivity to loan rates	-1.025	[0.054]	-1.501	[0.082]
q_d^d	Convenience of holding deposits	3.465	[0.358]	2.324	[0.470]
q_c^d	Convenience of holding cash	1.763	[0.387]	-0.446	[0.430]
q_l^l	Convenience of borrowing through loans	-0.005	[0.088]	1.804	[0.212]
Panel D: Parameters Estimated via SMM					
γ	Banks' discount rate	0.046	[0.012]	0.046	[0.017]
W/K	Relative size of the deposit market	0.183	[0.051]	0.283	[0.022]
q_n^l	Value of firms' outside option	-7.030	[0.568]	-9.689	[1.025]
ϕ^N	Quadratic cost of non-reservable borrowing	0.010	[0.004]	0.010	[0.015]
ϕ^d	Bank's cost of taking deposits	0.009	[0.002]	0.009	[0.001]
ϕ^l	Bank's cost of servicing loans	0.005	[0.004]	0.008	[0.005]
ψ	Net fixed operating cost	0.039	[0.001]	0.009	[0.001]

Panel A: Deposit Spread
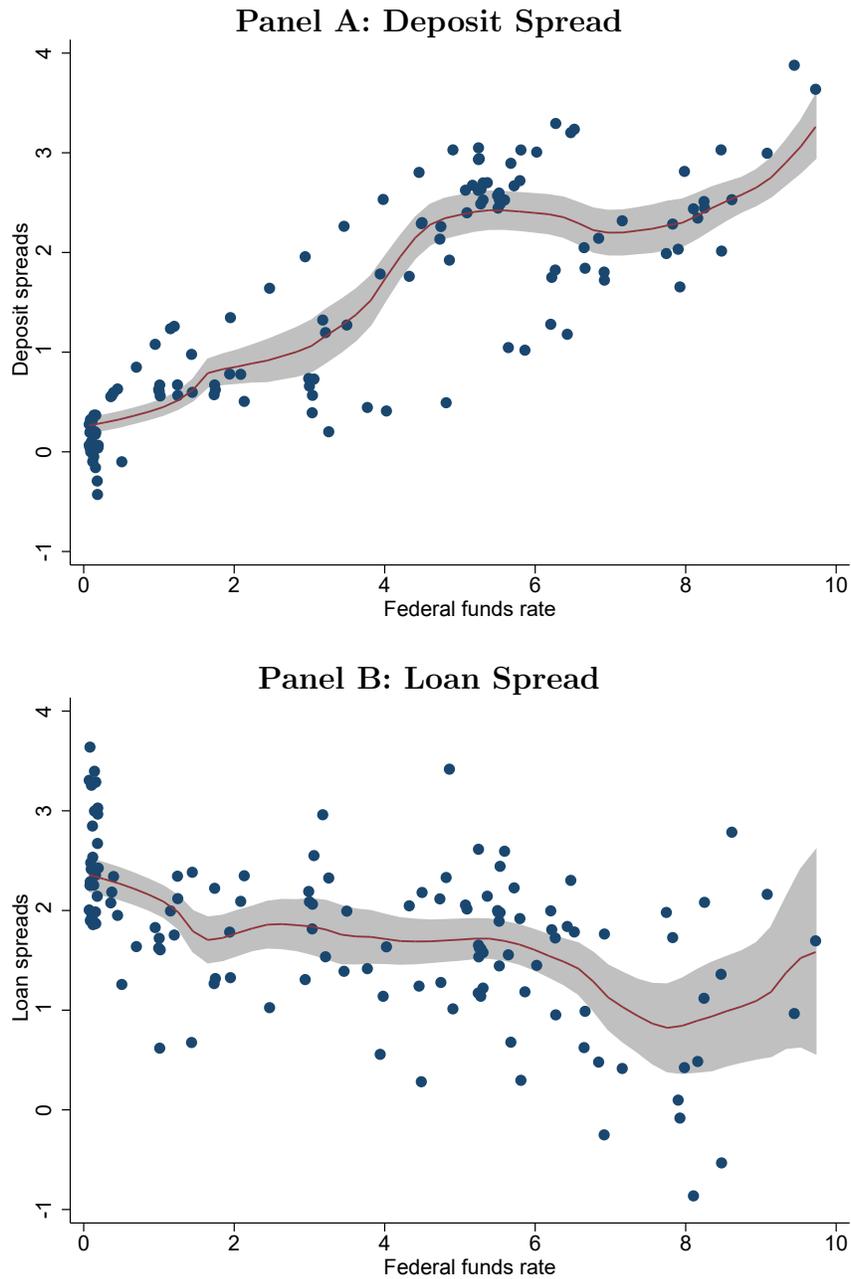

Panel B: Loan Spread

Figure 1: Deposit Spread, Loan Spread, and the Federal Funds Rate

In this figure, we plot kernel regressions of average deposit and loan spreads for U.S. banks on the federal funds rate. We use an Epanechnikov kernel with a bandwidth of 0.66 for deposits and 0.61 for loans. The sample period is 1985–2017. The data frequency is quarterly. The deposit and loan rates are constructed using the Call Reports, and the federal funds rate is from the FRED database.

Federal funds rate: f_t

Charge-offs: δ_t

Banks set deposit rates: $r_{j,t}^d$,
loan rates: $r_{j,t}^l$

Banks choose reserves, government securities, non-reservable borrowing

Depositors choose $D\left(r_{j,t}^d|f_t\right)$

Borrowers choose $B\left(r_{j,t}^l|f_t\right)$

$\eta < 1$ loans mature

Banks collect profits and distribute dividends
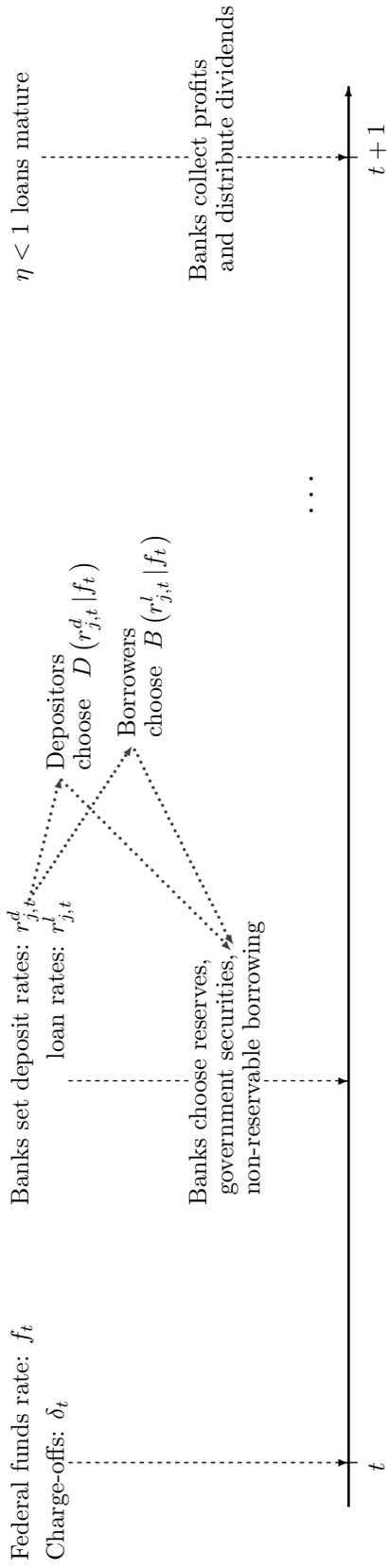
t

