The Reserve Supply Channel of Unconventional Monetary Policy*

William Diamond
Wharton

Zhengyang Jiang
Northwestern Kellogg

Yiming Ma
Columbia Business School

April 5, 2021

Abstract

We document the "reserve supply channel" of Quantitative Easing (QE) that has the unintended consequence of reducing bank lending to firms. Each dollar of central bank reserves created by QE crowds out 13 cents of bank lending. We reach this conclusion using a structural model that is estimated with instrumental variables for deposit and loan demand across regions of the country. Our results depend on two key estimates: the elasticity of demand for bank loans and how the cost of supplying loans is impacted by a bank's holding of reserves. We find that each $1 trillion of reserves in the banking system raises the cost of capital for loans by 1.5 basis points, leading to a 3.42% reduction in the quantity of corporate loans demanded. In a counterfactual simulation, we show that the $2.7 trillion of reserves created by QE reduced bank lending to firms by over $500 billion but had modest impacts on deposit and mortgage quantities. Our results imply that forcing the banking sector to hold the large quantity of reserves created by QE reduces QE's ability to stimulate the economy. This unintended consequence of QE could potentially be allevaited by the relaxation of the Supplementary Leverage Ratio (SLR) regulation.

*We thank Olivier Darmouni, Arvind Krishnamurthy, Gregor Matvos, Luke Taylor, Quentin Vandeweyer, Stijn van Nieuwerburgh, Jessica Wachter, Yao Zeng, and audiences at Columbia Business School, Durham University, Kellogg, Johns Hopkins, University of Amsterdam, University of Utah, Wharton, and the Junior Macro-Finance Workshop 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."

1 Introduction

In the aftermath of the 2020 Covid-19 pandemic and the 2008-2009 financial crisis, the U.S. Federal Reserve purchased trillions of dollars of assets in its Quantiative Easing (QE) program to stimulate the economy. When the Federal Reserve buys an asset from a bank in QE, it pays by crediting the bank's so-called reserve balance, an interest-bearing account at the Federal Reserve that is available only to banks. This has the effect of creating a special interest-bearing asset that can only be held within the banking system. As seen in figure 1, QE increased the quantity of reserves held by banks from less than $50 billion in 2006 to a peak of $2.8 trillion in 2015 and then a maximum of $3.3 trillion in 2021. While there is a large empirical literature[1] on QE, this paper is the first to analyze QE's unintended consequence of forcing the banking system to hold trillions of dollars of newly created bank reserves. We show that the "reserve supply channel" of QE leads to a reduction in the supply of bank loans to firms.

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 like the supplementary leverage ratio 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. 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 $1.97 trillion in 2020Q1, the proportion of illiquid assets on bank balance sheets declined from 83% to 67% (see Figure 1).[2] However, because QE is a policy used only in severe crises, this substitution away from holding illiquid assets may reflect poor economic fundamentals (such as low loan demand) and may not necessarily be caused by the policy itself.

[1] See for example Krishnamurthy and Vissing-Jorgensen (2011); Di Maggio et al. (2020); Rodnyansky and Darmouni (2017); Chakraborty et al. (2020)

[2] 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.

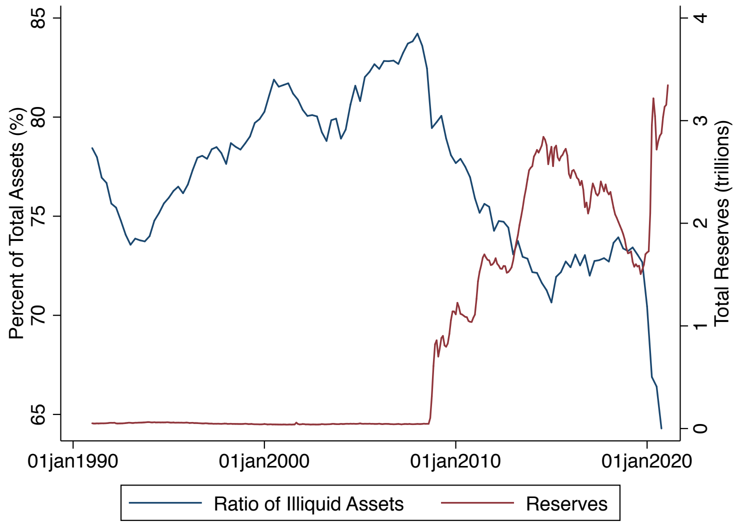

Figure 1: Supply of Central Bank Reserves and Bank Asset Illiquidity

To empirically estimate the impact of QE on the banking system without relying on aggregate time series data, we use a structural model with two key ingredients. First, in each region of the country, banks compete in imperfectly competitive markets to provide deposits, loans, and mortgages. Second, a bank's cost of capital to lend depends on the quantity of deposits, mortgages, loans, and reserves on its balance sheet.[3] Of particular relevance to QE, a bank's cost of lending depends on the quantity of reserves it holds. There are two key quantities we need to estimate to quantify the reserve supply channel of QE. 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 the additional reserves created by QE. With our estimated model, we show that a \$4.76 trillion increase in the supply of central bank reserves crowds out bank lending by \$555.9 billion, so the reserve supply channel of QE counterproductively reduces the supply of bank loans.

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 construct such an exogenous

[3] A bank's cost of providing deposits and mortgages also depends on the composition of the bank's balance sheet in an analgous manner.

shock by observing how banks reallocate funds in their internal capital markets after a natural disaster. This follows Cortés and Strahan (2017), who show that loan demand in a region increases after it is hit by a disaster. Banks reallocate funds away from non-disaster regions to provide 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.[4]

Our demand estimates show that the total demand for bank loans is considerably 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, the quantity of corporate loans demanded falls by 22.8%. For comparison, a 10 basis point increase in rates would raise deposit demand by 1.3% and would lower mortgage demand by 4.4% If banks change their deposit, loan, and mortgage interest rates by similar amounts, their loan quantities will respond by a much larger amount than their mortgage or deposit quantities. This explains why we find that corporate loan quantities respond most to the increase in reserve supply caused by QE.

Next, we estimate how a bank's cost of providing loans (and mortgages and deposits) depends on the composition of its balance sheet. We first use our demand estimates to infer what a bank's marginal cost of lending must be to rationalize the loan interest rate it chooses. We then observe how this marginal cost varies when a bank adjusts the composition of its balance sheet in response to an exogenous shock to deposit and/or loan demand. Because a bank can adjust several components of its balance sheet in response to a demand shock, our problem is analogous to estimating a regression with multiple endogenous variables. Solving this endogeneity problem requires multiple exogenous shocks to loan and/or deposit demand. In addition to the Cortés and Strahan (2017) 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.

[4] We estimate demand curves for mortgages and deposits similarly.

Our estimates imply that increasing a bank's reserve holdings crowds out mortgage and corporate lending and crowds in deposit issuance. We find that a $1 trillion increase in central bank reserves divided acrosss the banks we observe in 2007 leads to a 1.1 basis point increase in mortgage costs, a 1.49 basis point increase in loan costs, a 2.3 basis point reduction in deposit costs, and a 3.73 basis point increase in the required return on reserves. This implies that the reserves created by QE make it more expensive for banks to provide mortgages and loans.

We use our estimated model to run a counterfactual analysis of the increase in reserve supply caused by QE on the banking system. We allow each bank to adjust its deposit, loan, and mortgage rates at each bank branch as well as to trade the newly created reserves with other banks.[5] We find that in 2007, a $ 4.25 trillion injection of reserves by the Federal Reserve pushes up the reserve interest rate by 15 basis points. 15 basis points is comparable to the maximum size of the spread between the interest on excess reserves (IOER) and the federal funds rate, a measure of the extra interest paid by reserves over a rate available to all investors. This suggests that much of the IOER-fed funds rate spread was induced by QE.

sol In our counterfactual, loan quantities fall by much more than the change in mortgage and desposit quantities. However, the change in loan rates is comparable to the changes in mortgage and deposit rates. An additional $4.25 trillion in the supply of reserves induces a pass-through of 6.19 basis points in deposit rates, 3.86 basis points in mortgage rates, and 5.20 basis points in corporate loan rates. These rate changes imply a $15.4 billion increase in deposits, a $6.1 billion decrease in mortgages, and a $555.9 billion decrease in corporate loans. The pass-through of QE to deposit and mortgage quantities through the reserve supply channel we analyze is small, but the crowding-out effect reduces corporate loan quantities by 13% of the size of the reserve supply increase.

Our findings imply that requiring banks to hold the trillions in reserves created by QE caused a significant reduction in bank lending to firms. Reducing the regulatory constraints banks face when

[5] The vector of all new rates and portfolio choices in our simulation is over 10,000 dimensions, and the symbolic Jacobian for our model provided in the appendix is crucial to make this numerically tractable.

holding reserves could avoid this crowding out of bank lending. One example is the recent relaxation of the Supplemenetary Leverage Ratio (SLR) requirement, which allows reserves and Treasuries held on bank balance sheets to be exempt from the calculation of a bank's leverage. Alternatively, allowing non-bank investors to hold reserves could also reduce this crowding out of bank lending. Such a change in the implementation of QE could increase its ability to provide economic stimulus.

Literature Review This paper contributes first to the empirical literature on how quantiative easing impacts the banking system with a new structural methdology and focus on reserve creation . Existing reduced-form work in this literature, such as Rodnyansky and Darmouni (2017) and Chakraborty et al. (2020), focuses on the mortgage-backed securities purchased in QE rather than the reserves created to fund the purchase. They show that after QE a "treatment group" of banks that held more mortgage-backed securities (and are more impacted by the purchase of these assets) increase their mortgage lending relative to a "control group" of those that hold fewer. Our paper is the first to study the effects on banks of forcing them to hold trillions in newly-created central bank reserves. Because reserves are traded by all banks in a single market, a treatment group-control group methodology is not available to us, so we take the alternative approach of estimating a structural model.

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 . Closest to our work is Wang et al. (2020), who use a structural model of banking in the presence of imperfect competition to study conventional monetary policy transmission. Our work shows that the degree of competition across markets is crucial for the transmission of the reserve supply channel of QE, since highly price elastic corporate loan demand is impacted much more by reserve supply than deposit and mortgage demand.

Our structural model places us in a growing recent literature on structural estimation in banking. Many papers estimate models of imperfect competition similar to ours (Egan, Hortaçsu, and Matvos, 2017; Buchak, 2018; Wang, Whited, Wu, and Xiao, 2020; Xiao, 2020; Buchak, Matvos, Piskorski, and

Seru, 2018), while others estimate models of networks and matching (Akkus, Cookson, and Hortacsu, 2016; Schwert, 2018; Craig and Ma, 2018). Our use of branch-level data allows us to be the first to use exogenous variation across regions of the country, common in the reduced form literature (e.g. Cortés and Strahan (2017)). Our estimates of the synergies between the different components of bank balance sheets are also new and quantify a core idea in banking theory.[6]

2 A Model of Banks Balance Sheets

The purpose of our model is to quantify how the banking system responds to policy interventions, such as an increase in reserve supply caused by QE. Because policy interventions like QE tend to be implemented after large recessions, using data on banks directly after QE may reveal the effect of the recession rather than that of QE. This motivates a structural approach to policy analysis. We impose a theoretical model of banks' decisions and estimate the model's parameters with data that is directly related to QE. We introduce our model in Subsection 2.1. Subsection 2.2 shows that the effect of increasing the supply of central bank reserves depends on two things: the slopes of the demand curves banks face and the "balance sheet costs" banks face in holding deposits, loans, mortgages, and liquid reserves. In sections 3 and 4, we estimate these demand curves and supply costs using exogenous cross-sectional shocks to the demand for mortgages, loans, and deposits across the different regions in the country.

