The Journal of FINANCE
The Journal of THE AMERICAN FINANCE ASSOCIATION

THE JOURNAL OF FINANCE • VOL. LXXVII, NO. 4 • AUGUST 2022

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

YIFEI WANG, TONI M. WHITED, YUFENG WU, and KAIRONG XIAO*

ABSTRACT

We quantify the impact of bank market power on monetary policy transmission through banks to borrowers. We estimate a dynamic banking model in which monetary policy affects imperfectly competitive banks' funding costs. Banks optimize the pass-through of these costs to borrowers and depositors, while facing capital and reserve regulation. We find that bank market power explains much of the transmission of monetary policy to borrowers, with an effect comparable to that of bank capital regulation. 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.

WE EXAMINE THE QUANTITATIVE IMPACT of bank market power on the transmission of monetary policy through the banking system. This transmission channel is potentially important, given three decades of consolidation in the banking industry that has softened competitive pressure. Moreover,

*Yifei Wang is from Cornerstone Research. Toni M. Whited is from the University of Michigan and NBER. Yufeng Wu is from the University of Illinois. Kairong Xiao is from Columbia University. None of the authors have any relevant or material financial interests related to the research described in this paper. We would like to thank Stefan Nagel (Editor), an anonymous Associate Editor, two anonymous referees, Harjoat Bhamra, Marco Bonomo, Yasser Boualam, Dean Corbae, Olivier Darmouni, Itamar Dreschler, Mark Egan, Brent Glover, Valentin Haddad, Ali Hortaçsu, Frank de Jong, Erica Li, Gregor Mavtos, Patricia Mosser, Stijn van Nieuwerburgh, Neil Pearson, George Pennacchi, Alexi Savov, David Sraer, Olivier Wang, and Pavel Zryumov for helpful comments and discussions. We also thank participants at the CICF, EFA, FIRS, the Chicago Booth Asset Pricing Conference, the Financial Innovation and Risk Management Conference, the FMA Wine Country Finance Conference, the FRBSF Conference on Advances in Financial Research, the Macro-Finance Society Workshop, the NBER Summer Institute, the Northeastern University Finance Conference, the RCFS Conference, the Short-Term Funding Markets Conference, the University of Connecticut Conference, the UBC Summer Conference, the Stanford Institute for Theoretical Economics (SITE) Summer Workshop, and the SFS Cavalcade, as well as seminar participants at the Bank of Canada, Chicago Booth, Columbia University, CUHK, Georgetown University, Johns Hopkins University, Northwestern University, the FDIC, the Federal Reserve Bank of New York, the Federal Reserve Board, ITAM, INSPER, the University of Lausanne, the University of Michigan, the University of Rochester, UIUC, Wharton, Harvard University, MIT, and Wirtschaftsuniversität Wien.

Correspondence: Toni M. Whited, Ross School of Business, University of Michigan and NBER, Ann Arbor, MI 48109; e-mail: twhited@umich.edu

DOI: 10.1111/jofi.13159

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)). Yet, market power is not the only friction in the banking system that influences pass-through. For example, traditional analysis of monetary policy transmission focuses on regulatory constraints, such as bank reserve or capital requirements, as the central frictions that influence monetary policy transmission (Bernanke and Blinder (1988), Kashyap and Stein (1995)). However, the qualitative nature of this evidence leaves open the question of the relative magnitude of traditional versus market-power transmission channels.

To do so, we use 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 identify the relative magnitudes 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 to those of bank capital regulation, while the effects of bank reserve requirements are limited.

In further analysis, we first show that banks face nontrivial costs when they access external financial markets. These frictions play a pivotal role in connecting banks' deposit market power to their lending decisions, as external financing costs link banks' sources and uses of funds. In addition, these frictions help explain the differential sensitivity of lending to the policy rate between 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 find that when the federal funds rate is below 0.9%, further cuts in the policy rate 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 a dynamic industry equilibrium, imperfectly competitive banks act as intermediaries between borrowers and depositors. Banks' lending decisions are dynamic for two reasons: financial frictions that induce precautionary capital accumulation and a maturity mismatch between short-term deposits and long-term loans. In this setting, monetary policy moves 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 balance-sheet 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, Savov, and Schnabl (2017)). The fourth is the loan market power channel, in which banks reduce markups to mitigate the effects of monetary tightening on loan demand (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 loan and deposit demand to interest rates. We then plug these estimates into our model and use simulated minimum distance (SMD) 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 assess 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, Savov, and Schnabl (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, Savov, and Schnabl (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.

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), Campos (2019), Wang (2019)). While these studies are insightful, 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 rate policy (Demiralp, Eisenschmidt, and Vlassopoulos (2017), Basten and Mariathasan (2018), Heider, Saidi, and Schepens (2019)). These studies show that negative policy rates 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, Hortaçsu, and Matvos (2017), Buchak et al. (2018), Xiao (2020), Egan, Lewellen, and Sunderam (2022)). 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 tangentially related to recent work that uses dynamic banking models to study optimal capital regulation (Begenau (2018), Begenau and Landvoigt (2021), Landvoigt, Van Nieuwerburgh, and Elenev (2021)). In particular, Corbae and D'Erasmo (2021) develop a dynamic dominant-fringe model to study the quantitative impact of regulatory policies on bank risk-taking and market structure. Our paper stands apart from this literature in that we emphasize the role of imperfect competition in monetary transmission. Moreover, our approach is more empirical in nature, as we estimate, rather than calibrate, all of our model parameters.

The paper proceeds as follows. Section I describes our data. Section II presents the model. Section III discusses the estimation procedure and results. Section IV presents the results of our counterfactuals. Section V examines subsample heterogeneity and model robustness. Section VI concludes.

I. Data and Stylized Facts

Our main data set is the Consolidated Reports of Condition and Income (Call Reports). This data set 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 Federal Deposit Insurance Corporation (FDIC) Summary of Deposits, which provides branch-level information on each bank since 1994 at an annual frequency. The sample period is 1994 to 2017.

We also use several other data sources. First, we retrieve publicly listed bank returns from the Center for Research in Security Prices (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 (FOMC) meeting dates from the FOMC meeting calendar. Finally, we obtain the following time series from the Federal Reserve Economic Data (FRED) database: National Bureau of Economic Research (NBER) recession dates, the effective federal funds rate, two- 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 I.

Table II 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 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, Savov, and Schnabl (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

Table I
Variable Definitions

Variable	Construction
Deposit market share	Deposits of a bank divided by the sum of deposits, cash, Treasury bills, and money market funds in the U.S. economy.
Loan market share	Loans of a bank divided by the sum of U.S. household debt, corporate debt, and corporate equity.
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	Noninterest 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. Nonmortgage loan maturity is the repricing maturity and average prepayment adjusted mortgage duration comes from Landvoigt, Van Nieuwerburgh, and Elenev (2021).
Nonreservable borrowing share	Nonreservable 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.

the five-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 loan 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.

II. Model

While this evidence suggests interesting equilibrium 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

Table II

Summary Statistics

In this table, we report summary statistics for our sample. The sample period is 1994 to 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 nonfinancial 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. Expenses 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

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 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 nondeposit 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.

[1] In reality, banks accept deposits from firms and extend loans to households. Therefore, in our model, households should be interpreted broadly as savers, and firms should be interpreted broadly as borrowers.

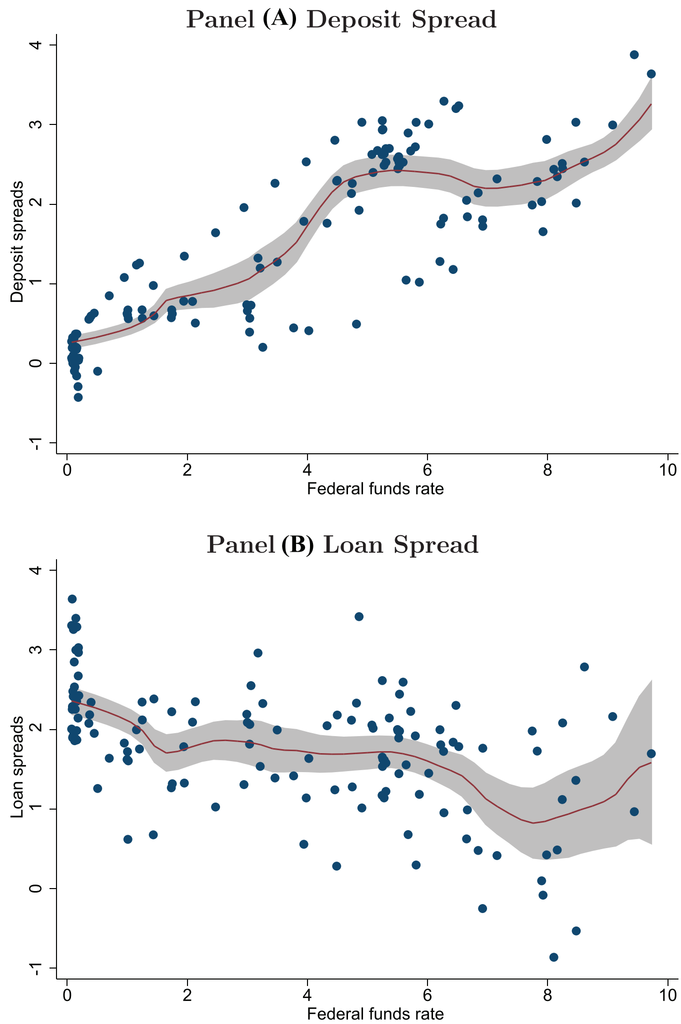

Figure 1. Deposit spread, loan spread, and the federal funds rate. In this figure, we plot kernel regressions of average deposit and loan spreads for U.S. banks on the federal funds rate. We use an Epanechnikov kernel with a bandwidth of 0.66 for deposits and 0.61 for loans. The sample period is 1985 to 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. (Color figure can be viewed at wileyonlinelibrary.com)

