The Reserve Supply Channel of Unconventional Monetary Policy*

William Diamond[†] Zhengyang Jiang[‡] Yiming Ma[§]

August 6, 2022

Abstract

We find that central bank reserves injected by QE crowd out bank lending. We estimate a structural model with cross-sectional instrumental variables for deposit and loan demand. Our results are determined by the elasticity of loan demand and the impact of reserve holdings on the cost of supplying loans. We find that the reserves injected by QE raise loan rates by 15.6 bps, and each dollar of reserves reduces bank lending by 19 cents. Our results imply that a large injection of central bank reserves has the unintended consequence of crowding out bank loans because of bank balance sheet costs.

Keywords: Financial Intermediation, Quantitative Easing, Deposit Competition, Balance Sheet Cost, Structural Estimation
JEL Codes: G21, G28

*We thank Olivier Darmouni, Mark Egan (discussant), Arvind Krishnamurthy, Giorgia Piacentino, Luke Taylor, Quentin Vandeweyer, Stijn van Nieuwerburgh, Jessica Wachter, Yufeng Wu (discussant), Yao Zeng, and audiences at Columbia Business School, Durham University, Kellogg, Johns Hopkins, London Business School, Michigan Ross, NYU Stern, University of Amsterdam, Virgnia Darden Business School, University of Utah, Wharton, ASSA, Chicago Junior Macro and Finance Conference, Columbia Workshop on New Empirical Methods, EFA, Midwest Finance Assocatiation Conference, NYU Stern New York Fed Intermediation Conference, Yale Junior Macro-Finance Workshop, and WFA for feedback and helpful comments. We thank Naz Koont for excellent research assistance. This paper was previously circulated under the title "Monetary Transmission Through Bank Balance Sheet Synergies."

[†]Finance Department, Wharton School of the University of Pennsylvania. wfdiamond27@gmail.com

[‡]Finance Department, Kellogg School of Management, Northwestern University, and NBER. zhengyang.jiang@kellogg.northwestern.edu

[§]Finance Division, Columbia Business School. 665 W 130th St, New York, NY 10027. ym2701@gsb.columbia.edu

1 Introduction

There has been a massive expansion of central bank reserves issued by the Federal Reserve in the last 15 years. As seen in Figure 1, reserves on bank balance sheets amounted to less than \$50 billion in 2008Q1, but reached \$2.8 trillion in 2015 and exceeded \$3 trillion in 2021. These reserves were created in the aftermath of the 2008–2009 financial crisis and the 2020 Covid-19 pandemic, when the Federal Reserve purchased trillions of dollars of assets in its Quantiative Easing (QE) program to stimulate the economy. In QE, the Federal Reserve buys assets such as Treasuries, which are primarily held outside of the banking sector. In 2008Q1, only 1.1% of Treasuries outstanding were held by US banks. The Federal Reserve then pays with central bank reserves, which are a special interest-bearing asset that can only be held within the banking system. QE therefore results in a net injection of liquid assets to bank balance sheets. While a large literature has studied the effect of QE's asset purchases, less is known about how QE's injection of trillions of dollars of reserves impacted the key functions of the banking system—lending and deposit-taking. We show the unintended consequence that central bank reserves crowds out bank lending to the real economy, i.e., the "reserve supply channel".

In principle, an increase in the supply of central bank reserves could either increase or decrease bank lending. If a mismatch between holding illiquid assets (mortgages and loans) and issuing liquid liabilities (deposits) raises the risk of a bank run, increasing the supply of liquid reserves could increase banks' willingness to lend. Conversely, regulatory constraints can make it costly for banks to expand their asset holdings so that a bank which holds more reserves will want to reduce its holdings of other assets such as loans.

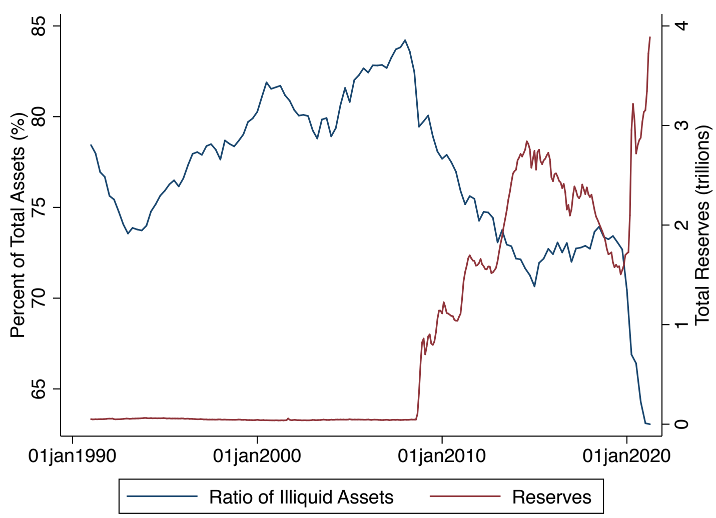

Figure 1: Supply of Central Bank Reserves and Bank Asset Illiquidity

In aggregate time series data, we find suggestive evidence that reserves crowd out bank lending. As reserves increased from \$0.02 trillion in 2006Q1 to \$3.88 trillion in 2021Q1, the proportion of illiquid assets on bank balance sheets declined from 83% to 63% (see Figure 1).[1] However, reserves were injected by QE to simulate the economy during recessions so the observed substitution away from holding illiquid assets may reflect poor economic fundamentals such as low demand for bank loans and may not necessarily be caused by the increase in reserve supply itself.

To quantify the impact of the reserve injection on the banking system without relying on aggregate time series data, we develop a structural model that can be transparently estimated with a series of linear regressions using plausibly exogenous cross-sectional instruments. On the demand side, in each region of the country, banks compete in imperfectly competitive markets to provide deposits, loans, and mortgages. Our demand-side setting is similar to other structural banking papers such as Egan, Hortaçsu, and Matvos (2017) and Wang, Whited, Wu, and Xiao (2020). The supply side of our model is novel in quantifying how a bank's costs depend on its holdings of deposits, mortgages, loans, and reserves. Our specification allows a bank's cost of lending to depend on the quantity of reserves it holds, so our framework is uniquely suited for quantifying the impact of a reserve injection.

There are two key quantities we need to estimate to quantify the reserve supply channel. First, we estimate how the quantity of loans demanded changes when the banking system changes loan interest rates. Second, we estimate how the banking system's overall cost of providing loans changes when it is forced to hold additional reserves. With our estimated model, we show that the reserves injected during QE from 2007 to 2018 crowded out a total of \$611 billion of bank lending, so the reserve supply channel suggests a counterproductive reduction in the supply of bank loans. We note that this crowding out exists in addition to the effects of asset purchases that have been identified in the literature. Hence, the reserve supply channel is important in understanding the true effectiveness of QE. Indeed, Elenev, Landvoigt, Shultz, and Van Nieuwerburgh (2021) find evidence consistent with the quantitative importance of the reserve supply channel in a DSGE model.

To estimate the demand for bank loans, we need to observe how the quantity of loans demanded from a bank varies when it exogenously changes its loan interest rate. We apply an instrument from the reduced-form literature based on banks' reallocation of funds in their internal capital markets after a natural disaster. As Cortés and Strahan (2017) show, loan demand in a region increases after it is hit by a disaster. Banks reallocate funds away from non-disaster regions to provide funds

[1] Illiquid assets include assets except for cash, reserves, Fed funds, repos, Treasury securities and agency securities. Data is for all U.S.-Chartered Depository Institutions from the Flow of Funds.

to the disaster region, and this creates an exogenous shock to the interest rates the bank chooses in non-disaster regions. This reallocation provides precisely the exogenous interest rate shock needed to estimate a bank's loan demand curve under the assumption that natural disasters do not impact the demand for borrowing and lending far away from the regions where they occur. We estimate demand curves for mortgages and deposits in a similar way.

Our demand estimates show that the total demand for bank loans is more interest-rate sensitive than the demand for deposits and mortgages. If all banks in a market raise their corporate loan interest rates by 10 basis points in 2007, the quantity of corporate loans demanded falls by 16.1%. In comparison, a 10 basis point increase in rates would raise deposit demand by 1.3% and would lower mortgage demand by 4.0%. If banks change their deposit, loan, and mortgage interest rates by similar amounts, their loan quantities will respond by a larger amount than their mortgage or deposit quantities. This explains why we find that corporate loan quantities respond most to a larger reserve supply.

Next, we estimate how a bank's cost of providing loans, mortgages, and deposits depends on the composition of its balance sheet, i.e., the quantity of loans, mortgages, deposits, and liquid securities. The interdependence between the components of a bank's balance sheet poses a difficult identification problem because a bank can adjust several components of its balance sheet in response to a demand shock. We solve this problem by first running a series of reduced-form regressions of a bank's marginal costs and balance sheet quantities on two distinct exogenous demand shocks. In addition to the disaster instrument mentioned above, we use a Bartik-style instrument for deposit demand using cross-sectional variation in deposit growth across regions of the country to provide the needed exogenous variation.[2] We then choose our cost function parameters to match these reduced form regressions, providing a relatively transparent approach to estimation.

Our estimates imply that increasing a bank's reserve holdings crowds out mortgage and corporate lending and crowds in deposit issuance. In other words, reserves and bank lending are substitutes rather than complements for banks. One reason could be that bank balance sheet space is costly due to regulation. Acharya and Rajan (2021) argue that reserves may amplify liquidity strains during stress epsiodes, which may also render lending more costly. Quantitatively, a $100 million increase in reserves held by a bank per branch increases its marginal cost of providing mortgages and loans by 31.7 bps and 26.4 bps, respectively. At the same time, the marginal cost of deposits decreases by 21.9 bps.

We use our estimated model to quantify how an increase in reserve supply affects lending

[2] In appendix G, we show that for a firm that sells goods in multiple markets, demand shocks in one market can be used to estimate both the demand curves the firm faces in other markets as well as the firm's marginal cost curve.

and deposit-taking by the banking system. We first shock the reserve supply, allow each bank to trade reserves and adjust its deposit, loan, and mortgage rates, and then determine its quantities of loans, mortgages, and deposits using the demand system.[3] We find that reserve injections affect the interest rates on loans, mortgages, and deposits to a similar extent. However, a larger reserve supply predominantly crowds out bank lending to firms, while the effect on mortgage and deposit quantities is more muted. We estimate that the reserve injections due to QE from 2008 to 2017 crowded out 19 cents of bank lending per dollar of reserves injected. Further, we find that the spread between the interest paid on reserves and risk-free rates available to non-bank investors generated by the model approximately matches the dynamics of a proxy for this spread in the data.

Our findings imply that requiring banks to hold the trillions in reserves created by QE causes a counterproductive reduction in bank lending to firms. We conclude by discussing two potential ways to reduce the crowding out of loans: reducing banks' cost of holding reserves and lowering the proportion of reservce injections trapped in the banking sector.

Our structural model belongs to a growing recent literature on structural estimation in banking. Closely related is Wang, Whited, Wu, and Xiao (2020), who use a structural model of banking to study conventional monetary policy transmission, while our structural model estimates the effect of reserve injections from unconventional monetary policy on the banking system. Several other papers estimate models of imperfect competition in banking similar to ours (Egan, Hortaçsu, and Matvos, 2017; Xiao, 2020; Buchak, Matvos, Piskorski, and Seru, 2018; Albertazzi, Burlon, Jankauskas, and Pavanini, 2022), while others estimate models of networks and matching (Akkus, Cookson, and Hortacsu, 2016; Schwert, 2018; Craig and Ma, 2018). Our application of demand systems in banking complements work that applies demand systems in other financial markets (Koijen and Yogo, 2019, 2020; Koijen, Richmond, and Yogo, 2020; Bretscher, Schmid, Sen, and Sharma, 2020; Jiang, Richmond, and Zhang, 2020). In particular, Koijen, Koulischer, Nguyen, and Yogo (2021) quantify the effect of asset purchases from QE using a demand system, whereas our focus is on the reserves injected by QE.

This paper also contributes to the empirical literature on how quantiative easing impacts the banking system. Existing work in this literature has mostly focused on the effect of asset purchases. For example, Rodnyansky and Darmouni (2017) and Chakraborty, Goldstein, and MacKinlay (2020), focus on the mortgage-backed securities purchased in QE and show that banks with more mortgage-backed securities increase their mortgage lending by more relative to those that hold fewer mortgage-backed securities. Another set of papers study the effect of asset purchases on

[3]The vector of all new rates and portfolio choices in our simulation is over 38,000 dimensions, and the symbolic Jacobian for our model provided in the appendix is crucial to make this numerically tractable.

flattening the long-term yield curve (Gagnon, Raskin, Remache, and Sack, 2010; Krishnamurthy and Vissing-Jorgensen, 2011). The effect of reserve expansion has received little attention. One exception is theoretical work by Acharya and Rajan (2021) who point to another unintended consequence of reserve injection in exacerbating liquidity strains during stress epsiodes. Christensen and Krogstrup (2019) find that long-term government yields are lowered even when short-term assets are purchased. Kandrac and Schlusche (2021) show that loan growth at foreign banks was higher than at domestic banks after a change in regulation caused a redistribution of reserves from domestic to foreign banks. Our paper is the first to quantify the aggregate effect of central bank reserve injection in the US banking system on bank lending and deposit-taking. Our "reserve supply channel" points to an important unintended consequence of central bank reserves in crowding out bank lending from bank balance sheets that complements the transmission channels from asset purchases in the literature. Elenev, Landvoigt, Shultz, and Van Nieuwerburgh (2021) study the effect of QE in a DSGE model and find evidence consistent with the importance of our reserve supply channel.

Our work also relates to a recent literature demonstrating the role of imperfect competition in deposit markets (Drechsler, Savov, and Schnabl, 2017; Li, Ma, and Zhao, 2019) and mortgage markets (Scharfstein and Sunderam, 2016) in the transmission of conventional monetary policy. Our work shows that demand elasticity is an important determinant of the reserve supply channel, since highly price-elastic corporate loan demand is impacted much more by reserve supply than deposit and mortgage demand.

Finally, our estimates of the synergies between the different components of bank balance sheets is novel and sheds lights on how banks' marginal cost of lending is shaped by a number of seminal banking theories, see e.g. Diamond and Rajan (2000); Kashyap, Rajan, and Stein (2002); Hanson, Shleifer, Stein, and Vishny (2015); Diamond (2019).

2 A Model of Bank Balance Sheets

This section introduces the theoretical framework that guides our structural analysis. The goal of the model is to quantify how the banking system responds to policy interventions, such as an increase in reserve supply caused by QE. This response depends on two key model components: the demand that banks face and the balance sheet costs that banks incur in supplying loans, mortgages, and deposits. We first provide a graphical illustration of the demand and supply systems in Subsection 2.1. Then, in Subsection 2.2, we formally set up the model and derive the banking sector's response to an increase in reserve supply.

2.1 A Graphical Illustration

We first present a simplified, visual depiction of our model using a single bank as an illustration. In the model, banks provide loans, mortgages, and deposits in imperfectly competitive markets. Each bank faces a demand curve that determines quantities given the interest rate they choose in each market. In Figure 2, the loan demand for a given bank, i.e., the green line, pins down the loan quantity, Q_L, based on the loan rate it chooses. Like any firm facing a downward-sloping demand curve, banks choose their interest rate so that the marginal cost equals the marginal revenue. In Figure 2, the bank chooses the loan rate R^L at which its marginal cost of supplying loans, i.e., the red line, equals the marginal revenue from loans lending, i.e., the blue line.

Banks' holdings of liquid reserves may impact their marginal cost of lending. For example, having more liquid assets may prevent fire-sales of illiquid assets and help comply with liquidity regulations. However, a larger supply of reserves also uses up balance sheet space and may add to the cost of meeting capital requirements when bank equity is costly. We will formally set up and estimate a cost function, but for now, suppose that the increase in reserve supply shifts the marginal cost of lending as illustrated in Figure 2, then the bank would raise its loan rate to $R^{L'}$, at which the new marginal cost meets the marginal revenue. In the new equilibrium, the quantity of loans supplied by the bank would drop to Q'_L, as implied by the loan demand curve at the new loan rate, $R^{L'}$. Hence, the increase in loan rate and the drop in loan volume as a result of reserve injection would be $R^{L'} - R^L$ and $Q_L - Q'_L$, respectively.

Our empirical approach to quantify the banking system's response to an increase in reserves is similar to the framework in Figure 2. We first estimate the loan demand curve, which determines the marginal revenue. In the full model, banks compete with each other so we extend the loan demand curve for a single bank to a demand system that captures how banks' chosen loan rates affect their own and each other's quantities. Next, we estimate a bank's cost of lending as a function of its balance sheet composition, i.e., its volume of loans, mortgages, securities, and deposits. Then, we can infer how an incrasease in reserve supply shifts the bank's marginal cost curve to trace out the equilibrium change in loan rates and volumes. The same estimation is performed for deposits and mortgages.

2.2 Model

We consider a set of banks indexed by m that operate in a set of markets indexed by n at each time t. Banks invest in loans, L, mortgages, M, and liquid securities, S, backed by deposits, D. Each bank m chooses market-specific rates $R_{D,nmt}$, $R_{M,nmt}$, and $R_{L,nmt}$ for its deposits, mortgages, and loans, respectively. Taking the vector of rates chosen by their competitor banks as given, banks

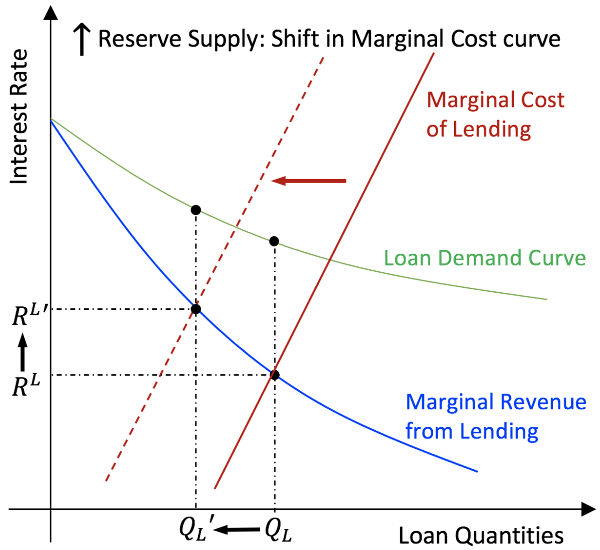

Figure 2: This figure illustrates the effect of an increase in reserves on the loan market. An increase in reserve supply shifts the bank's marginal cost curve for lending. This results in a new intersection with the marginal revenue curve, yielding a new interest rate, $R^{L'}$. The new loan quantity, Q'_L, is then pinned down by the demand curve.

choose their own rates to maximize their profits. In terms of loans, for example, bank m takes the rates of its competitor banks $-m$, $R_{L,n(-m)t}$, as given. The quantity of funds it lends is given by the residual demand curve $Q_{L,nmt}(R_{L,nmt}, R_{L,n(-m)t})$. Similarly, its residual demand curves for mortgages and deposits are $Q_{M,nmt}(R_{M,nmt}, R_{M,n(-m)t})$ and $Q_{D,nmt}(R_{D,nmt}, R_{D,n(-m)t})$. For simplicity, we supress the arguments of the residual demand functions, writing $Q_{L,nmt}$, $Q_{M,nmt}$, and $Q_{D,nmt}$ going forward. Liquid securities, $Q_{S,mt}$, trade in a competitive market at an interest rate $R_{S,t}$. Loans, mortgages, deposits, and securities have cash flows that are discounted at rates $R_t^{L,m}$, $R_t^{M,m}$, $R_t^{D,m}$, and $R_t^{S,m}$ reflecting their respective riskiness.