We first present a simplified, visual depiction of our model, with the formal statement in section 2.1. Our banks provide loans, mortgages, and deposits in imperfectly competitive markets. They take as given a loan demand curve (and similarly for deposits and mortgages) that determines the quantity of loans they issue given the interest rate they choose. This loan demand curve determines the marginal revenue the bank earns from changing its interest rate. Like any firm facing a downward-sloping demand curve, banks choose their interest rate so that the marginal cost of providing loans

[6]See e.g. Diamond and Rajan (2000); Kashyap et al. (2002); Hanson et al. (2015); Diamond (2019).

equals the marginal revenue they earn. Given this optimally chosen interest rate, the loan demand curve pins down the quantity of loans they can issue. This is depicted in figure 2.

If a bank's holdings of liquid reserves supplied by QE impacts its marginal cost of lending, QE shifts a bank's marginal cost curve for lending, as shown in figure 3. The intersection of the marginal revenue curve (blue) with the new marginal cost curve (dotted red) determines the new equilibrium interest rate. Plugging this new rate into the loan demand curve (green) yields the new equilibrium quantity of loans issued by the bank. To use this graphical model as a framework for quantifying the effects of a reserve supply increase on the lending market, we must estimate the green loan demand curve (which in turn determines the blue marginal revenue curve). Because banks compete with

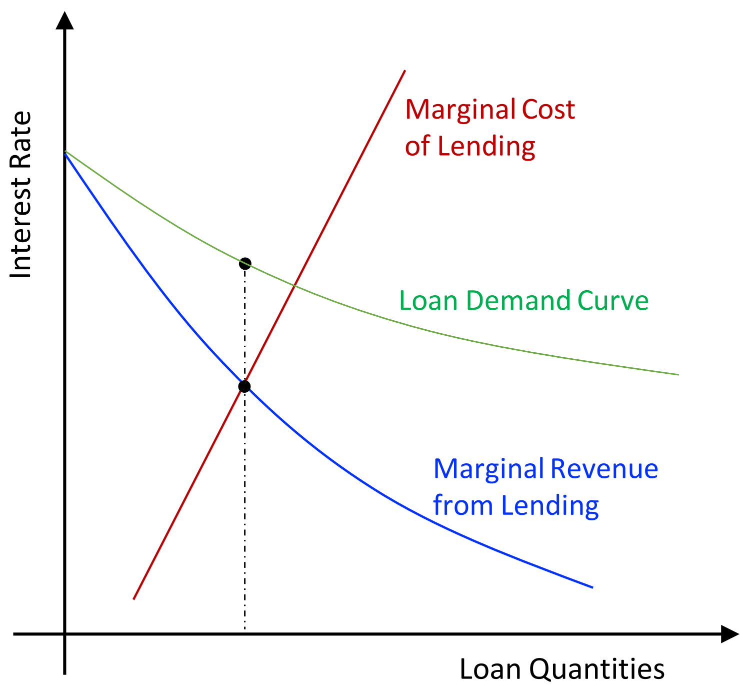

Figure 2: Rates and Quantities in Imperfectly Competitive Loan Markets. The intersection of the marginal cost and marginal revenue curve determines the equilibrium interest rate. Plugging this interest rate into the demand curve determines the equilibrium quantity.

each other, we estimate a demand system that quantities how all banks' chosen interest rates impact each other's quantities instead of just a single demand curve. This is done seperately for deposits,

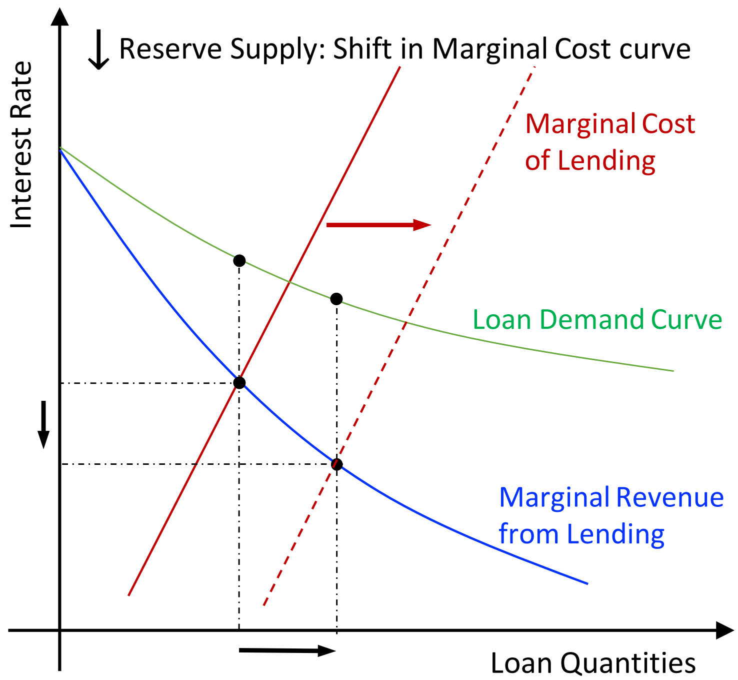

Figure 3: Pass Through of Reserve Supply Increase to Loan Markets. A change in reserve supply shifts the bank's marginal cost curve for lending. This results in a new intersection with the marginal revenue curve, yielding a new interest rate. The new loan quantity comes from plugging this new rate into the demand curve.

mortgages, and loans. Next, we have to estimate how an increase in reserve supply shifts the red marginal cost curve. Because a bank's cost of lending (and also providing mortgages and deposits) can depend on the composition of its entire balance sheet, we must estimate an entire cost function rather than a one dimensional cost curve. While our demand system and cost function estimation does not just yield one dimensional curves presented above, figures 2 and 3 illustrate their importance to quantifying the effects of QE on the banking system.

2.1 Model Set-Up

We consider a set of banks indexed by m that operated in a set of markets indexed by n at each time t. Each bank m chooses market-specific rates $R_{D,nmt}, R_{M,nmt}, R_{L,nmt}$ for, deposits D, mortgages M, and loans L. For each product $P \in (D, M, L)$ bank m takes as given the vector of rates $R_{P,n(-m)t}$ chosen by its competitor banks in market n, and the quantity $Q_{P,nmt}$ it sells is given by a residual demand curve $Q_{P,nmt}(R_{P,nmt}, R_{P,n(-m)t})$. 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 $R_{S,t}$. Loans, mortgages, securities, and deposits held by the bank have cash flows that are discounted at rates $R_t^{L,m}, R_t^{M,m}, R_t^{D,m}, R_t^{S,m}$ reflecting their riskiness. The bank chooses its rates at time t to maximize the expected present value of its profit at time $t+1$.[7]

The bank's objective function can be written as

$$\begin{aligned}\textstyle\sum_n Q_{L,nmt}(R_{L,nmt}, R_{L,n(-m)t})(R_{L,nmt} - R_t^{L,m}) + \sum_n Q_{M,nmt}(R_{M,nmt}, R_{M,n(-m)t})(R_{M,nmt} - R_t^{L,m})\\ +Q_{S,mt}(R_{S,t} - R_t^{S,m}) + \textstyle\sum_n Q_{D,nmt}(R_{D,nmt}, R_{D,n(-m)t})(R_t^{D,m} - R_{D,nmt}) - C(\Theta_{mt}),\end{aligned} \tag{1}$$

where $C(\Theta_{mt})$ is a "balance sheet cost" incurred by the bank at time $t+1$ that depends on the composition of the bank's balance sheet at time t. Specifically, the argument Θ_{mt} is a vector that contains bank m's balance sheet items $Q_{D,nmt}, Q_{M,nmt}, Q_{L,nmt}, Q_{S,mt}$ for all markets n, as well as the exogenous shocks ω_t.

The first order conditions for the choice variables $R_{D,nmt}, R_{M,nmt}, R_{L,nmt}, Q_{S,mt}$ are

$$\frac{1}{\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}}\frac{\partial}{\partial R_{D,nmt}}[Q_{D,nmt}(R_{D,nmt}, R_{D,n(-m)t})(R_t^{D,m} - R_{D,nmt})] = \frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}} \tag{2}$$

$$\frac{1}{\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}}\frac{\partial}{\partial R_{M,nmt}}[Q_{M,nmt}(R_{M,nmt}, R_{M,n(-m)t})(R_{M,nmt} - R_t^{L,m})] = \frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}} \tag{3}$$

$$\frac{1}{\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}}\frac{\partial}{\partial R_{L,nmt}}[Q_{L,nmt}(R_{L,nmt}, R_{L,n(-m)t})(R_{L,nmt} - R_t^{L,m})] = \frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}} \tag{4}$$

$$R_{S,t} - R_t^{S,m} = \frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}. \tag{5}$$

Equations 2-4 are each a version of the standard first order condition for price setting in an imperfectly competitive market. On the left hand side is the "marginal revenue" from changing an inter-

[7]The appendix shows 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.

est rate chosen by the bank, divided by the quantity change induced by that interest rate change. For example, the expression $Q_{L,nmt}(R_{L,nmt}, R_{L,n(-m)t})(R_{L,nmt} - R_t^{L,m})$ is the quantity of loans $Q_{L,nmt}(R_{L,nmt}, R_{L,n(-m)t})$ issued by the bank times the “profit" per unit $(R_{L,nmt} - R_t^{L,m})$- the spread betwen the interest rate charged on the mortgage and the fair rate of return $R_t^{L,m}$ for discounting the mortgage's cash flows. This yields the marginal revenue from adding one marginal unit of quantity deposits, loans, or mortgages.On the right hand side of these first order conditions is the marginal cost of changing the composition of the bank's balance sheet. Equations 2-4 can be rewritten as

$$\begin{aligned}
R_t^{D,m} - R_{D,nmt} &= \frac{Q_{D,nmt}}{\partial Q_{D,nmt}/\partial R_{D,nmt}} + \frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}};\\
R_t^{M,m} - R_{M,nmt} &= \frac{Q_{M,nmt}}{\partial Q_{M,nmt}/\partial R_{M,nmt}} - \frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}};\\
R_t^{L,m} - R_{L,nmt} &= \frac{Q_{L,nmt}}{\partial Q_{L,nmt}/\partial R_{L,nmt}} - \frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}
\end{aligned}$$

The model has two key features that have to be estimated from data: the “demand systems" $(Q_{L,nmt}(R_{L,nmt}, R_{L,n(-m)t})$, $Q_{M,nmt}(R_{M,nmt}, R_{M,n(-m)t})$, $Q_{D,nmt}(R_{D,nmt}, R_{D,n(-m)t}))$ and the “cost function" $C(\Theta_{mt})$. The demand systems are mappings from the interest rates chosen by all banks in a given market to the quantities each bank provides. Taking the rates chosen by other banks as given, these are simply the demand curves faced by an individual bank. Our first empirical step is to estimate these demand systems, by observing how each bank's quantities respond to shocks to the interest rates they and other banks choose. All terms on the left hand side of the equation depend only on the demand systems banks face for their products, which can be estimated first. Having estimated the demand system, these first order conditions yield values of the bank's marginal cost, which can then be used to estimate the bank's cost function.

The cost $C(\Theta_{mt})$ is a reduced-form function that accounts for the cost synergies between the various borrowing and lending businesses of a bank. For example, having more liquid assets on balance sheets may reduce the cost of fire-sales in the event of a bank-run and render bank runs less likely to begin with (Diamond and Dybvig, 1983). The use of demandable deposits may also serve as a commitment device in reducing fire sales (Diamond and Rajan, 2000). We do not need to take a stand on the specific source of the cost ex ante. Instead, our framework can uncover how the overall costs vary with the relative magnitudes of various balance sheet components. For instance, if the marginal cost of providing loans in market n drops with the bank's supply of deposits in market n', $\partial^2 C/(\partial Q_{D,nmt}\partial Q_{L,n'mt})$ would be negative. Estimating this cost function is the second empirical step of the paper, after which the model is completely parametrized and a counterfactual simulation of the impact of QE is possible.