$t+1$
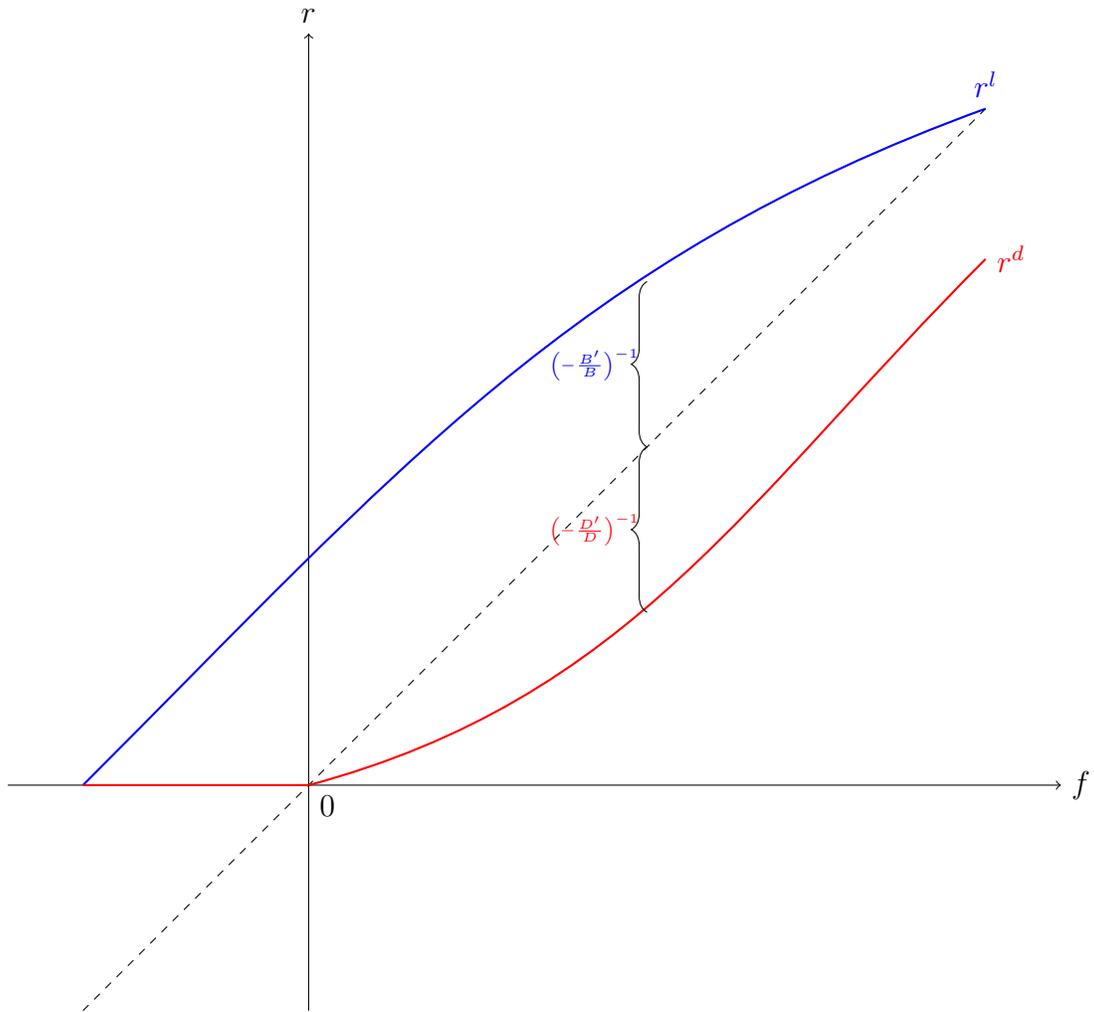
Figure 2: Timeline within a Period

Figure 3: Deposit Rates, Loan Rates, and the Federal Funds Rate

This figure illustrates the relationship between deposit rates, loan rates, and the federal funds rates when banks have market power in the deposit and loan markets. The federal funds rate is plotted on the x-axis, and the deposit and lending rates are plotted on the y-axis.

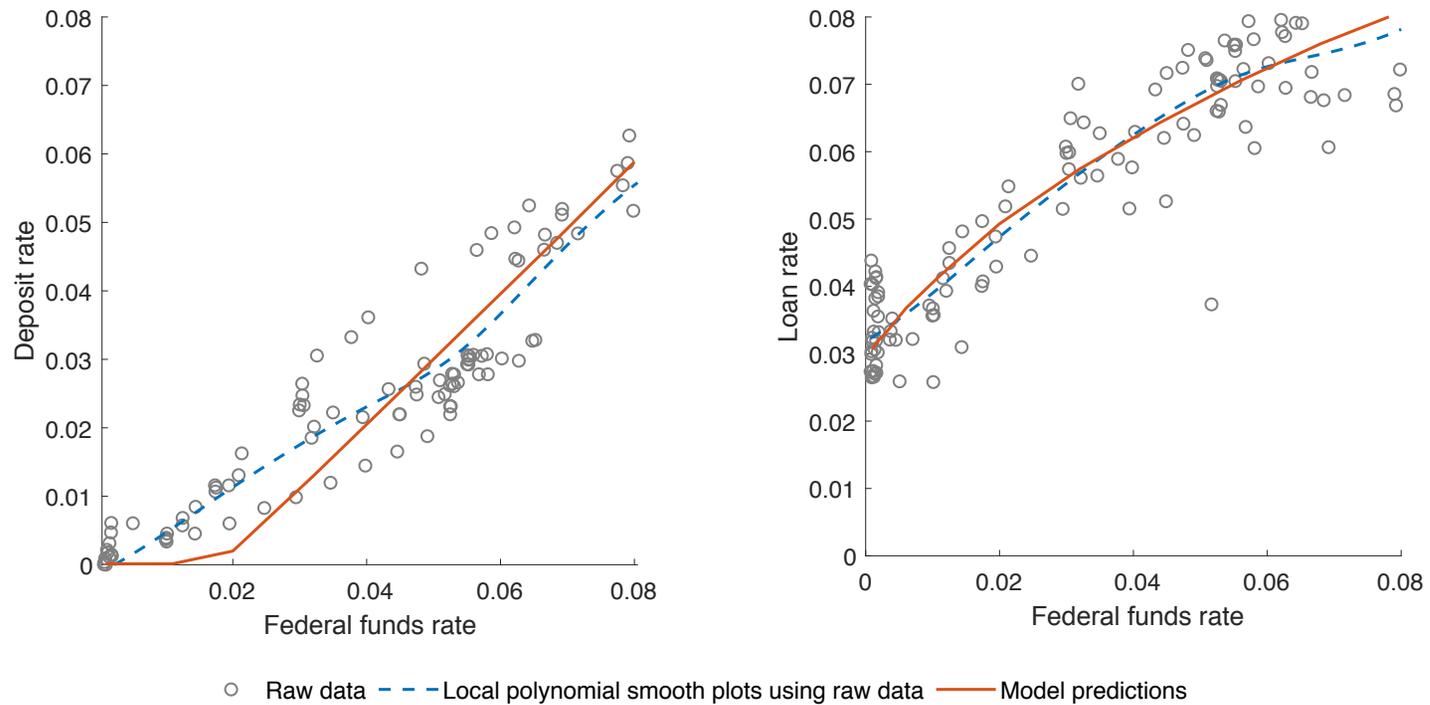Figure 4: Model-predicted versus Actual Rates