A. 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. Households allocate their endowments in across three investment options: cash, bonds, and deposits. 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. Because the households' problem is static, we drop the t subscript hereafter for convenience. 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, 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 zero, 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. All interest rates are quoted in real terms, as we assume that inflation expectations are anchored at zero, but in the general equilibrium version of the model in Section XI of the Internet Appendix, we allow for an endogenously determined inflation rate.[2]

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 (2020)). We model the distribution of depositors' yield sensitivity as a uniform distribution with mean α^d and standard deviation σ_α^d. The coefficients β^d are sensitivities to the nonrate product characteristics x_j^d, and ξ_j is the unobservable product-level demand shock. We let $\epsilon_{i,j}^d$ represent a relationship-specific shock to the choice of option j by household i and $\epsilon_{i,j}^d$ capture 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 the 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

[2] The Internet Appendix may be found in the online version of this article.

market share, s_j^d, as

$$
\begin{aligned}
s_j^d\left(r_j^d|f\right) &\equiv \int \mathbb{I}_{i,j}^d dF(\epsilon) \\
&= \sum_{i=1}^{I} \mu_i^d \frac{\exp\left(\alpha_i^d r_j^d + \beta^d x_j^d + \xi_j^d\right)}{\exp\left(\alpha_i^d f + \beta^d x_{J+1}^d + \xi_{J+1}^d\right) + \exp\left(\beta^d x_c^d + \xi_c^d\right) + \sum_{m=1}^{J} \exp\left(\alpha_i^d r_m^d + \beta^d x_m^d + \xi_m^d\right)},
\end{aligned} \tag{3}
$$

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, $\exp(\alpha_i^d f + \beta^d x_{J+1}^d + \xi_{J+1}^d)$, represents the utility of holding Treasury bills, and the second term, $\exp(\beta^d x_c^d + \xi_c^d)$, 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,

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

B. Firms

There is a mass K of firms, where we again drop the time subscript. Each firm wants to borrow one dollar, so aggregate borrowing demand is K. Firms can borrow by issuing long-term bonds or by 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 are long term and that a fraction η of the outstanding balance is due every year. Thus, if the borrower obtains one dollar, this debt has a maturity of $\frac{1}{\eta}$ years, on average. If a firm obtains a loan from bank j, it will be charged a fixed interest rate of r_j^l. If a firm issues long-term bonds, the interest rate is given by an expected default cost, δ, plus the expected weighted average of future federal funds rates, f_t, which is given by

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

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,

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

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 mean α^l and standard deviation σ_α^l. The sensitivities to nonrate characteristics, x_j^l, are given by β^l, and ξ_j^l is the unobservable product-level demand shock. We let $\epsilon_{i,j}^l$ represent an idiosyncratic shock when firm i borrows from bank j. The optimal choice of firm i is given by the 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, as

$$\begin{aligned} s_j^l\left(r_j^l|f\right) &\equiv \int \mathbb{I}_{i,j}^l dF(\epsilon) \\ &= \sum_{i=1}^I \mu_i^l \frac{\exp\left(\alpha_i^l r_j^l + \beta^l x_j^l + \xi_j^l\right)}{\exp\left(\alpha_i^l(\overline{f} + \overline{\delta}) + \beta^l x_{J+1}^l + \xi_{J+1}^l\right) + \exp\left(\beta^l x_n^l + \xi_n^l\right) + \sum_{m=1}^J \exp\left(\alpha_i^l r_m^l + \beta^l x_m^l + \xi_m^l\right)}, \end{aligned} \tag{8}$$

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, $\exp(\alpha_i^l(\overline{f} + \overline{\delta}) + \beta^l x_{J+1}^l + \xi_{J+1}^l)$, represents the utility from issuing bonds, and the second term, $\exp(\beta^l x_n^l + \xi_n^l)$, 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,

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

C. 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 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 return of zero, 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 $r^l_{j,t}$, firms solve their profit-maximization problem, which yields the quantity of loans borrowed from bank j, $B_j(r^l_{j,t})$, given by equation (9). To simplify notation, we suppress the dependence of loans and deposits on the relevant interest rates, denoting them simply by $D_{j,t}$ 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 $r^l_{j,t}$. The new loans, once issued, have the same maturity structure as existing loans, 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

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

where γ is the bank's discount factor. To simplify model computation, we assume that the borrower repays a fraction η of the principal plus $I_{j,t}$ at the end of the first year, and from the second year onward the borrower repays a fraction η of the remaining principal. These assumptions about maturity structure imply that a bank's outstanding loans evolve according to

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

We assume that in each period, a random fraction of maturing loans, $\delta_t \in [0, 1]$, becomes delinquent. The bank takes δ as an exogenous state variable in its decision-making problem. We assume that the bank writes off delinquent payments, with charge-offs equal to $\delta_t \times \eta \times (L_t + B_t)$, so default depresses the bank's current-period cash flows. However, defaulting on a payment in one period does not exonerate the borrower from future payments, so 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 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 nonreservable borrowing, N_t. A typical example of nonreservable borrowing is large-denomination CDs. As argued by Kashyap and Stein (1995), because nonreservable borrowing is not insured by FDIC deposit insurance, purchasers of this debt must concern themselves with the default risk of the issuing bank and therefore with any possible deadweight default costs. These considerations imply that the marginal cost of nonreservable borrowing is likely an increasing function of the amount raised. Thus, we assume that nonreservable borrowing incurs a quadratic financing cost

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

which exceeds the prevailing federal funds rate.

The cost in equation (13) represents an important friction because, in its absence, banks could always raise nonreservable funding to compensate for any deposit shortfalls. The availability of such funding 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,

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

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 nonreservable 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, the bank's beginning-of-period book equity, evolves according to

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

In equation (17), 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, to bank capital regulation. The profits in equation (17) are then given by

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

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

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

This constraint reflects a bank's limited liability and thus implies that banks cannot raise equity capital to replace deposits or nonreservable borrowing. In Section V, we replace this assumption with costly equity issuance, finding only a limited impact on our results, as banks' equity issuances are 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 noninterest-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, in Section IV, we show that modeling this newer policy has a limited impact on our model solution.

D. 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

$$\begin{bmatrix} \ln \delta_{t+1} - \mathbb{E}(\ln \delta) \\ \ln f_{t+1} - \mathbb{E}(\ln f) \end{bmatrix} = \begin{bmatrix} \rho_\delta & \rho_{\delta f} \\ 0 & \rho_f \end{bmatrix} \cdot \begin{bmatrix} \ln \delta_t - \mathbb{E}(\ln \delta) \\ \ln f_t - \mathbb{E}(\ln f) \end{bmatrix} + \begin{bmatrix} \sigma_\delta & 0 \\ 0 & \sigma_f \end{bmatrix} \varepsilon_{t+1}, \tag{22}$$

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 nonreservable 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.

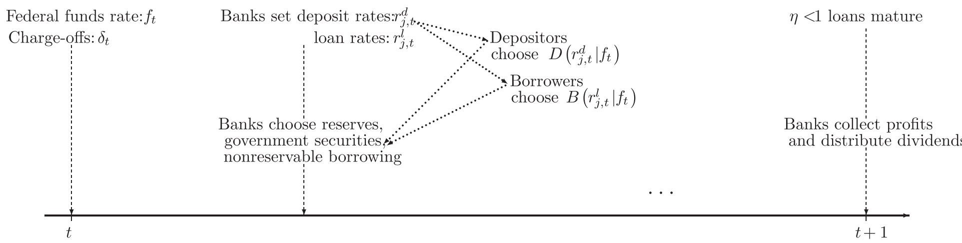

Figure 2. Timeline within a period.

E. Bank's Problem and Equilibrium

Figure 2 shows the sequence of events in each time period. The bank enters the period and observes both 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 the amount of loans to firms. Depending on the extent of these activities, the banks adjust their reserves, holdings of government securities, and nonreservable borrowing. Finally, banks collect profits at the end of the period and distribute dividends to shareholders.

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^d_t and r^l_t, as well as nonreservable borrowing N_t and investment in government securities G_t, it rationally considers the choices made by other banks in both 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,

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

$$s.t. (4), (9), (10), (11), (12), (16), (17), (18), (19), (20), (21), (22), (23).$$

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 consider 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 Section I of the Internet Appendix, 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 banks' beliefs. The R^2s for these approximations are over 95% for the deposit market and 97% 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, not on 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.

F. Monetary Policy Transmission in a Static Model

In this subsection, we use a simplified, static version of the model to highlight the intuition behind the effects of frictions on monetary policy transmission. We assume that both deposits and loans have a maturity of one year, thus precluding a maturity transformation. Because the model is then static, we drop the t subscripts. Banks face no idiosyncratic uncertainty, and their ex ante identical problems produce ex post identical solutions. In addition, all depositors have the same rate sensitivities: $\alpha_i^d = \alpha^d, \forall i$. Finally, we assume that the operating costs, ϕ^d and ϕ^l, are zero. Derivations of all of the statements in this subsection can be found in Section II of the Internet Appendix.

F.1. Frictionless Benchmark

First, we consider a frictionless benchmark in which the bank has no market power in either the deposit or the loan market, so the deposit and loan demand elasticities are infinite. Moreover, nonreservable borrowing is frictionless, so $\phi^N = 0$, and banks face neither a capital requirement nor a reserve requirement, so $\kappa = \theta = 0$.

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

$$\Pi_j = \max_{\{r_j^l, r_j^d\}} \left\{ r_j^l B_j - r_j^d D_j - \left(B_j - D_j\right) f \right\}, s.t.\ r_j^d \geq 0. \tag{27}$$

In the absence of balance sheet frictions, the bank optimizes its choices of deposits and loans separately. For example, when deposits fall short of loans, a bank can make up any funding shortfall, $B_j - D_j$, with nonreservable borrowing at a cost that equals the federal funds rate, f, with no additional financing costs. Conversely, when there are excess deposits, the bank can invest any of this surplus, $D_j - B_j$, in government securities and earn the federal funds rate, f. Furthermore, because competition is perfect, both r_j^l and r_j^d equal f.

