The Network Structure of Money Multiplier*  
Ye Li† Yi Li‡ Huijun Sun§

First Draft: December 29, 2021 | Current Version: December 14, 2022

Abstract

The money view of banking emphasizes deposits as part of money supply that central banks can manage to influence the macroeconomy (Friedman and Schwartz, 1963). The credit view focuses instead on the asset side of bank balance sheets and proposes a bank lending channel (Bernanke, 1983). We show that by connecting monetary base and credit supply, the monetary role of deposits is in fact key to the credit channel. The circulation of deposits as means of payment reshuffles reserves—the settlement assets—among banks. A bank’s position in the payment network determines its liquidity risk and willingness to invest in illiquid loans. We develop a structural model of liquidity percolation in the payment system and characterize an equilibrium money multiplier that links the churn of reserves and the creation of bank credit and deposits. Based on granular payment data, our estimation reveals a small set of systemically important banks that have disproportionately large impact on the equilibrium outcome.

Keywords: Credit supply, money multiplier, payment, network, externalities, systemic risk

*We are grateful to helpful comments from Tobias Berg, Juliane Begenau, Patrick Bolton, Martin Brown, Yasser Boualam, Markus Brunnermeier, Dean Corbae, Mark Egan, Co-Pierre Georg, Ben Golub, Simon Gilchrist, Michael Gofman, Florian Heider, Bernard Herskovic, Peter Hoffmann, Yannick Kalantzis, Arvind Krishnamurthy, Ricardo Lagos, Hanno Lustig, Ouarda Merrouche, Andreas Schaab, Martin Schneider, René Stulz, Avanidhar (Subra) Subrahmanymam, Pierre-Olivier Weill, Haoxiang Zhu, and conference and seminar participants at Advances in Macro-Finance at CMU Tepper-LAEF, Barcelona School of Economics Summer Forum, Boston College, CESifo Conference on Macro Money and International Finance, European Banking Center Network Conference at Tilburg University, European Finance Association, European Central Bank, Frankfurt School of Finance & Management, Imperial College London, NYU Econ/Stern macro seminar, Stanford GSB finance seminar, UCLA Anderson finance seminar, UNC Junior Roundtable, University of Zurich. The views expressed herein are those of the authors and do not necessarily reflect those of the Federal Reserve Board or its staff. This paper was previously circulated under the title “Bank Credit and Money Creation on Payment Networks: A Structural Analysis of Externalities and Key Players”.

†University of Washington, Foster School of Business. E-mail: liye@uw.edu
‡Board of Governors of the Federal Reserve System. E-mail: yi.li@frb.gov
§Columbia University. E-mail: hsun25@gsb.columbia.edu
1 Introduction

Does money supply matter? To answer this question, the classic approach examines connections between changes in the stock of money and macroeconomic outcome (Friedman and Schwartz, 1963). However, this approach faces the great challenge of defining money (Tobin, 1965). What assets should we include? For monetary policy, the literature shifted focus away from monetary quantities and towards interest rate. From Patinkin’s *Money, Interest, and Prices* (1965) to Woodford’s *Interest and Prices* (2003), “money” is dropped as we struggled for measurement. However, with quantitative easing and tightening becoming regular policy tools and demand for monetary assets rising due to institutional and regulatory changes, it is important to revisit the questions on money supply. How does money supplied by central banks influence the macroeconomy? What is the connection between money supply and credit intermediation by the banking sector?

We tackle these questions from a new angle. Our focus is on the circulation of money rather than quantity, and we examine the unique role of reserves as settlement assets in the payment system. When depositors make payments (i.e., transfer deposits) to one another, settlement requires the payers’ banks to transfer reserves to the payees’ banks. Therefore, banks face liquidity risk as they cannot control their depositors’ payment activities. Payment activities are large in magnitude. The average weekly volume in Fedwire, the primary payment system in the U.S., exceeds the GDP.