This figure illustrates the relation between the federal funds rate and banks' deposit and loan rates. The circles represent a scatter of the raw data from 1984 to 2017, aggregated at the quarterly frequency. The dashed lines represent local polynomial smoothed plots based on the raw data, and the solid lines represent the relations predicted using the model.
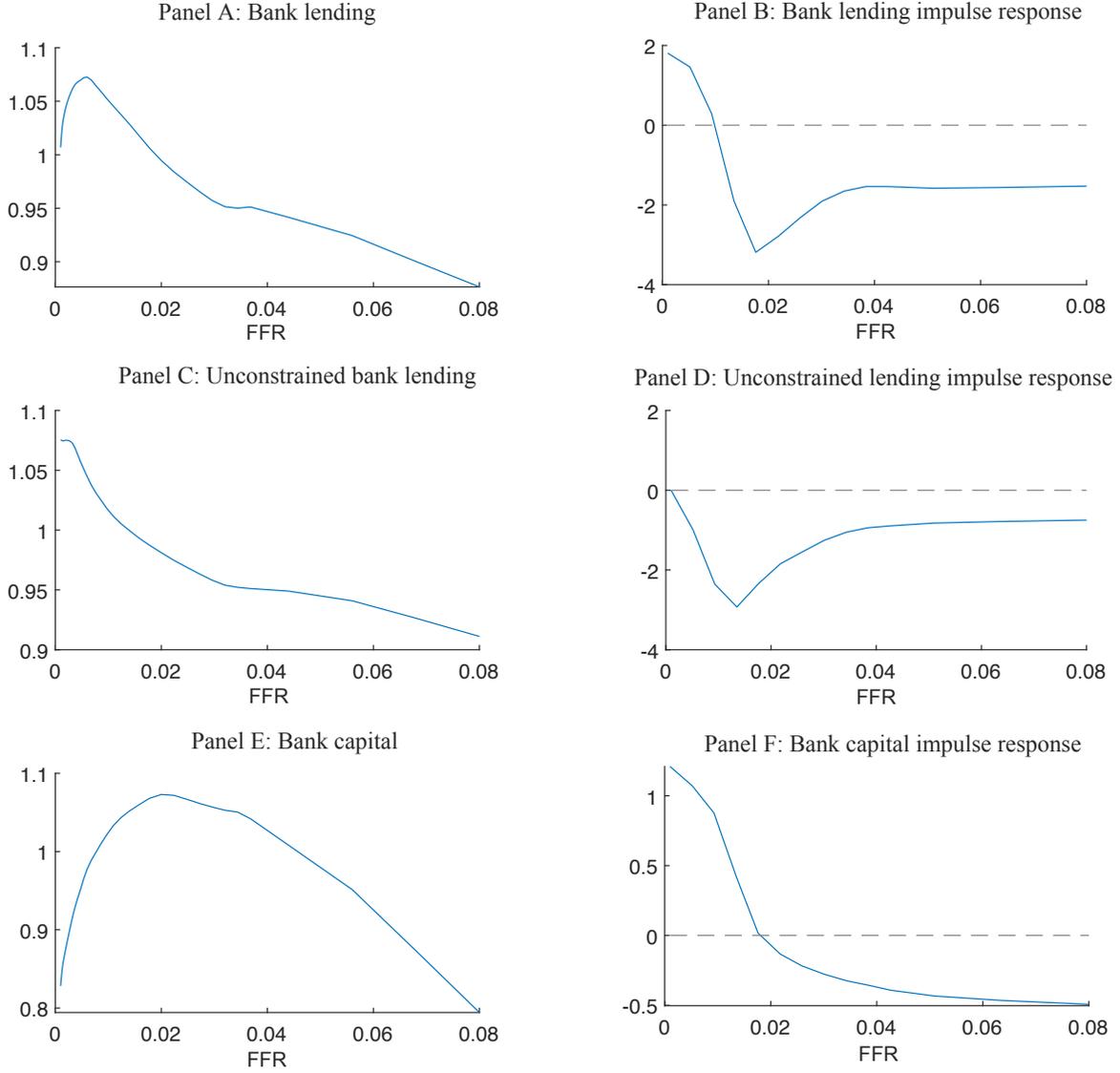
Figure 5: Bank Capital, Bank Lending, and the Federal Funds Rate

This figure illustrates how bank capital and optimal lending vary with the federal funds rate. In all panels, the federal funds rate is on the x-axis. In Panels A, C, and E, bank characteristics, scaled by their respective unconditional average values, are on the y-axis. In Panels B, D, and F, the y-axis measures the impulse response of bank characteristics to a one-percentage-point federal funds rate shock.

Panel A: Impulse response to increases in the Federal funds rate

Panel B: Impulse response to decreases in the Federal funds rate

Bank loans ◆   Deposits ●   Bank capital ✳   Non-reservables +
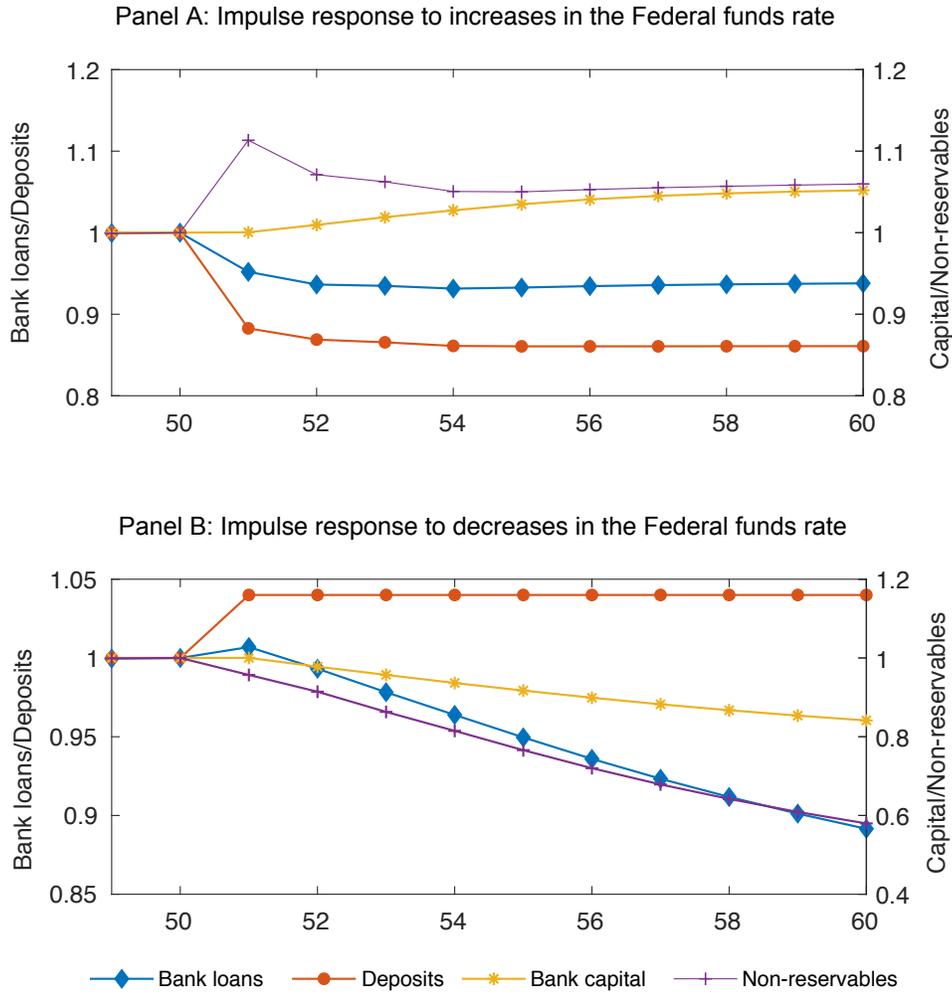
Figure 6: Impulse Response to Federal Funds Rate Shocks