2.2 Response of Banking System to a Reserve Supply Increase

To illustrate how an increased supply of central bank reserves would impact the banks in our model, we compute a comparative static where our bank m's liquid security holdings $Q_{S,mt}$ exogenously increases. The bank continues to choose its deposit, loan, and mortgage rates optimally that satisfy the following first-order conditions.

We note that the marginal liquidity costs of borrowing or lending may depend on the bank's entire balance sheet. If we parametrize this term by the bank's market level quantity (which implies an interest rate by inverting the demand curve) and add one unit of securities $Q_{S,mt}$ to the bank's balance sheet, we have that

$$\begin{aligned}
\frac{\partial Q_{D,nmt}}{\partial Q_{S,mt}} \cdot \frac{\partial}{\partial Q_{D,nmt}} \left(R_t^{D,m} - R_{D,nmt} - \frac{Q_{D,nmt}}{\partial Q_{D,nmt}/\partial R_{D,nmt}} \right) &= \frac{\partial^2 C(\Theta_{mt})}{\partial Q_{D,nmt}\partial \vec{Q}_{mt}} \cdot \frac{\partial \vec{Q}_{mt}}{\partial Q_{S,mt}} \\
\frac{\partial Q_{M,nmt}}{\partial Q_{S,mt}} \cdot \frac{\partial}{\partial Q_{M,nmt}} \left(R_t^{D,m} - R_{D,nmt} - \frac{Q_{D,nmt}}{\partial Q_{M,nmt}/\partial R_{D,nmt}} \right) &= -\frac{\partial^2 C(\Theta_{mt})}{\partial Q_{M,nmt}\partial \vec{Q}_{mt}} \cdot \frac{\partial \vec{Q}_{mt}}{\partial Q_{S,mt}} \\
\frac{\partial Q_{L,nmt}}{\partial Q_{S,mt}} \cdot \frac{\partial}{\partial Q_{L,nmt}} \left(R_t^{D,m} - R_{D,nmt} - \frac{Q_{D,nmt}}{\partial Q_{L,nmt}/\partial R_{D,nmt}} \right) &= -\frac{\partial^2 C(\Theta_{mt})}{\partial Q_{L,nmt}\partial \vec{Q}_{mt}} \cdot \frac{\partial \vec{Q}_{mt}}{\partial Q_{S,mt}}
\end{aligned}$$

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 response of each individual bank branch quantity to the reserve increase, and $\vec{Q}_{mt}$ is the vector of the banks' balance sheet quantities $(Q_{D,nmt}, Q_{M,nmt}, Q_{L,nmt}, Q_{S,mt})$.

Therefore, $\frac{\partial \vec{Q}_{mt}}{\partial Q_{S,mt}}$ is a vector of how all of the bank's balance sheet quantities respond to the reserve increase. By construction, the term representing securities satisfies $\frac{\partial Q_{S,mt}}{\partial Q_{S,mt}} = 1$, and the remainder of this vector is determined by solving this system of equations.

This system of equations determines how all of a bank's deposit, mortgage and loan quantities change if reserves are added to its balance sheet. On the left hand side is a term determined only by the demand curve a bank faces in an individual market. We estimate this term with an industrial organization style demand system.[8] On the right hand 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 the *cost synergies* between the different components of a bank's balance sheet (e.g., the synergy between borrowing from depositors and lending to homeowners or firms, a central concept in banking theory). We develop and apply a novel econometric approach to estimating these cost synergies that requires two separate instrumental variables for the demand for a bank's services. Together, our estimates of the demand for a bank's services and its cost of providing them allows us to compute the aggregate effect of an increased supply of reserves—the policy we intend to analyze.

3 Demand Systems

This section estimates the demand systems for deposits, mortgages, and loans. Subsection 3.1 introduces the logit demand system curves and their estimation strategy. Subsection 3.2 and 3.3 explain the data and instruments we use. The estimation results on demand elasticities, size of outside options, and implied mark-ups are shown in Subsection 3.4.

[8]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

Depositors in each market n at time t have a total supply of funds $F_{D,nt}$ that they choose how to invest. They can either invest in deposits at each bank m which has branches in the market or can invest in an unobserved outside option 0. This outside option allows for the possibility for consumers to substitute between deposits and other savings vehicles such as money market fund shares that are not in our data. An observed quantity $Q_{D,nmt}$ of deposits are 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 loans and mortgages have total funding needs of $F_{M,nt}$ and $F_{L,nt}$, respectively. They can either borrow from banks or resort to the outside option, which includes borrowing from non-banks or not obtaining funding altogether. $Q_{M,nmt}$ and $Q_{L,nmt}$ denote the observed quantities of mortgages and loans borrowed from bank m in market n in time t, while $Q_{M,n0t}$ and $Q_{L,n0t}$ denote the unobserved quantity of the respective outside option.

Preferences of depositors (firm borrowers or mortgage borrowers) follow a standard logit demand system (Berry, 1994)

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

The utility for customer j investing in bank m is made up of four components. The first is the interest rate $R_{D,nmt}$ paid on deposits times the customer's preference for the interest rate α_D. Notice that depositors prefer a higher interest rate while borrowers prefer a lower cost of funding so that α_D is positive. Customer utility is also affected by the desirability of bank m's deposits, which depends on a vector of observed characteristics X_{nmt}, the customer's preferences β_D for observed charecteristics, and unobservable characteristics $\delta_{D,nmt}$. Finally, the error term $\varepsilon_{D,jnmt}$ is assumed to be i.i.d. and follow a standard logit distribution. We normalize outside options to zero without loss of generality since only differences in utility across the choices available to a customer impact her decisions.

Under the assumptions of logit demand systems, the quantity of deposits invested in branches of bank m in market n at time t satisfies

$$Q_{D,nmt} = F_{D,nt} \frac{\exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{1 + \sum_{m'} \exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})}. \tag{6}$$

Since the denominator is common across all banks in market n at time t, this demand system implies

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

This linear specification with a market-time specific constant $\zeta_{D,nt}$ allows us to transparently estimate α_D and β_D using market-time fixed effects, which pin down the price disutility parameters required for the demand side of our model. Nevertheless, directly regressing log market shares $Q_{D,nmt}$ on interest rates $R_{D,nmt}$ and observable characteristics $X_{D,nmt}$ may yield biased estimates of the price disutility parameter because a bank with high quality banking services $\delta_{D,nmt}$ may rationally pay a lower deposit rate on deposits than a bank with low quality banking services. This implies that $R_{D,nmt}$ may likely be correlated with $\delta_{D,nmt}$. However, if we have an instrumental variable $z_{D,nmt}$ that only affects a bank's choice of interest rates but is uncorrelated with its unobserved quality characteristics $\delta_{D,nmt}$, the model can be consistently estimated using two-stage least squares. That is, we can obtain the price disutility parameters α_D and β_D by running the following two-stage least squares regression

$$R_{D,nmt} = \gamma_{D,nt} + \gamma_D z_{D,nmt} + X_{D,mt}\gamma_D + e_{D,nmt}, \tag{7}$$

$$\log Q_{D,nmt} = (\zeta_{D,nt} + \mathbb{E}_{D,nt}\delta_{D,nmt}) + \alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + (\delta_{D,nmt} - \mathbb{E}_{D,nt}\delta_{D,nmt}). \tag{8}$$

Note that the mean of the latent demand term $\delta_{D,nmt}$ is absorbed by the market-year fixed effect in Eq. (8), so that the residual in this equation is $\delta_{D,nmt} - \mathbb{E}_{D,nt}\delta_{D,nmt}$ rather than $\delta_{D,nmt}$. While we assume that unobserved product quality $\delta_{D,nmt}$ is uncorrelated with our instrument $z_{D,nmt}$ within each market-year, its market-year specific mean $\mathbb{E}_{D,nt}\delta_{D,nmt}$ need not be everywhere zero. Some markets may have unobservably better banking services provided than others, and this will impact the size of the market-year fixed effect $\zeta_{D,nt} + E_{D,nt}\delta_{D,nmt}$.

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

3.1.2 Market Size

Our two-stage least squares procedure, where market-time-specific means are differenced out through market-time fixed effects, relies on observing how the difference in two bank's log-quantities responded to the difference in their interest rates. It does not tell us how the overall quantity of deposits in a deposit market would respond if every bank in the market raised its interest rates. Similarly, we cannot tell how the overall quantity of mortgages would change if every bank raised its mortgage rates. 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.

For loans, we obtain the outside option size by directly multiplying the number of potential borrowers by the average loan size. For the number of potential borrowers, 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. 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.

For deposits and mortgages, we do not observe an analogous population of those who choose not to take out a mortgage or to hold bank deposits. The overall size of the market is therefore unobserved.

This leaves our demand system not entirely identified based on the price disutility parameters obtained in Subsection 3.1.1 alone.

Below, we describe our procedure to estimate the local elasticity between the deposits/mortgages and the outside options. We again use deposits as an example. We use $\mathcal{Q}_{D,nt}$ in a different font to denote the total quantity of deposit in a market n:

$$\mathcal{Q}_{D,nt} = \sum_m Q_{D,nmt}.$$

Summing equation 6 across all branches in a market, we have

$$\mathcal{Q}_{D,nt} = \mathcal{F}_{D,nt} \frac{\sum_m \exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt})}{1 + \sum_{m'} \exp(\alpha_D R_{D,nm't} + X_{D,nm't}\beta_D + \delta_{D,nm't})}. \tag{9}$$

We define $\delta_{D,nt} = \log(\sum_m \exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt}))$, which can be interpreted as the desirability of a composite deposit representing all banks operating in a market. Then, $\mathcal{Q}_{D,nt} = \mathcal{F}_{D,nt} \frac{\exp(\delta_{D,nt})}{1+\exp(\delta_{D,nt})}$, and using a log-linear approximation,

$$\log \mathcal{Q}_{D,nt} \approx \log \mathcal{F}_{D,nt} + \beta_{D,o}\delta_{D,nt}.$$

This equation allows us to estimate how $\log \mathcal{Q}_{D,nt}$ changes with the value of $\delta_{D,nt}$ 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.

We apply an instrumental variables approach to consistently estimate parameter $\beta_{D,o}$. From our estimation of the price disutility parameters in Eq. (7) and (8), we can observe all terms in the expression for $\delta_{D,nt}$ except the mean of $\delta_{D,nmt}$. We therefore decompose into an "observable" and an "unobservable" desirability component $\delta_{D,nt} = \delta^o_{D,nt} + \delta^u_{D,nt}$, where

$$\delta^u_{D,nt} = \frac{1}{N_{nt}} \sum_m \delta_{D,nmt}. \tag{10}$$

$$\delta^o_{D,nt} = \log \left(\sum_{m'} \exp(\alpha_D R_{D,nmt} + X_{D,nmt}\beta_D + \delta_{D,nmt} - \delta^u_{D,nt}) \right). \tag{11}$$

If we have an instrumental variable $z_{D,nt}$ that is uncorrelated with $(\log F_{D,nt} + \beta_{D,o}\delta^u_{D,nt})$ conditional on a vector $\chi_{D,nt}$ of controls, we can estimate $\beta_{D,o}$ by two-stage least squares as

$$\delta^o_{D,nt} = \rho_{D,t} + \theta_D z_{D,nt} + \chi_{D,nt}\theta_D + \varepsilon^o_{D,nt}, \tag{12}$$

$$\log Q_{D,nt} = \alpha_{D,t} + \beta_{D,o}\delta^o_{D,nt} + \chi_{D,nt}\rho_D + \eta_{D,nt}. \tag{13}$$

To construct this market-year level instrument, we take our market-bank-time level instrumental variable we used previously $z_{D,mnt}$, and construct the equal-weighted average at the market-year level:

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

which measures how exposed a region is to indirect rate changes coming through internal capital markets. Recall that under our log-linear approximation, $\log Q_{D,nt} = \log F_{D,nt} + \beta_{D,o}\delta^o_{D,nt} + \beta_{D,o}\delta^u_{D,nt}$. The identifying assumption is that these 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 ($\delta^u_{D,nt}$). In the appendix, we show that together with our previous estimates of the rate sensitivity coefficient α_D at an individual bank, the aggregate quantity's sensitivity to rate changes $\beta_{D,o}$ yields the following expressions for banks' demand curves.[9]

$$\frac{\partial \log Q_{D,nmt}}{\partial R_{D,nmt}} = \alpha_D + \alpha_D(\beta_{D,o} - 1)\frac{Q_{D,nmt}}{Q_{D,nt}}. \tag{14}$$

[9] We provide additional expressions for how a bank's quantities depend on all banks' chosen rates in Appendix 8.4.1. We also discuss some details of how we implemented our construction of $\delta^o_{D,nt}$ in the presence of missing data in some markets in in Appendix 8.2.