The circulation of deposits as means of payment generates a network structure of interbank reserve flows through settlement. We find that the topology of this network is a key determinant of credit supply. A bank’s position in the network of payment flows determines its liquidity risk exposure and thereby affects its willingness to invest in illiquid loans. Moreover, strategic interaction arises in banks’ lending decisions. When one bank decides to finance lending with new deposits, other banks receive reserve and deposit inflows when the lending bank’s depositors make payments to other banks’ depositors. We develop and estimate a structural model that captures
banks’ liquidity risk from payment settlement and the strategic interaction in their lending decisions. In contrast to a reduced-form approach, our structural estimation incorporates the entire payment network rather than network statistics, such as cohesion and centrality measures, and it allows us to conduct counterfactual analysis to demonstrate the importance of network topology.

Our findings add more nuance to the historic debate between money view and credit view of banking. While the former emphasizes deposits as part of money supply (Friedman and Schwartz, 1963), the latter focuses on bank lending (Bernanke, 1983). As depositors pay one another, moving deposits between banks, reserves are transferred accordingly for settlement. The monetary role of deposits exposes banks to the uncertainty in reserve movements. The liquidity risk affects banks’ incentive to invest in illiquid loans. This intrinsic link between money and credit cautions against any dichotomy, particularly efforts to test which force dominates in macroeconomic dynamics.

Such connection between money and credit can be overlooked if one simply examines the relationship between reserve quantity and bank lending. Reserve quantity does not contain information on how reserves circulate in the banking system. And, reserve holdings do not correspond to accessible liquidity for payment settlement as a bank can sell other assets for reserves or borrow reserves from the central bank or other banks. Accessible liquidity depends on the bank’s reserve holdings, its reserve borrowing capacity, and the time-varying market liquidity of other assets.\(^1\)

The equilibrium of our structural model features a modern version of money multiplier. After a bank finances lending with new deposits, payments take place, causing reserves to flow from the lending bank to the payees’ banks. After gaining liquidity, the payees’ banks are willing to expand balance sheets, extending loans financed by deposits. A new round of payment propagates liquidity to other banks including the first bank and thereby creates echo effects. Liquidity percolation through payment has been known to be the foundation of the concept of money multiplier, but

---

\(^1\)Any study on monetary quantities cannot circumvent the substitutability or convertibility of near-money assets (Poterba and Rotemberg, 1986; Lucas and Nicolini, 2015; Krishnamurthy and Vissing-Jørgensen, 2015; Nagel, 2016).
little has been done to formalize and quantify the multiplier effects. We take a structural approach with detailed payment data and characterize the network structure of money multiplier.

Next, we explain the structural model and provide more details on our findings. The model has $N$ banks. Each bank is endowed certain amount of reserves and chooses the amount of loans financed by deposits. Before making lending decisions, the bank is hit by a shock that affects its incentive to lend. The model setup allows the shock to be generic. It may originate from the credit demand side, such as collateral values and the profitability and scalability of borrowers’ projects, or the credit supply side that depends on the bank’s market power and regulatory constraints.

After the bank extends loans and issues deposits, payments take place and a random fraction of deposits flow to the other banks, draining reserves in the process. The bank may receive payment inflows as a result of the other banks expanding balance sheets via lending and deposit-taking. Importantly, any loss of reserves leads to a quadratic cost (increasing and convex). It reflects the fact that it is costly to lose liquidity, which can potentially be saved to buffer future uncertainty unrelated to payment shocks, and the fact that it is costly to replenish liquidity through interbank borrowing or other funding sources. In fact, if the interbank markets operate frictionlessly, a bank in liquidity surplus can always lend to a bank in deficit, costlessly undoing the impact of payment shocks. Therefore, our work builds on the literature on frictions in the interbank markets.\(^2\)