This figure illustrates banks' impulse responses to federal funds rate shocks. The economy starts at Year 0 when it is in the old steady state with the federal funds rate equal to 0.9%. In Year 1, the federal funds rate either increases to 2% or decreases to 0.1%, 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 in which the federal funds rate is 0.9%.

Figure 7: Monetary Policy Shocks and Bank Equity Returns

This figure provides scatter plots of bank industry excess returns against monetary policy shocks on FOMC days from 1994 to 2017. The excess returns are defined as the difference between bank industry index returns and market returns. Monetary policy shocks are measured by one-day changes in two-year Treasury yields on FOMC days. The sample for the upper panel constitutes observations in which the starting level of the federal funds rate is above 2%. The sample the lower panel constitutes observations in which the starting level of the federal funds rate is below 2%. We exclude observations during the collapse of the dot-com bubble (2000-2001) and the financial crisis (2007-2009). Bank industry stock returns are from Kenneth French's website, and the two-year Treasury yield is from the FRED database.

56

Appendix A. Numerical Model Solution

In this Appendix, we summarize the numerical methods used to solve the model. As a preliminary step, we discretize the state variable space. Our state space ($\mathbf{s}$) consists of four state variables: $\mathbf{s} \equiv \{f, \delta, L, E\}$, where f is the federal funds rate, δ is the annual loan charge-off rate, L is the amount of outstanding loans, and E is bank equity. We set the federal funds rate to be in the range of 0.1% to 8%, that is, $\underline{f} = 0.1\%$ and $\bar{f} = 8\%$. We set the boundaries for the loan charge-off rate, $\underline{\delta}$ and $\bar{\delta}$, at 0.1% and 3%, respectively. We set the upper bound for outstanding loans ($\bar{L}$) at B^*/η, where B^* stands for the unconstrained optimal lending among banks in the static model in Subsection 4.6 as follows:

$$r^l = f + \phi^d \tag{A.1}$$

$$r^d = f - \phi^d. \tag{A.2}$$

We set the upper bound of bank equity ($\bar{E}$) at 10% of $\bar{L}$. We set the lower bounds of L and E to be small positive numbers.

Bank optimal policies are characterized by the set of choice variables, $\mathbf{p} \equiv (B(r^l), D(r^d), G, N, R, C)$, where $B(\cdot)$ is new loans issued, $D(\cdot)$ is deposits, G is government securities held, N is the amount of wholesale funding obtained, R is reserves, and C is cash dividends. Taking $B(r^l)$ as given, we consider the following two cases:

If $D^* \times (1 - \theta) \geq L + B - E$, where D^* stands for the unconstrained optimal deposit intake in a static model as specified in equation (A.2) above, banks will only hold deposits and government securities. Because external financing is costly, it is never profitable for them to issue non-reservables and hold government securities at the same time. We can verify that $D = D^*$ yields the highest possible profit for banks.

If $D^* \times (1 - \theta) < L + B - E$, banks will issue non-reservables subject to the linear-quadratic financing cost. At the same time, they will also choose higher deposit intake to help close

their funding shortage. Banks will keep increasing their deposits until the marginal cost of doing so (and thus deviating from the static optimum) is equal to the marginal cost of external financing. Banks' choice of deposits will solve the following equation:

$$\frac{\partial D \times r^d(D)}{\partial D} - f + \phi^d = -\Phi^N \frac{N}{D}. \tag{A.3}$$

After solving for D, we can substitute it into equations (16), (17), (18), and (21) to obtain other firm policies: $\{G, N, R\}$. Last, we do a grid search over C and B to find the set of bank policies that yields the highest expected value. Finding the expected value for banks involves iterating on the value function, the details of which are as follows:

1. We conjecture the equilibrium deposit rate $(\boldsymbol{r_0^d}(f))$ and loan rate $(r_0^l(f))$ as functions of the federal funds rate.

2. We conjecture the bank value, $V_0(f, \delta, E, L)$, as a function of the state variables.

3. For each set of (f, δ, E, L), we search for the bank's optimal policy $(\mathbf{p})$ using the algorithm described above, and we update the bank's Bellman equation using equation (24) in the main text, which yields the updated value function for the bank, $V_1(f, \delta, E, L)$.

4. We repeat Steps 2–3 until the initial and the updated value functions, $V_0(f, \delta, E, L)$ and $V_1(f, \delta, E, L)$, converge.

5. Taking the policy calculated in Step 3, and using the law of motion for the aggregate and idiosyncratic shocks described in equation (22), we can simulate a panel of banks, denoting the simulated deposit and loan rates by $r_t^{d,j}$ and $r_t^{l,j}$, respectively, where j indexes the bank and t indexes time. We use simulated data to calculate the aggregate average deposit rates and loan rates, $\boldsymbol{r_1^d}(f)$ and $r_1^l(f)$, using equations (25) and (26) in the main text. Notice that because we have savers with heterogeneous sensitivities to deposit rates (α_i^l), we need to forecast one average deposit rate for each α_i^l class. To

accomplish this task, we approximate the savers' rate sensitivities using a truncated uniform distribution, with mean α^l and standard deviation σ_{α^l}. We then discretize the domain of α_i^l using 10 equally spaced grid points.

6. We repeat Steps 1 through 5 using the updated deposit and loan rates as the banks' conjectured equilibrium outcome. We keep iterating until the initial guess, $\boldsymbol{r_0^d}(f)$ and $r_0^l(f)$, is consistent with the actual equilibrium outcome, $\boldsymbol{r_1^d}(f)$ and $r_1^l(f)$, in the simulated data.