Banks face a cost $C(\Theta_{mt})$ of providing loans, deposits, and mortgages that depends on all of the items Θ_{mt} on its balance sheet. Θ_{mt} is a vector of bank m's balance sheet components, $Q_{D,nmt}$, $Q_{M,nmt}$, $Q_{L,nmt}$, and $Q_{S,mt}$. In general, this cost function quantities the various ways that a bank's decisions for one part of its balance sheet can impact its costs for another. For example, having more liquid securities on balance sheets may reduce the cost of selling illiquid loans or mortgages in the event of large deposit withdrawals in a bank run (Diamond and Dybvig, 1983). In addition, bank regulations such as the SLR (which constrains a bank's leverage) and the Liquidity Coverage Ratio (which constrains the mismatch between a bank's holding of illiquid assets and issuance of liquid deposits) impose costs that depend on multiple balance sheet components. We show in Section 4 how the bank's overall cost depends on the composition of its balance sheet.

In this setting, bank m chooses its rates $R_{D,nmt}$, $R_{M,nmt}$, and $R_{L,nmt}$ and security quantities $Q_{S,mt}$ at time t to maximize the expected present value of its profits at $t+1$ in all markets n, which are given by[4]

$$\max_{(R_{D,nmt},R_{M,nmt},R_{L,nmt},Q_{S,mt})} \sum_n Q_{L,nmt}(R_{L,nmt} - R_t^{L,m}) + \sum_n Q_{M,nmt}(R_{M,nmt} - R_t^{L,m})$$
$$+Q_{S,mt}(R_{S,t} - R_t^{S,m}) - \sum_n Q_{D,nmt}(R_{D,nmt} - R_t^{D,m}) - C(\Theta_{mt}). \tag{1}$$

In words, bank m's profits are the sum of its revenue from loans, mortgages, and securities, less the nominal cost of deposit funding and the balance sheet costs $C(\Theta_{mt})$. The first order conditions of bank profits with respect to the choice variables, $R_{D,nmt}$, $R_{M,nmt}$, $R_{L,nmt}$, and $Q_{S,mt}$, are

$$\overbrace{\frac{\partial}{\partial R_{D,nmt}}[Q_{D,nmt}(R_t^{D,m} - R_{D,nmt})]}^{\text{Marginal Revenue}} = \overbrace{\frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}}\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}}^{\text{Marginal Cost}}, \tag{2}$$
$$\frac{\partial}{\partial R_{M,nmt}}[Q_{M,nmt}(R_{M,nmt} - R_t^{M,m})] = \frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}, \tag{3}$$
$$\frac{\partial}{\partial R_{L,nmt}}[Q_{L,nmt}(R_{L,nmt} - R_t^{L,m})] = \frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}, \tag{4}$$
$$R_{S,t} - R_t^{S,m} = \frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}. \tag{5}$$

The left hand side of equations (2)) to (4) is the marginal revenue from changing each of the bank's interest rates (the blue curve in Figure 2). This is because the bank's "revenue" from loans, for example, can be seen as the quantity $Q_{L,nmt}$ of loans times its interest rate spread $R_{L,nmt} - R_t^{L,m}$ above the loan discount rate $R_t^{L,m}$. On the right hand side of equations (2) to (5), we have the marginal costs from changing each of the bank's interest rates. Liquid securities are traded in a competitive market, so the first order condition for securities holdings $Q_{S,mt}$ in equations (5) sets price $R_{S,t} - R_t^{S,m}$ equal to marginal cost of holding these securities $\frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}$. Based on equation (5), we refer to $\frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}$ as the "reserve spread"—the difference between risk-free rates $R_{S,t}$ available only to banks and risk-free rates $R_t^{S,m}$ available to non-bank investors as well.

When the supply of liquid securities increases, as in the increase in reserve supply from QE, banks respond by optimally changing their interest rates in all markets as well as their securities holdings. The interest rates they choose still satisfy the first order conditions in equations (2) to (5), which allows us to solve for the equilibiurm quantities of loans, mortgages, and deposits. Specifically, the comparative statics with respect to a change in bank m's liquid security holdings

[4]In Online Appendix F, we show that this static optimization problem is consistent with a model in which the bank maximizes the expected present value of profits over an infinite horizon.

$Q_{S,mt}$ are

$$\frac{\partial\left(R_t^{D,m} - R_{D,nmt} - \frac{Q_{D,nmt}}{\partial Q_{D,nmt}/\partial R_{D,nmt}}\right)}{\partial Q_{D,nmt}} \frac{\partial Q_{D,nmt}}{\partial Q_{S,mt}} = \frac{\partial^2 C(\Theta_{mt})}{\partial Q_{D,nmt}\partial \Theta_{mt}} \cdot \frac{\partial \Theta_{mt}}{\partial Q_{S,mt}}, \tag{6}$$

$$\frac{\partial\left(R_t^{M,m} - R_{M,nmt} - \frac{Q_{M,nmt}}{\partial Q_{M,nmt}/\partial R_{M,nmt}}\right)}{\partial Q_{M,nmt}} \frac{\partial Q_{M,nmt}}{\partial Q_{S,mt}} = -\frac{\partial^2 C(\Theta_{mt})}{\partial Q_{M,nmt}\partial \Theta_{mt}} \cdot \frac{\partial \Theta_{mt}}{\partial Q_{S,mt}}, \tag{7}$$

$$\frac{\partial\left(R_t^{L,m} - R_{L,nmt} - \frac{Q_{L,nmt}}{\partial Q_{L,nmt}/\partial R_{L,nmt}}\right)}{\partial Q_{L,nmt}} \frac{\partial Q_{L,nmt}}{\partial Q_{S,mt}} = -\frac{\partial^2 C(\Theta_{mt})}{\partial Q_{L,nmt}\partial \Theta_{mt}} \cdot \frac{\partial \Theta_{mt}}{\partial Q_{S,mt}}, \tag{8}$$

$$\frac{\partial Q_{S,mt}}{\partial Q_{S,mt}} = 1, \tag{9}$$

where $\frac{\partial Q_{D,nmt}}{\partial Q_{S,mt}}, \frac{\partial Q_{M,nmt}}{\partial Q_{S,mt}}, \frac{\partial Q_{D,nmt}}{\partial Q_{S,mt}}$ are the responses of each individual bank branch quantity to the reserve increase, and Θ_{mt} is the vector of balance sheet quantities $(Q_{D,nmt}, Q_{M,nmt}, Q_{L,nmt}, Q_{S,mt})$. Please see Appendix A.1 for detailed derivations.

To determine the equilibrium response of the banking system to a change in the supply of liquid securities, we need empirical estimates of the components of equations (6) to (8). The left hand side is determined by the bank's loan, mortgages, and deposit demand curves. In Section 3, we estimate this term with an industrial organization style demand system by observing how each bank's quantities respond to shocks to the interest rates they and other banks choose.[5] On the right hand side is an expression reflecting how a bank's marginal cost of borrowing or lending in a market changes with the composition of its entire balance sheet. We therefore need to estimate how a bank's marginal costs of lending and borrowing depend on the different components its balance sheet. In Section 4, we develop and apply a novel econometric approach to estimate this cost function. Taken together, our estimates of the demand for a bank's services and its cost of providing them allow us to infer the aggregate effect of an increased supply of reserves caused by QE—the policy we intend to analyze.

3 Demand Systems

This section estimates the demand systems for deposits, mortgages, and loans. Section 3.1 introduces the logit demand systems and their estimation strategy. Section 3.2 details the data and instruments we use for estimating the demand systems. The estimation results are reported in Section 3.3.

[5]This section considers a single bank in isolation, while our full model allows for competition between banks. Thus, we need to estimate a demand system across all banks rather than just a demand curve faced by an individual bank.

3.1 Estimation Strategy

3.1.1 Demand Curves

Our first step is to estimate the demand curves that individual banks face in deposit, loan, and mortgage markets. For deposits, we let depositors have a total supply of funds $\overline{F}_{D,nt}$ in each market n at time t. They can either invest in deposits at banks m that have branches in the market or an unobserved outside option 0. This outside option reflects the availability of investment options other than deposits that are not in our data such as money market fund shares. An observed quantity $Q_{D,nmt}$ of deposits is invested in bank m's branches in market n in time t. In addition, an unobserved quantity $Q_{D,n0t}$ is invested in the outside option.

Similarly, borrowers of mortgages and loans have total funding needs $\overline{F}_{M,nt}$ and $\overline{F}_{L,nt}$, respectively. They can either borrow from banks or resort to the outside option, which includes borrowing from non-banks or not borrowing altogether. $Q_{M,nmt}$ and $Q_{L,nmt}$ denote the observed quantities of mortgages and loans borrowed from bank m in market n at time t, while $Q_{M,n0t}$ and $Q_{L,n0t}$ denote the unobserved quantities of the respective outside options.

Preferences of borrowers and depositors follow a standard logit demand system (Berry, 1994; McFadden, 1974). For example, depositor j investing in bank m in market n has the following utility:[6]

$$u_{D,jnmt} = \alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt} + \varepsilon_{D,jnmt}. \tag{10}$$

The first term is the the interest rate $R_{D,nmt}$ paid on deposits times the depositor's preference for receiving interest, α_D. We expect the price disutility parameter for deposits, α_D, to be positive because depositors should prefer a higher deposit rate, all else equal. In constrast, we expect the corresponding price disutility parameter for mortgage and loan borrowers to be negative because they prefer a lower funding cost. Depositors' utility is also affected by the desirability of bank m's deposits, which depends on a vector of observed characteristics, $X_{D,nmt}$, preferences for observed charecteristics, β_D, and unobserved characteristics, $\delta_{D,nmt}$. Finally, the error term, $\varepsilon_{D,jnmt}$, is assumed to be i.i.d. and follow a type one extreme value distrubtion. We normalize the utility of outside options to zero without loss of generality since only differences in utility across the choices available to a depositor impact her decisions. Each depositor chooses the good for which she has the highest realized utility.

[6]The demand curves for mortgages and loans are defined similarly. We use the subscript M for mortgages and L for loans to describe these demand systems.

According to equation (10) in McFadden (1974), the quantity of deposits invested in branches of bank m in market n at time t satisfies

$$Q_{D,nmt} = \overline{F}_{D,nt} \frac{\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{1 + \sum_{m'} \exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})} \tag{11}$$

under the standard assumptions in a logit demand system.

Taking the log of equation (11) yields

$$\log Q_{D,nmt} = \zeta_{D,nt} + \alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt}, \tag{12}$$

where $\zeta_{D,nt} = \log(\frac{\overline{F}_{D,nt}}{1+\sum_{m'} \exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})})$ is a market-time specific term that is the same across banks m. To estimate the deposit demand curve for bank m in market n, we need to know how its quantity $Q_{D,nmt}$ is impacted by changes in the rate $R_{D,nmt}$ that it pays. This relation is given by

$$\frac{\partial \log Q_{D,nmt}}{\partial R_{D,nmt}} = \frac{\partial \zeta_{D,nt}}{\partial R_{D,nmt}} + \alpha_D. \tag{13}$$

To estimate the demand curve in equation (13), we first estimate α_D by controlling for the impact of $\zeta_{D,nt}$ with market-time fixed effects. We then present a method for estimating the impact of changes in banks' interest rates on the term $\zeta_{D,nt}$ that impacts all banks in the market.

Our first goal is to estimate the price disutility parameter, α_D. We face the endogeneity problem that banks may choose an interest rate that is correlated with its unobserved charecteristics $\delta_{D,nmt}$. Directly regressing log market shares, $\log Q_{D,nmt}$, on interest rates, $R_{D,nmt}$, observable characteristics, $X_{D,nmt}$, and a market-time fixed effect to absorb $\zeta_{D,nt}$ may yield biased estimates because a bank with unobservably high-quality banking services, $\delta_{D,nmt}$, can pay a lower deposit rate than a bank with low quality banking services and still attract depositors. To address the endogeneity concern, we require an instrument $z_{D,nmt}$ that impacts banks' interest rate choices but is not correlated with unobserved quality $\delta_{D,nmt}$. With such an instrument, we can estimate α_D and β_D using the following two-stage least squares specification:

$$R_{D,nmt} = \gamma_{D,nt} + \gamma_D z_{D,nmt} + X_{D,mt}\gamma_D + e_{D,nmt}, \tag{14}$$

$$\log Q_{D,nmt} = \chi_{nt} + \mathbb{E}_{D,nt}\delta_{D,nmt} + \alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \varepsilon_{nmt}, \tag{15}$$

where $\chi_{nt} = \zeta_{D,nt} + \mathbb{E}_{D,nt}\delta_{D,nmt}$ and $\varepsilon_{nmt} = \delta_{D,nmt} - \mathbb{E}_{D,nt}\delta_{D,nmt}$. The term $\mathbb{E}_{D,nt}\delta_{D,nmt}$ is the market-specific mean of the unobserved characteristics, $\delta_{D,nmt}$. The mean $\mathbb{E}_{D,nt}\delta_{D,nmt}$ is added to $\zeta_{D,nt}$ because it will be absorbed by the market-time fixed effect in equations (15). Thus, the residual in the regression is $\delta_{D,nmt} - \mathbb{E}_{D,nt}\delta_{D,nmt}$ rather than $\delta_{D,nmt}$. The demand for mortgages and loans is defined similarly.

3.1.2 Market Size

To estimate the price disutility parameter, α_D, we used market-time fixed effects that control for any market-level shocks to the desirability of bank deposits. With this fixed effect added, we inferred α_D from observing how differences in banks' log-quantities are impacted by differences in the rates they pay. To fully characterize the demand curves banks face, we also need to know how the overall quantity of deposits in a market would respond if every bank in the market changed its interest rate. This section develops a novel approach to estimating how the overall quantity in a market changes with an aggregate change in rates, which is the final piece of information needed to complete the estimation of our demand systems.

We again use deposits as an example to illustrate how the elasticity of substitution between the market-level deposit volume and the outside options is estimated. Denoting the total quantity of deposits in a market as $\overline{Q}_{D,nt}$, we obtain

$$\overline{Q}_{D,nt} = \overline{F}_{D,nt} \frac{\sum_m \exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{1 + \sum_{m'} \exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})} \tag{16}$$

by summing the deposit quantities of bank branches operating in market n given in equation (11).

We define

$$\psi_{D,nt} = \log(\sum_m \exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})) \tag{17}$$

to represent the desirability of a "composite good" provided by all banks operating in the market. Then, $\overline{Q}_{D,nt} = \overline{F}_{D,nt} \frac{\exp(\psi_{D,nt})}{1+\exp(\psi_{D,nt})}$, and using a log-linear approximation,

$$\log \overline{Q}_{D,nt} \approx \log \overline{F}_{D,nt} + \beta_{D,o}\psi_{D,nt}. \tag{18}$$

From this equation, we can estimate how $\log \overline{Q}_{D,nt}$ changes with the value of $\psi_{D,nt}$ to learn the value of $\beta_{D,o}$. The parameter $\beta_{D,o}$ quantifies the sensitivity of total deposit quantities to changes in the overall desirability of deposits. In Online Appendix D, we present a slight modification of our logit demand system where equation (18) holds exactly.[7]

[7] This modification differs from our logit demand system only in that depositors have heterogeneous values of consuming the outside option.

We apply an instrumental variables approach to estimate parameter $\beta_{D,o}$. From our estimation of the price disutility parameters in equations (14) and (15), we can observe all terms in equation (17) that defines $\psi_{D,nt}$ except the mean of $\delta_{D,nmt}$. We therefore decompose $\psi_{D,nt}$ into an observable component, $\psi^o_{D,nt}$, and an unobservable component, $\psi^u_{D,nt}$, where[8]

$$\psi^u_{D,nt} = \frac{1}{N_{nt}} \sum_m \delta_{D,nmt}, \tag{19}$$

$$\psi^o_{D,nt} = \log\left(\sum_{m'} \exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt} - \psi^u_{D,nt})\right). \tag{20}$$

We need an instrumental variable $z_{D,nt}$ that is uncorrelated with the unobserved component, $\log \overline{F}_{D,nt} + \beta_{D,o}\psi^u_{D,nt}$, to consistently estimate $\beta_{D,o}$ using two-stage least squares. Specifically, we have

$$\psi^o_{D,nt} = \rho_{D,t} + \theta_D z_{D,nt} + \chi_{D,nt}\theta_D + \varepsilon^o_{D,nt}, \tag{21}$$

$$\log \overline{Q}_{D,nt} = \alpha_{D,t} + \beta_{D,o}\psi^o_{D,nt} + \chi_{D,nt}\rho_D + \eta_{D,nt}, \tag{22}$$

where $\chi_{D,nt}$ is a vector of controls.

3.2 Instruments and Data

Estimating our demand systems requires information on deposits, mortgages, loans, and bank characteristics. In addition we construct an instrumental variable from property damage data. We first introduce the data we use and then explain how our instruments are constructed. Summary statistics for the demand-side variables are reported in Table 1.

Deposits. County-level deposit volumes are obtained from the FDIC, which covers the universe of US bank branches at an annual frequency from June 2001 to June 2017. We exclude branches that consolidate deposits in another location, do not accept deposis, or are owned by foreign banks. We define each county-year as a deposit market and sum branch-level deposits at the bank-county-year level. Our sample is from 2001 to 2017. Table 1 reports the summary statistics.

County-level deposit rates are obtained from RateWatch, which collects weekly branch-level deposit rates by product. Data coverage varies by product, especially in the earlier years. To maximize the sample size, we focus on the most commonly available savings account type, which is the 10K money market account. We collapse the data at the bank-county-year level from June

[8] Equation (A12) in Appendix A.2.1 provides an expression for $\psi^o_{D,nt}$ in terms of observable data.

Table 1: Summary Statistics (Market-Bank-Year Level)

This table reports summary statistics of bank deposits, mortgages, and loans at the market-bank-year level. Rates are reported in basis points and volumes are in millions. The instrument refers to property losses due to natural disasters. The sample period is from 2001 to 2017.

	Num. of Obs.	Mean	25th Pct.	50th Pct.	75th Pct.	Std. Dev.
Log Deposit Market Share	74007	−2.67	−3.45	−2.33	−1.50	1.69
Deposit Volume	74007	188.47	23.05	47.82	103.16	2287.78
Deposit Rate	45894	58.04	10.00	20.00	80.00	77.98
Log Mortgage Market Share	38957	−4.12	−5.32	−3.73	−2.56	2.08
Mortgage Volume	38957	23.67	1.23	3.79	11.62	209.53
Mortgage Rate	11735	457.62	332.50	450.55	570.00	126.41
Log Loan Market Share	25943	−5.06	−6.62	−4.95	−3.45	2.09
Loan Volume	25989	977.24	40.25	132.00	553.78	3218.81
Loan Spread	25943	183.52	101.38	171.43	250.00	120.46

2001 to June 2017 to match with the reporting of the branch-level deposit volumes from the FDIC.