F.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. We start with the loan market, deriving the expression for the loan spread by taking the first-order conditions of equation (27) with respect to the loan rate, r_j^l. With the loan demand functions specified in equations (8) and (9), the optimal lending rates are given by the federal funds rate plus markups,

$$r_j^l = f + \left(\frac{-\partial B_j / \partial r_j^l}{B_j} \right)^{-1} = f + \left(-\alpha^l \left(1 - S^l / J \right) \right)^{-1}, \tag{28}$$

where $S^l \equiv \sum_{j=1}^{J} s_j^l$ is the total market share of all banks in the loan market. Similarly, the optimal deposit rates are given by the federal funds rate minus markups,

$$r_j^d = f + \left(\frac{-\partial D_j/\partial r_j^l}{D_j}\right)^{-1} = f - \left(\alpha^d\left(1 - S^d/J\right)\right)^{-1}, \tag{29}$$

where $S^d \equiv \sum_{j=1}^{J} s_j^d$ is the total market share of all banks in the deposit market.

Equations (28) and (29) emphasize that markups depend on both rate sensitivities and market concentration. These equations imply that markups rise when the rate sensitivities, α^l and α^d, fall in absolute value. Holding the total size of the banking sector fixed, a more concentrated market (lower J) also leads to higher markups. Note that market power does not disappear completely when the number of banks goes to infinity because idiosyncratic taste shocks, as represented by $\epsilon_{i,j}$ in equations (1) and (6), generate product differentiation. Only when the rate sensitivities go to infinity do markups converge to zero.

Crucially, banks' markups also depend on the federal funds rate because the latter affects the attractiveness of bank deposits and loans relative to households' or firms' outside options. To see this point, we take the derivative of the loan spread with respect to f,

$$\frac{d(r_j^l - f)}{df} = -\frac{(S^l/J)(1 - S^l - s_{J+1}^l)}{(1 - S^l/J)^2 + S^l(1 - S^l)/J} < 0, \tag{30}$$

where s_{J+1}^l is the share of firms that choose to borrow from the bond market. We can see that banks' loan spreads decline with the federal funds rate. In the lending market, an increase in the federal funds rate makes bank loans less attractive relative to the outside option of not borrowing. Therefore, lending shrinks, and banks optimally reduce markups on loans to mitigate the effects of lower loan demand.

Next, we derive an analogous relation between the deposit spread and f,

$$\frac{d(f - r^d)}{df} = \frac{(S^d/J)(1 - S^d - s_{J+1}^d)}{(1 - S^d/J)^2 + S^d(1 - S^d)/J} > 0, \tag{31}$$

where s_{J+1}^d is the market share of Treasury bills. We see that banks' deposit spreads increase with the federal funds rate. In the deposit market, an increase in the federal funds rate makes bank deposits more attractive relative to the outside option of holding cash, thus allowing banks to charge larger markups on their deposits (e.g., Drechsler, Savov, and Schnabl (2017)).

Bank market power also affects the relation between the quantity of loans and the policy rate. In this simplified setting, the equilibrium quantity of bank loans depends solely on banks' rate-setting decisions in the loan market because, as noted above, in the absence of financial or regulatory frictions, loan supply is independent of the quantity of deposits. Thus, the relation between bank loans and the policy rate is given by

$$\frac{d\log B_j}{df} = \alpha^l \frac{(1-S^l/J)^2(1-S^l-s^l_{J+1})}{(1-S^l/J)^2+S^l(1-S^l)/J} < 0, \tag{32}$$

which is negative. Moreover, the magnitude increases with rate sensitivity in the loan market α^l and the number of banks J. Intuitively, as bank market power rises, the loan spreads they can charge also rise, so the pass-through of the policy rate f to loan rates falls, thus dampening the impact of monetary policy on bank lending.

In this simplified setting with homogeneous depositor rate sensitivities, the effect of the policy rate on the quantity of bank deposits is positive, as households move cash proportionally into deposits and interest-bearing assets when f rises,

$$\frac{d\log D_j}{df} = \alpha^d \frac{(1-S^d/J)^2(1-S^d-s^d_{J+1})}{(1-S^d/J)^2+S^d(1-S^d)/J} > 0. \tag{33}$$

This result stands in contrast to the results in Drechsler, Savov, and Schnabl (2017) and in our full model, but the reason is instructive. Without heterogeneity in depositor rate sensitivities, logit demand implies the independence of irrelevant alternatives. This property prevents a negative relation between the federal funds rate and deposits in our static model because it precludes the strong substitution out of deposits and into other interest-bearing assets when the federal funds rate rises.

Specifically, with depositor heterogeneity, deposit quantity can fall in response to an increase in the federal funds rate. When depositors are heterogeneous, an increase in the federal funds rate lowers the average yield sensitivity of banks' clientele, as yield-insensitive depositors move disproportionately from cash to deposits, rather than from cash to bonds. In response to the reduction in the average yield sensitivity, banks charge higher deposit spreads, which further drive yield-sensitive clientele to bonds. Overall deposit demand can fall if depositor heterogeneity is sufficiently large to induce a sharp rise in deposit spreads. This force is absent when depositors are homogeneous.

F.3. Balance Sheet Frictions

Next, we consider banks' balance sheet frictions, which imply that they incur additional costs when engaging in nonreservable borrowing. In this case, the banks' optimization problem is

$$\Pi_j = \max_{\{r^l_j, r^d_j\}} \left\{ r^l_j B_j - r^d_j D_j - \Phi^N(N_j) \right\}, s.t.\ r^d_j \geq 0, \tag{34}$$

where $\Phi^N(N_j)$ is the cost of nonreservable borrowing and $N_j = B_j - D_j$ is the funding imbalance. The presence of $\Phi^N(N_j)$ implies that the bank cannot costlessly substitute between deposits and nonreservable borrowing as a funding source. Thus, any decrease in equilibrium deposits impacts bank lending, as banks need to turn to a more expensive funding source. In particular, together with the logit demand in equation (3), equation (31) implies that as market power rises, the ensuing rise in markups results in smaller deposit quantities in equilibrium, so banks need to use more nonreservable borrowing.

F.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 interest on reserves is zero, banks' optimization problem becomes

$$\Pi_j = \max_{\{r_j^l, r_j^d\}} \left\{ r_j^l B_j - r_j^d D_j - \left(B_j + R_j - D_j\right) f \right\}, s.t.\, R_j \geq \theta D_j. \tag{35}$$

Because the interest rate on reserves is zero, the reserve constraint is binding, and the opportunity cost of holding reserves is θf. When the federal funds rate increases, it increases the banks' opportunity cost of holding deposits, which leads to widened deposit spreads and a fall in the household's deposit demand. If banks face balance sheet frictions so that they cannot perfectly replace deposits with nonreservable funding, then the supply of loans falls together with deposits.

F.5. Capital Regulation

Next, we assume that banks face regulation that requires bank capital to exceed a certain fraction of bank assets. In this case, the banks' optimization problem becomes

$$\Pi_j = \max_{\{r_j^l, r_j^d\}} \left\{ r_j^l B_j - r_j^d D_j - \left(B_j - D_j\right) f \right\}, s.t.\, r_j^d \geq 0,\ E_j + \Pi \geq \kappa B_j, \tag{36}$$

where E_j is a bank's initial capital, Π_j is the bank's profit, and κ is the minimum capital requirement. Because we assume no dividends in our static model, the bank's end-of-period capital is given by $E_j + \Pi$.

Equation (36) shows that in the presence of capital regulation, movements in bank capital (E_j) affect lending capacity. Although it falls outside the scope of this static environment, one possible source of movements in capital is maturity mismatch. Because deposits are short term, deposit rates adjust instantaneously when the federal funds rate rises. Loans are long term, however, so only a fraction of loans matures, with the remaining loans carrying the same prefixed rate. Hence, an increase in the federal funds rate temporarily reduces bank capital and tightens the bank capital constraint.[3]

[3] Drechsler, Savov, and Schnabl (2021) note that deposit rate stickiness can dampen this effect.

A final way that monetary policy affects bank capital is through market power. Changes in the federal funds rate affect the attractiveness of bank deposits and loans relative to households' or firms' outside options. As a result, as equations (28) and (29) suggest, the markups that banks can charge on their deposits and loans depend crucially on the federal funds rate, as movements in the federal funds rate affect the profit that banks can generate from the deposit and loan markets. In turn, movements in profits affect the tightness of their capital constraints.

III. Estimation

In this section, we describe our estimation method, present results, and conduct counterfactuals to measure the relative importance of various banking frictions for monetary policy transmission.

A. 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 SMD to estimate the remaining parameters that describe banks' balance sheet frictions.

We estimate deposit and loan demand, given by equations (3) and (8), using the methods in Berry, Levinsohn, and Pakes (1995) (BLP) 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, and time fixed effects. We provide a brief outline of our implementation of this procedure below. A detailed explanation is in Section III of the Internet Appendix.

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 Section IA.I of the Internet Appendix, 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 Ho and Ishii (2011), we use salaries and noninterest expenses related to the use of fixed assets.