3.1.3 Mark-ups

After estimating the demand systems for deposits, mortgages, and loans, we proceed to infer bank's mark-ups in these markets. Mark-up estimates are not only interesting on their own. They also allow us to infer the marginal costs of producing deposits, loans and mortgages, which are essential for estimating the cost function parameters in Section 4.

To express mark-ups, we can simply rewrite the first-order conditions in Subsection 2.2. For deposits, we have

$$\nu_{D,nmt} = -\frac{Q_{D,nmt}}{\partial Q_{D,nmt}/\partial R_{D,nmt}} \quad = \quad R_{D,nmt} - R_t^{D,m} + \frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}} \tag{15}$$

which is the spread between the interest rate and the rate at which the bank would make zero profit, accouting both for its discount rate $R_t^{D,m}$ and for the marginal utility cost $\frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}}$ of an additional unit of deposit.

For mortgages and loans, we have

$$\nu_{M,nmt} = -\frac{Q_{M,nmt}}{\partial Q_{M,nmt}/\partial R_{M,nmt}} \quad = \quad R_{M,nmt} - R_t^{M,m} - \frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}, \tag{16}$$

$$\nu_{L,nmt} = -\frac{Q_{L,nmt}}{\partial Q_{L,nmt}/\partial R_{L,nmt}} \quad = \quad R_{L,nmt} - R_t^{L,m} - \frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}. \tag{17}$$

We expect deposit mark-ups to be negative because market power allows banks to offer depositors a lower return than they would have obtained in competitive markets. Loan and mortgage mark-ups should be positive because market power raises the cost of funding relative to a competitive benchmark. Security markets are competitive so that mark-ups are absent.

3.2 Data

3.2.1 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.

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 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).[10]

3.2.2 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

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

to be reported under this criterion.[11] 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.

We first merge bank-level identifiers in HMDA to the FDIC bank-level identifiers using the mapping file developed by Bob Avery.[12] 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).

3.2.3 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).[13]

[11] 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.

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

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

3.2.4 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.

3.3 Instruments

The Spatial Hazard Events and Losses Database for the United States (SHELDUS) records information on 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.

Our instrument z_{nmt} is constructed following Cortés and Strahan (2017). For deposits and mortgages, z_{nmt} measures for the branch of bank m in county n and year t the property losses from natural disasters accrued to the bank's branches in all other counties n':

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

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 rationale behind our instrument is that property losses from natural disasters create loan demand shocks in the regions they affect so that funds are allocated away from branches in county n

to branches in affected counties n' through banks' internal capital markets. Property losses to bank m's branches in regions n' therefore constitute a supply shock to bank m's branches in county n, which allows us to trace out the demand curve for deposits and mortgages. In all specifications, we include the log property damage to that county to account for direct effects of disaster losses on demand. The exclusion restriction requires that natural disasters do not directly influence local demand for deposits and mortgages in unaffected counties.

For commercial loans, we use the same instrument constructed at the bank-state-year level instead of the bank-county-year level.

3.4 Estimation Results

Table 2 reports the first-stage and second-stage results for the price disutility estimation for deposits, mortgages, and loans as in Equations 7 and 8. For all specifications, we include the ratio of loan loss provisions over total loans to remove any direct effects of natural disasters on the credit risk of bank assets. Since our deposit volume is a stock measure, whereas the 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. These signs are consistent with downward-sloping demand curves. Since the deposit rates are paid by the bank, 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 4.6%.[14] When the same bank lowers its mortgage and loan rates in one

[14]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, and hence the share of bank deposits relative to the outside option at the county level.

market by 10 basis points, its mortgage and loan volumes increase by 55.7% and 51.9%, respectively. The price disutility of deposits is an order of magnitude smaller than that for mortgages and loans, consistent with banks having much more market power in retail deposit markets than in mortgage and loan markets.

The outside option size can be directly obtained from the loans data, which we report in Table 4. For deposits and mortgages, we proceed to estimate the sensitivity of market-level quantities $Q_{P,nt}$ to the market-level desirability parameter $\delta^o_{P,nt}$ as in Equations 12 and 13. 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 $Q_{P,nt}$ to the market-level desirability parameter $\delta^o_{P,nt}$ to be 0.29 for deposits and 0.08 for mortgages. Hence, when all banks in a county raise their deposit rates by 10 basis points, the deposit quantity in that state increases by

$$\frac{\partial \log Q_{D,nt}}{\partial R_{D,nt}} = \frac{\partial \log Q_{D,nt}}{\partial \delta^o_{D,nt}} \frac{\partial \delta^o_{D,nt}}{\partial R_{D,nt}} = 0.29 \times 4.6\% = 1.3\%;$$

where 4.6% is the increase in the aggregate desirability of deposits relative to the outside option at the county-level. Similarly, when all banks in a county lower their mortgage rates by 10 basis points, the aggregate desirability of mortgages increases by 55.7% in that county relative to the outside option, and hence the mortgage quantity increases by $0.08 \times 55.7\% = 4.4\%$.

For loans, we report the outside option size at the state-year level in Table 4. On average, the implied β_o is 0.44. When all banks in a state lower their loan rates by 10 basis points, the aggregate desirability of loans in that state increases by 51.9% relative to their outside options, and hence the loan quantity in that county increases by $0.44 \times 51.9\% = 22.8\%$. Notice that the demand elasticity of loans is much higher than that of mortgages because although the sensitivity of their observed desirability to changes in interest rate is similar, the outside option of loans responds much more to changes in observed desirability than in the case of mortgages. One reason could be that corporate borrowers have more flxibility to borrow from other sources such as the bond market. 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 \$62.7 billion. if all banks lowered their mortgage rates by 10 basis points, the aggregate mortgage volume will increase by \$91.6 billion, and if all banks lowered their loan rates by 10 basis points, the aggregate loan volume would increase by \$1.03 trillion.

The behavior of an actual bank, due to its heterogeneity in bank characteristics and size, is different from that of a very small bank. Still, we can ask how an average bank's balance sheet quantities will change if it adjusts its deposit, mortgage, or loan rate. By Equation (14), the response of an average bank's deposit quantity in a given county is 4.3%, or 12.1 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 55.0%, or 4.1 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 51.7%, or 2.1 billion dollars, with respect to a 10 basis points decrease in the loan rate.

Table 5 reports summary statistics of the implied mark-ups, defined as the spread between the actual rate and the hypothetical competitive rate that incorporates liquidity cost. The average deposit mark-up is 2.50%, consistent with banks having high market power in deposit markets. In comparison, the average mortgage and loan mark-ups are 0.19% and 0.60%, which reflect more competitive lending markets.

4 Cost Function

This section specifies and estimates the bank's cost function for producing deposits, mortgages, and loans. We first use our estimated demand system to infer a bank's marginal cost in each market from the interest rate it chooses in that market. To identify the effects of a policy intervention like QE that impacts the composition of bank balance sheets, we need to know how these marginal costs change as bank balance sheets adjust. We begin with a reduced form analysis of how the quantities and marginal costs of these balance sheet components respond to cross-sectional instrumental variables that shock the demand for the bank's services. We then estimate the bank's cost function by choosing its parameters to be consistent with these reduced form natural experiments.

4.1 Cost Function Specification

We begin by specifying the bank's cost function and showing how it can be estimated using cross-sectional natural experiments. We assume that the bank's cost function for bank m at time t takes the form

$$\begin{aligned} C(\Theta_{mt}) &= H(Q_{D,mt}, Q_{M,mt}, Q_{L,mt}, Q_{S,mt}) \\ &+ \sum_n (Q_{M,nmt}\varepsilon^Q_{M,nmt} + Q_{L,nmt}\varepsilon^Q_{L,nmt} + Q_{D,nmt}\varepsilon^Q_{D,nmt}) + Q_{S,mt}\varepsilon^S_{mt}. \end{aligned} \tag{18}$$

This includes a term $H(Q_{D,mt}, Q_{M,mt}, Q_{L,mt}, Q_{S,mt})$ that can depend on the bank-level quantities of deposits, mortgages, loans, and securities. This allows, for example for the bank's holding of securities to impact its cost of mortgage lending, but not in a manner that depends on the specific mortgage market. 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. As shown above, the response of our model to external shocks depends entirely on the second derivatives of the bank's cost function, which are due only to the function H. Our cost function is therefore flexible enough to match the data with the ε_{nmt} shocks while ensuring that the cost synergies between a bank's borrowing, lending, and security holdings are the same across all branches.

To model the synergies between the bank's assets and liabilities, in a manner that is both flexible and yet restrictive enough to be identified from data, we assume the following functional form for H

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

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 loosely be interpreted as the bank's "equity" and measures the cost of expanding the size of the bank's balance sheet with non-deposit funding. This is because it equals the gap between the value of the assets we observe on the bank's balance sheet and its deposit financing.[15] The term $\mathcal{I}_{mt}$ we interpret as a measure of the "liquidity" of a bank's assets, where the coefficients ω_M and ω_L quantify 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 below, this is precisely the number of parameters that can be estimated by observing how our bank responds to two different cross-sectional instrumental variables.

[15] This term is not a perfect measure of a bank's equity capital since it ignores wholesale funding and other non-deposit debt financing as well as assets held on the bank's balance sheet that are not included in $Q_{S,mt}$. $Q_{S,mt}$ is a measure only of liquid securities such as reserves and treasuries held by a bank and does not include, for example, mortgage-backed securities.

4.2 Estimation Strategy

Differentiating Eq. (18) implies that the marginal costs of deposit, mortgage and loan for bank m in market n at time t is

$$\frac{\partial C}{\partial Q_{D,nmt}} = \mu_D - K_1\mathcal{E}_{mt} + K_3 Q_{D,mt} + K_4\mathcal{I}_{mt} + K_5(\mathcal{E}_{mt} - Q_{D,mt}) + \varepsilon^D_{nmt} \tag{19}$$

$$\frac{\partial C}{\partial Q_{M,nmt}} = \mu_M + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_M + K_4 Q_{D,mt}\omega_M + K_5 Q_{D,mt} + \varepsilon^M_{nmt} \tag{20}$$

$$\frac{\partial C}{\partial Q_{L,nmt}} = \mu_L + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_L + K_4 Q_{D,mt}\omega_L + K_5 Q_{D,mt} + \varepsilon^L_{nmt} \tag{21}$$

$$\frac{\partial C}{\partial Q_{S,mt}} = \mu_S + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt} + K_4 Q_{D,mt} + K_5 Q_{D,mt} + \varepsilon^S_{mt} \tag{22}$$

Recall that our markup estimate allowed us to recover $\partial C/\partial X_{nmt} - R^{m,X}_t$—a term that combines the discount rate for the balance sheet item of type X together with the marginal cost. If we replace the left hand sides of each of equations 19 to 22 with this observable counterpart, the right hand sides would change only in their intercept μ_X, since the discount rate does not depend on the composition of the bank's balance sheet. Averaging these equations across the markets n in which the bank operates yields

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{D,nmt}} - R_t\right) = \mu^*_D - K_1\mathcal{E}_{mt} + K_3 Q_{D,mt} + K_4\mathcal{I}_{mt} + K_5(\mathcal{E}_{mt} - Q_{D,mt}) + \varepsilon^D_{mt} \tag{23}$$

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{M,nmt}} + R^{M,m}_t\right) = \mu^*_M + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_M + K_4 Q_{D,mt}\omega_M + K_5 Q_{D,mt} + \varepsilon^M_{mt} \tag{24}$$

$$\frac{1}{N_{mt}}\sum_n\left(\frac{\partial C}{\partial Q_{L,nmt}} + R^{L,m}_t\right) = \mu^*_L + K_1\mathcal{E}_{mt} + K_2\mathcal{I}_{mt}\omega_L + K_4 Q_{D,mt}\omega_L + K_5 Q_{D,mt} + \varepsilon^L_{mt} \tag{25}$$

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

To estimate the parameters in these equations, 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 (and may choose to adjust their quantities in response to these shocks), we require exogenous variation in the quantities on the right hand side of each equation that is uncorrelated with the cost shocks ε^X_{mt}. Further complicating the problem, there are multiple endogenous variables on the right hand side of each equation. 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 are unable to tell 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^1_{mt}, z^2_{mt} that are both assumed to be uncorrelated with the cost shocks ε_{mt}.