The branch-level identifier in Ratewatch (accountnumber) is matched to the branch-level identifier in the FDIC data (uninumbr) using the mapping file developed by Bord (2017).[9]

Mortgages. We use data on mortgage originations made available under the Home Mortgage Disclosure Act (HMDA). The data available to us is at the annual frequency and includes information on the lender, loan size, location of the property, loan type, and loan purpose. Any depository institution with a home office or branch in a Central Business Statistical Area (CBSA) is required to report data under HMDA if it has made or refinanced a home purchase loan and if it has assets above $30 million. As explained by Cortés and Strahan (2017), the bulk of residential mortgage lending activity is likely to be reported under this criterion.[10] We define each county-year as a mortgage market and sum mortgage loan volumes at the bank-county-year level. Our sample is from 2001 to 2017.

County-level mortgage rates are obtained from RateWatch, which collects weekly branch-level mortgages rates by product. Data coverage varies by product, especially in the earlier years. To maximize the sample size, we focus on the most commonly available mortgage loan product, which is the 15-year Fixed Rate Mortgage. We collapse data at the bank-county-year level from 2001 to 2017 to match with the reporting of the mortgage volume data from the HMDA.

[9]Special thanks to Vitaly Bord for sharing the mapping file with us.

[10]Any non-depository institution with at least 10% of its loan portfolio composed of home purchase loans must also report HMDA data if its asset size is above $ million. These institutions are not included in our sample given our focus on deposit-taking commercial banks.

We first merge bank-level identifiers in HMDA to the FDIC bank-level identifiers using the mapping file developed by Bob Avery.[11] Then, the branch-level identifier in the FDIC data (uninumbr) is merged with the branch-level identifier in Ratewatch (accountnumber) using the mapping file developed by Bord (2017).

Loans. We use data on syndicated loans from Thomson Reuters Dealscan database. We select all loans originated by US banks and sum loan volumes at the bank-state-year level, where the location of the borrower is given in Dealscan. We define loan markets at the state-year level instead of the county-year level because firm borrowers tend to be less geographically confined than individual depositors. Similarly, we collapse loan spreads at the bank-state-year level. Our sample is from 2001 to 2017.

We build on the mapping file used in Chakraborty et al. (2018) to hand-match lenders in Dealscan to Call Report bank identifiers (RSSD).[12]

Bank Characteristics. We use Call Reports to obtain bank-level characteristics as control variables. Specifically, we calculate the ratio of insured deposits as insured deposits over total liabilities and the ratio of loan loss provision as loan loss provisions over total loans. We collapse the data at the bank-year level from 2001 to 2017.

Property losses from natural disasters. Property losses from natural disasters are obtained from the Spatial Hazard Events and Losses Database for the United States (SHELDUS). This dataset records the location, time, and damage brought about by natural disasters in the US . We include all reported disasters in the database and calculate the total property losses for each county-year from 2001 to 2017 for our instrument.

Instruments. We first explain our instrument, $z_{D,nmt}$, for estimating the price disutility of deposits. Recall that interest rates may be correlated with unobserved product quality. Hence, we use a two-stage least squares approach with a supply shock as the instrument to estimate the price disutility parameter of our demand systems as in equations (14) and (15). Following Cortés and Strahan (2017), we construct the instrument based on property losses from natural disasters and banks' branch networks. As Cortés and Strahan (2017) show, natural disasters increase the demand for loans in the local area, which means that banks present in the area allocate funds from branches in other branches to branches in affected counties through their internal capital markets. Hence, property losses to bank m's branches in regions n' constitute a supply shock to bank m's

[11]The version we used is available here https://sites.google.com/site/neilbhutta/data.

[12]Special thanks to Indraneel Chakraborty, Itay Goldstein, and Andrew MacKinlay for sharing the mapping file with us.

branches in county n, which allows us to trace out the demand curves. Appendix G formalizes this argument.

Formally, our natural disaster instrument $z_{D,nmt}$ measures for bank m's branches in county n and year t the property losses from natural disasters accrued to the bank's branches in all other counties n':

$$z_{D,nmt} \quad = \quad \frac{1}{N^u_{mt}} \log \left(\sum_{n'} damage_{n't} \cdot \frac{Q_{D,n'mt}}{\sum_{n_0} Q_{D,n_0 mt}} \right),$$

where N^u_{mt} is the number of branches of bank m that are not affected by natural disasters, and $damage_{n't}$ is the property loss in county n'. Following Cortés and Strahan (2017), we scale $damage_{n't}$ by the fraction of deposits belonging to branches of bank m in county n and take logs after summing the scaled damage losses. The former adjustment captures the portion of the demand shock in county n absorbed by branches of bank m, while the latter ensures that the largest shocks (e.g. Hurricane Katrina) do not drive the overall result. The instrument for mortgages follows that of deposits. For commercial loans, we use the same instrument constructed at the bank-state-year level instead of the bank-county-year level, where the state is determined by the location of the borrower's headquarters.

One concern for our identification could be that the effect of disasters spills over to affect local demand in unaffected counties. To this end, notice that our exclusion restriction does not require the absence of spillover effects altogether. It only requires that any potential influence of natural disasters on unobserved deposit characteristics in unaffected areas is not correlated with banks' branch networks. We include the log property damage to each county in all specifications to help account for any direct effects of disaster losses on demand. Another concern could be that loan losses from the disaster itself directly influence interest rates. To this end, we also include banks' loan loss provision as control variable in all specifications.

Next, to estimate the sensitivity of total deposit quantities to changes in the overall desirability of deposits as in equations (21) and (22), we average our market-bank-time level instrument, $z_{D,mnt}$, at the market-year level to construct

$$z_{D,nt} = \frac{1}{N_{nt}} \sum_m z_{D,nmt}.$$

This instrument captures how exposed a region is to indirect rate changes coming through internal capital markets. The identifying assumption is that the indirect shocks through banks' internal capital markets are uncorrelated with the log-size of each market, $\log F_{D,nt}$, and with the average unobservable quality, $\psi^u_{D,nt}$. The corresponding instrument for mortgages is constructed in the same way.

3.3 Estimation Results

Table 2 reports the first-stage and second-stage results for the price disutility estimation for deposits, mortgages, and loans in equations (14) and (15). Since our deposit volume is a stock measure, whereas new issuances of mortgages and loans are flow measures, we include the lagged deposit market share to account for persistence in the stock of deposits and the share of insured deposits to capture differences in the deposit base.

The price disutility parameters reported in the first row of panel (b) of Table 2 are positive for deposits and negative for mortgages and loans. Intuitively, deposit rates are paid by the bank so that raising deposit rate increases a bank's market share. In contrast, mortgage, and loan rates are paid by borrowers, so a bank can improve its market share by offering lower mortgage and loan rates. Quantitatively, the coefficients imply that when an infinitely small bank raises its deposit rate in one county by 10 basis points, its deposit volume will increase by 4.7%.[13] When the same bank lowers its mortgage and loan rates in one market by 10 basis points, its mortgage and loan volumes increase by 57.5% and 48.7%, respectively. The price disutility of deposits is an order of magnitude smaller than that for mortgages and loans, consistent with depositors being less attentive to interest rates than firm and mortgage borrowers.

Regarding the outside option, we estimate the sensitivity of market-level quantities $\overline{Q}_{P,nt}$ to the market-level desirability parameter $\psi^o_{P,nt}$ as in equations (21) and (22) for deposits and mortgages. We include the average age, average income, the share of residents with a college degree, log population, growth of house prices, log property damage due to natural disasters, and lagged quantities as county-level control variables.

Panel (b) in Table 3 reports the sensitivity of market-level quantities $\overline{Q}_{P,nt}$ to the market-level desirability parameter $\psi^o_{P,nt}$ to be 0.28 for deposits and 0.07 for mortgages. Hence, as we show in equation (A22), the increase in deposit quantity when all banks in a county raise their deposit rates by 10 basis points is given by

$$\frac{\partial \log \overline{Q}_{D,nt}}{\partial R_{D,nt}} = \frac{\partial \log \overline{Q}_{D,nt}}{\partial \psi^o_{D,nt}} \frac{\partial \psi^o_{D,nt}}{\partial R_{D,nt}} = 0.28 \times 4.7\% = 1.3\%.$$

[13]The magnitude of the price disutility parameters can be interpreted for an infinitely small bank because the interest rates of that bank will have a negligible impact on the observed desirability of the aggregate deposits at the county level, i.e., when $\frac{\partial \zeta_{D,nt}}{\partial R_{D,nmt}} = 0$ in equation (13).

Table 2: Demand System Estimates

This table reports the two-stage least squares results for estimating price disutility of deposit, mortgage, and loan demand systems as in equations (14) and (15). These regressions are run at the market-bank-year level. Loan loss provision is the ratio of loan loss provision over total loans, lag deposit market share is the deposit market share in the county lagged by 1 year, lag insured deposit ratio is the ratio of insured deposits over total liabilities lagged by 1 year, and log property damage is the direct property loss from natural disasters at the county level. For the deposit, mortgage and loan rates, 0.01 means 1%. The sample period is from 2001 to 2017. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors are clustered at the year level.

Panel (a): First Stage Panel Regression

	(1)	(2)	(3)
	Deposit Rate	Mortgage Rate	Loan Rate
IV	1.76***	10.82***	2.15***
	(0.22)	(1.66)	(0.28)
Loan Loss Provision	0.01***	−0.02***	0.02*
	(0.003)	(0.01)	(0.01)
Lag Insured Deposit Ratio	0.004***		
	(0.001)		
Log Property Damage	−0.0003***	−0.0003***	
	(0.0001)	(0.0000)	
Observations	217,623	77,329	25,115
R^2	0.82	0.91	0.19
Adjusted R^2	0.77	0.85	0.16
Market-Year F.E.	Y	Y	Y

Panel (b): 2SLS Panel Regression

	(1)	(2)	(3)
	Deposit Market Share	Mortgage Market Share	Loan Market Share
Rate (with IV)	11.26**	−574.89***	−487.30***
	(4.86)	(72.33)	(76.96)
Loan Loss Provision	−0.75*	−15.47***	8.41
	(0.39)	(5.21)	(5.23)
Lag Insured Deposit Ratio	−0.02		
	(0.03)		
Log Property Damage	−0.01***	0.77***	
	(0.003)	(0.04)	
Observations	217,623	77,329	25,115
Market-Year F.E.	Y	Y	Y

Table 3: Outside Option Estimates (Deposits and Mortgages)

This table reports two-stage least squares results for estimating the sensitivity of market-level quantities to the aggregate observed desirability parameter ψ^o_{nt} as in equations (21) and (22). The regression is run at the market-year level. We include market-year level controls, including average age and income of the population, fraction of residents college degree, log population, annual house price growth, log property loss due to natural disaster, and lag log deposit quantity. For the deposit and mortgage rates, 0.01 means 1%. The sample period is 2001-2017. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively. Standard errors are clustered by year.

Panel (a): First Stage		
	(1)	(2)
	Deposit Rate	Mortgage Rate
IV	0.01	-0.87^{***}
	(0.01)	(0.21)
Observations	39,053	25,234
R^2	0.69	0.88
Controls	Y	Y
Market-Year F.E.	Y	Y

Panel (b): 2SLS		
	(1)	(2)
	Deposit Share	Mortgage Share
ψ^o (with IV)	0.19	0.07
	(0.14)	(0.04)
Observations	39,053	25,234
R^2	0.99	0.91
Controls	Y	Y
Market-Year F.E.	Y	Y

Similarly, when all banks in a county lower their mortgage rates by 10 basis points, the mortgage quantity increases by $0.07 \times 57.5\% = 4.0\%$.

For loans, our estimate of the quantity of firms choosing the outside option uses the fact that we can observe both firms that do and do not borrow. We count the number of firms in the Dealscan database that did not borrow in a given year and state and the divide the number by four, which reflects the average loan maturity. We then multiply this number of firms times the average size of a loan in its market. The average loan size is linearly projected from the existing loans in that year with state fixed effect to account for state-level heterogeneity in the size of loans. The underlying assumption is that potential borrowers would have on average obtained a loan of the same size as the existing ones in the market that year.

We report the outside option size for firms at the state-year level in the Online Appendix Table OA1. In 2007, for example, the implied β_o is 0.33. This means that when all banks in a state lower their loan rates by 10 basis points, the loan quantity increases by $0.33 \times 48.7\% = 16.1\%$. Notice that the demand elasticity of loans is higher than that of mortgages because although their price disutility parameters are of similar magnitudes, the outside option of loans responds much more to changes in observed desirability than in the case of mortgages. One reason could be that borrowers in the syndicated loan market have more flexibility to borrow from other sources such as the bond market. Although our focus is on bank lending to firms, our framework could be extended to account for a potential substitution from loans to bonds with public firm-level data on bond financing. Deposits have a low sensitivity along both dimensions, which leads to a highly inelastic deposit demand curve.

In absolute terms, if all banks raise their deposit rates by 10 basis points based on 2007 levels, the aggregate deposit volume will increase by \$36.9 billion; if all banks lowered their mortgage rates by 10 basis points, the aggregate mortgage volume will increase by \$13.8 billion; and if all banks lowered their loan rates by 10 basis points, the aggregate loan volume would increase by \$236.8 billion.

When a bank adjusts its deposit rate, its impact on the observed desirability of aggregate deposits at the county level also influences its deposit demand. In equation (A19) of Appendix A.2.1, we show that the sensitivity of a bank's deposit quantity to its deposit rate is given by

$$\frac{\partial \log Q_{D,nmt}}{\partial R_{D,nmt}} = \alpha_D + \alpha_D(\beta_{D,o} - 1)\frac{Q_{D,nmt}}{\overline{Q}_{D,nt}}. \tag{23}$$

Using our previous estimates of α_D and $\beta_{D,o}$ and the sample of banks active in all 3 markets in 2007, we find by averaging this expression across banks that the response of an average bank's deposit quantity in a given county is 3.4%, or 17.6 million dollars, with respect to a 10 basis points increase in deposit rate. Similarly, the response of an average bank's mortgage quantity in a given county is 52.4%, or 5.1 million dollars, with respect to a 10 basis points decrease in the mortgage rate. The average response in the average bank's loan quantity in a given state is 47.1%, or 630 million dollars, with respect to a 10 basis points decrease in the loan rate.

Because our result that the reserves created by QE crowd out loans relies on our estimate of the elasticity of loan demand, we present an alternative estimation approach in Online Appendix B. This approach to estimating firms' demand for loans exploits the fact that firms tend to borrow persistently from the same bank multiple times, even though they have the option of borrowing from a new bank. Because of this persistence, we can identify firms' elasticity of credit demand by observing how much they decrease their total borrowing when there is a negative credit supply shock to a bank from which they have already borrowed. We find similar results to our benchmark estimates, as we discuss in Online Appendix B.

4 Cost Function

This section specifies and estimates the bank's cost function for producing deposits, mortgages, and loans. Quantifying the cost function is challenging because banks choose all their balance sheet components simultaneously. We propose a specification of how banks' costs depend on their balance sheet components that can be feasibly estimated using multiple instrumental variables. In our novel estimation method, we first perform a reduced form analysis of how banks' marginal costs respond to exogenous shocks to the demand for borrowing and lending across different markets. Next, we estimate the bank's cost function by choosing its parameters to replicate these reduced form results. This estimated cost function tells us how a bank's marginal costs are impacted when the supply of reserves in the banking system changes.

We first set up the cost function in Section 4.1 and then explain how we estimate it in Section 4.2. Section 4.3 describes the data and instruments we use and Section 4.4 reports our estimation results.

4.1 Cost Function Specification

We proceed to formally set up the the bank's cost function. For bank m at time t, we let the cost function be

$$\begin{aligned} C(\Theta_{mt}) &= H(Q_{D,mt}, Q_{M,mt}, Q_{L,mt}, Q_{S,mt}) \\ &+ \sum_n (Q_{M,nmt}\varepsilon^Q_{M,nmt} + Q_{L,nmt}\varepsilon^Q_{L,nmt} + Q_{D,nmt}\varepsilon^Q_{D,nmt}) + Q_{S,mt}\varepsilon^S_{mt}. \end{aligned} \tag{24}$$

This includes a term $H(Q_{D,mt}, Q_{M,mt}, Q_{L,mt}, Q_{S,mt})$, which depends on the bank-level quantities of deposits, mortgages, loans, and securities. This specification allows, for example, for the bank's holding of securities to impact its cost of mortgage lending.[14] In addition, the cost function features shocks to the cost of borrowing or lending in individual markets (given by each of the ε_{nmt} variables). These market-specific shocks are assumed to be linear in the bank's market-specific quantities and give us the flexibility to match observed rates and quantities. However, as shown in equations (6) to (8), the response of our model to external shocks such as an increase in reser supply depends only on the second derivatives of the bank's cost function, which are fully reflected in the function H that we estimate.

To model how a bank's costs depend on the composition of its balance sheet in a manner that is flexible and yet restrictive enough to be identified from data, we assume the following functional

[14]This dependence is allowed for at the bank-level rather than at the bank-market level, which is consistent with banks optimizing over their bank-level balance sheets considering their bank-level balance sheet composition. This bank-level dependence is also consistent with banking regulation applied to bank-level balance sheets.

form for H

$$H(Q_{D,mt}, Q_{M,mt}, Q_{L,mt}, Q_{S,mt}) = \mu_D Q_{D,mt} + \mu_M Q_{M,mt} + \mu_L Q_{L,mt} + \mu_Q Q_{S,mt}$$
$$+ \quad \frac{1}{2}(K_1 \mathcal{E}_{mt}^2 + K_2 \mathcal{I}_{mt}^2 + K_3 Q_{D,mt}^2 + 2K_4 \mathcal{I}_{mt} Q_{D,mt} + 2K_5 \mathcal{E}_{mt} Q_{D,mt}), \tag{25}$$

where $\mathcal{E}_{mt} = Q_{M,mt} + Q_{L,mt} + Q_{S,mt} - Q_{D,mt}$ and $\mathcal{I}_{mt} = Q_{S,mt} + \omega_M Q_{M,mt} + \omega_L Q_{L,mt}$. The term $\mathcal{E}_{mt}$ can be interpreted as the bank's equity and non-deposit debt funding because it equals the gap between the value of bank assets that we observe on the bank's balance sheet and its deposit financing.[15] The term $\mathcal{I}_{mt}$ measures the liquidity of a bank's assets, where the coefficients ω_M and ω_L capture how much less liquid mortgages and loans are than reserves.

This cost function has two key features. First, it is quadratic in all bank-level quantities, which implies that a bank's marginal costs of borrowing and lending are linear in the quantities on the bank's balance sheet. This will allow us to use linear instrumental-variable regressions as a straightforward tool for estimating its parameters. Second, the quadratic component of the cost function has 7 unknown parameters (ω_M,ω_L,K_1,K_2,K_3,K_4,K_5). As we show in Appendix A.3, this is precisely the number of parameters that can be estimated by observing how our bank responds to two different cross-sectional instrumental variables.