We need to assume that these supply shifters are orthogonal to unobservable demand shocks and thus shift the supply curve along the demand curve, allowing us to trace out the slope of the demand curve. One assumption that supports but does not guarantee identification is that customers do not care about these costs, holding product characteristics constant. 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),

$$D_j\left(r_j^d|f\right) = \sum_{i=1}^{I} \mu_i^d \frac{\exp\left(\hat{\alpha}_i^d r_j^d + q_j^d\right)}{\exp\left(\hat{\alpha}_i^d f\right) + \exp\left(q_c^d\right) + \sum_{m=1}^{\hat{J}} \exp\left(\hat{\alpha}_i^d r_m^d + q_m^d\right)} W \tag{37}$$

$$B_j\left(r_j^l|f\right) = \frac{\exp\left(\hat{\alpha}^l r_j^l + q_j^l\right)}{\exp\left(\hat{\alpha}^l\left(\overline{f} + \overline{\delta}\right)\right) + \exp\left(q_n^l\right) + \sum_{m=1}^{\hat{J}} \exp\left(\hat{\alpha}^l r_m^l + q_m^l\right)} K, \tag{38}$$

where $\hat{J}$ is the number of banks used in the second-stage estimation, and q generically represents an option's quality value, which is the utility derived from nonrate product characteristics. As shown in Section III of the Internet Appendix, 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 (37) and (38) 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 (37) and (38) 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 produces parameter estimates that minimize the distance between moments (or functions of moments) generated by the model and their analogs in the data. We use 10 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 nonreservable borrowing as a fraction of deposits to identify the cost of holding nonreservables, 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 noninterest expenses scaled by assets. This moment measures the costs that banks pay outside of their routine deposit-taking and loan-servicing businesses. 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 because 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 and 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 Section V of the Internet Appendix. These two moments suit this purpose for several reasons. Holding banks' market shares constant, when W/K increases, the value of deposits rises relative to the value of loans, leading to a higher deposit-to-asset ratio. In addition, the deposit-to-asset ratio is positively related to q_n^l because loan demand falls with this outside option. Fewer loans imply a smaller overall balance sheet, so ceteris paribus, the ratio of deposits to assets rises. The sensitivity of aggregate corporate borrowing to the federal funds rate also helps identify q_n^l because 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.

Finally, we include two extra moments that are important for ensuring that our counterfactuals are empirically relevant. To make our model accurately reflect bank valuation, we include banks' average market-to-book ratio. To ensure that our model correctly quantifies baseline monetary transmission from the federal funds rate to bank lending, we include the sensitivity of bank lending to the federal funds rate, again estimated using a VAR. Otherwise, it would be hard to claim that our model provides a sensible decomposition of the various monetary transmission mechanisms.

B. Baseline Estimation Results

Table III 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 the Basel III 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 six, which matches the average county-level banking concentration in the data.

Table III
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 with 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	0.35	
θ	The reserve ratio	0.024	
κ	The capital ratio	0.06	
$\hat{J}$	Number of representative banks	6	
	Panel B: Parameters Estimated Separately		
μ	Average loan maturity	3.448	[1.445]
$\bar{f}$	Log Federal funds rate mean	−4.305	[0.320]
σ_f	Std of Federal funds rate innovation	0.551	[0.723]
ρ_f	Log Federal funds rate persistence	0.900	[0.040]
$\bar{\delta}$	Log loan chargeoffs mean	−5.905	[0.177]
σ_δ	Std log loan chargeoffs innovation	0.960	[0.339]
ρ_δ	Log loan chargeoffs persistence	0.600	[0.055]
$\rho_{\delta f}$	Corr of Federal funds rate innovation and log loan chargeoffs	−0.110	[0.069]
	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	0.553	[0.116]
α^l	Borrowers' sensitivity to loan rates	−1.462	[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	1.151	[1.065]
	Panel D: Parameters Estimated via SMD		
γ	Banks' discount rate	0.048	[0.007]
W/K	Relative size of the deposit market	0.217	[0.011]
q_n^l	Value of firms' outside option	−9.631	[0.262]
ϕ^N	Quadratic cost of nonreservable borrowing	0.010	[0.001]
ϕ^d	Bank's cost of taking deposits	0.010	[0.001]
ϕ^l	Bank's cost of servicing loans	0.007	[0.001]
ψ	Net fixed operating cost	0.027	[0.006]

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 equation (22). Next, average loan maturity, defined in Table I, is approximately 3.5 years.

Panel C in Table III provides the demand parameters from the first-stage BLP estimation, with details of the estimation results presented in Table IA.II in Section VI of the Internet Appendix. 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 1 percentage point increase in the deposit rate increases a bank's market share by 0.968%, while a 1 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 nonrate bank characteristics. The estimates are also both statistically and economically significant. A 1 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 1 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. These estimated yield sensitivities lie close to similar estimates in the literature, with Dick (2008), Ho and Ishii (2011), and Egan, Hortaçsu, and Matvos (2017) finding deposit sensitivity estimates from 0.6 to 1.1. Comparable loan sensitivity estimates can only be found in the mortgage demand literature, which features estimates in the −1.1 to −5.2 range (DeFusco and Paciorek (2017), Buchak et al. (2018)).

One challenge to identification of this demand estimation is the possibility that when banks face more demand, they hire more higher-quality staff. Conversely, hiring higher-quality staff might spur demand. We confront these concerns in two ways. First, we emphasize that this issue is partially alleviated by the inclusion of bank fixed effects, which absorb the heterogeneity in demand in the cross-section of banks. Similarly, time fixed effects absorb the aggregate shocks to deposit demand. Second, because our instruments may nonetheless be correlated with bank-specific, time-varying unobservable demand shocks, as a robustness check, we also use an instrument from Dick (2008) that is less likely to capture bank-specific labor costs, namely, the weighted average of the local bank teller wages from the Bureau of Labor Statistics (BLS) over the markets in which the bank operates, where the weight is the bank's deposit share in each market relative to its total deposits. This instrument addresses the concern that the Call Report salary data likely contain a quality component that might influence demand. The results from using this instrument are in Tables IA.III and IA.IV in Section VII of the Internet Appendix. We find a deposit sensitivity of 0.668 and a loan sensitivity of −0.950. Both estimates are somewhat smaller than those in Table III, but their relative magnitudes remain intact, even though the sample period for the BLS data runs from only 1997 to 2017. Both estimates are also noisier than those in Table III, with the loan sensitivity being insignificantly different from zero, likely because loan markets are less local in nature than deposit markets. Because of the noise in these estimates, for our baseline analysis, we use the estimates in Table III.

Panel D of Table III presents the parameters from our second-stage SMD estimation. We find that banks have a subjective discount rate of 4.8%, 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,

Table IV
Moment Conditions

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

	Actual Moment	Simulated Moment	t-statistic
Dividend yield	3.38%	2.87%	−0.856
Nonreservable borrowing share	29.90%	25.90%	−1.913
Std of nonreservable borrowing	12.60%	14.73%	0.818
Deposit spread	1.29%	1.32%	0.328
Loan spread	2.03%	2.04%	0.066
Deposit-to-asset ratio	0.699	0.737	1.034
Net noninterest expenses	1.20%	1.03%	−1.654
Leverage	11.20	11.89	1.398
Market-to-book ratio	2.061	1.796	−1.094
Credit-FFR sensitivity	−0.995	−1.008	−0.100
Bank loan-FFR sensitivity	−1.592	−1.641	−0.129

the cost of nonreservable borrowing is both statistically and economically significant. At the average level of nonreservable borrowing (30% of total deposits), a marginal dollar of nonreservable borrowing costs a bank 30 basis points (bps) above the cost implied by the prevailing federal funds rate, where we calculate this cost as $\partial\Phi^N/\partial N = \phi^N N = 0.010 \times 0.3 = 0.003$, which implies that banks pay an extra 30 bps on each extra dollar of nonreservable borrowing. 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 1% cost of maintaining deposits and a slightly lower 0.7% cost of servicing their outstanding loans.

In Table IV, 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. Banks borrow nonreservable securities that amount to 30% of the deposit intake in the data (versus 26% in the model). 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.

C. External Validity

While Table IV provides evidence of model fit, we now proceed to several external model validation exercises. First, in Figure 3, 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

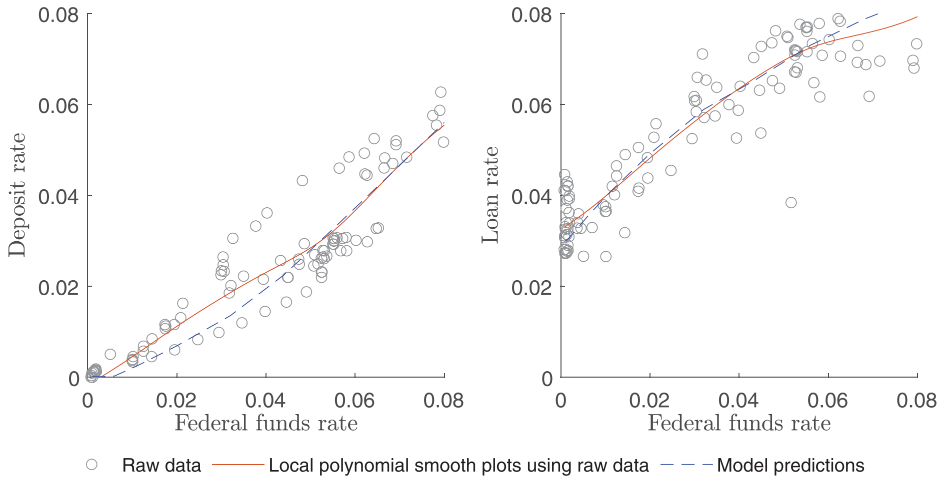

Figure 3. Model-predicted versus actual rates. This figure illustrates the relation between the federal funds rate and banks' deposit and loan rates. The circles represent a scatter of the raw data from 1994 to 2017, aggregated at the quarterly frequency. The dashed lines represent local polynomial smoothed plots based on the raw data. The solid lines represent the relations predicted using the model. (Color figure can be viewed at wileyonlinelibrary.com)

in our moment-matching exercise, as we target only the average levels of these rates. Figure 3 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 III and Table IA.II in Section VI of the Internet Appendix, 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 the loan markets.

Second, we compare our model predictions with the reduced-form evidence in Drechsler, Savov, and Schnabl (2017) on heterogeneity in monetary transmission. In particular, we consider their finding that the sensitivity of deposit spreads to the federal funds rate rises with the local market Herfindahl-Hirschman Index (HHI) by 3 to 4 bps when moving from a local market at the 25$^{\text{th}}$ percentile of the HHI distribution to one at the 75$^{\text{th}}$ (Drechsler, Savov, and Schnabl, 2017, figure IV). Using the summary statistics in Drechsler, Savov, and Schnabl (2017) and assuming that HHIs are normally distributed, we estimate this interquartile range to be (0.15, 0.34). We find quantitatively similar results in our model when we allow the HHI in our model to vary over this range, with Panel A of Figure 4 showing a fall in the deposit rate sensitivity of 3 bps, which implies that the deposit spread sensitivity rises by 3 bps.