So far, our setup suggests that banks’ lending decisions are strategic complements. If one bank expands balance sheet, its payment outflow has a liquidity spillover effect on the other banks, reducing their marginal cost of losing reserves and increasing their incentive to lend, and vice versa. However, banks’ lending decisions can also be strategic substitutes. As the payees receive payments and hold more liquidity in the form of deposits, their demand for credit from their banks

---

\(^2\)Bhattacharya and Gale (1987) show that interbank markets allow banks to share liquidity risk. However, trading frictions give rise an increasing and convex cost (Afonso and Lagos, 2015; Bigio and Sannikov, 2019). The freeze of interbank market can be interpreted as stronger convexity in the cost of reserve loss, which reduces bank lending in line with the evidence (Iyer, Peydró, da-Rocha-Lopes, and Schoar, 2013; Ippolito, Peydró, Polo, and Sette, 2016).
Thus, one bank’s decision to lend can reduce the amount of loans extended by the other (payees’) banks. These two opposing forces reflect the two-layer design of payment system: Payment between depositors (i.e., an electronic transfer of deposits) happens simultaneously with interbank transfer of reserves (Kahn and Roberds, 2007; Piazzesi and Schneider, 2016).

In sum, the strategic interaction in banks’ lending decisions is modelled through a linear-quadratic game on a random graph of payment flows. When a bank chooses loan volume, it takes into account the randomness in both outflows (reserve and deposit reductions) due to its own depositors’ payments and inflows (reserve and deposit additions) due to payments sent by other banks’ depositors. Under the quadratic objective function, the first and second moments of payment flows enter into banks’ lending decisions. The quadratic objective also implies a linear system of optimal responses. Bank $i$’s lending is linear in bank $j$’s lending. The coefficient of spillover effect can be decomposed into a network effect parameter, $\phi$, and the $ij$-th element of a network adjacency matrix that incorporates the first and second moments of payment flows. $\phi$ captures whether banks’ lending decisions exhibit strategic complementarity ($\phi > 0$) or substitution ($\phi < 0$).

The equilibrium conditions map to a spatial structure, and our estimation strategy follows Denbee, Julliard, Li, and Yuan (2021). When estimating the model, we quantify the probability distribution of reserve flows using data from Fedwire, the major real-time payment settlement system in the U.S. The network adjacency matrix depends on both the first and second moments of the joint distribution of reserve flows between each pair of banks. Without the structure from our model, it is not obvious how the probability distribution enters into banks’ decisions and transforms

---

3It is assumed that a bank’s depositor base at least partially overlaps with the pool of potential borrowers. We model bank depositors’ liquidity management under financial constraints (Froot, Scharfstein, and Stein, 1993; Riddick and Whited, 2009; Bolton, Chen, and Wang, 2011; Décamps, Mariotti, Rochet, and Villeneuve, 2011).

4While payment systems differ in netting efficiency, overdraft standards, and bilateral credit lines (Kahn and Roberds, 1998; Freixas and Parigi, 1998; Bech and Garratt, 2003), banks ultimately settle payments with reserves. The process can be viewed as the deposit holders withdrawing cash to pay and her payee depositing cash into the payee’s bank. The lending bank loses reserves and deposits, while the payee’s bank gains reserves and deposits.
into a network adjacency matrix that summarizes the payment-induced bilateral connections.

As previously discussed, it is challenging to define available reserves. When a bank needs reserves, it can potentially sell other assets. Available reserves depend on the actual amount of reserve holdings and the market liquidity (convertibility) of other bank assets. Therefore, we want to stay agnostic about the definition of available reserves. The model solution features a convenient two-step structure that allows us to do so. In the first step, each bank’s shock and its available reserves enter into the network-independent level of lending, which is the optimal lending when the bank is isolated from its peers. In the second step, the network propagation mechanism transforms banks’ network-independent lending into the equilibrium levels of lending. In our estimation, we treat each bank’s network-independent lending as a random variable, a black box that encapsulates a bank’s available reserves and its shock, and we directly estimate its mean and volatility.\(^5\) Our empirical results focus on the second step in the equilibrium formation — network propagation.