While this cost function can flexibly incorporate a range of different frictions, we present one consistent microfoundation for it in Appendix C based on Kashyap et al. (2002). In Kashyap et al. (2002), banks must meet liquidity demands due to both their runnable deposit liabilities as well as their assets such as credit lines that borrowers can draw down. We assume that liquidity demands from bank depositors are imperfectly correlated with liquidity demands from bank borrowers and that banks face a quadratic cost of raising new funds. This imperfect correlation provides banks with a degree of liquidity insurance that is reflected in the interaction terms $\mathcal{I}_{mt} Q_{D,mt}$ and $\mathcal{E}_{mt} Q_{D,mt}$. Alternatively, the cost function can also reflect frictions ranging from exogenous costs from bank regulation such as the leverage ratio requirements, the risk of having to sell illiquid assets during a bank run, and incentives to take excessive risk due to the availability of deposit insurance. For our purpose, it is only necessary to quantify how a bank's cost of capital is impacted by an increase in reserve supply rather than to take a stand on the specific mechanism at play.

4.2 Estimation Strategy

This section uses our quadratic cost function to derive linear expressions for banks' marginal costs that are tractable to estimate. Differentiating equation (24) implies that the marginal costs of

[15] $\mathcal{E}_{mt}$ is not a perfect measure of a bank's equity and non-deposit debt financing. It also incorporates the small amount of assets held on the bank's balance sheet that are not included in our measures of loans, mortgages, and securities.

deposits, mortgages and loans for bank m in market n at time t are

$$\frac{\partial C}{\partial Q_{D,nmt}} = \mu_D - K_1\mathcal{E}_{mt} + K_3 Q_{D,mt} + K_4\mathcal{I}_{mt} + K_5(\mathcal{E}_{mt} - Q_{D,mt}) + \varepsilon^D_{nmt}, \tag{26}$$

$$\frac{\partial C}{\partial Q_{M,nmt}} = \mu_M + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_M + K_4 Q_{D,mt}\omega_M + K_5 Q_{D,mt} + \varepsilon^M_{nmt}, \tag{27}$$

$$\frac{\partial C}{\partial Q_{L,nmt}} = \mu_L + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_L + K_4 Q_{D,mt}\omega_L + K_5 Q_{D,mt} + \varepsilon^L_{nmt}, \tag{28}$$

$$\frac{\partial C}{\partial Q_{S,mt}} = \mu_S + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt} + K_4 Q_{D,mt} + K_5 Q_{D,mt} + \varepsilon^S_{mt}. \tag{29}$$

Banks' marginal costs of providing deposits, mortgages, and loans on the left hand side of equations (26) to (29) can be obtained using our estimated demand systems. Recall the first order conditions for a bank's profit maximizing choices, equations (2) to (4). These first order conditions can be rewritten as

$$-\frac{Q_{D,nmt}}{\partial Q_{D,nmt}/\partial R_{D,nmt}} - R_{D,nmt} = -R^{D,m}_t + \frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}} \tag{30}$$

$$-\frac{Q_{M,nmt}}{\partial Q_{M,nmt}/\partial R_{M,nmt}} - R_{M,nmt} = -R^{M,m}_t - \frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}, \tag{31}$$

$$-\frac{Q_{L,nmt}}{\partial Q_{L,nmt}/\partial R_{L,nmt}} - R_{L,nmt} = -R^{L,m}_t - \frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}. \tag{32}$$

Equations (30) to (32) tell us that if banks are maximizing profits facing the demand systems we estimated, the interest rates they chose reveal what their marginal costs must be. The left hand side of equations (30) to (32) depend only on observed interest rates and markups so we can infer the marginal costs $\frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}}$, $\frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}$, and $\frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}$ up to the value of unknown constants $R^{D,m}_t, R^{M,m}_t, R^{L,m}_t$. These constants are market-wide discount rates reflecting the riskiness of cash flows from deposits, mortgages, and loans, so they do not depend on the composition Θ_{mt} of the bank's balance sheet. Hence, we can replace the marginal costs on the left hand sides of equations (26) to (29) with their observable counterparts from equations (30) to (32). The right hand sides would change only in their intercept since the discount rates $R^{D,m}_t, R^{M,m}_t, R^{L,m}_t$ does not depend on the composition of the bank's balance sheet. Averaging equations (26) to (28) across the markets

n in which the bank operates yields

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{D,nmt}}-R_t\right) = \mu_D^* - K_1\mathcal{E}_{mt} + K_3 Q_{D,mt} + K_4\mathcal{I}_{mt} + K_5(\mathcal{E}_{mt}-Q_{D,mt}) + \varepsilon_{mt}^D, \quad (33)$$

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{M,nmt}}+R_t^{M,m}\right) = \mu_M^* + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_M + K_4 Q_{D,mt}\omega_M + K_5 Q_{D,mt} + \varepsilon_{mt}^M, \quad (34)$$

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{L,nmt}}+R_t^{L,m}\right) = \mu_L^* + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_L + K_4 Q_{D,mt}\omega_L + K_5 Q_{D,mt} + \varepsilon_{mt}^L, \quad (35)$$

where each intercept μ is now some other constant μ^* due to the changed left hand side.

To estimate the parameters in equations (33) to (35), we need to see how the marginal costs on the left hand side of each equation respond to changes in the bank balance sheet quantities on the right hand side. Because banks may face unobservable shocks to their cost of borrowing or lending that may affect their choice of balance sheet quantities, we need exogenous variation in the quantities on the right hand side of each equation that is uncorrelated with the ε cost shocks. Further, there are multiple endogenous variables on the right hand side of each equation. For example, if we see how a bank's marginal cost of mortgage lending responds to an increase in both its deposit quantities and its mortgage quantities, we need to know how each of these two quantity changes individually impacted the bank's marginal cost. To overcome this problem, we use two cross-sectional instrumental variables z_{mt}^1, and z_{mt}^2 that are both uncorrelated with the cost shocks ε_{mt}.

We then regress a bank's marginal cost of providing deposits, mortgages, and loans on instruments z_{mt}^i, where $i = 1, 2$,

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{D,nmt}}-R_t\right) = \theta_t^D + \kappa^{i,D} z_{mt}^i + u_{D,mt}^Q, \quad (36)$$

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{M,nmt}}+R_t^{M,m}\right) = \theta_t^M + \kappa^{i,M} z_{mt}^i + u_{L,mt}^Q, \quad (37)$$

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{L,nmt}}+R_t^{L,m}\right) = \theta_t^L + \kappa^{i,L} z_{mt}^i + u_{L,mt}^Q. \quad (38)$$

In addition, we regress each of the bank's balance sheet quantities on each instrument as well:

$$Q_{D,mt} = \alpha_t^D + \gamma^{i,D} z_{mt}^i + \varepsilon_{D,mt}^Q \quad (39)$$

$$Q_{M,mt} = \alpha_t^M + \gamma^{i,M} z_{mt}^i + \varepsilon_{M,mt}^Q \quad (40)$$

$$Q_{L,mt} = \alpha_t^L + \gamma^{i,L} z_{mt}^i + \varepsilon_{L,mt}^Q \quad (41)$$

$$Q_{S,mt} = \alpha_t^S + \gamma^{i,S} z_{mt}^i + \varepsilon_{S,mt}^Q. \quad (42)$$

If our instruments are uncorrelated with all unobserved error terms, these regressions tell us that adding $\gamma^{i,D}$ deposits, $\gamma^{i,M}$ mortgages, $\gamma^{i,L}$ loans, and $\gamma^{i,S}$ securities to the bank's balance sheet changes its marginal cost of providing deposits by $\kappa^{i,D}$. This provides one relationship inferred using reduced form regressions that our cost function H must satisfy. Similarly, adding $\gamma^{i,D}$, $\gamma^{i,M}$, $\gamma^{i,L}$, and $\gamma^{i,S}$ in deposit, mortgage, loan, and security quantities result in mortgage and loan cost changes of $\kappa^{i,M}$ and $\kappa^{i,L}$, respectively. We include time fixed effects in all of these regressions so that we rely only on cross-sectional variation across banks for identification. We assume that our instruments do not trigger cross-sectional changes in the costs of securities, i.e., $\kappa^{i,S} = 0$, because securities trade in a competitive market where marginal costs are equalized across banks.

We must use all of these reduced form results jointly to identify our cost function. No cost function parameter would be identified from just observing the response of our costs and quantiies to just a single instrument z_{mt}^1. For example, if we observed that a bank's marginal cost of providing deposits increases after being hit by instrument z_{mt}^1 but that several quantities on the bank's balance sheet change, we do not know how much of this cost change is due to specifically to the change $\gamma^{i,D}$ in deposit quantities and how much is due to the change in the banks' other balance sheet quantities. In effect, the bank's balance sheet quantities $Q_{D,mt}, Q_{M,mt}, Q_{L,mt}, Q_{S,mt}$ are all endogenous variables that respond to our exogenous instruments. Multiple instruments are needed for identification when there are multiple endogenous variables. Appendix A.3 presents a system of equations that uniquely determines our cost function H so that it is consistent with our reduced-form regressions in equations (37) to (42).

4.3 Data and Instruments

Marginal Costs As explained in section 4.2, we infer the marginal costs for mortgages, deposits, and loans using our demand system estimates in section 3.3. We then average these estimated marginal costs to the bank-year level.

Deposits, Loans, Mortgages, and Securities We obtain bank-level quantities from Call Reports, which allows us to keep track of the volume of deposits, loans, mortgages, and securities that are actually retained on bank balance sheets to affect banks' costs. Mortgages loans are mapped to residential loans and commercial loans make up the remainder of loans from Call Reports. We further include bank-level securities from Call Reports, which is the sum of cash, Fed funds, repos, Treasury securities, and agency securities. Finally, we normalize all bank-level volume variables by the number of counties in which the bank operates in. This normalization prevents our estimates from being driven mostly by a few banks with a very large number of deposits.

Instruments We need two instruments, z^1_{mt} and z^2_{mt}, to identify the cost function parameters. These instruments are at the bank-level and must be independent of banks' cost shocks in the cross-section.

Our first instrument is simply the natural disaster losses that a bank's branches are directly exposed to. Unlike in the instrument for demand systems, we are no longer in need of a branch-level supply shock. Rather, disaster losses to an area increase the need to rebuild and repair local infrastructure and housing, which directly comprise a bank-level demand shock for bank lending. These disaster losses are also plausibly unrelated to variations in banks' marginal costs in the cross-section. Hence, we construct the instrument by adding up the disaster losses that each bank is exposed to through its branches. Specifically, for bank m at time t, we have

$$z^1_{mt} \quad = \quad \frac{1}{N_{mt}} \log\left(\sum_n damage_{nt} \cdot \frac{Q_{D,nmt}}{\sum_{n_0} Q_{D,n_0mt}}\right),$$

where $\sum_n damage_{nt} \cdot \frac{Q_{D,nmt}}{\sum_{n_0} Q_{D,n_0mt}}$ is the sum of disaster losses accrued to branches of bank m in county n and N_{mt} is the number of branches of bank m. We mathematically show in appendix G that for a firm that sells a good in several markets, a demand shock in one market can be used to estimate the firm's marginal cost curve.

Our second instrument is a Bartik deposit instrument. Following Bartik (1991), we construct our instrument based on the average growth rates of deposits in markets where banks have branches in. Intuitively, we make use of the fact that counties experience different rates of deposit growth and that banks operate branches in different counties to construct our Bartik deposit instrument. The identifying assumption is that the deposit growth rates in different counties that a bank is exposed to arise from county-level economic conditions rather than shocks to the bank's cost of supplying deposits, mortgages, and loans. Specifically, for bank m in year t, we have

$$z^2_{mt} \quad = \quad \frac{1}{N_{mt}} \left(\sum_n \frac{\overline{Q}_{D,nt} - \overline{Q}_{D,nt-1}}{\overline{Q}_{D,nt-1}}\right),$$

where $\frac{\overline{Q}_{D,nt}-\overline{Q}_{D,nt-1}}{\overline{Q}_{D,nt-1}}$ is the deposit market growth rate in county n and N_{mt} is the number of branches. In the baseline specification, we use a simple average to compute the bank-level exposure to county-level deposit growth, but our qualitative results are robust to using value-weighted exposures as well.

While the Bartik instrument should primarily serve as a shock to deposit demand while the natural disasters instrument should primarily serve as a shock to loan demand, we do not require them to only affect deposit or loan demand. The exclusion restiction should continue to hold even if the intruments affect the demand for other bank balence sheet components. The only requirement for identifiying our parameters is that the effects of the two instruments are not perfectly collinear, which we show is satisfied in the data.

4.4 Estimation Results

Table 4 reports the coefficients from regressing marginal costs and bank-level quantities on each of the two instruments, i.e., $(\kappa^{i,D}, \kappa^{i,M}, \kappa^{i,L}, \gamma^{i,D}, \gamma^{i,M}, \gamma^{i,L}, \gamma^{i,Q})$. Since these parameters are instrument-specific, we report the parameter values corresponding to the bank-level natural disaster shock in Panel (a) and the parameter values corresponding to the bank-level Bartik deposit shock in Panel (b).

According to Panel (a), banks with branches in areas with larger natural disaster losses also increase the volume of deposits, mortgages, loans, and securities on their balance sheets. At the same time, mortgage and loans become more costly to provide while deposits become less costly to provide for these banks. Taken together, we infer that the increase in volumes is consistent with an increase in loan and mortgage demand following natural disasters. From Panel (b), banks experiencing a positive Bartik deposit shock also increase their deposits, mortgages, loans, and securities. Deposit costs become less negative, implying that deposits become more costly to provide. At the same time, the costs of lending to firms and issuing mortgage loans declines as deposits become more abundant. Hence, the increase in balance sheet size in this case is aligned with a positive deposit demand shock, as expected from the Bartik deposit instrument.

Based on these coefficient estimates, we solve for the cost function's Hessian H and present the results in Table 5. First notice that all diagonal terms are positive, which means that a higher stock of deposits leads to a higher marginal cost on deposits, a higher mortgage stock leads to a higher marginal cost on mortgages, etc. Regarding the off-diagonal terms, the marginal cost of mortgages, loans, and securities are decreasing in deposits, which reflects a lower cost of lending and holding securities when deposit funding is more abundant. In other words, there are cost synergies between banks' deposit-taking and lending that support the joint provision of deposits and

Table 4: Cost Function Estimates

This table reports the sensitivity of bank-level costs and quantities to losses from natural disasters and a Bartik deposit shock as in equations (36) to (42). Sheldus Instrument refers to property losses due to natural disasters as explained in Section 4.3. Bartik Deposit Instrument refers to a bartik-style instrument of deposit growth as explained in Section 4.3. Rates are in basis points and quantities are in millions. The sample period is from 2001 to 2017. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Panel (a): Results using Natural Disaster Instrument

	Dep Cost	Mtg Cost	Loan Cost	Dep Vol	Mtg Vol	Loan Vol	Sec Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sheldus Instrument	-1.04^{***}	1.24^{***}	2.14^{***}	11.11^{***}	1.09^{***}	8.84^{***}	3.62^{***}
	(0.10)	(0.19)	(0.70)	(1.77)	(0.33)	(1.40)	(0.81)
Loan Loss Provision	-1.28	-16.55^{***}	4.86^{*}	8.10^{**}	27.00^{***}	536.38^{***}	1.13
	(1.30)	(2.83)	(2.59)	(3.81)	(4.18)	(17.48)	(1.74)
Observations	52,752	12,208	2,953	118,942	119,236	119,236	118,923
R^2	0.59	0.77	0.18	0.002	0.002	0.01	0.001

Panel (b): Results using Bartik Deposit Shock

	Dep Cost	Mtg Cost	Loan Cost	Dep Vol	Mtg Vol	Loan Vol	Sec Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bartik Instrument	64.39^{***}	-52.06^{***}	1.44	$1,414.31^{***}$	345.34^{***}	315.13^{***}	439.36^{***}
	(5.30)	(12.14)	(45.04)	(173.95)	(17.37)	(43.36)	(86.24)
Loan Loss Provision	-0.15	-16.68^{***}	3.99	31.07	24.35^{***}	161.54^{***}	-16.77
	(1.27)	(3.00)	(8.69)	(36.38)	(4.17)	(10.41)	(18.04)
Observations	49,265	10,446	2,512	62,352	66,839	66,839	62,346
R^2	0.46	0.77	0.18	0.002	0.01	0.005	0.001

loans by the same institution. However, the marginal cost of loans and mortgages are increasing in securities holdings, which suggests that banks' holdings of reserves and other liquid assets make it not cheaper but more costly to give out loans and mortgages. This increase in the marginal cost of lending will be a crucial determinant for the crowding out of loans following resserve injections in our counterfactual. It also implies that the crowding out effect of reserves on lending dominates potential cost synergies between liquid assets like reserves and illiquid assets like bank loans on bank balance sheets.

The coefficients imply that a \$100 million increase in reserves for each bank branch would increase the marginal cost for mortgages and loans by 100×0.317 =31.7 bps and $100 \times 0.264 =$ 26.4 bps, respectively. At the same time, the marginal cost of deposits decreases by 21.9 bps. To put these numbers in context, if \$1 trillion in reserves were injected and equally distributed across bank branches in 2008, each bank branch would receive \$39.4 million in reserves, which

Table 5: Cost Function Estimate

This table reports the cost function estimates including parameters K and ω, and the implied Hessian matrix H. Please refer to Section 4 for a detailed description of the estimation. The Hessian matrix reports impact of an extra \$1 million dollars per branch of a balance sheet quantity on the number of basis points by which a bank's marginal cost changes.

Parameter Estimates						
K_1	K_2	K_3	K_4	K_5	ω_M	ω_L
0.283	−0.018	0.043	0.014	0.050	−1.900	1.080

Implied Hessian H	$\frac{\partial C}{\partial Q_D}$	$\frac{\partial C}{\partial Q_M}$	$\frac{\partial C}{\partial Q_L}$	$\frac{\partial C}{\partial Q_S}$
Q_D	0.225	−0.260	−0.218	−0.219
Q_M	−0.260	0.220	0.319	0.317
Q_L	−0.218	0.319	0.263	0.264
Q_S	−0.219	0.317	0.264	0.266

would increase the marginal costs of mortgages and loans by 12.5 bps and 10.4 bps, respectively. Of course, the change in marginal costs may differ in equilibrium, where reserves are not equally distributed. The eventual effect on equilibrium quantities of mortgages and loans also depends on our estimated demand elasticities. To quantify the equilibrium impact of reserve injections in the banking system, we present a counterfactual analysis using both our estimated cost function and demand system in the next section.

5 Counterfactual Exercise

We use our estimated model to compute the effect of an increase in the supply of central bank reserves, as was caused by the Federal Reserve's QE Programs. These reserves are safe, liquid assets that must only be held by banks, so this increased supply forces banks to hold a larger portfolio of safe assets. While QE is an exchange between Treasuries and reserves, commercial banks only hold a very small proportion of Treasuries on bank balance sheets. Thus, the reserve injection comprises a net increase in banks' liquid asset holdings in our counterfactual.

The impact of this increased reserve supply has two main effects. First, an increase in reserve holdings changes banks' marginal cost of providing deposits, mortgages, and loans. This change in marginal cost is quantified by our estimated cost function in equation (24). Second, because of these cost changes, banks change the interest rates they choose to for deposits, loans, and mortgages. Given our estimated demand systems, we can compute how the equilibrium quantities of deposits, loans, and mortgages respond to these changes in the rates that banks choose. As a result, our model tells us how an increase in the supply of central bank reserves passes through to changes in both interest rates and quantities of deposits, mortgages, and loans provided by the banking system.