Third, Drechsler, Savov, and Schnabl (2017) find that an increase in the deposit market HHI over its interquartile range accompanies a rise in the sensitivity of deposits to the federal funds rate by 66 bps (Table III). Panel

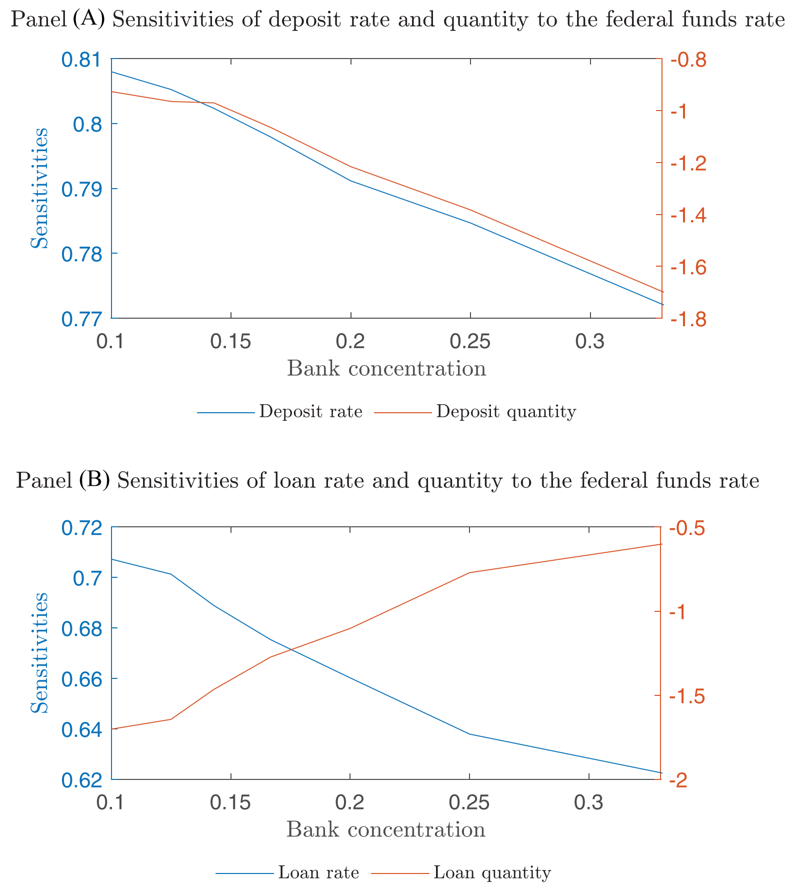

Figure 4. Bank concentration and monetary transmission. This figure plots model-generated sensitivities of deposit and loan rates and quantities to the federal funds rate. Panel A plots the deposit rate and deposit quantity sensitivities, each as a function of market concentration, which we measure as the market Herfindahl index (HHI). Panel B plots the loan rate and loan quantity sensitivities, each as a function of the HHI. We calculate rate sensitivities by regressing changes in the deposit or loan rate on changes in the federal funds rate. We calculate quantity sensitivities by regressing changes in the log quantity on changes in the federal funds rate. (Color figure can be viewed at wileyonlinelibrary.com)

A of Figure 4 shows that in our model, an equivalent increase in the deposit market HHI accompanies a rise the sensitivity of deposits to the federal funds rate by 76 bps, which is close to the estimate in Drechsler, Savov, and Schnabl (2017). Panel A also shows that an increase the HHI over its interquartile range raises the sensitivity of deposits to the federal funds rate by 70 bps.

Fourth, our model also generates the pattern in Drechsler, Savov, and Schnabl (2017, figure IV) whereby even in regions with an extremely low HHI, the sensitivity of the deposit spread to the federal funds rate is still significantly different from zero. Drechsler, Savov, and Schnabl (2017) argue that factors beyond market concentration, such as depositor sophistication, also play a role in determining banks' market power. We find a similar pattern in our estimated model—even when the HHI is 0.1, a 100 bp increase in the federal funds rate leads to only a 80 bp increase in the deposit rate, as the deposit spread increases by 20 bps. Our model generates this result through a relatively large estimate of the yield dispersion, σ_{α}^{d}. High dispersion implies that even in less concentrated markets, a significant fraction of depositors are rate-insensitive. They tolerate higher spreads charged by banks, thus giving banks market power.

Fifth, we compare the results in Scharfstein and Sunderam (2016) with predictions from our model. This comparison cannot be quantitative because Scharfstein and Sunderam (2016) examine the mortgage market, while we examine the total loan market. Nonetheless, as shown in Panel B of Figure 4, we find that the sensitivity of the loan spread to the federal funds rate falls in absolute value as the HHI rises, which implies that the transmission of the federal funds rate to lending rates becomes lower when the market becomes more concentrated. Additionally, Panel B shows that an increase in the loan market HII over its interquartile range decreases the sensitivity of loans to the federal funds rate by 82 bps. In comparison, Scharfstein and Sunderam (2016, figure 3) show that the transmission of changes in yields on mortgage-backed securities, which they use as a proxy for banks' cost of funds, is lower in counties with more concentrated loan markets. Panel B also shows that the sensitivity of the quantity of loans to the federal funds rate decreases in absolute value with the model HHI. Scharfstein and Sunderam (2016) show that the sensitivity of mortgage refinancing to the yield on mortgage-backed securities decreases with market concentration.

Sixth, we check the external validity of the estimation results in Table IV by ascertaining whether our model can match the impulse responses of several variables not used for estimation. To calculate the VARs in our data, we use the high-frequency monetary shocks (Gertler and Karadi (2015)) as external instruments. Details are in Section V of the Internet Appendix. This high-frequency approach is not possible for our actual estimation because our model is not sufficiently rich to produce these shocks. However, using it in our actual data as an external validity check is informative because the federal funds rate in our model is assumed to be exogenous, so it is natural to compare model impulse responses to well-identified data estimates.

Table V presents the simple VAR coefficients generated by our model, along with the well-identified VAR coefficients estimated from our data. Panels A and B of Table V present the impulses of prices and quantities, respectively, over a two-year horizon. We find that we can closely match the sensitivities between the federal funds rate and both deposit and loan spreads. We also find that in the data, net noninterest expense is insensitive to the federal funds rate, which is also consistent with our model because we model this expense as a constant, so its sensitivity to federal funds rate shocks is zero by definition. Finally, our model generates impulse responses that match their empirical counterparts for other important bank balance sheet variables such as deposits, loans, borrowing, and securities.

Table V

Sensitivities of Additional Moments to Monetary Shocks

This table reports the sensitivities of bank deposit spreads, loan spreads, noninterest income, assets, loans, securities, deposits, and borrowing to the federal funds rate. In the second column, we report the sensitivities estimated in the data. The data are from the Call Reports and include all U.S. commercial banks from 1994 to 2008 at a quarterly frequency. We estimate impulse responses to a 1% change in the federal funds rate using the local projections method of Jordà (2005) at an eight-quarter horizon. We instrument the federal funds rate with the unexpected monetary policy shocks constructed from changes in the federal funds futures price on FOMC announcement dates, as in Bernanke and Kuttner (2005). We follow Gertler and Karadi (2015) and English, Van den Heuvel, and Zakrajšek (2018) to sum up the surprises within a quarter to form the instrument. In the third column, we report the sensitivities predicted by the model, which correspond to changes in the variables of interest over a two-year horizon scaled by their respective steady-state levels following a 1% change in the federal funds rate.

Panel A: Sensitivity of Prices		
	Empirical	Model
Deposit spread	0.386	0.310
Loan spread	−0.024	−0.011
Net noninterest expense	−0.001	0.000
Panel B: Sensitivity of Quantities		
	Empirical	Model
Deposit	−2.491	−2.482
Loan	−1.949	−1.641
Borrowing	−3.195	−5.335
Securities	−1.232	−3.052
Assets	−2.587	−2.773

Finally, we examine how the market value of bank equity reacts to an unexpected federal funds rate shock. In the data, a 1 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 1.44% drop in equity value. This result is important because models without market power can overpredict this response.

This small equity market response occurs in our model because of deposit market power, which makes deposit rates insensitive to the policy rate and thereby implies that deposits effectively have long duration (Drechsler, Savov, and Schnabl (2021)). To emphasize this point, we examine the equity response in our model when we force deposit rates to move one-for-one with the federal funds rate, finding a sharp 11% drop.

IV. Counterfactuals

A. 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

Table VI

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 sensitivity captures changes in loans over a two-year horizon scaled by the steady-state loan-level following a 1% change in the federal funds rate. The third column presents the percent change in the sensitivity relative to the baseline case in row (1). All model solutions are under the same set of parameters reported in Table III.

		Sensitivity of Loans to FFR	Change Relative to Baseline (%)
(1)	All frictions are present	−1.641	/
(2)	– Reserve regulation	−1.499	8.65%
(3)	– Capital regulation	−1.248	23.91%
(4)	– Deposit market power	−1.132	31.00%
(5)	– Loan market power	−1.951	−18.91%

rate. Table VI depicts, for different versions of our model, the percentage change in aggregate bank lending in response to a 1 percentage point change in the federal funds rate over a two-year horizon. We show the results of an experiment in which we start with our model exactly as specified in Section II, with all frictions as estimated in Table III. We then eliminate from the model 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 (1) corresponds to the baseline model and shows a response of 1.64%, which is insignificantly different from the response of 1.59% that we find in the data. Row (2) 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 8.6%. This modest magnitude reflects the small amount of noninterest-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 argues that "by paying banks not to lend, the central bank diminished its ability to control interest rates.[5]" Also, a June 22, 2019 article in the *American Banker* states that "It was thanks to interest on excess reserves that the Fed ended up stimulating so little in the economy, despite its efforts

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

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

to ease so much.[6]" 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 (2020), who shows that the reserve requirement is not a quantitatively important feature that distinguishes commercial banks from shadow banks.