We find that the force of strategic complementarity dominates (i.e., \(\phi > 0\)) and the payment-flow network becomes a shock amplification mechanism. Consider a positive shock to a bank that triggers more lending and deposit issuances. Under \(\phi > 0\), the other banks increase lending in response, which in turn triggers another round of shock propagation. Our estimate of \(\phi\) is statistically significant and large in magnitude. It implies that when all banks are hit by unitary shocks, the payment network amplifies the shock by 17\% to 1.17 for each bank.

To examine the importance of network topology, we compare the mean and volatility of aggregate credit supply in equilibrium with the mean and volatility of credit supply solved under a hypothetical network where banks are equally connected. Under \(\phi > 0\), both our data network and the counterfactual network amplify shocks. The expected levels of credit generated by the two networks are similar with the uniform network slightly outperforming. However, the two networks

\(^5\)Therefore, the shock to each bank can be interpreted as any shock that contributes to the randomness of network-independent lending, for example, a liquidity shock that depletes reserves before the lending decision is made.
differ significantly in generating the volatility of credit supply (20% higher for the data network).

The volatility of aggregate credit supply can be decomposed into individual banks’ contributions, and each bank’s contribution is a product of a network centrality measure, which summarizes the shock propagation routes through the payment linkages, and the volatility of its network-independent lending. We identify banks that contribute the most to the volatility of credit supply and highlight the role of network topology in generating the cross-sectional heterogeneity in banks’ systemic importance. Less than 10% of banks contribute to more than 90% of credit-supply volatility. For these banks, their special positions in the network amplify the impact of their shocks.

Finally, we calculate the planner’s solution. The planner’s objective function is simply the equal-weighted sum of banks’ profits without incorporating the depositors’ and borrowers’ utilities. Therefore, our analysis does not aim for welfare implications but rather focuses on quantifying the impact of network externalities. By comparing the planner’s solution and the market equilibrium, we identify three payment-network externalities. First, individual banks do not internalize the expected costs and benefits of their payment outflows for the payees’ banks. Second, when a bank increases lending, the associated payment-flow uncertainty increases for the payees’ banks. Third, depending on the pair-wise correlation of payment flows, a bank’s lending may offer a hedge against (or a multiplier for) the other banks’ payment-flow risk.

The three forces contribute to the differences in mean and volatility of credit supply between the market equilibrium and the planner’s solution. The planner’s expected level of credit supply is 8.6% higher than that of the market equilibrium, and the planner’s volatility is 20% lower. To the extent that the real sector benefits from a more favorable risk-return trade-off in the supply of bank credit, our analysis indicates that policy interventions, aiming at correcting the payment-flow externalities, can benefit both the borrowers in the real economy and the banks. Our rolling

---

6Payment system reforms involve the design of netting mechanisms, bilateral credit lines between banks, and overdraft at the central bank (see, e.g., Calomiris and Kahn, 1996; Freixas and Parigi, 1998; Kahn and Roberds, 1998;
estimation shows that the mean and volatility wedges between the market equilibrium and planner’s solution are both wider in periods with stronger network effects (higher values of $\phi$).

The planner’s solution and market equilibrium also differ in the distribution of credit provision across banks. Because many borrowers rely on relationship lending, the distribution of credit across banks affects the real economy.\(^7\) Relative to the planner’s solution, the market equilibrium features more dispersed distributions of both the mean and volatility of bank lending. If a borrower can switch between lenders, she may prefer moving towards a lender with a higher expected level of credit provision and less volatility. Our finding suggests that payment-flow externalities amplify the cross-sectional dispersion in the risk-return profiles of banks’ credit provision, making any frictions limiting borrowers’ mobility more costly.\(^8\)