7. As the last step, we assess the goodness-of-fit of this algorithm. Using the simulated panel of banks constructed in Step 5, for each bank in each period, we calculate the average deposit and loan rates offered by its competitors in the economy:

$$\exp(\alpha_i^d \bar{r}_{j,t}^{d,i} + \beta^d x^d + \xi^d) \equiv \frac{1}{\hat{J} - 1} \sum_{m \neq j} \left[\exp(\alpha_i^d r_{m,t}^d + \beta^d x^d + \xi^d) \right], \qquad \text{(A.4)}$$

and

$$\exp(\alpha^l \bar{r}_{j,t}^l + \beta^l x^l + \xi^l) \equiv \frac{1}{\hat{J} - 1} \sum_{m \neq j} \left[\exp(\alpha^l r_{m,t}^l + \beta^l x^l + \xi^l) \right]. \qquad \text{(A.5)}$$

where $\hat{J}$ is the number of banks we include in our simulation in a given period. Note that these rates, $\{\bar{r}_{j,t}^{d,i}, \bar{r}_{j,t}^l\}$ can differ across banks even within a given period because we exclude the bank's own rates when computing the averages. We then run a regression of $\{\bar{r}_{j,t}^{d,i}, \bar{r}_{j,t}^l\}$ on the banks' forecasted equilibrium rates, $\boldsymbol{r_1^d}(f)$ and $r_1^l(f)$, calculated in Step 6, and the R-squareds for these regressions are reported in Subsection 4.5.

Appendix B. Demand Estimation Details

In this Appendix, we describe our demand estimation in detail. We start with a data definition. We combine the data from tiny local banks, which we define as banks with fewer than ten domestic branches or with market shares less than 0.001%. This aggregation is useful because disaggregated data substantially slows down our demand estimation. It is also innocuous because these tiny banks have limited influence on the equilibrium. After combining tiny local banks, we have 753 banks on average in each year in our final sample.

Next, we describe deposit demand, which is characterized by the following preference parameters, $\Theta^d = \left(\alpha^d, \sigma_\alpha^d, \beta^d\right)$, where α^d and σ_α^d are the mean and standard deviation of the sensitivity to deposit rates, and the vector β^d contains the sensitivities to non-rate characteristics.

Following Berry, Levinsohn, and Pakes (1995), we construct a nonlinear GMM estimator for the preference parameters by exploiting a moment condition that is the product of instrumental variables, Z, and the unobservable demand shocks, ξ^d. Formally, the estimator is:

$$\hat{\Theta}^d = \arg \min_{\Theta^d} \xi\left(\Theta^d\right)' Z' W^{-1} Z \xi\left(\Theta^d\right), \tag{B.1}$$

where W is a consistent estimate of $\mathbb{E}\left[Z'\xi\xi'Z\right]$.

We compute the unobservable demand shocks, ξ^d, using the nested fixed-point algorithm described in Nevo (2001). Specifically, for a given set of demand parameters Θ^d and the actual market shares in the data, s_0, we can solve for the shocks, ξ^d, as:

$$\xi^d\left(\Theta^d\right) = s^{-1}\left(s_0|\sigma_{\alpha^d}\right) - \left(\alpha^d r_j^d + \beta^d x_j^d\right), \tag{B.2}$$

where $s^{-1}\left(\cdot\right)$ is the inverse of the demand function specified by equation (3) in the main text.

A key challenge in identifying the demand parameters is the natural correlation between deposit rates r_j^d and unobservable demand shocks ξ_j^d. For example, banks might lower deposit rates if they observe a positive demand shock. To identify the associated yield sensitivity, we use a set of supply shifters, c_j, as instrumental variables. Our particular supply shifters are salaries and non-interest expenses related to the use of fixed assets. These shifters have been used in previous studies, such as Dick (2007) and Ho and Ishii (2011). Our identifying assumption is that customers do not care about these costs, holding product characteristics constant. Therefore, 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, the vector of instrumental variables Z is defined as follows:

$$Z = [x, c], \tag{B.3}$$

where c is a vector of supply shifters including salaries and non-interest expenses related to the use of fixed assets, and x is a vector of non-rate bank characteristics including the number of branches, the number of employees per branch, bank fixed effects, and time fixed effects.

Using the demand parameters estimated in the first stage, we can construct the empirical demand system. Note that our data contain a large number of banks even after we combine tiny local banks into one option in the demand estimation. This feature of the data poses a challenge for the second-stage SMD estimation because estimating a dynamic model with a large number of heterogeneous banks would be intractable. Therefore, we use the estimated demand parameters in the first stage to construct a demand system with a small number, $\hat{J}$, of ex ante symmetric representative banks. We calibrate the number of representative banks, $\hat{J}$, to match the average local banking concentration in the data, as measured by the HHI. 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 in the model close

5

to what we observe in the data.

Because $\hat{J}$ is much smaller than the sample size of 753 banks, we need to adjust the quality value of the non-rate product characteristics for each of the $\hat{J}$ representative banks. We let q generically denote the utility from the non-rate product characteristics. These non-rate product characteristics include the number of branches, the number of employees per branch, bank fixed effects, and time fixed effects. We then choose a quality value, q, such that the sum of the exponential utility of the $\hat{J}$ symmetric banks is the same as that of $J = 753$ banks in the sample, as follows:

$$\hat{J}\exp(q) = \sum_{j=1}^{J} \exp\left(\hat{\beta}x_j + \hat{\xi}_j\right). \tag{B.4}$$

Equation (B.4) transforms the heterogeneous non-rate characteristics into symmetric quality values while preserving market concentration and the sum of exponential utility. We use (B.4) to obtain the quality values for cash and the deposits at bank m, which are given by q_c^d and q_m^d, respectively. We normalize the quality value of Treasury bills to zero.

With the quality values in hand, we parameterize the deposit demand functions as:

$$D_j(r_j^d|f) = \sum_{i=1}^{I} \mu_i^d \frac{\exp\left(\hat{\alpha}_i^d r_j^d + q_j^d\right)}{\exp\left(\hat{\alpha}_i^d f\right) + \exp\left(q_c^d\right) + \sum_{m=1}^{\hat{J}} \exp\left(\hat{\alpha}_i^d r_m^d + q_m^d\right)} W, \tag{B.5}$$

where μ_i^d represents the fraction of each type, i, of depositors. Finally, we draw $\hat{\alpha}_i^d$ from a discretized uniform distribution with a mean of $\hat{\alpha}$ and a standard deviation of $\hat{\sigma}_\alpha$.

We estimate the loan demand function in a procedure that is similar to the one we used to estimate deposit demand except that we assume homogeneous sensitivity to loan rates, as we find that introducing heterogeneity in loan-rate sensitivity considerably slows down the estimation but has a limited impact on banks' rate-setting decisions. We include the same set of supply shifters and non-rate characteristics as in the deposit market but allow the sensitivities to these characteristics to differ from those in the deposit market.

As in the case of deposit demand, we can use equation (B.4) to construct quality values for not borrowing and for loans from bank m, which are given by q_n^l and q_m^l, respectively. We normalize the quality value of borrowing from the bond market to zero.