We regress all of the marginal costs and quantities in these equations on our two instruments z^i_{mt} (indexed by $i = 1, 2$). For deposits, mortgages, loans, and securities, the response of the marginal cost to an instrument must equal the response of the quantity times the associated cost function parameter.[16] We can show that the regression coefficients solve a system of 8 equations, which identify the 7 parameters of our cost function. For the specific equations and details on how we average two of our equations to obtain a just-identified system, we refer the reader to Appendix 8.3.

While this cost function estimation procedure necessarily relies on a simultaneous system of equations, it builds directly on our reduced-form instrumental variable analysis. Our procedure matches the causal effects we estimated of how changes in a bank's balance sheet quantities impact its marginal costs of borrowing and lending.[17] Our approach is an application of using multiple instrumental variables to estimate models with multiple endogenous variables (e.g., a bank's quantities of deposits, mortgages, loans, and securities).

[16]Notice that there is no cross-sectional variation in the return on securities. Hence, the sensitivity of securities' marginal costs to the instrument is zero.

[17]To resolve the overidentification problem, we average two of our equations to obtain a just identified system as shown in Appendix 8.3.

4.3 Data

Data for the cost function is at the bank level. Specifically, mortgage, deposit, and loan costs are obtained from interest rates and the mark-up estimates in Subsection 3.4. They are averaged at the bank level and merged to the respective bank-level volumes from Call Reports. 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, Treasury securities, and agency securities. Finally, we normalize all volume variables by the number of counties in which the bank operates to align with the definition of the bank-level instruments.

4.4 Instruments

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

The first instrument is simply the natural disaster loss instrument taken to the bank level. For bank m at time t, we have

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

where N_{mt} is the number of branches and $\sum_n damage_{nt} \cdot \frac{Q_{D,nmt}}{\sum_{n_0} Q_{D,n_0mt}}$ is the sum of disaster losses accrued to branches of bank m in county n. Notice that unlike in the instrument for demand systems, we are no longer in need of a branch-level supply shock. Rather, losses from disasters predominantly comprise a bank-level demand shock for loans, and their distribution is plausibly exogenous to unobserved variation in banks' liquidity cost in the cross-section.

We also use a Bartik deposit instrument based on the average growth rates of deposits in markets where bank m has branches:

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

where where N_{mt} is the number of branches and $\frac{Q_{D,nt}-Q_{D,nt-1}}{Q_{D,nt-1}}$ is the deposit market growth rate in county n. To remove outliers, we winsorize $\frac{Q_{D,nt}-Q_{D,nt-1}}{Q_{D,nt-1}}$ at the 1% level.

Intuitively, a bank's deposit size may very well be a result of shocks to its cost of supplying deposits. Instead, 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 Specifically, the identifying assumption is that banks' differential exposure to the deposit growth rates in the counties they have branches in is not correlated with shocks to their cost of supplying deposits, mortgages, and loans. 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.

4.5 Estimation Results

Table 6 reports the parameter estimates for $(\kappa^{i,D}, \kappa^{i,M}, \kappa^{i,L}, \gamma^{i,D}, \gamma^{i,M}, \gamma^{i,L}, \gamma^{i,Q})$. 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 incurring larger losses from natural disasters also increase their deposits, mortgages, loans, and securities. Based on the effect on costs, we infer that the increase in volumes is consistent with an increase in loan and mortgage demand following natural disasters (e.g., to meet reconstruction needs). Specifically, mortgage and loan costs both increase, while deposit costs become more negative (i.e., deposits become more valuable for the bank). From Panel (b), banks experiencing a positive Bartik deposit shock also increase their deposits, mortgages, loans, and securities. In this case, the increase in balance sheet size is aligned with a positive deposit demand shock, as expected from the Bartik instrument. Deposit costs become less negative, implying that they are less valuable to the bank. At the same time, the costs of lending to firms and issuing mortgage loans declines as deposits become more abundant.

Based on these coefficient estimates, Table 7 reports the cost function's Hessian H. 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.[18] 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. Notice also that 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 expensive to give out loans and mortgages.

Lastly, we consider the change to marginal costs when we distribute \$1 trillion in reserves across banks. In 2007, there are 5,445 bank-counties in our sample. If bank branches in each county receive the same amount of reserves, our cost function parameters imply there would be a $0.0125 \times 184 = 2.30$ basis point decrease in the marginal cost of deposits, a $0.0060 \times 184 = 1.10$ basis point increase in the marginal cost of mortgages, a $0.0081 \times 184 = 1.49$ basis point increase in the marginal cost of loans, and a $0.0203 \times 184 = 3.73$ basis point decrease in the marginal benefit of securities. To map these cost changes to the equilibrium impact of QE on the banking system, we present a counterfactual analysis using both our estimated cost function and demand systems.

[18] Notice that based on our cost function estimates in the Hessian H, a \$1 billion increase in deposit quantity per county is associated with a 62 bps change in the marginal cost of deposits. In comparison, the Bartik deposit shock raises the deposit cost by 63 bps and the deposit quantity per branch by \$1.4 billion. The similarity in magnitudes confirm that the Bartik shock is predominantly an exogenous shock to deposit demand.

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 Quantitative Easing 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.[19] 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 (18). Second, because of these cost changes, banks change the interest rates they choose to charge on loans and mortgages and choose to pay on deposits. 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 the quantities of deposits, mortgages, and loans provided by the banking system as well as the rates charges on these products.

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. This is an over 38,000-dimensional problem. Nevertheless, dimensionality can be considerably reduced and the model is tractable to solve. We define a function 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 increase

[19]While there are other safe and liquid assets that are held in practice both by banks and other investors (such as Treasury securities), in our counterfactual we assume that banks simply increase their holdings of reserves without selling any other securities.

this interest rate increase. Let $Q^i_{D,mt}, Q^i_{M,mt}, Q^i_{L,mt}, Q^i_{S,mt}$ (where i stands for initial) be the bank level quantities of deposits, mortgages, loans, and securities actually 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 8.4 for further details.

5.2 Counterfactual Analysis: The Reserve Supply Channel of QE

In our benchmark counterfactual, we use data on the state of the banking system in 2007 to compute the effects of providing \$4.24 trillion of added reserves. This quantity was chosen so that it would increase the interest rate paid on reserves by exactly 15 basis points, which is roughly the average spread between the interest paid on excess reserves above the federal funds rate in the post-crisis period of QE. Because only banks can hold reserves while non-banks can invest at the federal funds rate, this spread is an ideal measure of the degree to which banks can earn a higher rate of return than other market participants due to the increase in reserves caused by QE. During QE, the supply of excess reserves peaked at \$2.7 trillion, which is of the same order of magnitude as our quantity increase. This quantitative similarity is not mechanical; our model is identified entirely from cross-sectional variation in how banks respond to natural disaster shocks and Bartik shocks to deposit demand. No data directly from the implementation of QE or on the excess reserves spread was used in estimation. Nevertheless, the model yields estimates of how the excess reserves spread responds to the quantity of reserves that are in the same ballpark as a casual eyeballing of data on reserve rates and reserve quantities.

One salient feature of our results is that mortgages and corporate loans are crowded out by increases in central bank reserves in QE, which suggests that empirically, the synergies between liquid and illiquid assets on bank balance sheets suggested by the theoretical literature are limited. On net, liquid securities and illiquid loans are substitutes rather than complements for commercial banks as shown by the negative coefficients for loans and mortgages in Table 8. While QE may certainly have other channels of transmission, it is important to consider the "reserves channel" we find, by which central bank reserves take up balance sheet space to reduce, rather than expand, the capacity for bank lending to the real economy. [20] The potential crowding out of lending to firms is especially important in light of QE's renewed expansion in the aftermath of the Covid-19 crisis, where reserves increased from \$1.72 trillion in February to \$3.22 trillion in May 2020.

Quantitatively, the response in corporate loans makes up 13% of the size of the reserve supply increase. The response in deposits and mortgages change considerably less than the \$4.24 trillion increase in reserve holdings, even though much of the 15 basis point increase in reserve yields are passed through to the interest rates banks choose. In table 8 we report these changes in rates and quantities. Deposit, mortgage, and loan quantities increase by \$15.4 billion, decrease by \$6.1 billion, and decrease by \$555.9 billion, respectively. The branch-level average of deposit, mortgage, and loan rates increase by 6.193 basis points, 3.857 basis points, and 5.195 basis points, respectively.

One key driver of these magnitudes is that the demand for corporate loans is more price-elastic than that for mortgages and deposits. While the rate drop in corporate loans is only 1 basis point more than that of mortgages, and their price disutility parameters are similar at -556 and -519 respectively, the outside option parameter for mortgages is only .08 while for corporate loans it is .351 in 2007. This implies that the same rate increase leads to a 35/8=4.375 larger change in corporate loan quantities

[20]For example, QE may reduce the yields on long-maturity bonds, which passed through to lower mortgage rates. This reduction in long term yields happens through general equilibrium forces in asset markets that are outside of our model.

than mortgage quantities. Even though the rates charge on loans and mortgages change by similar amounts in the counterfactual, the quantity of corporate loans responds considerably more because of its more elastic demand curve. In addition, although deposit rates move the most (6.193 basis points), their quantities change very modestly. This is because of the inelastic deposit demand curve, which results from both a small price disutility parameter (46), and a small outside option parameter (.29).

This counterfactual also suggests that the traditional business model of commercial banks like deposit-taking and loan-making are relatively disconnected from their activities in the reserves market. In terms of the variables we track, the increased supply of reserves is larger than the changes in any other quantity.[21] A large expansion or contraction of banks' activities in securities markets (e.g., arbitrage trade of borrowing at the Fed funds rate and lending at the IOER rate) can occur with minimal impact on the traditional functions of the banking system. This is consistent with the finding of Anderson et al. (2019) that banks' securities positions or arbitrage trades are primarily financed by borrowing from money market funds.

5.3 Impact of Reserve Injections Year by Year

This section presents the impact on deposit, mortgage, and loan markets of the amount of reserves actually injected by QE in each year. In each year 2008-2017, we present the results of a counterfactual in which the quantity of reserves added is equal to the quantity of reserves actually held by U.S. commerical banks. In figure 4 we show the impact of this reserve injection on bank loan quantities, and in table 9 we report the quanty changes in deposits, mortgages, loans, and reserves.

As in the previous section, the response of mortgage and deposit quantities to an increase in reserve supply is model. However, the reduction in bank lending peaks at over 500 billion dollars in 2015. The gradual growth of this lending reduction from 2008-2015 reflects the fact that reserve balances grew over this period from 43 billion to 2.7 trillion dollars in 2015. While the amount of

[21] Banks may do a mix of selling securities (which are less liquid and money-like than reserves), raising wholesale funding or other debt financing, retaining payouts to equity, and issuing equity.