**Literature.** Our paper contributes to the literature on banks as inside money creators. Deposits are financing instruments for the lending bank and means of payment for its customers. Our paper is most related to Gersbach (1998), Freixas, Parigi, and Rochet (2000), Bianchi and Bigio (2014), Parlour, Rajan, and Walden (2020), and Garratt, Yu, and Zhu (2021). These papers theoretically analyze how payment activities affect banks’ liquidity management and generate spillover effects of one bank’s deposit creation through lending on other banks. We make three contributions. First, we use detailed payment data to quantify the network topology and structurally estimate its impact on credit supply, while the existing studies are mostly theoretical or calibrate models with aggregate statistics. Second, we model customer-initiated payment flows as random graphs. The existing studies do not model the randomness or the network structure. As shown theoretically by Martin and McAndrews, 2008; Bech, Chapman, and Garratt, 2010; Bech, Martin, and McAndrews, 2012; Chapman, Gofman, and Jafri, 2019). These measures potentially reshape the payment-flow topology and affect bank lending.

\(^7\)There is a large literature on relationship lending (e.g., Berger and Udell, 1995; Berlin and Mester, 1999; Dahiya, Saunders, and Srinivasan, 2003; Degryse and Ongena, 2005; Bolton, Freixas, Gambacorta, and Mistrulli, 2016).

\(^8\)Firms with less mobility may benefit from building more bank relationships. Firms maintain more bank relationships in countries with inefficient judicial systems (Ongena and Smith, 2000). Loan terms are often better at the onset of a lending relationship and toughen as time goes (Ioannidou and Ongena, 2010).
Bolton, Li, Wang, and Yang (2020) and empirically by Li and Li (2021), payment risk has a strong impact on bank lending. Third, following Piazzesi and Schneider (2016), our model emphasizes the two-layer design of payment system: Payments move both reserves and deposits. As a result, banks’ lending decisions exhibit both strategic complementarity and substitution. Our structural estimation allows us to identify the dominant force and how the relative strength varies over time.

Our paper offers direct evidence that ties banks’ roles as lenders and issuers of inside money (transferable debts as means of payment), providing support to the largely theoretical literature on banks’ dual role of credit and money creators (Gorton and Pennacchi, 1990; Freeman, 1996; Cavalcanti and Wallace, 1999; Azariadis, Bullard, and Smith, 2001; Kiyotaki and Moore, 2000, 2002, 2005; Monnet, 2006; Kahn and Roberds, 2007; Skeie, 2008; Stein, 2012; Gu, Mattesini, Monnet, and Wright, 2013; Bianchi and Bigio, 2014; Hart and Zingales, 2014; Quadrini, 2017; DeAngelo and Stulz, 2015; Jakab and Kumhof, 2015; Krishnamurthy and Vissing-Jørgensen, 2015; Monnet and Sanches, 2015; Brunnermeier and Sannikov, 2016; Li, 2016; Donaldson, Piacentino, and Thakor, 2018; Begena, 2020; Begena, Bigio, Majerovitz, and Vieyra, 2019; Bigio and Sannikov, 2019; d’Avernas, Vandeweyer, and Pariès, 2019; Donaldson and Piacentino, 2019; Wang, 2019; Piazzesi, Rogers, and Schneider, 2019; Faure and Gersbach, 2021).

The monetary role of deposits exposes banks to liquidity risk. After the global financial crisis, banks’ reserve holdings increased dramatically through a variety of mechanisms, for example, quantitative easing. One may argue that reserves are abundant and liquidity risk is no longer a concern. However, recent evidence shows that liquidity shortage in the payment system can still arise and disrupted the short-term funding markets. Several explanations have been suggested, such as liquidity regulations, variations in the Treasury account, changes in central bank behavior.