With the quality values in hand, we parameterize the loan demand functions as:

$$B_j(r_j^l \mid f) \;=\; \frac{\exp\left(\hat{\alpha}^l r_j^l + q_{lj}^l\right)}{\exp\left(\hat{\alpha}^l(\overline{f} + \overline{\delta})\right) + \exp\left(q_n^l\right) + \sum_{m=1}^{J} \exp\left(\hat{\alpha}^l r_m^l + q_m^l\right)} K, \qquad \text{(B.6)}$$

Note that the quality value of not borrowing, q_n^l, cannot be estimated from the demand estimation because we do not observe its share. While this data limitation does not bias the demand parameters, as this term is absorbed by fixed effects, it poses an issue for the SMD part of our estimation because we need to plug the totality of B.6 into our dynamic model. Therefore, we relegate this parameter to our second-stage estimation.

Appendix C. Demand Estimation Using Local Market Shares

In this Appendix, we check the robustness of the deposit demand estimation using a different level of aggregation. For the estimation of the demand parameters in Table 2 in the main text, market shares are defined at the U.S. national level. We now examine whether the assumption of a national market affects the estimated demand parameters. We discuss an alternative approach of using local deposit market shares and show that, both conceptually and practically, using local market shares should lead to similar results.

For illustration, we consider the simplest case of logit demand. In the context of logit demand, the demand estimation in the main text is equivalent to the following regression:

$$\ln s_{j,t} = \alpha r_{j,t} + \beta x_{j,t} + \xi_{j,t}, \tag{C.1}$$

where $s_{j,t}$ is the market share of bank j in year t, $r_{j,t}$ is the deposit rate, $x_{j,t}$ represents other non-rate characteristics. α is the yield sensitivity for the average depositor in the national market.

Alternatively, we can also estimate the demand using a more disaggregate market definition. This is feasible for the deposit market because branch-level deposit data are available from the FDIC Summary of Deposits data. For instance, we can define the market as a county-year combination and estimate the following regression:

$$\ln s_{j,m,t} = \alpha r_{j,m,t} + \beta x_{j,m,t} + \xi_{j,m,t}, \tag{C.2}$$

where $s_{j,m,t}$ is the market share of bank j in county m in year t, $r_{j,m,t}$ is the deposit rate and $x_{j,m,t}$ represents other non-rate characteristics, and α is the yield sensitivity of depositors. Note that the yield sensitivity, α, estimated using disaggregate market shares has the

8

same economic interpretation as the one estimated using aggregate market shares. Both parameters measure the average yield sensitivities of depositors across all the local markets. Therefore, conceptually, these two methods should give rise to similar results.

Another way to understand the relation between these two levels of aggregation is to think about the estimation in two steps. First, we estimate yield sensitivity county by county. Second, we take a weighted average of the county-specific yield-sensitivity to calculate the average yield sensitivity across all local markets. Again, this method is equivalent to imposing the restriction that α is the same across markets in a pooled regression.

Although the estimates from these two levels of aggregation are conceptually similar, it is nonetheless possible that these two methods could lead to quite different outcomes. To alleviate this concern, we re-estimate the banks' deposit demand using market shares defined at the county level. The estimation results reported in Table C.1 are quantitatively similar to those in Table 2 in the main text. For instance, the estimated yield sensitivity is 0.903 with local market shares and 0.968 with national market shares.

Table C.1: Demand Estimation: Local Deposit Market

| | (1) |
	Deposit
Yield sensitivity	0.903***
	[0.199]
Log number of branches	0.509***
	[0.048]
Log number of employees	0.322***
	[0.021]
Bank F.E.	Yes
Year-Sector F.E.	Yes
Year-County F.E.	Yes
Observations	377,309
Adj. R²	0.399

This table reports the estimated deposit demand parameters using county-level market shares. Yield sensitivity refers to the average sensitivity of the depositors to deposit rates. Log number of branches refers to the sensitivity of the depositors to the log number of each bank's branches. Log number of employees refers the sensitivity of the depositors to the log number of employees per branch. The sample includes all U.S. commercial banks from 1994 to 2017. The data is from the Call Reports and the Summary of Deposits.

Appendix D. Impulse Responses to Monetary Policy Shocks

In this Appendix, we report the impulse responses of bank loans and total credit in the economy to monetary policy shocks. Following Bernanke and Blinder (1992), we estimate a VAR including the federal funds rate, the unemployment rate, the log of the CPI, the log level of bank loans, and the log of total credit. The frequency is monthly. We use six lags of each variable.

We find that at a three-year horizon, a one percentage point monetary policy shock leads to a 0.995% reduction in total credit in the economy. We use the impulse response result to measure the sensitivity between total credit and the federal funds rates, which we report in Table 3. The bank loans are more sensitive to monetary policy shocks. A one percentage point increase in the federal funds rate translates into a 1.6% decrease in aggregate bank loans. We do not directly target this impulse response result in our estimation, but we report it as an untargeted moment in Subsection 6.1 and show that it lies in close proximity with the model prediction.
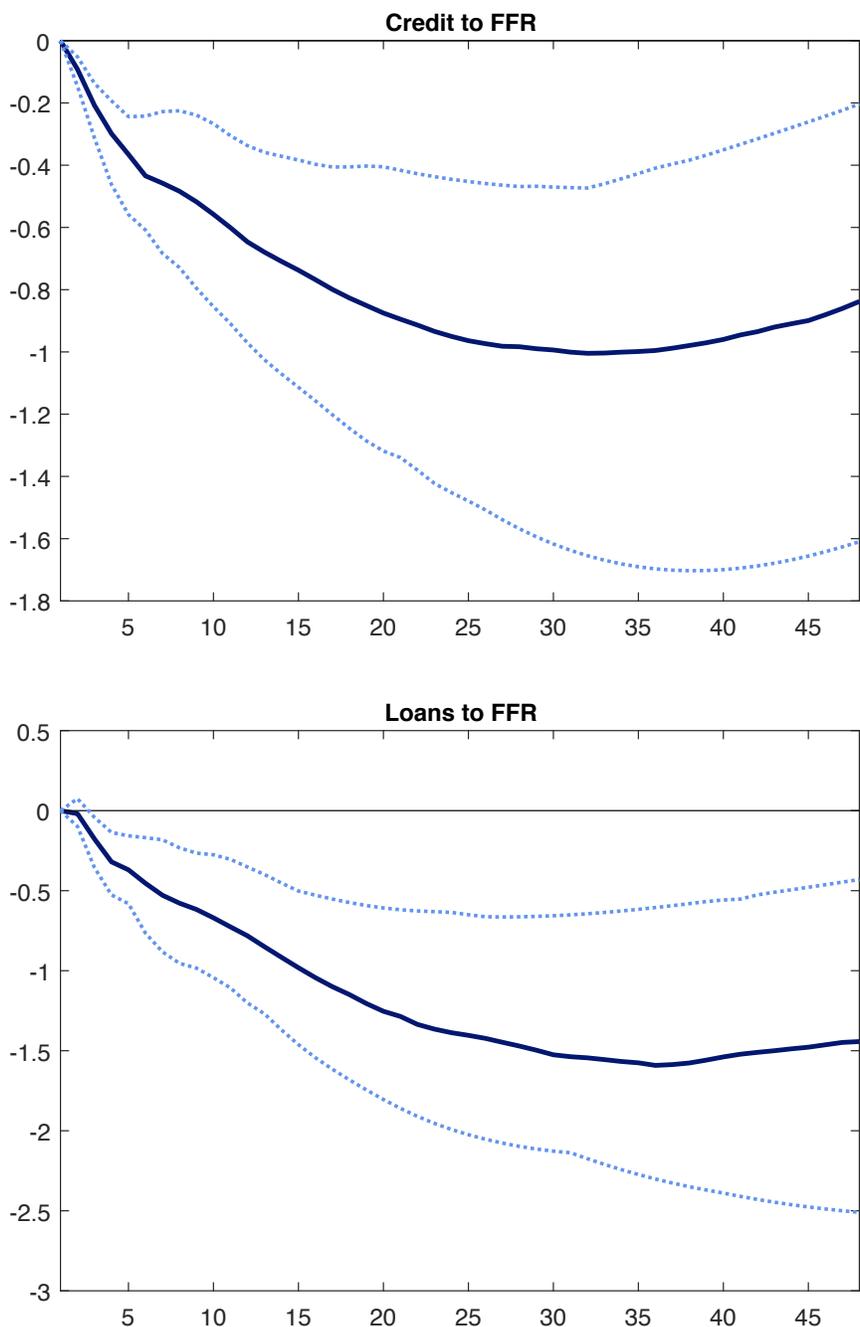
Figure D.1: Response of Total Credit and Bank Loans to a Shock to the Federal Funds Rate
The figure shows the impulse response of total credit (upper panel) and bank loans (lower panel) to a 1% shock to the federal funds rate estimated by a VAR. The units of the x-axis are months. The units of the y-axis are percent. The VAR includes the federal funds rate, the unemployment rate, the log of the CPI, the log level of bank loans, and the log of total credit. Both bank loans and total credit are deflated by the CPI. Total credit is defined as the sum of corporate and household debt. The frequency is monthly, and we use six lags of each variable. The data are from the FRED database from the Federal Reserve Bank of St. Louis.