We note that our results focus on the effect of reserve injection on the banking system, which is an integral part of QE. Our results complement other transmission channels of QE that have been analyzed in the literature.[16] These transmission channels primarily depend on the effect of asset purchases, while our work is novel in that we zoom in on the effect of reserves that the central bank uses to finance asset purchases.

5.1 Computational Strategy

To compute our counterfactual, we need to determine each bank's holdings of reserves as well as the quantity and interest rate each bank charges for loans, deposits, and mortgages in each market after an injection in reserves. Formally, we need to compute an equilibrium set of interest rates and quantities that solve the bank's first order conditions in equations (2)-(5) with an increased supply of reserves. This is an over 38,000-dimensional problem, since we need to solve for interest rates for mortgages, deposits, and loans at every branch of every bank. Nevertheless, we can reduce the dimensionality considerably, and the model is tractable to solve. We define a function in equation (OA55) of the Appendix that maps the set of bank-level deposit, mortgage, and loan quantities to itself whose fixed point yields the equilibrium of our model.

We posit an increase R in the interest paid on securities above the yield earned in the data. We then compute the quantity of reserves the central bank must add to the financial system to attain this interest rate increase. Let $Q^i_{D,mt}, Q^i_{M,mt}, Q^i_{L,mt}$, and $Q^i_{S,mt}$, where i stands for initial, be the bank level quantities of deposits, mortgages, loans, and securities observed in the data. First, we start with a hypothesized vector of bank-level quantities $Q_{D,mt}, Q_{M,mt}, Q_{L,mt}$. Second, for each bank, we compute a security quantity $Q_{S,mt}$ so that the bank's marginal cost of holding securities is consistent with the rise R in the yield on securities. Third, given the vector of bank-level quantities $Q_{D,mt}, Q_{M,mt}, Q_{L,mt}$, $Q_{S,mt}$, we use our estimated cost function to compute a bank's marginal cost of holding deposits, mortgages, loans, and securities. Fourth, we compute the optimal interest rates banks choose that are jointly consistent with all of their marginal costs. Fifth, given the rates chosen in each market, we compute the bank-market-level quantities demanded by depositors/borrowers. Finally, we sum up the bank-market level quantities from the previous step and compute the difference from the hypothesized bank-level quantities $Q_{D,mt}, Q_{M,mt}, Q_{L,mt}$. The market is in equilibrium when this difference is 0. Please refer to Appendix E for further details.

[16] For example, Gagnon et al. (2010); Krishnamurthy and Vissing-Jorgensen (2011); Christensen and Krogstrup (2019); Rodnyansky and Darmouni (2017); Chakraborty et al. (2020), as discussed in our literature review.

Table 6: Counterfactual Results: QE

This table reports the results of our counterfactual analysis that injects the actual amount of reserves QE supplied for each year from 2008 to 2017. We compute the effects on rates and quantities, and report the average across years.

Average Change in Rates (in Basis Points)				Average Change in Quantities (in Trn Dollars)			
Deposits	Mortages	Loans	Securities	Deposits	Mortages	Loans	Securities
12.6728	18.8151	15.6384	15.9824	0.1224	-0.0218	-0.3530	1.8258

5.2 Counterfactual Results: The Reserve Supply Channel of QE

We conduct a year-by-year counterfactual with the amount of reserves supplied by QE in each year from 2008 to 2017. On average, this amounted to a reserve supply increase of $1.8 trillion per year. The average changes in interest rates and quantities that resulted are shown in Table 6. From the table, we observe that the interest rate paid on reserves increases by an average of 16.0 bps. The increase in reserve yields are passed through to the interest rates on deposits, mortgages, and loans by 12.7 bps, 18.8 bps, and 15.6 bps, respectively. In terms of quantities, bank loans extended to firms respond the most with an average decline of $353.0 billion, which implies that each $1 of reserves injected crowds out 19 cents of lending to firms. Mortgage and deposit volumes respond by less with an average annual drop of $21.8 billion and an annual increase of $122.4 billion, respectively.

We then zoom in on lending to firms where the impact of the reserve injection is largest. In Figure 3 we show the volume of reserves that were in the banking system each year and our estimated impact on bank loan quantities. The volume of reserves injected from QE increased from 2008 to 2014 and remained at elevated levels until 2017. The reduction in firm loans extended follows a very similar trend, reaching a maximum annual volume of $557.7 billion in 2015.

Discussion Our first key finding is that reserves crowd out bank lending to firms and that mortgage and deposit quantities respond by less. A larger reserve supply reduces lending because holding reserves raises the cost of providing loans, as our cost function estimates in Table 5 show. This implies that in a plentiful reserve environment, the benefits from holding additional liquid assets on bank balance sheets proposed by the theoretical literature are limited. On net, central bank reserves take up balance sheet space to crowd out bank lending capacity to the real economy. Therefore, our findings suggest that the increase in reserve supply following QE may bring about a counterproductive effect on the banking system.

Lending to firms is crowded out the most because the aggregate elasticity of loan demand

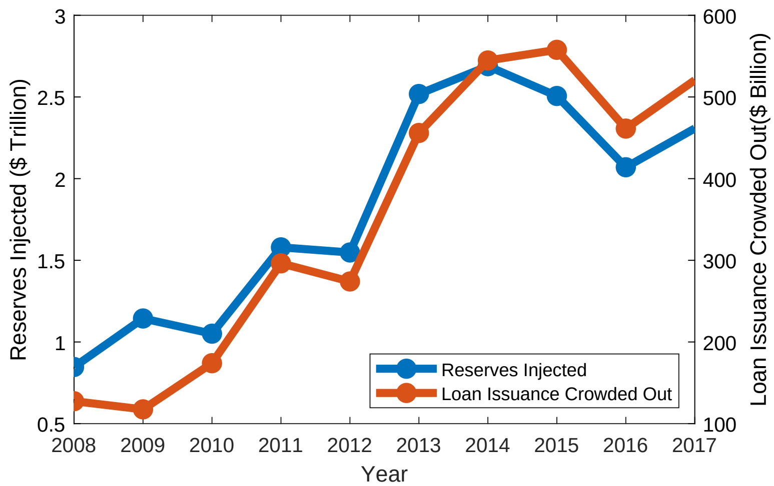

Figure 3: Reserve Supply and Reduction in Corporate Loan Issuance

is considerably larger than the aggregate elasticity of mortgage or deposit demand, as discussed in Section 3.3. While loan, mortgage, and deposit rates increase by similar amounts, this higher elasticity implies that corporate loan quantities are much more sensitive than deposit or mortgage quantities to an increase in reserve injections.

Another take-away from our results is that the model-implied mapping between reserve supply and the reserve spread largely follows that in the data. Adding the same volume of reserves as were injected through QE leads to a 17.0 bps increase in the model-implied reserve yield. In comparison, the average spread between the interest on excess reserves (IOER) and the federal funds rate was in the same ballpark at 11.6 bps.[17] Changes in our reserve yield also appear to move in tandem withchanges in the IOER-Fed funds spread with a correlation of 0.69. This comovement provides evidence in support of our model and estimation. The IOER-fed funds spread measures the higher risk-free yield available to banks, who can earn the interest rate paid by excess reserves, than is available to other market participants, who can invest at the federal funds rates. Economically, it should move in tandem with the spread at which banks hold reserves relative to the risk free rate, which is the reserve spread in our model. Magnitude wise, the reserve spread and the IOER-fed funds spread are in the same ballpark. The IOER-fed funds spread is lower likely because the federal funds rate also provides a small amount of convenience premium relative to the theoretical risk-free rate in our model. Note that the comovement and resemblance are not assumed nor mechnical. Our model is identified from cross-sectional variation in how banks respond to natural

[17] We calculate the IOER-fed funds spread as the median spread in December of each year because of year-end volatility in the federal funds market.

disaster shocks and Bartik shocks to deposit demand. No data directly from the implementation of QE or on IOER is used in our estimation.

Policy Implications While the focus of our model and estimation is on the banking system's response to an increase in reserve resully, it has some qualitative policy implications. We propose two ways to reduce the crowding out of loans: reducing banks' cost of holding reserves and lowering the proportion of reserves trapped in the banking sector.

To the extent that bank loans to firms are useful for stimulating the economy, the reserve supply channel of QE may detract from QE's goal of providing stimulus. This crowding out of loans could be reduced by lowering banks' cost of holding reserves. One possibility would be through relaxing the SLR, which constrains banks' leverage ratio that increases with reserve holdings financed by deposits or other forms of debt. As a result, this regulation could induce banks to hold fewer loans to offset their increased holding of reserves. Exempting reserves from the calcuation of the SLR as was temporarily implemented following the Covid-19 crisis, may be one way to reduce the crowding out of bank lending. With that said, our cost function captures the overall costs depending on banks' balance sheet composition. Future work may examine bank balance sheets and specific regulatory requirements at a more granular level.

An alternative way to reduce the crowding out of loans we document is for the Federal Reserve to reduce the quantity of reserves that QE forces banks to hold. In particular, the Federeal Reserve could extend reserve access to non-banks, which would reduce the pressure for the banking system to absorb the entire reserve injection. One step in this direction is the Reverse Repo (RPP) Facility that allows money market funds to hold secured deposits with the Federal Reserve. Expanding the size and eligibility of the RRP Facility is one potential approach to preserve bank balance sheet space during future rounds of QE. A more ambitious approach would be to use a Central Bank Digital Currency that could be traded outside of the banking system to pay for assets in future rounds of QE.

6 Conclusion

There has been a large expansion in the amount of central bank reserves outstanding following multiple rounds of QE. This paper develops and estimates a structural model of the U.S. banking system to analyze the effect of an increase in central bank reserve supply on bank lending and deposit taking. Our framework has two key factors that determine the impact of reserve injections on the banking system. The first one is the demand elasticity banks face in their respective deposit and loan markets. The second one is how banks' cost of capital depends on their balance sheet composition, where the effect of reserve holdings on the cost of capital is of particular importance.

One main challenge in estimating our model is that reserve supply increases are endogenus. In particular, reserve supply increased due to QE, which was implemented in response to the 2008 financial crisis and the Covid-19 pandemic. To avoid confounding by the direct effect of these crises, we estimate our structural model only using cross-sectional variation unrelated to QE in the time series.

In our estimated model, the increase in reserve supply from 2008 to 2017 reduces firm loans extended by an average of $353.0 billion, which amounts to 19 cents in bank lending crowded out per dollar of reserves injected. The impact on mortgage lending and deposit taking is more attenuated. Our model-generated reserve spread is similar to the observed IOER-Fed funds spread in the data. Importantly, the reduction in bank lending to firms following reserve increases may counteract the stimulative impacts of QE's asset purchases. This counterproductive effect of the reserve supply channel we document is important to consider when thinking about the design of unconventional monetary policy and bank regulation going forward.

References

Viral V Acharya and Raghuram Rajan. Liquidity, liquidity everywhere, not a drop to use. 2021.

Oktay Akkus, J Anthony Cookson, and Ali Hortacsu. The determinants of bank mergers: A revealed preference analysis. *Management Science*, 62(8):2241–2258, 2016.

Ugo Albertazzi, Lorenzo Burlon, Tomas Jankauskas, and Nicola Pavanini. The shadow value of unconventional monetary policy. 2022.

Timothy J Bartik. Who benefits from state and local economic development policies? 1991.

Steven T Berry. Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262, 1994.

Vitaly M Bord. Bank consolidation and financial inclusion: the adverse effects of bank mergers on depositors. *Unpublished Manuscript, Harvard University*, 2017.

Lorenzo Bretscher, Lukas Schmid, Ishita Sen, and Varun Sharma. Institutional corporate bond demand. *Available at SSRN 3756280*, 2020.

Greg Buchak, Gregor Matvos, Tomasz Piskorski, and Amit Seru. Fintech, regulatory arbitrage, and the rise of shadow banks. *Journal of Financial Economics*, 130(3):453–483, 2018.

Indraneel Chakraborty, Itay Goldstein, and Andrew MacKinlay. Housing price booms and crowding-out effects in bank lending. *The Review of Financial Studies*, 31(7):2806–2853, 2018.

Indraneel Chakraborty, Itay Goldstein, and Andrew MacKinlay. Monetary stimulus and bank lending. *Journal of Financial Economics*, 136(1):189–218, 2020.

Jens HE Christensen and Signe Krogstrup. Transmission of quantitative easing: The role of central bank reserves. *The Economic Journal*, 129(617):249–272, 2019.

Kristle Romero Cortés and Philip E Strahan. Tracing out capital flows: How financially integrated banks respond to natural disasters. *Journal of Financial Economics*, 125(1):182–199, 2017.

Ben R Craig and Yiming Ma. Intermediation in the interbank lending market. *Available at SSRN 3432001*, 2018.

Douglas W Diamond and Philip H Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of political economy*, 91(3):401–419, 1983.

Douglas W Diamond and Raghuram G Rajan. A theory of bank capital. *the Journal of Finance*, 55(6):2431–2465, 2000.

William Diamond. Safety transformation and the structure of the financial system. *Available at SSRN 3219332*, 2019.

Itamar Drechsler, Alexi Savov, and Philipp Schnabl. The deposits channel of monetary policy. *The Quarterly Journal of Economics*, 132(4):1819–1876, 2017.

Mark Egan, Ali Hortaçsu, and Gregor Matvos. Deposit competition and financial fragility: Evidence from the us banking sector. *American Economic Review*, 107(1):169–216, 2017.

Vadim Elenev, Tim Landvoigt, Patrick J Shultz, and Stijn Van Nieuwerburgh. Can monetary policy create fiscal capacity? Technical report, National Bureau of Economic Research, 2021.

Joseph Gagnon, Matthew Raskin, Julie Remache, and Brian P Sack. Large-scale asset purchases by the federal reserve: did they work? *FRB of New York Staff Report*, (441), 2010.

Samuel G Hanson, Andrei Shleifer, Jeremy C Stein, and Robert W Vishny. Banks as patient fixed-income investors. *Journal of Financial Economics*, 117(3):449–469, 2015.

Zhengyang Jiang, Robert Richmond, and Tony Zhang. A portfolio approach to global imbalances. *Available at SSRN*, 2020.

John Kandrac and Bernd Schlusche. Quantitative easing and bank risk taking: evidence from lending. *Journal of Money, Credit and Banking*, 53(4):635–676, 2021.

Anil K Kashyap, Raghuram Rajan, and Jeremy C Stein. Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. *The Journal of finance*, 57(1):33–73, 2002.

Ralph SJ Koijen and Motohiro Yogo. A demand system approach to asset pricing. *Journal of Political Economy*, 127(4):1475–1515, 2019.

Ralph SJ Koijen and Motohiro Yogo. Exchange rates and asset prices in a global demand system. Technical report, National Bureau of Economic Research, 2020.

Ralph SJ Koijen, Robert J Richmond, and Motohiro Yogo. Which investors matter for equity valuations and expected returns? Technical report, National Bureau of Economic Research, 2020.

Ralph SJ Koijen, François Koulischer, Benoît Nguyen, and Motohiro Yogo. Inspecting the mechanism of quantitative easing in the euro area. *Journal of Financial Economics*, 140(1):1–20, 2021.

Arvind Krishnamurthy and Annette Vissing-Jorgensen. The effects of quantitative easing on interest rates: channels and implications for policy. Technical report, National Bureau of Economic Research, 2011.

Wenhao Li, Yiming Ma, and Yang Zhao. The passthrough of treasury supply to bank deposit funding. *Columbia Business School Research Paper, Forthcoming*, 2019.

Daniel McFadden. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, pages 105–142, 1974.

Alexander Rodnyansky and Olivier M Darmouni. The effects of quantitative easing on bank lending behavior. *The Review of Financial Studies*, 30(11):3858–3887, 2017.

David Scharfstein and Adi Sunderam. Market power in mortgage lending and the transmission of monetary policy. *Unpublished working paper. Harvard University*, 2, 2016.

Michael Schwert. Bank capital and lending relationships. *The Journal of Finance*, 73(2):787–830, 2018.

Yifei Wang, Toni M Whited, Yufeng Wu, and Kairong Xiao. Bank market power and monetary policy transmission: Evidence from a structural estimation. Technical report, National Bureau of Economic Research, 2020.

Kairong Xiao. Monetary transmission through shadow banks. *The Review of Financial Studies*, 33 (6):2379–2420, 2020.

Appendix

A Additional Proofs and Derivations

A.1 Derivation of Equations (6)-(8) in Section 2

Taking the derivatives on the left hand side of equations (2) to (4) yields

$$\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}(R_t^{D,m} - R_{D,nmt}) - Q_{D,nmt} = \frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}}\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}, \tag{A1}$$

$$\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}(R_{M,nmt} - R_t^{M,m}) + Q_{M,nmt} = \frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}, \tag{A2}$$

$$\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}(R_{L,nmt} - R_t^{L,m}) + Q_{L,nmt} = \frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}, \tag{A3}$$

$$R_{S,t} - R_t^{S,m} = \frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}. \tag{A4}$$

Dividing equations (A1)-(A3) respectively by $\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}$, $\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}$, and $\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}$ yields

$$R_t^{D,m} - R_{D,nmt} - \frac{Q_{D,nmt}}{\partial Q_{D,nmt}/\partial R_{D,nmt}} = \frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}}, \tag{A5}$$

$$R_{M,nmt} - R_t^{M,m} + \frac{Q_{M,nmt}}{\partial Q_{M,nmt}/\partial R_{M,nmt}} = \frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}, \tag{A6}$$

$$R_{L,nmt} - R_t^{L,m} + \frac{Q_{L,nmt}}{\partial Q_{L,nmt}/\partial R_{L,nmt}} = \frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}, \tag{A7}$$

$$R_{S,t} - R_t^{S,m} = \frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}. \tag{A8}$$

If we take the left hand side of equations (A5) - (A7) as a function respectively of the quantities $Q_{D,nmt}$, $Q_{M,nmt}$ and $Q_{L,nmt}$, we can implicitly differentiate this system of equations to see how the bank responds to an exogenous increase in its its security holdings $Q_{S,mt}$. If we differentiate this system with respect to $Q_{S,mt}$, we obtain equations (6)- (8) in the main text.