---

9 Theoretical literature on bank liquidity creation studies the broad implications on risk sharing and intertemporal resource allocation (Bryant, 1980; Diamond and Dybvig, 1983; Diamond, 1984; Ramakrishnan and Thakor, 1984; Millon and Thakor, 1985; Jacklin, 1987; Postlewaite and Vives, 1987; Gorton and Pennacchi, 1990; Allen and Gale, 2004; Goldstein and Pauzner, 2005; Allen, Carletti, and Gale, 2014; Krishnamurthy and Vissing-Jørgensen, 2015).
in accommodating overdrafts, or banks issuing more liquidity claims (such as demand deposits and lines of credit) in response to an increase in reserve supply (Correa, Du, and Liao, 2020; Copeland, Duffie, and Yang, 2021; d’Avernas and Vandeweyer, 2021; Acharya, Chauhan, Rajan, and Steffen, 2022; Acharya and Rajan, 2022; Afonso, Duffie, Rigon, and Shin, 2022; Yang, 2022). Our analysis demonstrates again the importance of bank liquidity management even when the central bank supplies a large quantity of liquidity, and, in particular, shows for the first time that liquidity percolation and redistribution in the payment system drive the aggregate supply of bank credit.

This paper contributes to the literature on network analysis by providing the first empirical analysis of how the network of bank depositors’ payment flows affects credit supply. The existing literature focuses on the network of banks’ transactions (rather than their depositors). These two types of networks are related. Depositors’ payments induce liquidity risk for banks (unexpected reserve loss) that can be mitigated by interbank reserve borrowing/lending (Bhattacharya and Gale, 1987). The resultant trading network has been the focus of the network literature (see Allen and Babus (2009), Glasserman and Young (2016), and Jackson and Pernoud (2021) for a review). In

---

10 Banks trade reserves to smooth out liquidity shocks (e.g., Freixas, Parigi, and Rochet, 2000; Dasgupta, 2004; Afonso and Lagos, 2015; Castiglionesi, Feriozzi, and Lorenzoni, 2019; Parlour, Rajan, and Walden, 2020).

practice, payment activities and reserve transfers take place before interbank reserve trade under real-time gross settlement (RTGS) in most advanced economies (Bech and Hobijn, 2007). Therefore, our paper departs from the focus of the literature on the interbank network of reserve trade and takes a step back to analyze the more primitive network of bank depositors’ payment flows.\textsuperscript{12}

There are three common challenges in network analysis. First, inferring the network linkages requires bilateral transaction data. Second, network linkages are often endogenous to banks’ choices, and it is difficult to model and structurally estimate equilibrium with endogenous network. Third, network linkages may vary over time and exhibit randomness. In our paper, customer-initiated payment flows are directly observed from Fedwire. Moreover, the network of interest is not endogenous to banks’ choices (unlike, for example, the commonly studied interbank network of reserve borrowing and lending) but rather arises from depositors’ payment activities. Finally, we directly model a random network and are able to quantify the joint probability distribution of depositor-initiated payment flows from any given bank to other banks. In fact, our emphasis is precisely on banks’ lack of control over the random directions of payment flows.

In this paper, we find that what matters for bank lending is not only the volatility of individual banks’ payment outflows (as documented by Li and Li (2021)) but the complete network of random payment flows across banks. Our findings on how payment-flow networks affect bank lending contribute to the literature on funding stability and credit supply (Loutskina and Strahan, 2009; Ivashina and Scharfstein, 2010; Cornett, McNutt, Strahan, and Tehranian, 2011; Ritz and Walther, 2015; Dagher and Kazimov, 2015; Carletti, De Marco, Ioannidou, and Sette, 2021).\textsuperscript{13} In terms of

\textsuperscript{12}There are possibly two reasons behind the exclusive focus of literature on interbank networks rather than the network of customers’ payment flows. First, it is difficult to obtain customers’ payment data. Second, before the wide adoption of RTGS, settlement does not necessarily require reserve transfer. For example, in the old deferred net settlement (DNS) system, interbank borrowing/lending relationships can happen simultaneously as customers make payments (banks experiencing payment outflows borrow reserves to settle with banks experiencing inflows).