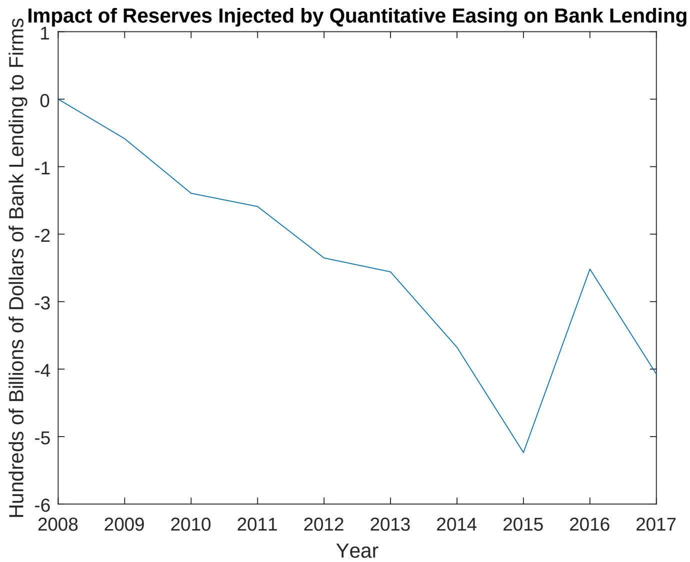

Figure 4: Supply of Central Bank Reserves and Bank Asset Illiquidity

bank lending reduced per dollar of reserves varies somewhat year by year compared to the 13 cents per dollar we state as a baseline result, the average of the ratio of the change in bank lending and the quantity of reserves created is $.1336$. This ratio peaks at 19 cents per dollar in 2015 and is relatively constant from year to year. Our results suggest that ths quantity of reserves injected is the main determinant of the reduction in bank lending. Thus, further expansions in central bank reserve supply may also lead to similar loan contractions as in our baseline estimates.

5.4 Counterfactual Results: The Effect of Demand Shocks

We run additional counterfactual exercises to study the effects of demand shocks to deposits, mortgages, or loans. Specifically, for the counterfactual deposit demand shock, we simulate a 10% increase in the log quantity $\log F_{D,nt}$ of the deposit market in each county. Recall that this quantity includes the households' holdings of both deposits and the outside option. This shock will change the equilibrium rates and quantities of deposits through our estimated demand system, as well as the mortgage and loan quantities through balance sheet synergies. We hold the banks' security holdings constant because the supply of reserves does not change.

Table 10 reports the equilibrium result in the counterfactual scenario obtained from 2007 data. A positive demand shock in the deposit market raises equilibrium deposit quantities and lowers equilibrium deposit rates. Due to the positive synergies between deposits and mortgages/loans, banks are also encouraged to hold more mortgages and loans, which push down the equilibrium mortgage and loan rates.

Similarly, we simulate a 10% increase in the log quantity $\log F_{M,nt}$ of the mortgage market in each county, as well as a 10% increase in the log quantity $\log F_{L,nt}$ of the loan market in each state. They generate qualitatively similar effects, but the loan demand shock generates a quantitatively larger effect because of greater market sizes. Take the loan demand shock for an example. A positive demand shock in the loan market raises equilibrium loan quantities and raises equilibrium loan rates. Note that a lower deposit rate leads to more profits for the bank, whereas a higher loan rate leads to more profits for the bank. Due to the positive synergies between deposits and loans, banks also hold more deposits, which push up the equilibrium deposit rates.

6 Conclusion

This paper develops and estimates a structural model of the U.S. banking system and uses the model to analyze the transmission of central bank policies, such as quantitative easing. We provide the first framework that captures two important determinants of the policy impact on the quantity and price of loans, mortgages, and deposits supplied by the banking sector to the real economy. The first one concerns the demand elasticity banks face in their respective deposit and loan markets. The second one is the synergy between the various components of bank balance sheets motivated by a large theoretical literature.[22] We find that a \$4.76 trillion increase in the supply of central bank reserves increases deposit supply by \$15.4 billion but crowds out lending by \$562 billion. Our findings suggest that the synergies between liquid and illiquid assets on bank balance sheets are limited so that

[22] See for example Diamond and Rajan (2000); Kashyap et al. (2002); Hanson et al. (2015) and Diamond (2019).

an increase in the supply of reserves crowds out rather than crowds in illiquid assets such as loans and mortgages. The limited increase in deposits further reflects how a highly inelastic retail deposit demand constrains the expansion in funding for banks.

One main challenge in the evaluation of central bank policy is their endogenous nature. For example, quantitative easing by the Federal Reserve was implemented in response to the 2008 financial crisis, which directly affected banks through the demand for loans and mortgages, amongst others. To this end, the identification of our structural model only relies on cross-sectional variation exogenous to changes in the time series. The demand systems are identified using demand shocks from natural disasters to bank branches in one region, which transmit through banks' internal capital markets to become supply shocks for branches in other regions. For estimating the supply-side cost function, we use shocks from natural disasters at the bank level as well as a Bartik instrument for deposit demand.

Imperfect competition in deposit and loan markets and the synergies between banks' assets and liabilities not only affect the transmission of quantitative easing but influence banks' decision making in general. Our framework can be further applied and extended to address a number of important questions. Future work may explore the effect of dynamic considerations, especially regarding the costly issuance of bank equity. With available data, bank balance sheets may also be studied at a more granular level. For example, different types of securities may bear varying degrees of liquidity. The composition of the outside option to borrowing from banks (e.g. not borrowing versus borrowing via bond markets) is another promising avenue of future research.

References

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

Alyssa Anderson, Wenxin Du, and Bernd Schlusche. Arbitrage capital of global banks. Technical report, Working Paper University of Chicago, 2019.

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.

Greg Buchak. Financing the gig economy. *University of Chicago Booth Working Paper*, 2018.

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.

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.

Marco Di Maggio, Amir Kermani, and Christopher Palmer. How Quantitative Easing Works: Evidence on the Refinancing Channel. *Review of Economic Studies*, 87(3):1498–1528, 2020.

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.

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.

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.

Arvind Krishnamurthy and Annette Vissing-Jorgensen. The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy (Brookings Papers on Economic Activity, 2011, No. 2). *Brookings Papers on Economic Activity*, (2):215–287, 2011.

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

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.

7 Appendix: Tables

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 as explained in Section 3.3. The sample period is from 2001 to 2017.

	Num. of Obs.	Mean	25th Pct.	50th Pct.	75th Pct.	Std. Dev.
Log Deposit Market Share	80441	−2.67	−3.45	−2.33	−1.50	1.69
Deposit Volume	80441	202.41	23.54	48.80	105.85	2691.91
Deposit Rate	50897	53.62	9.00	19.50	71.71	75.36
Deposit Instrument	80441	2.54	0.00	0.81	3.66	3.58
Log Mortgage Market Share	35316	−3.90	−5.07	−3.50	−2.35	2.08
Mortgage Volume	35316	24.53	1.25	3.86	11.81	219.03
Mortgage Rate	10603	469.23	337.50	476.67	578.13	126.01
Mortgage Instrument	35316	2.94	0.16	1.46	4.15	3.69
Log Loan Market Share	27761	−4.96	−6.46	−4.67	−3.23	2.16
Loan Volume	27761	5022.61	206.16	900.00	3703.75	13846.27
Loan Spread	27761	175.42	99.15	161.56	235.00	117.04
Loan Instrument	27761	7.09	1.95	5.02	12.42	5.60

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 (7) and (8). 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.

Panel (a): First Stage Panel Regression

	(1) Deposit Rate	(2) Mortgage Rate	(3) Loan Rate
IV	1.64***	10.93***	1.88***
	(0.22)	(2.15)	(0.29)
Loan Loss Provision	106.58***	−161.53	133.83*
	(33.46)	(118.57)	(75.87)
Lag Deposit Market Share	1.57***		
	(0.44)		
Lag Insured Deposit Ratio	44.07***		
	(10.02)		
Log Property Damage	−4.58***	−2.74***	
	(1.00)	(0.62)	
Observations	234,857	70,519	23,829
R^2	0.83	0.91	0.20
Adjusted R^2	0.78	0.85	0.16

Panel (b): 2SLS Panel Regression

	(1) Deposit Market Share	(2) Mortgage Market Share	(3) Loan Market Share
Rate (with IV)	46.45***	−556.81***	−519.04***
	(9.49)	(96.72)	(82.94)
Loan Loss Provision	−1.58***	−9.80	7.15
	(0.27)	(8.38)	(4.86)
Lag Deposit Market Share	0.91***		
	(0.01)		
Lag Insured Deposit Ratio	−0.32***		
	(0.05)		
Log Property Damage	0.11***	0.76***	
	(0.01)	(0.06)	
Observations	234,857	70,519	23,829
R^2	0.98	-0.78	-5.65
Adjusted R^2	0.97	-1.96	-5.92

Table 3: Outside Option Estimates (Deposits and Mortgages)

This table reports the two-stage least squares results for estimating the sensitivity of market-level quantities to the aggregate observed desirability parameter $\delta^{o}_{p,nt}$ for deposits and mortgages as in Equations (12) and (13). The regression is run at the market-year level. We include market-year level controls, including the average age and income of the population, the fraction of residents college degree, the log population, the 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 from 2001 to 2017. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Panel (a): First Stage Panel Regression

	(1)	(2)
	Deposit Rate	Mortgage Rate
IV	0.01^{***}	-0.78^{***}
	(0.002)	(0.20)
Age	0.003^{**}	0.09^{***}
	(0.001)	(0.03)
Income	0.003	0.47^{***}
	(0.003)	(0.02)
College	-0.001	0.08^{***}
	(0.003)	(0.03)
Population	0.001	0.40^{***}
	(0.01)	(0.03)
House Price Growth	0.0001	0.12^{***}
	(0.003)	(0.04)
Log Property Dmg	0.10^{***}	0.63^{***}
	(0.0003)	(0.01)
Lag Log Deposit	0.01^{***}	
	(0.003)	
Observations	41,773	24,701
R^2	0.93	0.87

Panel (b): 2SLS Panel Regression

	(1)	(2)
	Deposit Share	Mortgage Share
δ^o (with IV)	0.29^{**}	0.08^{*}
	(0.13)	(0.04)
Age	0.003^{*}	-0.02^{***}
	(0.002)	(0.01)
Income	0.002	0.36^{***}
	(0.001)	(0.02)
College	0.01^{***}	0.17^{***}
	(0.002)	(0.02)
Population	0.03^{***}	1.25^{***}
	(0.01)	(0.01)
House Price Growth	0.01^{***}	0.06^{*}
	(0.002)	(0.04)
Log Property Dmg	-0.03^{**}	-0.06^{**}
	(0.01)	(0.03)
Lag Log Deposit	0.98^{***}	
	(0.01)	
Observations	41,773	24,701
R^2	1.00	0.90

Table 4: Outside Option estimates (Loans)

This table reports the outside option size for loans as described in Subsection 3.2 in trillions of dollars. The Implied β_o is obtained following Subsection 3.1.2.

Year	Size of Outside Option	Implied β_o
2001	0.75	0.42
2002	0.79	0.46
2003	0.85	0.50
2004	0.75	0.37
2005	0.76	0.34
2006	0.83	0.33
2007	1.00	0.35
2008	1.61	0.66
2009	1.90	0.78
2010	1.56	0.59
2011	1.18	0.39
2012	1.30	0.46
2013	1.15	0.35
2014	1.23	0.37
2015	1.51	0.42
2016	1.58	0.43
2017	1.56	0.39
2018	1.50	0.36

Table 5: Summary Statistics (Bank-Year Level)

This table reports summary statistics for deposits, mortgages, and loans at the bank level. Markups and marginal costs are defined in Equations (15) respectively. Marginal costs and mark-up are in basis points. Bank-level volumes are normalized by the number of markets and denoted in millions. The instrument refers to property losses due to natural disasters as explained in Subsection 4.4. The sample period is from 2001 to 2017. *, **, and *** denote significance at the 10%, 5%, and 1% level, respectively.