Row (3) in Table VI 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 23.9% ($1 - 1.248/1.641$). This result connects two long-standing sets of reduced-form evidence on the bank capital channel. The first set shows 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 (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 bank capital is endogenous, 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 the quantitative magnitude of this long-established channel.

Row (4) of Table VI 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.[7] Note that we do not change the parameters governing investors' preferences (the yield sensitivities and dispersion) in this counterfactual analysis. 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 31% ($1 - 1.132/1.641$) smaller than the 1.641% 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 and other interest-bearing assets. Banks react by charging higher deposit spreads, so households substitute into other interest-bearing assets and the equilibrium quantity of deposits falls. This fall in deposits impacts banks' lending decisions because they need to use

[6] Michel, Norbert, and George Selgin, 2019, Fed must stop rewarding banks for not lending, *AmericanBanker*.

[7] Alternatively, we can impose a zero markup by letting depositor rate sensitivities approach infinity. However, in this case, depositor substitution patterns between different investment vehicles change, and we want to hold these patterns constant in our counterfactuals.

expensive nonreservable borrowing to finance their loans when the amount of loans exceeds deposits. Thus, bank market power, combined with the nonreservable 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 VI 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 that 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 -18.9% $(1 - 1.951/1.641)$. 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.

In Table IA.VIII in Section X of the Internet Appendix, we present the results of the converse experiment, in which we start with a version of the model without regulatory frictions or market power and then add these frictions back in one at a time. We find that the relative importance of the different channels remains intact, but the quantitative effects of both capital regulation and market power are weaker when we start from the frictionless case, suggesting significant complementarity between the two mechanisms.

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

B. 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 0.9%, a rate increase actually expands lending. We call this region a reversal-rate environment.

To understand the mechanism behind the reversal rate, in Panel B, we plot the amount of desired bank lending in a world with no capital requirements, and in Panel C, we plot the level of bank capital. 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%

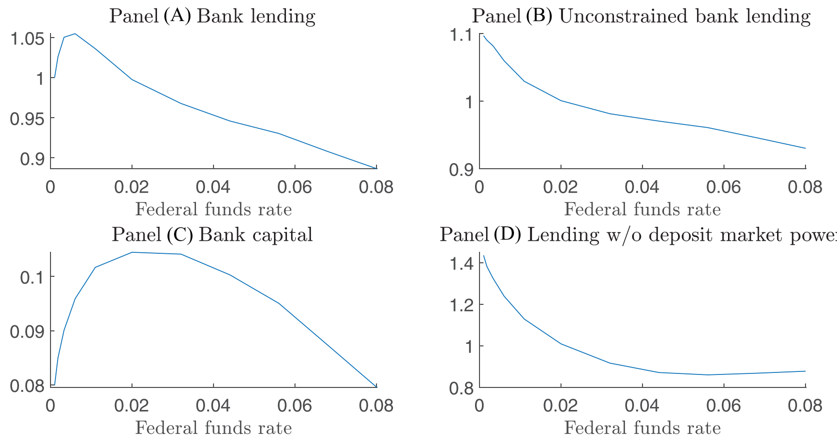

Figure 5. Bank capital, bank lending, and the federal funds rate. This figure illustrates how bank capital and optimal lending vary with the federal funds rate. In all panels, the federal funds rate is on the x-axis. Bank characteristics, scaled by the level of steady-state bank lending, is on the y-axis. (Color figure can be viewed at wileyonlinelibrary.com)

turning point at which the relation between bank equity and the federal funds rate flips sign.

Two properties of the relations shown in Panels B and C underlie the reversal rate. The intuition behind Panel B is straightforward, as high funding costs deter firms from borrowing in equilibrium. The intuition behind Panel C is more nuanced and depends on the relative profitability of lending and deposit taking. First, as the federal funds rate rises, depositors find holding cash to be increasingly 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 banks use to bolster their equity capital base. Banks accumulate equity capital instead of paying out their profits to shareholders as a precaution against 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-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

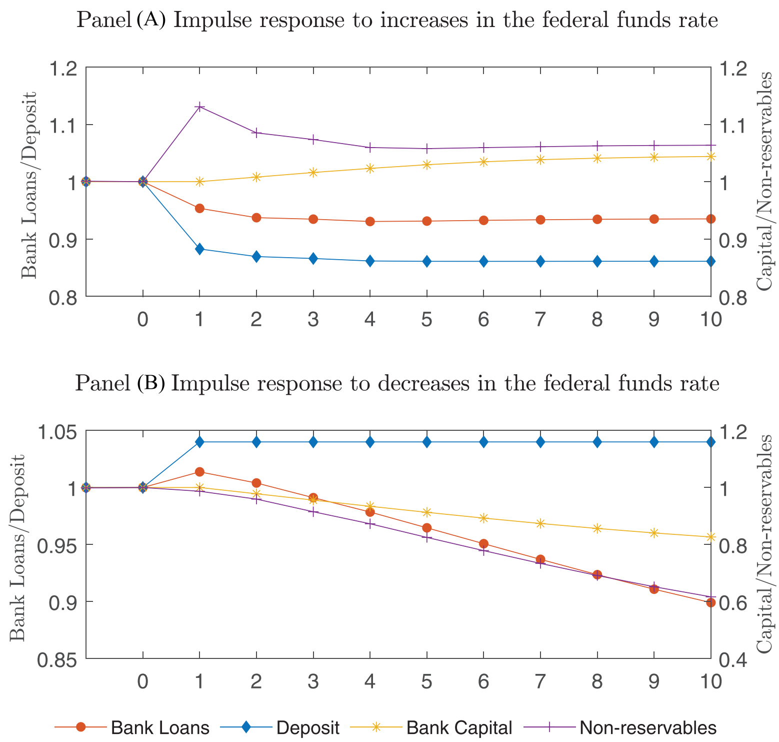

Figure 6. Impulse response to federal funds rate shocks. This figure illustrates banks' impulse responses to federal funds rate shocks. The economy starts at Year 0 when it is in the old steady state with the federal funds rate equal to 0.9%. In Year 1, the federal funds rate either increases to 2% or decreases to 0.1%, and it stays at that level afterward 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%. (Color figure can be viewed at wileyonlinelibrary.com)

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 bolster the intuition that the reversal rate stems from the interaction of deposit market power and capital regulation, in Panel D of Figure 5, we plot the relation between bank lending and the federal funds rate under the assumption that banks face perfectly elastic demand for deposits when deposits are priced at the federal funds rate minus the banks' marginal cost to service depositors. We find no reversal effect in this setting.

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 0 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 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 face 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 nonreservable 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 duration 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. Nonreservable 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, in turn, 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 when the federal funds rate is extremely low. Because total lending decreases, banks need less external financing and thus less nonreservable borrowing.

Note that in Figure 6, lending falls when the federal funds rate changes in either direction. Although 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 nonreservable 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.

This discussion highlights two differences between the reversal rate in our model and the reversal rate in Brunnermeier and Koby (2016). First, they assume an exogenous relation between the deposit rate and the federal funds rate, $r_t^d = \eta_1 + \eta_2 \exp(\eta_3 f_t)$, which generates a sticky deposit rate, which, in turn, drives the reversal mechanism. We provide a microfoundation for this assumption by modeling the interaction between imperfect competition in the deposit market and regulatory frictions. Second, Brunnermeier and Koby (2016) calibrate their model, while we estimate ours, thus providing an empirical estimate of the turning point.

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, cumulative bank lending had increased only about 25% from its low in August 2009. In contrast, from trough to peak, in all recessions since 1974, bank lending grew by 60% to 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.

C. External Validation

We check external validation of the reversal-rate prediction in two ways. First, we estimate a reduced-form regression to 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)).[8] 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 to 2001) and the financial crisis (2007 to 2009) because, in these crisis times, information other than conventional monetary policy news could also be released in FOMC meetings.[9]

Because Panel C of Figure 5 shows that the relation between bank capital and the federal funds rate changes sign around 2%, we split our data sample using a 2% cutoff for the federal funds rate. In Table VII, 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 a positive significant bank equity return, consistent

[8] As shown in Section VIII of the Internet Appendix, our results are robust to using use one-year Treasury yields.

[9] For instance, on January 28, 2009, the FOMC expressed 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 the Fed's willingness to support the banking system, leading to 12% one-day banking stock returns. Thus, the Fed affected stock prices through the Fed put channel (Cieslak and Vissing-Jorgensen (2021)), which is outside our model.

Table VII

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 as 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. Low is a dummy variable that equals 1 when the starting level of the federal funds rate (FFR) is below 2%. The control variables include market returns and term spreads. The sample includes all publicly traded U.S. banks from 1994 to 2017. The sample for columns (1) and (4) comprises observations in which the starting level of the federal funds rate (FFR) is above 2%. The sample for columns (2) and (5) comprises observations in which the starting level of the federal funds rate is below 2%. The sample for columns (3) and (6) comprises all observations. We exclude observations during the collapse of the dot-com bubble (2000 to 2001) and the subprime financial crisis (2007 to 2009). Standard errors are clustered by time.