A.2 Detailed Derivations for Section 3

A.2.1 Explicit Expression for $\psi^o_{D,nt}$

Using the expressions for $Q_{D,nmt}$ in equation (11) and for $\overline{Q}_{D,nt}$ in equation (16) yields

$$\frac{Q_{D,nmt}}{\overline{Q}_{D,nt}} = \frac{\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{\sum_{m'} \exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})}. \tag{A9}$$

With the definition of $\psi_{D,nt}$ in equation (17), this becomes

$$\log \frac{Q_{D,nmt}}{\overline{Q}_{D,nt}} = \alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt} - \psi_{D,nt}. \tag{A10}$$

Averaging this expression across the N_{nt} different banks m in market n at time t yields

$$\frac{1}{N_{nt}} \sum_m \log \frac{Q_{D,nmt}}{\overline{Q}_{D,nt}} = \frac{1}{N_{nt}} \sum_m (\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D) - \psi^o_{D,nt}, \tag{A11}$$

since the market-specific mean of $\delta_{D,nmt}$ is $\psi^u_{D,nt}$ and $\psi^o_{D,nt} + \psi^u_{D,nt} = \psi_{D,nt}$. This yields the expression we use for $\psi^o_{D,nt}$

$$\psi^o_{D,nt} = \frac{1}{N_{nt}} \sum_m (\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D) - \frac{1}{N_{nt}} \sum_m \log \frac{Q_{D,nmt}}{\overline{Q}_{D,nt}}. \tag{A12}$$

A.2.2 Derivation of Individual Bank and Market-Level Demand Curves

From equations (11) and (16) we have that

$$
\begin{aligned}
Q_{D,nmt} &= \overline{Q}_{D,nt}\frac{\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{\sum_{m'}\exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})} && \text{(A13)}\\
\log Q_{D,nmt} &= \log \overline{Q}_{D,nt} + \log \frac{\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{\sum_{m'}\exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})} && \text{(A14)}\\
&= \log \overline{Q}_{D,nt} + \alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt} - \psi^o_{D,nt}. && \text{(A15)}
\end{aligned}
$$

This implies that an individual bank's demand curve is given by

$$
\begin{aligned}
\frac{\partial \log Q_{D,nmt}}{\partial R_{D,nmt}} &= \alpha_D + \frac{\partial \log \overline{Q}_{D,nt}}{\partial R_{D,nmt}} - \frac{\partial \psi^o_{D,nt}}{\partial R_{D,nmt}} && \text{(A16)}\\
&= \alpha_D + \left(\frac{\partial \log \overline{Q}_{D,nt}}{\partial \psi^o_{D,nt}} - 1\right)\frac{\partial \psi^o_{D,nt}}{\partial R_{D,nmt}} && \text{(A17)}\\
&= \alpha_D + \alpha_D\left(\frac{\partial \log \overline{Q}_{D,nt}}{\partial \psi^o_{D,nt}} - 1\right)\frac{Q_{D,nmt}}{\overline{Q}_{D,nt}} && \text{(A18)}\\
&= \alpha_D + \alpha_D(\beta_{D,o} - 1)\frac{Q_{D,nmt}}{\overline{Q}_{D,nt}}. && \text{(A19)}
\end{aligned}
$$

In equation (A19), we apply our log-linear approximation in equation (18) which yields $\beta_{D,o} = \frac{\partial \log \overline{Q}_{D,nt}}{\partial \psi^o_{D,nt}}$.

This same log-linear approximation also allows us to derive an expression for the impact of an individual bank's rates on market-level quantities:

$$
\begin{aligned}
\log \overline{Q}_{D,nt} &= \log \overline{F}_{D,nt} + \beta_{D,o}(\psi^o_{D,nt} + \psi^u_{D,nt}) && \text{(A20)}\\
\frac{\partial \log \overline{Q}_{D,nt}}{\partial R_{D,nmt}} &= \beta_{D,o}\frac{\partial \psi^o_{D,nt}}{\partial R_{D,nmt}} = \frac{1}{N_{nt}}\alpha_D\beta_{D,o} && \text{(A21)}
\end{aligned}
$$

where in the last equality we use the expression for $\psi^o_{D,nt}$ in equation (A12). Summing this across all N_{nt} banks in the market gives an expression for how total quantites respond when all banks raise their rates:

$$
\frac{\partial \log \overline{Q}_{D,nt}}{\partial R_{D,nt}} = \alpha_D\beta_{D,o} \qquad \text{(A22)}
$$

A.3 Cost Function Estimation

This appendix shows how to use our regression results from estimating equations (33)-(35) and (37)-(42) to identify the bank's cost function. These regression results tell us that a one unit change in instrument z^i_{mt} changes bank deposit, mortgage, loan, and security quantities respectively by $\gamma^{i,D}$, $\gamma^{i,M}$, $\gamma^{i,L}$, $\gamma^{i,s}$. This same one unit change in instrument z^i_{mt} results in changes in the marginal costs of providing deposits, loans, and securities of $\kappa^{i,D}, \kappa^{i,M}, \kappa^{i,L}$. Because our cost function specification only depends on the deposit quantity $Q_{D,mt}$ and the aggregates $\mathcal{E}_{mt} = Q_{M,mt} + Q_{L,mt} + Q_{S,mt} - Q_{D,mt}$ and $\mathcal{I}_{mt} = Q_{S,mt} + \omega_M Q_{M,mt} + \omega_L Q_{L,mt}$, it is useful to define $\gamma^{i,\mathcal{E}} = \gamma^{i,Q} + \gamma^{i,M} + \gamma^{i,L} - \gamma^{i,D}$ and $\gamma^{i,\mathcal{I}} = \gamma^{i,Q} + \omega_M \gamma^{i,M} + \omega_L \gamma^{i,L}$. The terms $\gamma^{i,\mathcal{E}}$ and $\gamma^{i,\mathcal{I}}$ tells us how much $\mathcal{E}_{mt}$ and $\mathcal{I}_{mt}$ are impacted by a one unit change in the instrument z^i_{mt}.

Using the expressions for a bank's marginal costs in equations (26)-(29), these regression coefficients must satisfy

$$\kappa^{i,D} = -K_1\gamma^{i,\mathcal{E}} + K_3\gamma^{i,D} + K_4\gamma^{i,\mathcal{I}} + K_5[\gamma^{i,\mathcal{E}} - \gamma^{i,D}] \tag{A23}$$

$$\kappa^{i,M} = K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}}\omega_M + K_4\gamma^{i,D}\omega_M + K_5\gamma^{i,D} \tag{A24}$$

$$\kappa^{i,L} = K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}}\omega_L + K_4\gamma^{i,D}\omega_L + K_5\gamma^{i,D} \tag{A25}$$

$$0 = K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}} + K_4\gamma^{i,D} + K_5\gamma^{i,D}, \tag{A26}$$

where every marginal cost in equations (26)-(29) is replaced by its associated κ variable and every quantity is replaced by its associated γ variable. The zero on the left hand size of equation (A26) reflects the fact that securities trade in a competitive market, so in equilibrium there cannot be any cross-sectional dispersion in the marginal cost of holding securities across banks.

Using two instruments, we have two sets of these equations, or 8 equations in total. However, the rank of this system of equations is only 7, so we are only able to identify a 7-parameter cost funtion[18]. To estimate the 7 parameters (K_1 through K_5 and ω_L and ω_M), we minimize the sum of squared deviations between the left and right hand size of Eq. (A23) through (A26). To do so, we first pick a value of (ω_L, ω_M) and solve for $K_1, \ldots, K_5$ in an inner loop. With (ω_L, ω_M) fixed, choosing $K_1, \ldots K_5$ to minimize our loss function is a least squares solution of 8 linear equations with 5 unknowns, for which there is a closed-form solution. We then globally search over the parameters (ω_L, ω_M) to find the value of all 7 parameters which jointly minimize our loss function. We verified that there is a unique set of parameter values for which our loss function is minimized,

[18]To see this rank deficiency, note that our model implies $\omega_L = 1 + \frac{\kappa^{i,L}}{\kappa^{i,M}}(\omega_M - 1)$ for the values from both of the instruments $i = 1, 2$, which cannot be satisfied simultaneously. This rank deficiency is not specific to the functional form of our cost function. It is a consequence of the fact that the Hessian of any cost function is a symmetric matrix.

which we report in the main text and use to compute the implied Hessian matrix H.

Online Appendix

This document contains additional theoretical and empirical results. It contains two sections on (A) additional empirical results, (B) an extended demand system of bank loans with relationship banking, (C) a microfoundation for the log-linear modification of our logit demand system, (D) computation details for our counterfactuals, and (E) an extension of our model to infinite horizon.

A Loan Outside Option

Table OA1: Outside Option estimates (Loans)

This table reports the outside option size for loans in trillions of dollars. The implied β_o is obtained using equation (OA4) below.

Year	Size of Outside Option	Implied β_o
2001	0.32	0.50
2002	0.32	0.53
2003	0.34	0.56
2004	0.29	0.42
2005	0.29	0.36
2006	0.31	0.33
2007	0.36	0.33
2008	0.56	0.63
2009	0.65	0.76
2010	0.53	0.59
2011	0.40	0.41
2012	0.43	0.46
2013	0.37	0.34
2014	0.40	0.37
2015	0.48	0.44
2016	0.50	0.43
2017	0.49	0.37
2018	0.46	0.35

The estimate of the implied β_o can be inferred from knowing the overall market size $\overline{F}_{D,nt}$, which is possible once the outside option is observed. The total quantity of deposits and total market size are related by

$$\overline{Q}_{D,nt} = \overline{F}_{D,nt} \frac{\exp(\psi_{D,nt})}{1+\exp(\psi_{D,nt})}. \tag{OA1}$$

Equation (OA1) can be used to solve for $\psi_{D,nt}$ from observed data. Moreover, in this exact model,

$$\frac{\partial log(\overline{Q}_{D,nt})}{\partial \psi_{D,nt}} = \psi_{D,nt} - \frac{\exp(\psi_{D,nt}}{1+\psi_{D,nt}} = \frac{1}{1+exp(\psi_{D,nt})} \tag{OA2}$$

while in our log-linear model

$$\log \overline{Q}_{D,nt} \approx \log \overline{F}_{D,nt} + \beta_{D,o}\psi_{D,nt} \tag{OA3}$$

which implies $\frac{\partial log(\overline{Q}_{D,nt})}{\partial \psi_{D,nt}} = \beta_{D,o}$.

Our approximation (derived here for deposits D but also identically for loans L) is therefore

$$\beta_{D,o} \;\; = \;\; \frac{1}{1+exp(\psi_{D,nt})}. \tag{OA4}$$

B A Model of Relationship-Based Bank Loans

This appendix presents a modification of our logit demand system that allows us to exploit data on firm-bank relationships in the DealScan dataset. In the main text, we relied on a proxy for the number of firms in the market that did not borrow from a bank to infer the overall elasticity of loan demand. In contrast, we were able to estimate this aggregate elasticity of demand directly from the data for mortgages and deposits. Exploiting data on firm-bank relationships in DealScan allows us to estimate the aggregate elasticity of loan demand in a manner somewhat similar to our approach for deposits and mortgages.

Our modified model is identical to a logit demand system except that firms get extra utility k from borrowing from a bank they previously borrowed from. At time t, firm i in market n has the choice to borrow from banks indexed by m. We assume each firm i has the same expected amount $F_{L,nt}$ as all other firms in the market. Firm i gets utility

$$\alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + \delta_{L,nmt} + kl_{i,mt} + \epsilon_{i,nmt} \tag{OA5}$$

from choosing to borrow from bank m and selects its bank to maximize utility, where $kl_{i,mt}$ is the new term that describes the utility from prior relationships: if this firm borrowed from this bank in the past, then, $l_{i,mt} = 1$; otherwise $l_{i,mt} = 0$. We expect $k > 0$ so that firms prefer to borrow from banks with prior relationships. $\epsilon_{i,nmt}$ are standard logit draws, i.i.d. across firms. Firms also have an outside option yielding utility 0 from not borrowing. The probability firm i borrows from bank m is

$$\frac{\exp(\alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + \delta_{L,nmt} + kl_{i,mt})}{1 + \sum_{m'} \exp(\alpha_L R_{L,nm't} + X_{L,nm't}\beta_L + \delta_{L,nm't} + kl_{i,m't})}, \tag{OA6}$$

and the expected amount it borrows from this bank is

$$Q_{L,i,nmt} = F_{L,nt}\frac{\exp(\alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + \delta_{L,nmt} + kl_{i,mt})}{1 + \sum_{m'} \exp(\alpha_L R_{L,nm't} + X_{L,nm't}\beta_L + \delta_{L,nm't} + kl_{i,m't})}. \tag{OA7}$$

The crucial feature of this modified demand system is that if firms prefer to borrow from the same banks the have previously, a firm's total borrowing quantity is more sensitive to the rates charges by its relationship banks than the rates charge by other banks. That is, if $k > 0$, $\frac{\partial Q_{L,i,nmt}}{\partial R_{L,nmt}}$ is larger if $l_{i,m't} = 1$ than if $l_{i,m't} = 0$. Once we have estimated k, we can infer the aggregate elasticity of loan demand by observing how a shock to a bank's rates impacts the borrowing quantities of its relationship firms relative to the borrowing quantities of other firms.

B.1 Estimating the price disutility parameter α_L from new borrowers

We first consider the borrowing decisions of those firms which have no previous banking relationships. If there a large number N of such firms, then (by the law of large numbers) the amount borrowed from bank m by firms with no previous borrowing relationships is

$$Q^{new}_{nmt} = NF_{L,i,nt}\frac{\exp(\alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + \delta_{L,nmt})}{1+\sum_{m'}\exp(\alpha_L R_{L,nm't} + X_{L,nm't}\beta_L + \delta_{L,nm't})}. \tag{OA8}$$

Taking the log of equation (OA8) yields

$$\log(Q^{new}_{nmt}) - \log(Q^{new}_{nm't}) = \alpha_L(R_{L,nmt} + R_{L,nm't}) + (X_{L,nmt} - X_{L,nm't})\beta_L + (\delta_{L,nmt} - \delta_{L,nm't}). \tag{OA9}$$

This log-linear expression allows us to estimate the demand parameters of our model exactly in the same manner as in our baseline setting, which is a standard logit demand system. The only difference is that here we use the quantities of borrowing by the subset of firms with no previous relationships. Because the latent demand term $\delta_{L,nmt}$ may be correlated with the lending rate, we use the two stage least squares specification (with the same instrument following Cortés and Strahan (2017) as in the main text)

$$\begin{aligned}
\log(Q^{new}_{nmt}) &= (\zeta_{L,nt} + E_{L,nt}\delta_{L,nmt}) + \alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + (\delta_{L,nmt} - E_{L,nt}\delta_{L,nmt}) \\
R_{L,nmt} &= \gamma_{L,nt} + \gamma_L z_{L,nmt} + X_{L,nmt}\gamma_D + e_{L,nmt}.
\end{aligned}$$

This provides us with a consistent estimate of α_L. Table OA2 reports the result from this regression. The price disutility parameter $\alpha_l = -489.5$ is very similar to our main text estimate $\alpha_L = -487.3$.

B.2 Estimating the relationship stickiness parameter k

Having estimated α_L and β_L, we next infer how much firms value borrowing from a bank with which they have a past relationship. We do so by comparing the distribution of borrowing by new firms to the borrowing of firms that have past relationships. Because the unobserved latent demand δ shows up in the borrowing choices of all firms, we can use data on the borrowing decisions of new firms to control for the endogeneity problem.

It is useful to classify firms by their "relationship vector" l, which equals 1 for every bank m the firm has previously borrowed from and equals 0 otherwise. For such a firm with relationship vector l, its probability of borrowing from a bank with which it already has a relationship conditional on

Table OA2: Demand System Estimates

This table reports the two-stage least squares results for the estimating price disutility parameter α_L of our loan demand system. These regressions are run at the market-bank-year level, with a control for loan loss provision as in the main text. The sample period is from 2001 to 2017. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	Dependent variable:
	Loan Market Share
Rate (with IV)	-489.46^{***}
	(117.14)
Loan Loss Provision	27.98^{**}
	(10.20)
Observations	13,184

borrowing at all is

$$\frac{\sum_{m\in l}\exp(\alpha_L R_{L,nmt}+X_{L,nmt}\beta_L+\delta_{L,nmt}+k)}{\sum_{m'}\exp(\alpha_L R_{L,nm't}+X_{L,nm't}\beta_L+\delta_{L,nm't}+k1_{m'\in l})}=\frac{\sum_{m\in l}\exp(\log\sum_n Q^{new}_{nm't}+k)}{\sum_{m'}\exp(\log\sum_n Q^{new}_{nm't}+k1_{m'\in l})}, \tag{OA10}$$

where $\sum_n Q^{new}_{nm't}$ is the aggregate quantity of lending from bank m' to only new firms with no previous relationships. This equation follows from equation (OA9).

The ratio in equation (OA9) can be compared to the observed data $P_{lt|borrow}$, the fraction of observed borrowing by firms with relationship vector l that is from banks the firms have previous relationships with. Choosing the parameter k to make $P_{lt|borrow}-\frac{\sum_{m\in l}\exp(\log\sum_n Q^{new}_{nm't}+k)}{\sum_{m'}\exp(\log\sum_n Q^{new}_{nm't}+k1_{m'\in l})}$ as close as possible to 0 provides an estimate of k. We use this to construct a moment condition that aggregates across all possible relationship vectors, weighted by their total loan quantities:

$$\sum_l Q_{l,rel}\left[P_{lt|borrow}-\frac{\sum_{m\in l}\exp(\log\sum_n Q^{new}_{nm't}+k)}{\sum_{m'}\exp(\log\sum_n Q^{new}_{nm't}+k1_{m'\in l})}\right]=0. \tag{OA11}$$

Solving equation (OA11) yields our estimate of k. Our solution is $k=2.27$, which means that holding other characteristics constant, firms are $\exp(k)=9.66$ times more likely to borrow from banks with past relationships relative to the decisions of new borrowers.

B.3 Estimating the aggregate elasticity of loan demand

To estimate the aggregate elasticity of loan demand, we need to observe how the quantity that a firm borrows is impacted by shocks to the supply of credit. For a firm with relationship vector l, the probability that it borrows from a bank instead of choosing the outside option is

$$\frac{\sum_{m'} \exp(\alpha_L R_{L,nm't} + X_{L,nm't}\beta_L + \delta_{L,nm't} + kl_{i,m't})}{1 + \sum_{m'} \exp(\alpha_L R_{L,nm't} + X_{L,nm't}\beta_L + \delta_{L,nm't} + kl_{i,m't})} = H(\psi^{i,nt})$$

where $H(x) = \exp(x)/(1 + \exp(x))$ and $\psi^{i,nt} = \log(\sum_{m'} \exp(\alpha_L R_{L,nm't} + X_{L,nm't}\beta_L + \delta_{L,nm't} + kl_{i,m't}))$. If k were equal to 0, then $\psi^{i,nt}$ would not depend on the firm i and would be identical to $\psi_{D,nt}$ in the main text. As for $\psi_{D,nt}$ in the main text, every term in $\psi^{i,nt}$ except for the mean of $\delta_{L,nm't}$, so we define

$$\psi^u_{i,nt} = \psi^u_{nt} = \frac{1}{N_{nt}} \sum_m \delta_{L,nmt}. \tag{OA12}$$

$$\psi^o_{i,nt} = \log\left(\sum_m \exp(\alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + \delta_{L,nmt} - \psi^u_{i,nt} + kl_{i,mt})\right). \tag{OA13}$$

Here, $\psi^o_{i,nt}$ is observable but ψ^u_{nt} is not and $\psi^o_{i,nt} + \psi^u_{nt} = \psi_{i,nt}$. The expected quantity $\sum_{i\in l} Q_{i,t}$ borrowed by a firm with relationship vector l is

$$\log(\sum_{i\in l} Q_{i,t}) = \log(F_{L,nt}) + \log(H(\psi^o_{i,nt} + \psi^u_{i,nt})) \approx \log(F_{L,nt}) + \beta_{L,o}(\psi^o_{i,nt} + \psi^u_{nt})$$

after log-linearizing.