	Num. of Obs.	Mean	25th Pct.	50th Pct.	75th Pct.	Std. Dev.
Deposit Volume per Branch	119860	198.87	35.96	64.28	117.81	2413.72
Deposit Markup	52564	−250.24	−257.72	−234.85	−222.14	52.00
Deposit Cost	52564	−336.25	−387.40	−302.97	−253.55	108.15
Mortgage Volume per Branch	119874	78.26	15.21	34.40	71.67	462.43
Mortgage Markup	11113	19.30	18.14	18.60	19.62	2.21
Mortgage Cost	11113	474.57	333.01	501.69	584.09	136.45
Loan Volume per Branch	119874	93.46	6.79	12.86	24.07	1927.00
Loan Markup	2841	59.64	40.54	49.59	63.36	65.44
Loan Cost	2841	162.10	93.33	155.59	225.33	126.47
Securities Volume per Branch	119874	60.40	6.63	12.28	23.46	1165.65
Sheldus Instrument	119874	5.58	2.40	5.10	8.83	4.02
Bartik Instrument	62281	1.05	1.02	1.04	1.08	0.06

Table 6: Cost Function Estimate

This table reports the sensitivity of bank-level costs and quantities to losses from natural disasters and a bartik deposit shock as in Equations 28 to 34. Sheldus Instrument refers to property losses due to natural disasters as explained in Subsection 4.4. Bartik Deposit Instrument refers to a bartik-style instrument of deposit growth as explained in Subsection 4.4. 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

	Deposit Cost	Mtg Cost	Loan Cost	Deposit Vol	Mtg Vol	Loan Vol	Security Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sheldus Instrument	-1.09^{***}	1.18^{***}	1.98^{***}	11.11^{***}	1.09^{***}	8.84^{***}	3.62^{***}
	(0.10)	(0.20)	(0.66)	(1.77)	(0.33)	(1.40)	(0.81)
Loan Loss Provision	-1.13	-16.52^{***}	5.06^{**}	8.10^{**}	27.00^{***}	536.38^{***}	1.13
	(1.27)	(2.96)	(2.39)	(3.81)	(4.18)	(17.48)	(1.74)
Observations	52,564	11,113	2,841	118,942	119,236	119,236	118,923
R^2	0.60	0.76	0.20	0.002	0.002	0.01	0.001

Panel (b): Results using Bartik Deposit Shock

	Deposit Cost	Mtg Cost	Loan Cost	Deposit Vol	Mtg Vol	Loan Vol	Security Vol
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Bartik Deposit Instrument	62.98^{***}	-44.41^{***}	-1.70	$1,402.46^{***}$	369.69^{***}	332.85^{***}	432.46^{***}
	(5.18)	(13.38)	(42.17)	(174.67)	(18.68)	(45.82)	(86.59)
Loan Loss Provision	-0.03	-16.29^{***}	6.10	30.97	26.16^{***}	161.98^{***}	-16.84
	(1.24)	(3.21)	(7.84)	(36.48)	(4.43)	(10.86)	(18.08)
Observations	49,095	9,074	2,273	62,104	62,209	62,209	62,098
R^2	0.47	0.74	0.22	0.002	0.01	0.005	0.001

Table 7: 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.

K_1	K_2	K_3	K_4	K_5	ω_M	ω_L
0.0060	0.0143	−0.0053	0.0239	−0.0304	−0.0022	0.1444

Implied Hessian $H(D, M, L, Q)$			
0.0616	−0.0365	−0.0330	−0.0125
−0.0365	0.0060	0.0060	0.0060
−0.0330	0.0060	0.0063	0.0081
−0.0125	0.0060	0.0081	0.0203

Table 8: Counterfactual Results: QE

This table reports the results of the counterfactual analysis in Section 5, where we compute the equilibrium response to a hypothetical \$4.24 trillion increase in central bank reserves in the U.S. banking system in 2007. Rates are in basis points and quantities are in trillions.

Average Change in Rates (in Basis Points)				Total Change in Quantities (in Trn Dollars)			
Deposits	Mortages	Loans	Securities	Deposits	Mortages	Loans	Securities
6.297	3.868	5.207	15.0000	0.0162	−0.0060	−0.5343	4.2387

Table 9: Impacts of QE reserve injections year by year

This table reports the impact in each year of the quantity of reserves existing in each year relative to a world in which reserve balances where 0 in trillions of dollars.

Deposits	Mortages	Loans	Reserves	Year
0.0000	0.0000	0.0000	0.0430	2008
0.0060	-0.0011	-0.0586	0.8604	2009
0.0049	-0.0013	-0.1395	1.1128	2010
0.0043	-0.0001	-0.1591	1.1106	2011
0.0060	-0.0002	-0.2353	1.6198	2012
0.0029	-0.0001	-0.2559	1.6376	2013
0.0199	-0.0001	-0.3677	2.5573	2014
0.0029	-0.0001	-0.5237	2.7497	2015
0.0061	-0.0001	-0.2518	1.4406	2016
0.0004	-0.0001	-0.4075	2.1614	2017

Table 10: Counterfactual Results: Demand Shock

This table reports the results of the counterfactual analysis in Section 5, where we compute the equilibrium response to a hypothetical $\exp(10\%)$ increase in the market size of deposits, mortgages, or loans in 2007. Rates are in basis points and quantities are in trillions.

Panel (a) Deposit Shock							
Average Change in Rates (in Basis Points)				Total Change in Quantities (in Trn Dollars)			
Deposits	Mortages	Loans	Securities	Deposits	Mortages	Loans	Securities
-1.6048	-1.8552	-1.6309	-0.2626	0.2904	0.0041	0.2711	0.0000
Panel (b) Mortgage Shock							
Average Change in Rates (in Basis Points)				Total Change in Quantities (in Trn Dollars)			
Deposits	Mortages	Loans	Securities	Deposits	Mortages	Loans	Securities
0.2201	0.0442	0.0308	0.0337	0.0012	0.0385	-0.0039	0.0000
Panel (c) Loan Shock							
Average Change in Rates (in Basis Points)				Total Change in Quantities (in Trn Dollars)			
Deposits	Mortages	Loans	Securities	Deposits	Mortages	Loans	Securities
3.0394	0.5039	0.7977	0.7512	0.0190	-0.0009	0.5370	0.0000

8 Appendix: Estimation Strategy and Derivations

8.1 Mortgage and Loan FOC's

$$-\frac{\partial Q_{M,nmt}}{\partial Q_{S,mt}}\frac{\partial}{\partial Q_{M,nmt}}(R_t^{M,m} - R_{M,nmt} - \frac{Q_{M,nmt}}{\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}}) = \frac{\partial^2 C(\Theta_{mt})}{\partial Q_{M,nmt}\partial Q_{S,mt}} + \sum_{Y,n'}\frac{\partial^2 C(\Theta_{mt})}{\partial Q_{M,nmt}\,\partial Q_{Y,n'mt}}\frac{\partial Q_{Y,n'mt}}{\partial Q_{S,mt}} \quad (26)$$

$$-\frac{\partial Q_{L,nmt}}{\partial Q_{S,mt}}\frac{\partial}{\partial Q_{L,nmt}}(R_t^{L,m} - R_{L,nmt} - \frac{Q_{L,nmt}}{\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}}) = \frac{\partial^2 C(\Theta_{mt})}{\partial Q_{L,nmt}\partial Q_{S,mt}} + \sum_{Y,n'}\frac{\partial^2 C(\Theta_{mt})}{\partial Q_{L,nmt}\,\partial Q_{Y,n'mt}}\frac{\partial Q_{Y,n'mt}}{\partial Q_{S,mt}} \quad (27)$$

8.2 Market Size

This section provides an alternative expression for $\delta^o_{P,nt}$, with which we can compute it in the presence of missing data. Dividing equation 6 by equation 9 yields

$$\frac{Q_{P,nmt}}{Q_{P,nt}} = \frac{\exp(\alpha_P R_{P,nmt} + X_{P,nmt}\beta_P + \delta_{P,nmt})}{\sum_{m'}\exp(\alpha_P R_{P,nm't} + X_{P,nm't}\beta_P + \delta_{P,nm't})}$$

However, $\delta_{P,nt} = \delta^o_{P,nt} + \delta^u_{P,nt}$ was defined so that $exp(\delta_{P,nt}) = \sum_{m'}\exp(\alpha_P R_{P,nm't} + X_{P,nm't}\beta_P + \delta_{P,nm't})$. It follows that

$$\log\frac{Q_{P,nmt}}{Q_{P,nt}} \quad = \quad \alpha_P R_{P,nmt} + X_{nmt}\beta_P + \delta_{P,nmt} - \delta^u_{P,nt} - \delta^o_{P,nt}.$$

If we average this expression across all observations in a market we get

$$\frac{1}{M_{P,nt}}\sum_m \log(\frac{Q_{P,nmt}}{Q_{P,nt}}) = \frac{1}{M_{nt}}\sum_m(\alpha_P R_{P,nmt} + X_{nmt}\beta_P) - \delta^o_{P,nt}$$

since $\delta^u_{P,nt}$ is defined to equal the mean of the $\delta_{P,nmt}$ in its market, so $(\delta_{D,nmt} - \delta^u_{D,nt})$ is mean zero. This implies

$$\delta^o_{P,nt} = \frac{1}{M_{P,nt}} \sum_m (\alpha_P R_{P,nmt} + X_{nmt}\beta_P) - \frac{1}{M_{P,nt}} \sum_m \log(\frac{P_{nmt}}{P_{nt}}).$$

The two terms in this depression foor expressions are averages of quantities within a market. We average the first term over only observations that have data on interest rates and covariates. We average the second term over all observations, including those missing interest rates or covariates.

8.3 Cost Function

We regress all of the marginal costs and quantities on our two instruments z^i_{mt} (indexed by $i = 1, 2$).

$$\frac{1}{N_{mt}} \sum_n \left(\frac{\partial C}{\partial Q_{D,nmt}} - R_t \right) = \theta^D_t + \kappa^{i,D} z^i_{mt} + u^Q_{D,mt} \quad (28)$$

$$\frac{1}{N_{mt}} \sum_n \left(\frac{\partial C}{\partial Q_{M,nmt}} + R^{M,m}_t \right) = \theta^M_t + \kappa^{i,M} z^i_{mt} + u^Q_{L,mt} \quad (29)$$

$$\frac{1}{N_{mt}} \sum_n \left(\frac{\partial C}{\partial Q_{L,nmt}} + R^{L,m}_t \right) = \theta^L_t + \kappa^{i,L} z^i_{mt} + u^Q_{L,mt} \quad (30)$$

and

$$Q_{D,mt} = \alpha^D_t + \gamma^{i,D} z^i_{mt} + \varepsilon^Q_{D,mt} \quad (31)$$

$$Q_{M,mt} = \alpha^M_t + \gamma^{i,M} z^i_{mt} + \varepsilon^Q_{M,mt} \quad (32)$$

$$Q_{L,mt} = \alpha^L_t + \gamma^{i,L} z^i_{mt} + \varepsilon^Q_{L,mt} \quad (33)$$

$$Q_{S,mt} = \alpha^S_t + \gamma^{i,S} z^i_{mt} + \varepsilon^Q_{S,mt} \quad (34)$$

The coefficients from regressions 28 to 34 solve a system of equations that identifies our cost function:

$$\kappa^{i,D} = -K_1\gamma^{i,\mathcal{E}} + K_3\gamma^{i,D} + K_4\gamma^{i,\mathcal{I}} + K_5[\gamma^{i,\mathcal{E}} - \gamma^{i,D}] \quad (35)$$

$$\kappa^{i,M} = K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}}\omega_M + K_4\gamma^{i,D}\omega_M + K_5\gamma^{i,D} \quad (36)$$

$$\kappa^{i,L} = K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}}\omega_L + K_4\gamma^{i,D}\omega_L + K_5\gamma^{i,D} \quad (37)$$

$$0 = K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}} + K_4\gamma^{i,D} + K_5\gamma^{i,D} \quad (38)$$

where $\gamma^{i,\mathcal{E}} = \gamma^{i,Q} + \gamma^{i,M} + \gamma^{i,L} - \gamma^{i,D}$ and $\gamma^{i,\mathcal{I}} = \gamma^{i,Q} + \omega_M\gamma^{i,M} + \omega_L\gamma^{i,L}$. The final equation has a left hand side of 0 because it represents the rate of return that banks earn on a securities investment, for which there is no cross-sectional variation across banks.