	High (1)	Low (2)	All (3)	High (4)	Low (5)	All (6)
Policy shock	-1.292^{**}	2.202^{**}	-1.292^{**}	−0.639	−1.393	−0.639
	[0.615]	[0.879]	[0.612]	[0.653]	[0.852]	[0.649]
Low*Policy shock			3.494^{***}			−0.754
			[1.069]			[1.069]
HHI*Policy shock				−0.134	0.562^{***}	−0.134
				[0.145]	[0.153]	[0.144]
Low*HHI*Policy shock						0.696^{***}
						[0.210]
Control	Yes	Yes	Yes	Yes	Yes	Yes
Observations	27,257	33,805	61,062	27,257	33,805	61,062
Adj. R^2	0.015	0.123	0.074	0.016	0.125	0.075

with market expectations 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. To examine the statistical significance of the difference between columns (1) and (2), in column (3), we report results from a regression in which we pool all sample observations and include a term for the interaction between the monetary policy shock and a dummy for a low federal funds rate. We find a highly significant positive coefficient on this interaction term, which is consistent with a strong reversal effect. In summary, we find that monetary policy has a nonmonotonic effect on bank capital. When the federal funds rate is high, the relation between the short-term rates and bank capital is negative, but when this rate is low, the relation is positive.

In columns (4) and (5) of Table VII, we report the results of interacting changes in the federal funds rate with the HHI of the local deposit market (county) in which a bank operates. 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-rate environment, banks with greater deposit market power experience higher positive returns. Moreover,

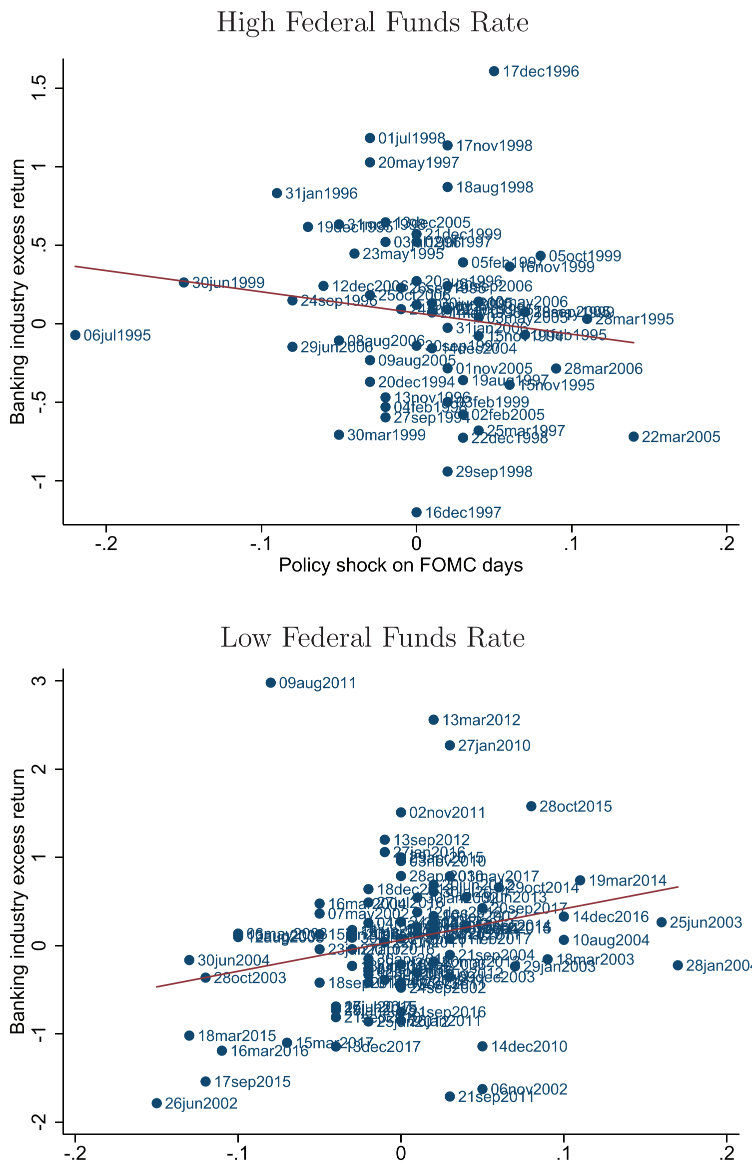

Figure 7. Monetary policy shocks and bank equity returns. This figure provides scatter plots of bank industry excess returns against monetary policy shocks on FOMC days from 1994 to 2017. The excess returns are defined as the difference between bank industry index returns and market returns. Monetary policy shocks are measured as one-day changes in two-year Treasury yields on FOMC days. The sample for the upper panel comprises observations in which the starting level of the federal funds rate is above 2%. The sample for the lower panel comprises 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 to 2001) and the financial crisis (2007 to 2009). Bank industry stock returns are from Kenneth French's website, and the two-year Treasury yield is from the FRED database. (Color figure can be viewed at wileyonlinelibrary.com)

this evidence that the reversal effect is closely linked to banks' deposit market power is reinforced by the finding in column (6) that the triple interaction between the low federal funds rate dummy, the HHI, and the policy shock is significantly positive.

Next, we explore whether the positive relation between monetary policy shocks and bank stock returns during low-rate environments is driven by the central bank's economic outlook (Nakamura and Steinsson (2018)). In Section VIII of the Internet Appendix, we present 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 (2020), who disentangle the potential information shocks from monetary policy shocks. Our results are robust to this alternative measure.

For our second external validation exercise, we conduct a difference-in-differences analysis that compares banks with different deposit market power when the policy rate enters the reversal region. Following Heider, Saidi, and Schepens (2019), we estimate

$$y_{i,t+4} = \beta \mathrm{HHI}_i \times \mathrm{Low}_t + \gamma' x_{i,t} + \eta_i + \eta_t + \epsilon_{i,t}, \tag{39}$$

where $y_{i,t+4}$ is log bank lending or bank capital in quarter $t+4$, HHI_i is the HHI for the local deposit market in which a bank operates, and Low_t is a dummy variable that equals 1 when the federal funds rate falls below 2%. The control variable vector, $x_{i,t}$, includes deposit and loan rates, as well as log bank assets, lending, deposits, and equity. We denote bank and time fixed effects as η_i and η_t. Because the policy rate is endogenous to economic conditions, we adopt the parallel trends assumption in Heider, Saidi, and Schepens (2019) that both high-HHI and low-HHI banks face the same deterioration in economic conditions during recessions. Finally, the federal funds rate falls into the reversal region twice in our sample period, early in the 2000 recession and during the 2008 financial crisis. We restrict the sample to 2000Q1 to 2004Q1 because this recession did not originate from the banking system.

The results are in Table VIII. Columns (1) and (2) show the effects of low interest rates on bank equity, and columns (3) and (4) show the effects of low interest rates on bank lending. Consistent with our model predictions, book equity and lending decrease more for high-HHI banks than for low-HHI banks when the federal funds rate falls into the reversal region.

V. Extensions and Robustness

A. Heterogeneous Transmission through Large and Small Banks

We extend our analysis by examining whether monetary transmission depends on bank size, motivated by the finding in Kashyap and Stein (1995) that monetary policy particularly affects lending by small banks, as they cannot replace deposits with frictionless access to nonreservable funding. However, one

Table VIII

Effects of Low Rates on Banks with Different Deposit Market Power

We report the estimates of the effect of low interest rates on banks with different deposit market power. HHI is the average Herfindahl-Hirschman Index of the local deposit markets in which a bank operates. Low is a dummy variable that equals 1 when the federal funds rate is below 2%. Each column is identified by the dependent variable. The control variables include lagged assets, lending, deposits, equity, deposit rate, loan rate, bank fixed effects, and time fixed effects. The sample includes all U.S. banks from 2000Q1 to 2004Q1. Standard errors are clustered by time.

	Equity (1)	Equity (2)	Loan (3)	Loan (4)
HHI*Low	−0.232***	−0.083***	−0.252***	−0.055***
	(0.007)	(0.006)	(0.008)	(0.006)
Control	No	Yes	No	Yes
Bank F.E.	Yes	Yes	Yes	Yes
Time F.E.	Yes	Yes	Yes	Yes
Observations	129,950	127,912	129,339	127,885
Adj. R^2	0.984	0.990	0.981	0.990

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 10$^{\text{th}}$ percentile of the bank size distribution and estimate 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 reestimate all of the remaining parameters.

The results are in Table IX. 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 nonreservables, and ψ is identified from the banks' net noninterest expense and leverage ratios. These parameter differences across samples imply that small-bank loans are 40% more sensitive to the federal funds rate, with a −1.742 sensitivity for small banks and a −1.235 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 increase the big banks' financing cost ϕ^N to the level estimated for the small banks (0.015), the big banks' sensitivity of loans to the federal funds rate rises to −1.471, an increase

Table IX

Large Banks versus Small Banks

In this table, we report SMD estimation results for subsamples of large and small banks. Panel A contains the simulated versus actual moments, along with 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 nonreservables. 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-statistic	Actual	Simulated	t-statistic
Dividends	3.60%	2.89%	−3.568	/	/	/
Nonreservable borrowing share	35.50%	36.66%	0.612	15.30%	16.09%	1.574
Std of nonreservable borrowing	13.30%	19.86%	2.626	8.70%	10.83%	2.363
Deposit spread	1.32%	1.37%	0.520	1.25%	1.26%	0.130
Loan spread	1.77%	1.98%	2.068	2.71%	2.42%	−2.979
Deposit-to-asset ratio	0.666	0.688	0.692	0.784	0.779	−0.382
Net noninterest expense	0.96%	0.85%	−1.117	1.90%	1.76%	−1.436
Leverage	11.36	12.54	2.026	10.78	10.03	−3.796
Market-to-book ratio	2.06	1.97	−1.217	/	/	/
Total credit-FFR sensitivity	−0.995	−0.998	−0.010	−0.995	−1.027	−0.108

Panel B: Parameter Estimates

	ϕ^N	ϕ^d	ϕ^l	ψ	W/K	Sensitivity of Loans to FFR
Large banks	0.006	0.010	0.006	0.013	0.228	−1.235
	[0.001]	[0.001]	[0.001]	[0.003]	[0.007]	
Small banks	0.015	0.008	0.009	0.075	0.140	−1.742
	[0.001]	[0.001]	[0.001]	[0.002]	[0.001]	

representing 47% of the difference in loan sensitivity between large and small banks. Because we allow only five 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 also contributes to the stronger monetary transmission among smaller banks. Small banks have relatively few sources of noninterest 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.