We can estimate $\beta_{L,o}$ with a firm level instrument $z_{i,nt}$ by

$$\log(\sum_{i\in l} Q_{i,t}) = \log(F_{L,nt}) + \beta(\psi^o_{i,nt} + \psi^u_{nt}) + \epsilon_{int} \tag{OA14}$$

$$\psi^o_{i,nt} = \kappa_o + \lambda_o z_{i,nt} + \eta_{int}. \tag{OA15}$$

To construct our instrument, we take for each firm the average amount of natural disaster damage done in regions where its relationship banks have branches.

The above expressions depend on knowing observed value of $\psi^o_{i,nt}$, which were first initially inferred from looking at the quantity of borrowing from new firms. While $\psi^o_{i,nt}$, can be inferred from observed data, this step has to be done jointly with the estimate of $\beta_{L,o}$.

The fraction of observed borrowing, under our log-linear approximation to the function $H()$ that goes to bank m is

$$\frac{\sum_i Q_{L,i,nmt}}{\sum_m \sum_i Q_{L,i,nmt}} = \frac{\sum_i \frac{\exp(\alpha_L R_{L,nmt}+X_{L,nmt}\beta_L+\delta_{L,nmt}+kl_{i,mt})}{\sum_{m'}\exp(\alpha_L R_{L,nm't}+X_{L,nm't}\beta_L+\delta_{L,nm't}+kl_{i,mt})}\exp(\beta_{L,o}(\psi^o_{i,nt}+\psi^u_{i,nt}))}{\sum_i \sum_m \frac{\exp(\alpha_L R_{L,nmt}+X_{L,nmt}\beta_L+\delta_{L,nmt}+kl_{i,mt})}{\sum_{m'}\exp(\alpha_L R_{L,nm't}+X_{L,nm't}\beta_L+\delta_{L,nm't}+kl_{i,mt})}\exp(\beta_{L,o}(\psi^o_{i,nt}+\psi^u_{i,nt}))}. \quad \text{(OA16)}$$

This expression depends on the unobserved variable $\beta_{L,o}$ that we aim to estimate. For a postulated value of $\beta_{L,o}$, we solve this system of equations given by equation (OA16) for the values of $\delta_{L,nmt}$. There is one fewer equation than unknown, so we solve for the $\delta_{L,nmt}$ up to the value of their unknown mean. We then use these values to construct the variable $\psi^o_{i,nt}$ (which does not depend on the mean of the $\delta_{L,nmt}$'s) and run the 2 stage least square regressions (OA14) and (OA15) above to get a new estimate of $\beta_{L,o}$. We iterate this procedure until $\beta_{L,o}$ reaches a fixed point.

The result is reported in Table OA3. We find $\beta = 0.30$. Compared to Table 4 in the main text, this estimate is close to the implied value of 0.33 as in the year of 2007, which is the input for our counterfactual analysis.

Table OA3: Outside Option Estimates

This table reports the two-stage least squares results for estimating the outside option parameter $\beta_{L,o}$ in the loan market. The sample period is from 2001 to 2017.

$\beta_{L,o}$	0.30
Observations	44,529

Finally, we use our parameters to compute the implied aggregate elasticity of loan demand. For this, we need an expression for how the quantity of total loans varies when every bank raises its loan rates by the same amount. We will denote the total quantity borrowed by Q and a change in every bank's rate a derivative with respect to R.

Equation (OA14) implies that the total quantity of borrowing is

$$Q = \sum_l N_l F_{L,nt} \exp(\beta_{L,o}(\psi^o_{i,nt} + \psi^u_{nt})). \quad \text{(OA17)}$$

Note that

$$\frac{\partial \psi^o_{i,nt}}{\partial R} = \frac{\left(\alpha_L \sum_m \exp(\alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + \delta_{L,nmt} - \psi^u_{i,nt} + kl_{i,mt})\right)}{\left(\sum_m \exp(\alpha_L R_{L,nmt} + X_{L,nmt}\beta_L + \delta_{L,nmt} - \psi^u_{i,nt} + kl_{i,mt})\right)} = \alpha_L \quad \text{(OA18)}$$

is the same for all firms i. The derivative of the borrowing quantity in equation (OA17) with respect to an equal increase in the rates of every bank is

$$\frac{dQ}{dR} = \sum_l [N_l F_{L,nt} exp(\beta_{L,o}(\psi^o_{i,nt} + \psi^u_{nt}))]\beta_{L,o}\frac{\partial \psi^o_{i,nt}}{\partial R} = Q\beta_{L,o}\alpha_L \tag{OA19}$$

It follows that $\frac{dlogQ}{dR} = \alpha_L \beta_{L,o}$ just like in the demand systems in the main text. Plugging in our estimates $\alpha_L = -489, \beta_{L,o} = .30$, we get that a 10 basis point increase in all bank rates leads to a decline in total loan quantities of $.30 \times 48.9\% = 14.7\%$. This is close to the main text estimate of $0.33 \times 48.7\% = 16.1\%$ using our main text estimate of $\alpha_L = -487$ and the year 2007 outside option estimate of 0.33 on which we based our benchmark counterfactual.

C Microfoundation for Bank's Cost Function

This section presents an economic model that yields the bank's cost function in equation (25). We first rewrite this cost function as

$$\begin{aligned} H \quad & (Q_{D,mt}, Q_{M,mt}, Q_{L,mt}, Q_{S,mt}) = \mu_D Q_{D,mt} + \mu_M Q_{M,mt} + \mu_L Q_{L,mt} + \mu_Q Q_{S,mt} \\ + \quad & \frac{1}{2}(K_1(\mathcal{A}_{mt} - Q_{D,mt})^2 + K_2 \mathcal{I}^2_{mt} + K_3 Q^2_{D,mt} + 2K_4 \mathcal{I}_{mt} Q_{D,mt} + 2K_5(\mathcal{A}_{mt} - Q_{D,mt})Q_{D,mt}), \end{aligned} \tag{OA20}$$

where $A_{mt} = Q_{M,mt} + Q_{L,mt} + Q_{S,mt} = \mathcal{E}_{mt} + Q_{D,mt}$. The quadratic terms of H in equation (OA20) can be rewritten as

$$\begin{aligned} & \frac{1}{2}(K_1[\mathcal{A}^2_{mt} - 2A_{mt}Q_{D,mt} + Q^2_{D,mt}] + K_2 \mathcal{I}^2_{mt} + K_3 Q^2_{D,mt} \\ & +2K_4 \mathcal{I}_{mt} Q_{D,mt} + 2K_5(\mathcal{A}_{mt} Q_{D,mt} - Q^2_{D,mt})) \\ = \quad & \frac{1}{2}(K_1 \mathcal{A}^2_{mt} + K_2 \mathcal{I}^2_{mt} + (K_1 + K_3 - 2K_5)Q^2_{D,mt} + 2K_4 \mathcal{I}_{mt} Q_{D,mt} + 2(K_5 - K_1)\mathcal{A}_{mt} Q_{D,mt}) \end{aligned} \tag{OA21}$$

According to our estimates in table 5 , we have $K_1 > 0, K_2 > 0, K_1 + K_3 - 2K_5 > 0, K_4 > 0, 2(K_5 - K_1) < 0$.

We present a simple microfoundation for this cost function based on the need of banks to manage liquidity risk. For simplicity, we abstract from costs due to regulation even though these likely play an important role in practice. First, banks face a standard quadratic cost of issuing equity ex ante before shocks are realized. This cost is, for some constant $Q_0 > 0$,

$$Q_0(\mathcal{A}_{mt} - Q_{D,mt})^2 = Q_0 \mathcal{A}^2_{mt} - 2Q_0 \mathcal{A}_{mt} Q_{D,mt} + Q_0 Q^2_{D,mt}. \tag{OA22}$$

Second, following Kashyap et al. (2002), banks jointly face the risk of outflows on both the asset and liability sides of their balance sheets. This is motivated by the fact that depositors can always withdraw from a bank, and borrowers with revolving lines of credit can demand cash at any time as well. In addition, term loans could potentially be renegotiated as well. For simplicity, we assume that in a withdrawal event, banks have a total funding outflow of $(k_1\mathcal{A}_{mt} - k_2 Q_{D,mt})$. Banks have again a quadratic cost of raising new funds to meet this redemption, but with a different constant Q_1 reflecting the fact that raising funds in a redemption crisis pay be particularly expensive. This yields a cost of

$$Q_1(k_1^2\mathcal{A}_{mt}^2 + k_2^2 Q_{D,mt}^2 - 2k_1k_2\mathcal{A}_{mt}Q_{D,mt}). \tag{OA23}$$

Third, the bank faces the risk of an identical outflow shock, except that the outflows on its asset side are proportional to its liquidity weighted assets I_{mt} instead of its overall assets A_{mt}. Here, the bank faces a quadratic cost of raising new funds afther the redemption but with a new constant Q_2. The total withdrawal is now $(k_3\mathcal{I}_{mt} - k_4 Q_{D,mt})$ resulting in a total cost of

$$Q_2(k_3^2\mathcal{I}_{mt}^2 + k_4^2 Q_{D,mt}^2 - 2k_3k_4\mathcal{I}_{mt}Q_{D,mt}). \tag{OA24}$$

Summing up equations (OA22) -(OA24), we get the bank's cost function

$$(Q_0+Q_1k_1^2)\mathcal{A}_{mt}^2+Q_2k_3^2I_{mt}^2+(Q_0+Q_1k_2^2+Q_2k_4^2)Q_{D,mt}^2-2Q_2k_3k_4Q_{D,mt}I_{mt}-(2Q_0+2k_1k_2)\mathcal{A}_{mt}Q_{D,mt}. \tag{OA25}$$

Note that depending on whether k_1 has the same sign as k_2 and whether k_3 has the same sign as k_4, the coefficients on $Q_{D,mt}I_{mt}$ and $\mathcal{A}_{mt}Q_{D,mt}$ can be either positive or negative. This reflects the fact that in some financial crises, such as 2008, banks faced drawdowns on their credit lines while also facing a large inflow of deposits. It is also possible, however, for a bank to have a run where depositors and creditors all want cash at the same time. However, the first three terms for $\mathcal{A}_{mt}^2, I_{mt}^2, Q_{D,mt}^2$ must be positive (as we find in our estimates). This expression is therefore flexible enough to match our estimated cost function parameters.

D Microfoundation for Log-Linear Modification of Logit Demand System

This appendix provides a specification of consumer preferences that exactly yields the log-linear approximation to a logit demand system we use in the paper's main text (starting with equation (18)). We again use the deposit market as an example. The model is identical to a logit demand system except that there is heterogeneity in how much depositors value the outside good. As in equation (10), depositor j gets from any good $m > 0$ in region n at time t utility

$$u_{D,jnmt} \quad = \quad \alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt} + \varepsilon_{D,jnmt}. \tag{OA26}$$

In addition, there is an outside good m=0 for which our depositor gets utillity

$$u_{jn0t} \quad = \quad \delta_{D,jn0t} + \varepsilon_{D,jn0t}. \tag{OA27}$$

In a standard logit demand system we would have $\delta_{D,jn0t} = 0$ for all depositors. In our modified demand system, $\delta_{D,jn0t}$ varies across depositors and is distributed according to a measure μ which has density f. The realization of $\delta_{D,jn0t}$ is assumed to be independent of all $\varepsilon_{D,jnmt}$. Conditional on a realized $\delta_{D,jn0t}$, our model is a logit demand system where the depositor's probability of choosing good m is

$$\frac{\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{\exp(\delta_{D,jn0t}) + \sum_{m>0}\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}$$

The quantity purchased of good m is, integrating the measure μ over the value of the outside good,

$$Q_{D,nmt} = \overline{F}_{D,nt}\int_{-\infty}^{\infty}\frac{\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{\exp(\delta_{D,jn0t}) + \sum_{m>0}\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}d\mu(\delta_{D,jn0t}) \tag{OA28}$$

This implies

$$\log(Q_{D,nmt}) - \log(Q_{D,nm't}) = \alpha_D(R_{D,nmt} - R_{D,nm't}) + (X_{D,nmt} - X_{D,nmt})\beta_D + (\delta_{D,nmt} - \delta_{D,n't}) \tag{OA29}$$

as in a logit demand system. Equation (OA29) implies that of the funds invested in goods other than the outside option, our demand system yields the same distribution of these funds between the individual goods available. If in addition, our model yields the same total quantity of funds invested as in equation (18), it will exactly match the log-linear approximation to a logit demand system used in the main text. The total quantity of funds invested in our model is, summing equation (OA28) over all $m > 0$ is

$$\overline{Q}_{D,nt} = \overline{F}_{D,nt}\int_{-\infty}^{\infty}\frac{\exp(\psi_{D,nt})}{\exp(\delta_{D,jn0t}) + \exp(\psi_{D,nt})}d\mu(\delta_{D,jn0t}),$$

where $\psi_{D,nt} = \log(\sum_{m>0}\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt}))$.

To exactly match our log-linear approximation for $\overline{Q}_{D,nt}$ in equation (18), we must have that

$$log\left[\int_{-\infty}^{\infty}\frac{\exp(\psi_{D,nt})}{\exp(\delta_{D,jn0t})+\exp(\psi_{D,nt})}d\mu(\delta_{D,jn0t})\right]=\beta_{D,o}\psi_{D,nt},$$

for a constant $0<\beta_{D,o}<1$. We want to choose our distrubtion μ of outside option utilities to yield this expression. That is, we must pick the distribution μ so that

$$\exp(\beta_{D,o}\psi_{D,nt})=\int_{-\infty}^{\infty}\frac{\exp(\psi_{D,nt})}{\exp(\delta_{D,jn0t})+\exp(\psi_{D,nt})}d\mu(\delta_{D,jn0t}) \tag{OA30}$$

Denote $K=\exp(\psi_{D,nt})$ and $k=\exp(\delta_{D,jn0t})$, equation (OA30) can be written as

$$K^{\beta_{D,o}-1}=\int_{0}^{\infty}\frac{d\mu^e(k)}{k+K} \tag{OA31}$$

where μ^e is the measure induced on $\exp(\delta_{D,jn0t})$ by having a measure μ over $\delta_{D,jn0t}$.

As a special case, if $K=1$,

$$1=\int_{0}^{\infty}\frac{d\mu^e(k)}{k+1} \tag{OA32}$$

which implies that for all $K>0$,

$$\int_{0}^{\infty}\frac{K^{1-\beta_{D,o}}d\mu^e(k)}{k+K}=\int_{0}^{\infty}\frac{d\mu^e(k)}{k+1}. \tag{OA33}$$

We first charecterize a class of measures μ^e that satisfy equation (OA33) and then choose one from this class for which $\int_0^\infty \frac{d\mu^e(k)}{k+1}=1$. If we write our measure μ^e in terms of its density f^e, i.e., $d\mu^e(k)=f^e(k)dk$, we have

$$\int_{0}^{\infty}\frac{K^{1-\beta_{D,o}}f^e(k)dk}{k+K}=\int_{0}^{\infty}\frac{f^e(k)dk}{k+1} \tag{OA34}$$

If we u-substitute $u=\frac{k}{K}, du=\frac{1}{K}dk$, we get

$$\int_{0}^{\infty}\frac{K^{1-\beta_{D,o}}Kf^e(Ku)du}{K(u+1)}=\int_{0}^{\infty}\frac{K^{1-\beta_{D,o}}f^e(Ku)du}{(u+1)}=\int_{0}^{\infty}\frac{f^e(k)dk}{k+1}. \tag{OA35}$$

Equation (OA35) holds if $K^{1-\beta_{D,o}}f^e(Ku)$ does not depend on the value of K, which is the case precisely when $f^e(u) = Cu^{\beta_{D,o}-1}$ for some constant C. The constant C is then determined by

$$1 = \int_0^\infty \frac{Ck^{\beta_{D,o}-1}dk}{k+1}, \tag{OA36}$$

noting that the integral $\int_0^\infty \frac{k^{\beta_{D,o}-1}dk}{k+1} \leq \int_0^1 k^{\beta_{D,o}-1}dk + \int_1^\infty k^{\beta_{D,o}-2}dk = \frac{1}{\beta_{D,o}} + \frac{1}{1-\beta_{D,o}}$ is convergent for any $0 < \beta_{D,o} < 1$. This implies that if all depositors have preferences following equations (OA26) and (OA27), with the exponent of the utility they get from the outside option having density

$$f^e(u) = \frac{u^{\beta-1}}{\int_0^\infty \frac{k^{\beta_{D,o}-1}dk}{k+1}}, \tag{OA37}$$

they choose exactly the quantities given by our log-linear approximation to a logit demand system.

E Computational Details for Counterfactual

E.1 Demand Systems under Log-linear Approximation

Each bank m has deposits $Q_{D,nmt}$ in region n at time t. The total quantity of deposits in the region is $\overline{Q}_{D,nt} = \sum_m Q_{D,nmt}$. Let δ_{nmt} denote the desirability of its deposit:

$$\delta_{nmt} = \alpha_D R_{D,nmt} + X_{nmt}\beta_D + \delta_{D,nmt} \tag{OA38}$$

and deposits $Q_{D,nmt}$ can be expressed as

$$Q_{D,nmt} = \overline{Q}_{D,nt}\frac{\exp(\delta_{nmt})}{\sum_{m'}\exp(\delta_{nm't})}. \tag{OA39}$$

Let $\overline{Q}^i_{D,nt}$ and $\delta^{o,i}_{nt}$ denote the actual value in the data (i for initial). Next, we approximate the variation in $\overline{Q}_{D,nt}$ by

$$\frac{\partial \log \overline{Q}_{D,nt}}{\partial \delta^o_{D,nt}} = \beta_o \tag{OA40}$$

which implies that

$$\begin{aligned} \overline{Q}_{D,nt} &= \overline{Q}^i_{D,nt}\exp(\Delta f_{D,nt})\exp(\beta_o(\delta^o_{D,nt} - \delta^{o,i}_{nt})) && \text{(OA41)} \\ &= \overline{Q}^i_{D,nt}\exp(\beta_o(\log\sum_{m'}\exp(\delta_{nm't}) - \log\sum_{m'}\exp(\delta^i_{nm't}))) && \text{(OA42)} \end{aligned}$$

Here we also consider a "demand shock" $\Delta f_{D,nt}$ that increases the total size of the deposit market uniformly.