Appendix E. Demand Estimate Results Details

Table E.1: Demand Estimation

In this table, we report the estimated deposit and loan demand parameters. The first column corresponds to the deposit demand parameter estimates. The second column contains the loan demand parameter estimates. Yield sensitivity (α) refers to the average sensitivity of depositors (borrowers) to deposit rates (loan rates). Log No. of Branches (β_1) refers to the sensitivity of depositors (borrowers) to the log number of branches that each bank operates. Log No. of Employees (β_1) refers to the sensitivity of depositors (borrowers) to the log number of employees per branch. Yield sensitivity (σ_α) refers to the dispersion in the sensitivity of depositors to deposit rates, with the dispersion set at 0 for firms. The sample includes all U.S. commercial banks from 1994 to 2017. The data sources are the Call Reports and the FDIC Summary of Deposits.

	Deposit	Loan
Yield Sensitivity (α)	0.968***	-1.424***
	[0.140]	[0.292]
Log No. of Branches (β_1)	0.804***	0.944***
	[0.012]	[0.000]
Log No. of Employees (β_2)	0.714***	0.630***
	[0.015]	[0.025]
Yield Sensitivity Dispersion (σ_α)	1.916***	
	[0.403]	
Sector F.E.	Y	Y
Time F.E.	Y	Y
Observations	18062	18062
Adj. R^2	0.966	0.570

Appendix F. Additional Analysis of Monetary Policy Shocks and Bank Equity Returns

In this Appendix, we provide additional evidence on the relation between monetary policy shocks and stock returns. First, we replicate the results in Table 5 in the main text using one-day changes in the one-year Treasury yield on FOMC days instead of using two-year Treasury yields. The results reported in Table F.1 confirm that this alternative measure of monetary policy shocks leads to similar results. We still find a non-monotonic relation between monetary policy shocks and the returns on bank equity, and the relation is more pronounced in markets with higher bank concentration.

We next examine the possibility that the positive announcement effects of bank stock returns in a low interest rate environment could be driven by the information channel proposed by Nakamura and Steinsson (2018). We perform two exercises to address this concern. First, in Figure F.1, we examine returns for all 49 Fama-French industries. If positive announcement effects are mainly a consequence of the central bank's assessment of the economic outlook, then we should see similar effects in other industries as well. However, 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 rate environment. Second, we use the monetary policy shocks constructed by Jarocinski and Karadi (2018). Their measure purges out the central bank information shocks from the surprises in federal funds futures on the FOMC days. Our results are robust to this alternative measure of a monetary policy shock. This last result mitigates the concern that our reversal rate result is driven by a central bank information effect.

14

Table F.1: Monetary Policy Shocks and Bank Equity Returns on FOMC Days

	(1) High FFR	(2) Low FFR	(3) High FFR	(4) Low FFR
Policy shock	-1.365**	1.290	-0.103	-2.623
	[0.597]	[1.326]	[1.019]	[2.108]
HHI*Policy shock			-0.303	0.740**
			[0.262]	[0.372]
Δ Term spread	-0.605	2.944***	-0.471	2.644**
	[0.688]	[1.103]	[0.696]	[1.070]
Market return	0.298***	0.727***	0.296***	0.730***
	[0.074]	[0.069]	[0.073]	[0.069]
Observations	27,257	33,805	27,257	33,805
Adj. R^2	0.015	0.124	0.016	0.125