This yields a system of 8 equations which we use to identify the 7 parameters of our cost function. To see why we only are able to estimate 7 parameters, re-organize these equations to get

$$\kappa^{i,M} = (K_2\gamma^{i,\mathcal{I}} + K_4\gamma^{i,D})(\omega_M - 1) \quad (39)$$

$$\kappa^{i,L} = (K_2\gamma^{i,\mathcal{I}} + K_4\gamma^{i,D})(\omega_L - 1) \quad (40)$$

which implies [23]

$$\omega_L = 1 + \frac{\kappa^{i,L}}{\kappa^{i,M}}(\omega_M - 1) \quad (41)$$

which yields a relationship between ω_L and ω_M separately from each instrument. We average these two equations and plug in

$$\omega_L = 1 + \frac{1}{2}\left(\frac{\kappa^{1,L}}{\kappa^{1,M}} + \frac{\kappa^{2,L}}{\kappa^{2,M}}\right)(\omega_M - 1) \quad (42)$$

The remaining 6 parameters of the model are now computed by solving an exact solution to the

[23] This overidentifying restriction 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.

remaining system of 6 equations

$$\begin{aligned}\kappa^{i,D} &= -K_1\gamma^{i,\mathcal{E}} + K_3\gamma^{i,D} + K_4\gamma^{i,\mathcal{I}} + K_5[\gamma^{i,\mathcal{E}} - \gamma^{i,D}] && (43)\\ \kappa^{i,M} &= K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}}\omega_M + K_4\gamma^{i,D}\omega_M + K_5\gamma^{i,D} && (44)\\ 0 &= K_1\gamma^{i,\mathcal{E}} + K_2\gamma^{i,\mathcal{I}} + K_4\gamma^{i,D} + K_5\gamma^{i,D}. && (45)\end{aligned}$$

8.4 Counterfactual

8.4.1 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 $\boldsymbol{Q}_{D,nt} = \sum_m Q_{D,nmt}$. Let δ_{nmt} denote the desirability of its deposit:

$$\delta_{nmt} = \alpha_D R_{D,nmt} + X_{nmt}\beta_D + \delta_{D,nmt} \tag{46}$$

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

$$Q_{D,nmt} = \boldsymbol{Q}_{D,nt}\frac{\exp(\delta_{nmt})}{\sum_{m'}\exp(\delta_{nm't})}. \tag{47}$$

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

$$\frac{\partial \log \boldsymbol{Q}_{D,nt}}{\partial \delta^o_{D,nt}} = \beta_o \tag{48}$$

which implies that

$$
\begin{aligned}
Q_{D,nt} &= Q^i_{D,nt} \exp(\Delta f_{D,nt}) \exp(\beta_o(\delta^o_{D,nt} - \delta^{o,i}_{nt})) && (49)\\
&= Q^i_{D,nt} \exp(\beta_o(\log \sum_{m'} \exp(\delta_{nm't}) - \log \sum_{m'} \exp(\delta^i_{nm't}))) && (50)
\end{aligned}
$$

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

Then,

$$
Q_{D,nmt} = Q_{D,nt} \frac{\exp(\delta_{nmt})}{\sum_{m'} \exp(\delta_{nm't})} = Q^i_{D,nt} \exp(\Delta f_{D,nt}) \frac{(\sum_{m'} \exp(\delta_{nm't}))^{\beta_o - 1}}{(\sum_{m'} \exp(\delta^i_{nm't}))^{\beta_o}} \exp(\delta_{nmt}). \qquad (51)
$$

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):

$$
\delta_{nmt} = \delta^i_{nmt} + \alpha_D(r_{nmt} - r^i_{nmt}) \qquad (52)
$$

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

$$
Q_{D,nmt} = Q^i_{D,nt} \exp(\Delta f_{D,nt}) \frac{(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})))^{\beta_o - 1}}{(\sum_{m'} \exp(\delta^i_{nm't}))^{\beta_o}} \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})). \qquad (53)
$$

8.4.2 Marginal Cost from Optimality Condition

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

$$R_{D,nmt} \quad = \quad R_t^D - \frac{Q_{D,nmt}(R_{D,nmt})}{Q'_{D,nmt}(R_{D,nmt})} - \frac{\partial C(Q_{D,nmt}(R_{D,nmt}), \ldots)}{\partial Q_{D,nmt}}. \tag{54}$$

Because

$$\log(Q_{D,nmt}) \quad = \quad \log(Q^i_{D,nt}) + \Delta f_{D,nt} + (\beta_o - 1)\log(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))) \tag{55}$$

$$- \quad \beta_o \log(\sum_{m'} \exp(\delta^i_{nm't})) + (\delta^i_{nmt} + \alpha(\Delta r_{nmt})). \tag{56}$$

we have

$$\frac{\partial \log(Q_{D,nmt})}{\partial \Delta r_{nmt}} = \alpha + \alpha(\beta_o - 1)\frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} \tag{57}$$

This implies

$$\frac{\partial C}{\partial Q_{D,nmt}} \quad = \quad R_t^D - \left[\frac{\partial \log(Q_{D,nmt})}{\partial r_{nmt}}\right]^{-1} - R_{D,nmt} \tag{58}$$

$$= \quad R_t^D - \left[\alpha + \alpha(\beta_o - 1)\frac{exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right]^{-1} - R_{D,nmt} \tag{59}$$

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

$$\frac{\partial C}{\partial Q_{D,nmt}} - \frac{\partial C^i}{\partial Q_{D,nmt}} \quad = \quad \left[\alpha + \frac{\alpha(\beta_o - 1)\exp(\delta^i_{nmt})}{\sum_{m'} \exp(\delta^i_{nm't})}\right]^{-1} \tag{60}$$

$$- \quad \left[\alpha + \frac{\alpha(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))}\right]^{-1} - \Delta r_{nmt}$$

8.4.3 Jacobian of marginal cost from optimality condition

For numerical accuracy, the Jacobian of Eq. (60) 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

$$\begin{aligned}
\frac{\partial}{\partial \Delta r_{nm^*t}} \frac{\partial C}{\partial Q_{D,nmt}} &= \frac{\partial}{\partial r_{nm^*t}} \left(- \left[\alpha + \frac{\alpha(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} \right]^{-1} - \Delta r_{nmt} \right) \\
&= - \left[1 + \frac{(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} \right]^{-2} \\
&\cdot \quad (\beta_o - 1) \left(\frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))\exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t}))}{(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})))^2} \right. \\
&+ \quad \left. 1_{\{m=m^*\}} \frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} \right) - 1_{\{m=m^*\}} \\
&= - \left[1 + \frac{(\beta_o - 1)\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} \right]^{-2} (\beta_o - 1) \\
&\cdot \quad \frac{\exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} \left(\frac{\exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t}))}{\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't}))} + 1_{\{m=m^*\}} \right) - 1_{\{m=m^*\}}.
\end{aligned} \tag{61}$$

8.4.4 Appendix: Computation of Counterfactual

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 60. 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 60 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 61.

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 53 across markets.

$$Q_{D,mt} = \sum_n Q_{D,nmt} = \sum_n Q^i_{D,nt} \frac{(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})))^{\beta_o - 1}}{(\sum_{m'} \exp(\delta^i_{nm't}))^{\beta_o}} \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})) \quad (62)$$

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

$$Q_{M,mt} = \sum_n Q_{M,nmt} = \sum_n Q^i_{M,nt} \frac{(\sum_{m'} \exp(\delta^{i,M}_{nm't} + \alpha^M(\Delta r_{nm't})))^{\beta^M_o - 1}}{(\sum_{m'} \exp(\delta^{i,M}_{nm't}))^{\beta^M_o}} \exp(\delta^{i,M}_{nmt} + \alpha^M(\Delta r^Q_{M,nmt})) \quad (63)$$

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

$$\begin{aligned} & \frac{\partial}{\partial \Delta r_{nm^*t}} D_{mt} \qquad (64) \\ = \; & (\beta_o - 1)\alpha Q^i_{D,nt} \frac{(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})))^{\beta_o - 2}}{(\sum_{m'} \exp(\delta^i_{nm't}))^{\beta_o}} \exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t})) \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})) \\ + \; & 1_{\{m=m^*\}} \alpha Q^i_{D,nt} \frac{(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})))^{\beta_o - 1}}{(\sum_{m'} \exp(\delta^i_{nm't}))^{\beta_o}} \exp(\delta^i_{nmt} + \alpha(\Delta r_{nmt})) \\ = \; & \alpha Q^i_{D,nt} \left((\beta_o - 1) \left(\sum_{m'} \exp(\delta^i_{nm't} + \alpha(\Delta r_{nm't})) \right)^{-1} \exp(\delta^i_{nm^*t} + \alpha(\Delta r_{nm^*t})) + 1_{\{m=m^*\}} \right). \end{aligned}$$

Thus, $f_R = f^R_1 \circ f_2 \circ f_3 \circ f_4$ 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.

8.5 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

$$\begin{aligned}
&\Pi_{m,t+1} = \\
&\sum_n Q_{L,nmt}(1+R_{L,nmt}) + \sum_n Q_{M,nmt}(1+R_{M,nmt}) + Q_{S,mt}(1+R_{S,t}) - \sum_n Q_{D,nmt}(1+R_{D,nmt}) \\
-&\left(\sum_n Q_{L,nm,t+1} + \sum_n Q_{M,nm,t+1} + Q_{S,m,t+1} - \sum_n Q_{D,nm,t+1}\right) - C(\Theta_{mt}),
\end{aligned} \tag{65}$$

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

$$\max_{(R_{D,nmt}, R_{M,nmt}, R_{L,nmt}, Q_{mt})} \sum_{j=0}^{\infty} \mathbb{E}_t[\Lambda_{t,t+j}\Pi_{m,t+j}] \tag{66}$$

subject to equation 65. 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 [24]

$$\begin{aligned}
\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}} &= \frac{1}{1+R_t^{D,m}}\left(\frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}(1+R_{D,nmt}) + Q_{D,nmt} + \frac{\partial Q_{D,nmt}}{\partial R_{D,nmt}}\frac{\partial C(\Theta_{mt})}{\partial Q_{D,nmt}}\right), \quad (67)\\
\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}} &= \frac{1}{1+R_t^{L,m}}\left(\frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}(1+R_{L,nmt}) + Q_{L,nmt} - \frac{\partial Q_{L,nmt}}{\partial R_{L,nmt}}\frac{\partial C(\Theta_{mt})}{\partial Q_{L,nmt}}\right), \quad (68)\\
\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}} &= \frac{1}{1+R_t^{M,m}}\left(\frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}(1+R_{M,nmt}) + Q_{M,nmt} - \frac{\partial Q_{M,nmt}}{\partial R_{M,nmt}}\frac{\partial C(\Theta_{mt})}{\partial Q_{M,nmt}}\right), \quad (69)\\
1 &= \frac{1}{1+R_t^{S,m}}\left((1+R_{S,t}) - \frac{\partial C(\Theta_{mt})}{\partial Q_{S,mt}}\right). \quad (70)
\end{aligned}$$

These are equivalent to equations 2-5.

[24]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 $\Lambda_{t,t+j}$.