Then,

$$Q_{D,nmt} = \overline{Q}_{D,nt} \frac{\exp(\delta_{nmt})}{\sum_{m'} \exp(\delta_{nm't})} = \overline{Q}^i_{D,nt} \exp(\Delta f_{D,nt}) \frac{(\sum_{m'} \exp(\delta_{nm't}))^{\beta_o - 1}}{(\sum_{m'} \exp(\delta^i_{nm't}))^{\beta_o}} \exp(\delta_{nmt}). \quad \text{(OA43)}$$

Note that the value of this expression is unchanged if we add a constant to all δ and δ^i variables in region n at time t. We also have the the difference between the δ of any two goods in the same market is the difference in their log quantities sold. It follows that we can simply use $\delta^i_{nmt} = \log(Q^i_{D,nmt})$ to compute it (since $\delta^i_{nmt} - \log(Q^i_{D,nmt})$ is the constant across all goods in each market):

$$\delta_{nmt} = \delta^i_{nmt} + \alpha_D(r_{nmt} - r^i_{nmt}) \quad \text{(OA44)}$$

Under our maintained assumption that only prices and not product qualities change in counterfactuals, we can write $\delta_{nmt} = \delta^i_{nmt} + \alpha(\Delta r_{nmt})$ where $\Delta r_{nmt} = R_{D,nmt} - R^i_{D,nmt}$ is the change in interest rates relative to the pre-counterfactual data. We can therefore write $Q_{D,nmt}$ as

$$Q_{D,nmt} = \overline{Q}^i_{D,nt} \exp(\Delta f_{D,nt}) \frac{(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})))^{\beta_o - 1}}{(\sum_{m'} \exp(\delta^i_{nm't}))^{\beta_o}} \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})). \text{(OA45)}$$

E.2 Marginal Cost from Optimality Condition

The optimal pricing-implied marginal cost comes from the first order condition is

$$R_{D,nmt} = R^D_t - \frac{Q_{D,nmt}(R_{D,nmt})}{Q'_{D,nmt}(R_{D,nmt})} - \frac{\partial C(Q_{D,nmt}(R_{D,nmt}), \ldots)}{\partial Q_{D,nmt}}. \quad \text{(OA46)}$$

Because

$$\begin{aligned} \log(Q_{D,nmt}) &= \log(\overline{Q}^i_{D,nt}) + \Delta f_{D,nt} + (\beta_o - 1)\log(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))) && \text{(OA47)} \\ &- \beta_o \log(\sum_{m'} \exp(\delta^i_{nm't})) + (\delta^i_{nmt} + \alpha(\Delta r_{nmt})). && \text{(OA48)} \end{aligned}$$

we have

$$\frac{\partial \log(Q_{D,nmt})}{\partial \Delta r_{nmt}} = \alpha + \alpha(\beta_o - 1)\frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} \tag{OA49}$$

This implies

$$\frac{\partial C}{\partial Q_{D,nmt}} = R^D_t - \left[\frac{\partial \log(Q_{D,nmt})}{\partial r_{nmt}}\right]^{-1} - R_{D,nmt} \tag{OA50}$$

$$= R^D_t - \left[\alpha + \alpha(\beta_o - 1)\frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right]^{-1} - R_{D,nmt} \tag{OA51}$$

and thus this demand system on its own implies a marginal cost of providing deposits coming from the optimal rate setting first order condition:

$$\begin{aligned} \frac{\partial C}{\partial Q_{D,nmt}} - \frac{\partial C^i}{\partial Q_{D,nmt}} &= \left[\alpha + \frac{\alpha(\beta_o - 1)\exp(\delta^i_{nmt})}{\sum_{m'} \exp(\delta^i_{nm't})}\right]^{-1} \\ &- \left[\alpha + \frac{\alpha(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right]^{-1} - \Delta r_{nmt} \end{aligned} \tag{OA52}$$

E.3 Jacobian of marginal cost from optimality condition

For numerical accuracy, the Jacobian of equation (OA52) is needed. The derivative of this marginal cost is only non-zero with respect to other rates in the same region and time. The change of bank m's marginal cost with respect to bank m^*'s rate is give by

$$\begin{aligned} \frac{\partial}{\partial \Delta r_{nm^*t}} \frac{\partial C}{\partial Q_{D,nmt}} &= \frac{\partial}{\partial r_{nm^*t}}\left(-\left[\alpha + \frac{\alpha(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right]^{-1} - \Delta r_{nmt}\right) \\ &= -\left[1 + \frac{(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right]^{-2} \\ &\cdot \quad (\beta_o - 1)\left(\frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))\exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t}))}{\left(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))\right)^2}\right. \\ &+ \quad \left. 1_{\{m=m^*\}}\frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right) - 1_{\{m=m^*\}} \\ &= -\left[1 + \frac{(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right]^{-2}(\beta_o - 1) \\ &\cdot \quad \frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\left(\frac{\exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} + 1_{\{m=m^*\}}\right) - 1_{\{m=m^*\}}. \end{aligned} \tag{OA53}$$

Let B be the number of banks and V be the space of 3B dimensional vectors representing each bank's deposit, loan, and mortgage quantities. We want to compute how these quantities change when the central bank raises the supply of reserves so that increases security yields by R. We define a function $f_R : V \to V$ that equals 0 after the economy equilibrates in response to this increased reserve supply.

First, we define a function $f_1^{*,R}$ from bank level deposit, mortgage, and loan quantities to an associated security quantity consistent with the rate rise R. For each bank, this function is given by (where B_i is the number of branches of the bank) $R = \frac{1}{B_i}\begin{pmatrix} \frac{\partial^2 C}{\partial Q_D \partial Q_S} & \frac{\partial^2 C}{\partial Q_M \partial Q_S} & \frac{\partial^2 C}{\partial Q_L \partial Q_S} & \frac{\partial^2 C}{\partial Q_S \partial Q_S}\end{pmatrix} * \begin{pmatrix} Q_{D,i} - Q^o_{D,i} \\ Q_{M,i} - Q^o_{M,i} \\ Q_{L,i} - Q^o_{L,i} \\ Q_{S,i} - Q^o_{S,i}\end{pmatrix}$ This implies $S_i = S_o + \frac{B_i}{\frac{\partial^2 C}{\partial Q_S \partial Q_S}}(R - \frac{1}{B_i}\begin{pmatrix} \frac{\partial^2 C}{\partial Q_D \partial Q_S} & \frac{\partial^2 C}{\partial Q_M \partial Q_S} & \frac{\partial^2 C}{\partial Q_L \partial Q_S}\end{pmatrix} * \begin{pmatrix} Q_{D,i} - Q^o_{D,i} \\ Q_{M,i} - Q^o_{M,i} \\ Q_{L,i} - Q^o_{L,i}\end{pmatrix})$ The Jacobian of this function is $\frac{-1}{\frac{\partial^2 C}{\partial Q_S \partial Q_S}}\begin{pmatrix} \frac{\partial^2 C}{\partial Q_D \partial Q_S} & \frac{\partial^2 C}{\partial Q_M \partial Q_S} & \frac{\partial^2 C}{\partial Q_L \partial Q_S}\end{pmatrix}$ for the effect of bank i's quantities on bank i's security quantity and 0 for the effect of any other bank j on bank i's quantities. Let f_1^R be given by ($id : V \to V$, $f_1^{*,R}$)- which maps each banks 3 given quantities to themselves together with this implied security quantity.

Next, we define a map f_2 from each bank's quantities D_i, M_i, L_i, S_i to the change in its marginal costs from those before the counterfactual. This change in marginal costs is given by $\begin{pmatrix} MC_{D,i} - MC^o_{D,i} \\ MC_{M,i} - MC^o_{M,i} \\ MC_{L,i} - MC^o_{L,i}\end{pmatrix} = \frac{1}{B_i}\begin{pmatrix} \frac{\partial^2 C}{\partial Q_D \partial Q_D} & \frac{\partial^2 C}{\partial Q_M \partial Q_D} & \frac{\partial^2 C}{\partial Q_L \partial Q_D} & \frac{\partial^2 C}{\partial Q_S \partial Q_D} \\ \frac{\partial^2 C}{\partial Q_D \partial Q_D} & \frac{\partial^2 C}{\partial Q_M \partial Q_M} & \frac{\partial^2 C}{\partial Q_L \partial Q_M} & \frac{\partial^2 C}{\partial Q_S \partial Q_M} \\ \frac{\partial^2 C}{\partial Q_D \partial Q_L} & \frac{\partial^2 C}{\partial Q_M \partial Q_L} & \frac{\partial^2 C}{\partial Q_L \partial Q_L} & \frac{\partial^2 C}{\partial Q_S \partial Q_L}\end{pmatrix} * \begin{pmatrix} Q_{D,i} - Q^o_{D,i} \\ Q_{M,i} - Q^o_{M,i} \\ Q_{L,i} - Q^o_{L,i} \\ Q_{S,i} - Q^o_{S,i}\end{pmatrix}$. The Jacobian of f_2 is $\frac{1}{B_i}\begin{pmatrix} \frac{\partial^2 C}{\partial Q_D \partial Q_D} & \frac{\partial^2 C}{\partial Q_M \partial Q_D} & \frac{\partial^2 C}{\partial Q_L \partial Q_D} & \frac{\partial^2 C}{\partial Q_S \partial Q_D} \\ \frac{\partial^2 C}{\partial Q_D \partial Q_D} & \frac{\partial^2 C}{\partial Q_M \partial Q_M} & \frac{\partial^2 C}{\partial Q_L \partial Q_M} & \frac{\partial^2 C}{\partial Q_S \partial Q_M} \\ \frac{\partial^2 C}{\partial Q_D \partial Q_L} & \frac{\partial^2 C}{\partial Q_M \partial Q_L} & \frac{\partial^2 C}{\partial Q_L \partial Q_L} & \frac{\partial^2 C}{\partial Q_S \partial Q_L}\end{pmatrix}$ from a bank's own quantities to its marginal cost changes and 0 for all other terms in the Jacobian matrix.

In each market, given the marginal cost changes of each bank in the market, we now compute the change in the bank's chosen interest rates that are consistent with the marginal cost changes. That is, each bank's change in interest rates Δr_{nmt} from that observed in the data is chosen so that they all solve equation (OA52). This system of equations must be solved numerically, but it is tractable since it can be solved seperately market by market. In market n, equation (OA52) defines a function g from a vector of rate changes for each bank in the market to an expression for that bank's change in marginal cost from that implied in the data. By solving g to equal our vector of marginal cost changes, we are computing the function $f_3 = g^{-1}$. The Jacobian of $f_3 = g^{-1}$ is the inverse of the Jacobian of g, which is given by equation (OA53).

Having solved in each market for the change in bank-market-level interest rate changes that are consistent with our marginal cost changes, we next compute the bank-level quantities implies by plugging these new interest rate changes into our demand system. The total quantity of deposits on a bank's balance sheet is, summing equation (OA45) across markets,

$$Q_{D,mt} = \sum_n Q_{D,nmt} = \sum_n \overline{Q}^i_{D,nt} \frac{\left(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))\right)^{\beta_o - 1}}{\left(\sum_{m'} \exp(\delta^i_{nm't})\right)^{\beta_o}} \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})).$$

Analogous expressions for the quantity of mortgages and loans also hold.

$$Q_{M,mt} = \sum_n Q_{M,nmt} = \sum_n \overline{Q}^i_{M,nt} \frac{\left(\sum_{m'} \exp(\delta^{i,M}_{nm't} + \alpha^M(\Delta r_{nm't}))\right)^{\beta^M_o - 1}}{\left(\sum_{m'} \exp(\delta^{i,M}_{nm't})\right)^{\beta^M_o}} \exp(\delta^{i,M}_{nmt} + \alpha^M(\Delta r^Q_{M,nmt}))$$

This defines a function f_4 from the rate changes we computed above back to a list of bank-level deposit, mortgage, and loan quantities. The Jacobian of this function is given by

$$\begin{aligned} & \frac{\partial}{\partial \Delta r_{nm^*t}} D_{mt} \qquad \text{(OA54)} \\ = \ & (\beta_o - 1)\alpha \overline{Q}^i_{D,nt} \frac{\left(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))\right)^{\beta_o - 2}}{\left(\sum_{m'} \exp(\delta^i_{nm't})\right)^{\beta_o}} \exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t})) \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})) \\ + \ & 1_{\{m=m^*\}} \alpha \overline{Q}^i_{D,nt} \frac{\left(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))\right)^{\beta_o - 1}}{\left(\sum_{m'} \exp(\delta^i_{nm't})\right)^{\beta_o}} \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})) \\ = \ & \alpha \overline{Q}^i_{D,nt} \left((\beta_o - 1) \left(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})) \right)^{-1} \exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t})) + 1_{\{m=m^*\}} \right). \end{aligned}$$

Thus,

$$f_R = f^R_1 \circ f_2 \circ f_3 \circ f_4 \qquad \text{(OA55)}$$

maps V to V, and a fixed point of f_R yields a counterfactual equilibrium of the economy. The Jacobian of this function is (by the expression for the Jacobian of composed functions) $J(f^R_1) \times J(f_2) \times J(f_3) \times J(f_4)$, where $J(.)$ denotes the Jacobian of each individual function. We provided closed form expressions for all of these Jacobians except f_3, which was a function defined by solving a system of equations (that must be computed numerically). However, f_3 is given by the inverse of our function g that does have a closed form Jacobian, which can be used to give the Jacobian of f_3 at its computed numerical solution. We compute our counterfactual by solving the equation $f_R(v) - v = 0$

numerically, using our analytic expression for its Jacobian to speed computation.

F Infinite-Horizon Model

This section presents an infinite-horizon profit maximization problem for each bank that results in the same optimal behaviour as the two-period model presented in the main text. Each bank m chooses market-specific rates $R_{P,nmt}$, where P corresponds to D, M, and L, for its deposits, mortgages and corporate loans in market n at time t. These markets are imperfectly competitive, and bank m faces demand curves that determine its quantities $Q_{P,nmt}(R_{P,nmt}, \omega_t)$ of deposits (D), mortgages (M), and loans (L) in market n at time t. These demand curves depend on the bank's own chosen rates as well as a vector ω_t of variables the bank does not choose, such as competitors' rates and exogenous shocks. In addition, bank m chooses its quantity $Q_{S,mt}$ of liquid securities at time t that trade in a competitive market paying an interest rate $R_{S,t}$.

In period $t+1$, bank m makes a payout to its equity holders of

$$\begin{aligned}
&\Pi_{m,t+1} = \qquad \text{(OA56)}\\
&\sum_n Q_{L,nmt}(1+R_{L,nmt}) + \sum_n Q_{M,nmt}(1+R_{M,nmt}) + Q_{S,mt}(1+R_{S,t}) - \sum_n Q_{D,nmt}(1+R_{D,nmt})\\
-\ &\left(\sum_n Q_{L,nm,t+1} + \sum_n Q_{M,nm,t+1} + Q_{S,m,t+1} - \sum_n Q_{D,nm,t+1}\right) - C(\Theta_{mt}),
\end{aligned}$$

The bank's equity holder has a pricing kernel $\Lambda_{t,t+j}$ and maximizes the present value of its payouts

$$\max_{(R_{D,nmt}, R_{M,nmt}, R_{L,nmt}, Q_{mt})} \sum_{j=0}^{\infty} \mathbb{E}_t[\Lambda_{t,t+j}\Pi_{m,t+j}] \qquad \text{(OA57)}$$

subject to equation (OA56). Note that each rate chosen at time $t+j$ only impacts $\Pi_{m,t+j}$ and $\Pi_{m,t+j+1}$

The first-order conditions for the bank's problem are [19]

$$\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}} = \frac{1}{1+R_t^{D,m}}\left(\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}(1+R_{D,nmt})+Q_{D,nmt}+\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}\frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}}\right) \quad \text{(OA58)}$$

$$\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}} = \frac{1}{1+R_t^{L,m}}\left(\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}(1+R_{L,nmt})+Q_{L,nmt}-\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}\frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}\right) \quad \text{(OA59)}$$

$$\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}} = \frac{1}{1+R_t^{M,m}}\left(\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}(1+R_{M,nmt})+Q_{M,nmt}-\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}\frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}\right) \quad \text{(OA60)}$$

$$1 = \frac{1}{1+R_t^{S,m}}\left((1+R_{S,t})-\frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}\right). \quad \text{(OA61)}$$

These are equivalent to equations (2)-(5).

G Estimating Demand and Cost Curves for Firms in Multiple Markets

This appendix analyzes the decisions of a firm that sells goods in multiple markets. The key result is that a demand shock in one market can be used both to identify the demand curves the firm faces in other markets as well as to identify the firm's marginal cost curve of production. A firm sets price P_n for the goods it sells in market n, facing demand curve $D_n(P_n, \lambda_n)$. The parameter λ_n is an exogenous shock that shifts demands for the good only in market n. There is a total of N markets. The firm faces a cost $C(\sum_n D_n(P_n, \lambda_n)) + \sum_n \epsilon_n D_n(P_n, \lambda_n)$ of production. The firm maximizes its profits

$$\sum_n P_n D_n(P_n, \lambda_n) - C(\sum_n D_n(P_n, \lambda_n)) - \sum_n \epsilon_n D_n(P_n, \lambda_n) \quad \text{(OA62)}$$

yielding first-order condition for P_n

$$D_n(P_n, \lambda_n) + P_n[\frac{\partial (D_n)(P_n, \lambda_n)}{\partial P_n}] - (C'(\sum_n D_n(P_n, \lambda_n)) + \epsilon_n)(D_n)'(P_n, \lambda_n) = 0 \quad \text{(OA63)}$$

$$\frac{D_n(P_n, \lambda_n)}{\frac{\partial (D_n)(P_n, \lambda_n)}{\partial P_n}} + P_n - C'(\sum_n D_n(P_n, \lambda_n)) + \epsilon_n) = 0 \quad \text{(OA64)}$$

When this system of equations has a unique solution, it implicitly defines a function $P(\lambda)$, mapping the vector of the λ_n demand shocks to a vector of prices, with price $P_n(\lambda)$ in market n. For j not

[19] For simplicity, we assume that the riskiness of a bank's entire deposit base is the same (and respectively all of its mortgages and all of its loans). This allows us to define bank-asset-specific discount rates ($R_t^{D,m}$,$R_t^{M,m}$,$R_t^{L,m}$,$R_t^{Q,m}$ in each first order condition implied by the pricing kernel $\Lambda_{t,t+j}$.

equal to n, we have that

$$\frac{d}{d\lambda_j}[D_n(P_n,\lambda_n)] = \frac{\partial D_n(P_n,\lambda_n)}{\partial P_n}\frac{\partial P_n}{\partial \lambda_j} \tag{OA65}$$

$$\frac{\frac{d}{d\lambda_j}[D_n(P_n,\lambda_n)]}{\frac{\partial P_n}{\partial \lambda_j}} = \frac{\partial D_n(P_n,\lambda_n)}{\partial P_n} \tag{OA66}$$

It follows that if we divide the response of quantities in market n to the demand shock λ_j by the response of prices in market n to the demand shock λ_j, we get the slope $\frac{\partial D_n(P_i,\lambda_i)}{\partial P_n}$ of the demand curve. This implies that a 2 stage least squares regression estimates the impact of P_n on D_n using the demand shock λ_j as an instrument identifies the slope of the demand curve in market i. This is the approach we take when using a natural disaster instrument to estimate our demand system.

Having estimated the demand curves D_i faced by the firm, we identify the average MC of the firm's marginal costs across markets by

$$MC = C'(\sum_n D_n(P_n,\lambda_n)) + \frac{1}{N}\sum_n \epsilon_n \tag{OA67}$$

The response of this marginal cost cost to a shock to any given λ_j is

$$\frac{dMC}{d\lambda_j} = C''(\sum_n D_n(P_n,\lambda_n))\frac{d[\sum_n D_n(P_n(\lambda),\lambda_n)]}{d\lambda_j} \tag{OA68}$$

It follows that if we regress the marginal cost MC on the demand shock λ_j and then regress the total quantity $\sum_n D_n(P_n(\lambda),\lambda_j)$ on the demand shock λ_j, the ratio of these regression coefficients identifies the slope $C''()$ of the firm's marginal cost curve. This shows how in a setting where firms are active in multiple markets, we can use a demand shock in a given market to identify both the demand curve the firm faces in other markets as well as the firm's marginal cost curve.