This table reports the estimates of the relation between bank equity returns and monetary policy shocks on FOMC days. Monetary policy shocks are measured by the one-day changes in the one-year Treasury yield on FOMC days. HHI is the Herfindahl-Hirschman Index of the local deposit market in which the bank operates. The sample for columns (1) and (3) constitutes observations in which the starting level of the federal funds rate is above 2%. The sample for columns (2) and (4) constitutes observations in which the starting level of the federal funds rate is below 2%. We exclude observations during the collapse of the dot-com bubble (2000-2001) and the financial crisis (2007-2009). The standard errors are clustered by time.
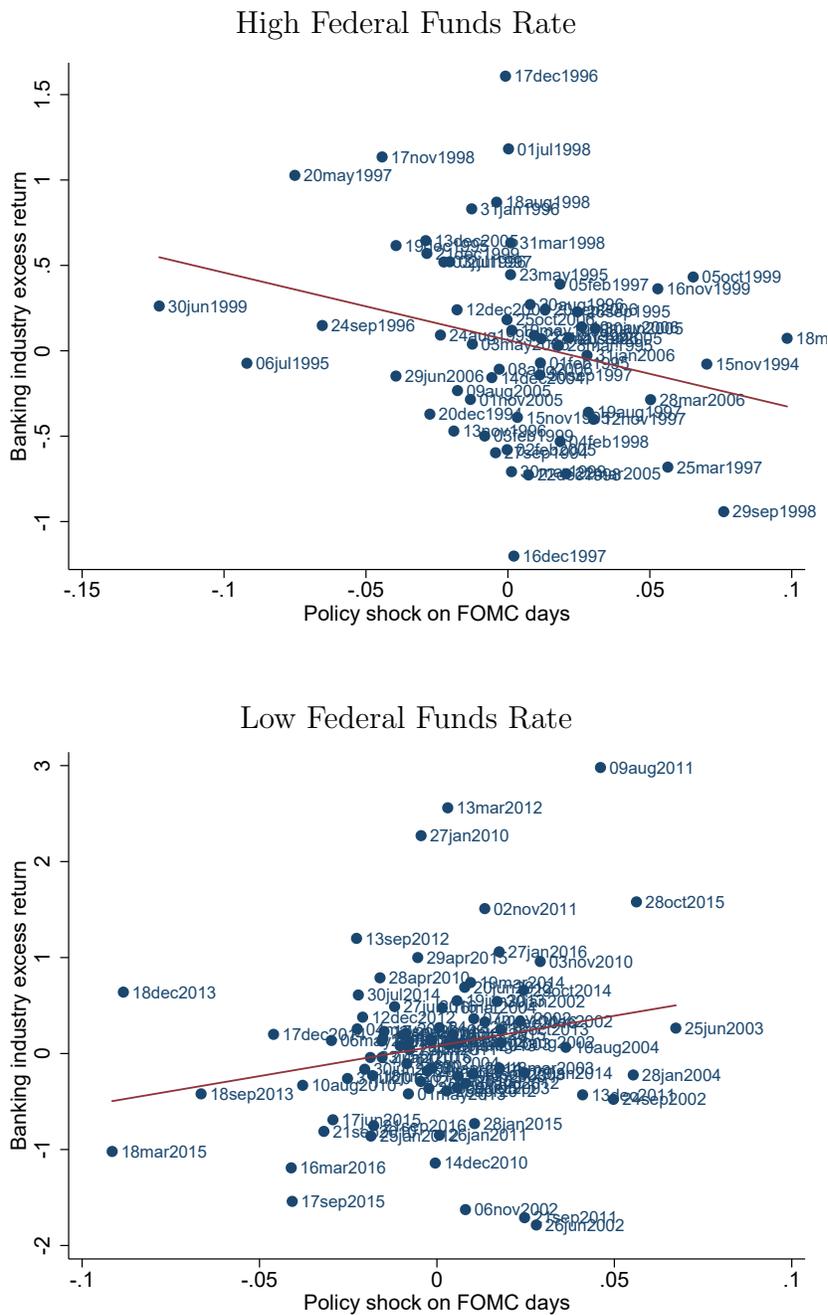
Figure F.1: Monetary Policy Shocks and Bank Equity Returns

This figure provides the scatter plot of bank industry excess returns against monetary policy shocks on FOMC days from 1994 to 2017. The excess returns are defined as the difference between bank industry index returns and the market returns. Monetary policy shocks are measured by the change in the 3-month federal funds future subtracting central bank information shocks (Jarocinski and Karadi 2018). The sample for the upper panel constitutes observations in which the starting level of the federal fund rate is above 2%. The sample for the lower panel constitutes observations in which the starting level of the federal fund rate is below 2%. We exclude observations during the collapse of the dot-com bubble (2000-2001) and the financial crisis (2007-2009). Bank industry stock returns are from Kenneth French's website and the monetary policy shocks are from the website of the *American Economic Journal: Macroeconomics*.
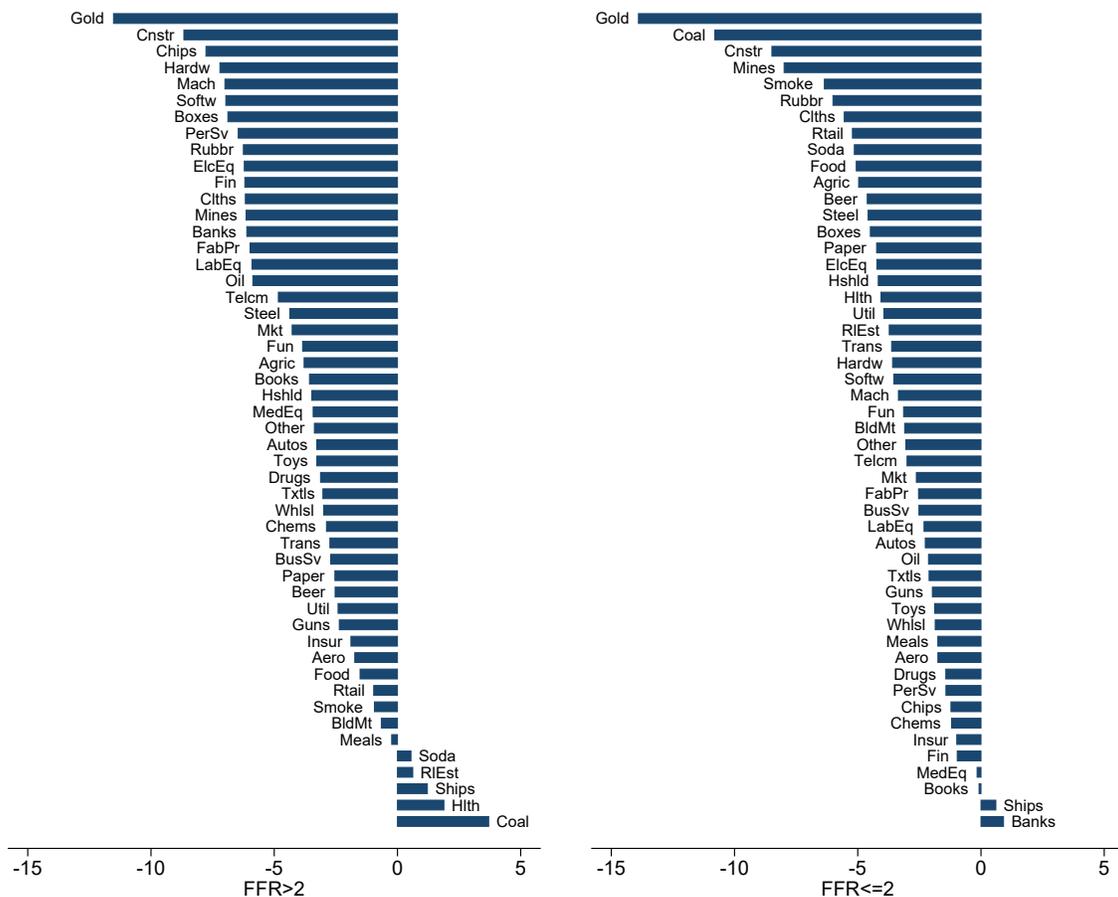
16

Figure F.2: Monetary Policy Shocks and Bank Equity Returns

The figure shows the sensitivity of bank and other industry stock portfolios to monetary policy shocks on the FOMC days. The bars present the coefficients from regressions of Fama-French 49 industry returns on the change in the two-year Treasury rate over a two-day window around FOMC meetings, as in Hanson and Stein (2015). The sample includes all scheduled FOMC meetings from 1994 to 2017. The left panel uses a sample in which the federal funds rate (FFR) is above 2%. The right panel uses a sample in which the federal funds rate is below or equal to 2%. The industry returns are from Kenneth French's website. The two-year Treasury rate is obtained from the FRED database.