B. 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. To this end, we split our sample into two subperiods—early (1994 to 2005) and late (2006 to 2017). We then reestimate all model parameters for the two subperiods.[10] The parameter estimates in Table X imply that the sensitivity of lending to the federal funds rate has declined from 1.53% to 1.22% 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 characterizes macroeconomic conditions—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—discount rates, operating costs, and external financing costs. The last group consists of parameters that govern banks' market power—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 X, banks' market concentration has risen over time, with the number of competing banks in the local market falling from seven to five. 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. 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 34%.

Table X 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 nonreservable 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 22%.

[10] We set the value of firms' outside option, q_n^l, to the baseline estimates reported in Table III. This parameter is identified by the sensitivity of total borrowing to the federal funds rate, which we estimate using a VAR. When we split the sample by time, the subsamples are too short to generate reliable VAR estimates.

Table X

Subsample Estimates: Early versus Late

In this table, we report the model parameter estimates for the early (1994 to 2005) and late (2006 to 2017) subsamples. Panel A presents calibrated parameters. Panel B presents values for parameters that can be calculated as simple averages or with 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	3.170	[1.402]	3.590	[1.448]
$\bar{f}$	Log Federal funds rate mean	−3.305	[0.170]	−5.605	[0.416]
σ_f	Std of log federal funds rate innovation	0.553	[0.148]	0.516	[0.885]
ρ_f	Log Federal funds rate persistence	0.700	[0.141]	0.900	[0.111]
$\bar{\delta}$	Log Loan chargeoffs mean	−6.005	[0.163]	−5.806	[0.258]
σ_δ	Std of log loan chargeoffs innovation	0.800	[0.255]	1.040	[0.383]
ρ_δ	Log Loan chargeoffs persistence	0.600	[0.060]	0.600	[0.068]
$\rho_{\delta f}$	Corr of federal funds rate innovation and chargeoffs	0.040	[0.435]	−0.160	[0.159]
	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	0.424	[0.144]	0.528	[0.279]
α^l	Borrowers' sensitivity to loan rates	−1.017	[0.054]	−1.454	[0.082]
q_d^d	Convenience of holding deposits	3.465	[0.358]	2.340	[0.470]
q_c^d	Convenience of holding cash	2.763	[0.387]	−0.444	[0.430]
q_l^l	Convenience of borrowing through loans	−0.016	[0.088]	1.804	[0.212]
	Panel D: Parameters Estimated via SMM				
γ	Banks' discount rate	0.047	[0.002]	0.044	[0.005]
W/K	Relative size of the deposit market	0.184	[0.005]	0.254	[0.007]
ϕ^N	Quadratic cost of nonreservable borrowing	0.010	[0.001]	0.010	[0.004]
ϕ^d	Bank's cost of taking deposits	0.009	[0.001]	0.009	[0.004]
ϕ^l	Bank's cost of servicing loans	0.005	[0.001]	0.008	[0.002]
ψ	Net fixed operating cost	0.048	[0.002]	0.010	[0.005]

The remaining 44% of the gap is attributable to changes in macroeconomic 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.

C. 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 reestimate our model, with the parameter ϕ^e being identified by an additional moment, namely, the ratio of bank equity issuance to total assets. In the data, this moment is 2%. As shown in Table IA.VII in Section IX of the Internet Appendix, we find that matching this moment yields an equity issuance cost of 11%, which is comparable to 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, which is the same moment that we use to identify the constant discount factor in our baseline model. As shown in Table IA.VII in Section IX of the Internet Appendix, we find that $\omega = 1.5\%$. 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.

Third, we address the concern 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 et al. (2011), who show that the credit risk premium in the bond market roughly equals the expected default loss. After adjusting the data moment, we reestimate the model. As shown in Table IA.VII in Section IX of the Internet Appendix, we find that the only notable difference in the new results lies in banks' estimated cost of servicing loans, which becomes insignificantly different from zero. This result suggests that omitting loan default risk in our original model causes this component to load onto loan servicing costs.

Fourth, we address the functional form of the nonreservable financing cost in equation (13), which is quadratic. We reestimate the model under the assumption that the nonreservable cost takes a more general power-function form: $\Phi^N(N_t) = \phi^c (N_t/D_t)^{\phi^p} D_t$. We identify the multiplicative coefficient and the exponent by matching the mean and standard deviation of banks' nonreservable borrowing. We find that the power term, ϕ^p, equals 2.177 and is insignificantly different from two, suggesting that our original functional form is a reasonable approximation. All other parameter estimates under this power financing cost specification also remain near the baseline parameter estimates.

Fifth, one simplifying assumption of the model is that the present value of all future interest payments is paid to the bank in the first period in which a loan is issued. Moreover, we implicitly assume that this payment is default free. To address these issues, we extend the model by introducing an additional state variable that represents the average contractual interest rate on outstanding loans. Details are in Section IX of the Internet Appendix. We solve the model using the parameters reported in Table III. In Table IA.VI, we show that the moment conditions are largely unchanged relative to the values in Table IV. In Figure IA.3, we show that the reversal rate result is nearly unchanged.

Finally, in Section XI of the Internet Appendix, we consider an extension of the model in which we endogenize the federal funds rate. Our extended model contains a standard New Keynesian block and a banking block, with the New Keynesian block determining the effects of productivity and monetary policy shocks on the nominal short-term rate and inflation. The banking block determines the transmission of the nominal short-term rate to the lending rate. In this setting, the federal funds rate process is pinned down by a Taylor rule, which depends on the monetary authority's policy as well as aggregate output and inflation.

We assume a continuum of ex ante homogeneous households with separable preferences over real consumption and real money balances. These households face a two-stage decision-making process. First, they choose the quantity of consumption and money holdings given aggregate prices. Second, they allocate their money demand across different options—cash or deposits in any bank. Similarly, firms decide their optimal demand for capital and whether to finance this demand via corporate bonds or bank loans. We model the banking sector as in our baseline partial equilibrium setting discussed in Section II. The detailed model setup and parameter calibrations are discussed in Section XI of the Internet Appendix.

With the calibrated general equilibrium model, we first confirm that the relation between the federal funds rate and aggregate bank lending is nonmonotonic and hump-shaped. We also repeat our decomposition of monetary policy transmission mechanisms using the general equilibrium framework. We find that the sensitivity of loans to the federal funds rate is lower than in our baseline model in Section IV.A because our New Keynesian model implies a much weaker relationship between the federal funds rate and long-term real rate than what we observe in the data. Nevertheless, our main conclusions remain valid. We find that the qualitative effect of the reserve regulation remains limited, while both the deposit and loan market power channels are quantitatively important in explaining monetary transmission.

VI. 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, Savov, and Schnabl (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 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.

Initial submission: May 19, 2020; Accepted: June 28, 2021
Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

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* 136, 355–378.

Begenau, Juliane, and Tim Landvoigt, 2021, Financial regulation in a quantitative model of the modern banking system, *Review of Economic Studies*, https://doi.org/10.1093/restud/rdab088

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

Bernanke, Ben S., and Kenneth N. Kuttner, 2005, What explains the stock market's reaction to Federal Reserve policy? *Journal of Finance* 60, 1221–1257.

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 Benjamin M. Friedman, and Michael Woodford, eds.: Handbook of Monetary Economics (Elsevier, Amsterdam).

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, Working Paper, Princeton University.

Brunnermeier, Markus K., and Yuliy Sannikov, 2016, The I theory of money, Working Paper, 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.

Campos, Mauricio, 2019, Going negative at the zero lower bound: The effects of negative nominal interest rates, Working Paper, Federal Reserve Bank of San Francisco.

Cieslak, Anna, and Annette Vissing-Jorgensen, 2021, The economics of the Fed put, *Review of Financial Studies* 34, 4045–4089.

Corbae, Dean, and Pablo D'Erasmo, 2021, Capital buffers in a quantitative model of banking industry dynamics, *Econometrica* 89, 2975–3023.

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

DeFusco, Anthony A., and Andrew Paciorek, 2017, The interest rate elasticity of mortgage demand: Evidence from bunching at the conforming loan limit, *American Economic Journal: Economic Policy* 9, 210–240.

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, Working Paper, ECB.

Dick, Astrid A., 2008, Demand estimation and consumer welfare in the banking industry, *Journal of Banking and Finance* 32, 1661–1676.

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

Drechsler, Itamar, Alexi Savov, and Philipp Schnabl, 2021, Banking on deposits: Maturity transformation without interest rate risk, *Journal of Finance* 76, 1091–1143.

Egan, Mark, Ali Hortaçsu, and Gregor Matvos, 2017, 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, 2022, The cross section of bank value, *Review of Financial Studies* 35, 2101–2143.

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

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.

Gertler, Mark, and Peter Karadi, 2015, Monetary policy surprises, credit costs, and economic activity, *American Economic Journal: Macroeconomics* 7, 44–76.

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.

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, 2019, Life below zero: Bank lending under negative policy rates, *Review of Financial Studies* 32, 3728–3761.

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, 2020, Deconstructing monetary policy surprises: The role of information shocks, *American Economic Journal: Macroeconomics* 12, 1–43.

Òscar, Jordà, 2005, Estimation and inference of impulse responses by local projections, *American Economic Review* 95, 161–182.

Kashyap, Anil K., and Jeremy C. Stein, 1995, The impact of monetary policy on bank balance sheets, *Journal of Monetary Economics* 42, 151–195.

Landvoigt, Tim, Stijn Van Nieuwerburgh, and Vadim Elenev, 2021, A macroeconomic model with financially constrained producers and intermediaries, *Econometrica* 89, 1361–1418.

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, Working Paper, Harvard Business School.

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

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

Xiao, Kairong, 2020, Monetary transmission through shadow banks, *Review of Financial Studies* 33, 2379–2420.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1: Internet Appendix.
Replication Code.