\textsuperscript{13}The broader literature on funding stability and credit supply includes studies on the impact of legal and regulatory frameworks that restrict banks’ funding access (Jayaratne and Strahan, 1996; Qian and Strahan, 2007; Adelino and Ferreira, 2016; Di Maggio and Kermani, 2017; Cortés, Demyanyk, Li, Loutskina, and Strahan, 2020).
funding instability, our paper complements the traditional literature on bank runs, as our focus is different and on the day-to-day operations of banks without the threat of insolvency. Our approach of measuring payment risk emphasizes banks’ regular operations rather than large deposit outflows at distressed banks (Iyer, Puri, and Ryan, 2016; Martin, Puri, and Ufier, 2018; Brown, Guin, and Morkoetter, 2020). Large deposit outflows are triggered by fundamental news or coordination failure (Gorton, 1988; Saunders and Wilson, 1996; Calomiris and Mason, 1997; Iyer and Puri, 2012; Artavanis, Paravisini, Robles-Garcia, Seru, and Tsountsoura, 2022). This literature also studies deposit withdrawal as discipline on banks (Park and Peristiani, 1998; Billett, Garfinkel, and O’Neal, 1998; Martinez Peria and Schmukler, 2001; Goldberg and Hudgins, 2002; Bennett, Hwa, and Kwast, 2015; Brown, Guin, and Morkoetter, 2020). In our model, what constrains banks’ balance-sheet capacity is the friction that make reserves costly, for example, interbank market frictions, rather than depositor discipline.

and has features unique to bank lending. Our empirical specification is a form of spatial econometric models (Anselin, 1988; LeSage and Pace, 2009; Elhorst, 2010; Lee and Yu, 2010). Spatial models have been adopted only recently in the finance literature in different settings (Buraschi and Porchia, 2012; Ozdagli and Weber, 2017; Lu and Luo, 2019; Herskovic, Kelly, Lustig, and Van Nieuwerburgh, 2020; Denbee, Julliard, Li, and Yuan, 2021; Jiang and Richmond, 2021).

2 Model

2.1 The model setup

Consider an economy with $N$ banks. At $t = 0$, bank $i$ ($i \in \{1, \ldots, N\}$) is endowed with $m_i$ amount of reserves. Bank $i$ lends at $t = 0$. Depositors make payments at $t = 1$. Loans are repaid at $t = 2$. The loans cannot be liquidated or sold at $t = 1$, so the bank covers payment outflows with reserves or incur a cost to cover any shortfall. The timing is in line with the banking literature (e.g., Diamond and Dybvig, 1983), and, in practice, payment settlement is done at a higher frequency (typically overnight) than loan book adjustment.

Bank $i$ extends $y_i$ amount of loans financed by a matching amount of deposits. Depositors make payments at $t = 1$ before the loans are repaid. If the payees hold accounts at other banks, bank $i$ has to send reserves to the payees’ banks to settle payments and deduct the corresponding amount of deposit liabilities, shrinking its balance sheet, while the payees’ banks receive reserves and credit the payees’ deposit accounts with new deposits, expanding their balance sheets.

---

16The first deposit holders are often the borrowers who naturally obtain loans for purchases and have payment needs. In practice, credit creation is a debt swap. Bank $i$ obtains the borrowers’ debts (loans) while the borrowers obtain bank $i$’s debts (new deposits). This practice of credit and money (deposit) creation has a long history and is kept in the modern banking system (Gurley and Shaw, 1960; Tobin, 1963; Bianchi and Bigio, 2014; McLeay, Radia, and Thomas, 2014) and throughout the history of banking (Wicksell, 1907; Donaldson, Piacentino, and Thakor, 2018).
Let $g_{ij}$ denote the fraction of payees at bank $j$ ($j \neq i$). Given the deposits $y_i$, we define

$$z_i \equiv \sum_{j \neq i} g_{ij} y_i$$

as the total reserve outflow to other banks due to the depositors’ payments. We capture the risk in payment flows by assuming that $g_{ij}$ is random with mean $\mu_{ij}$ and variance $\sigma^2_{ij}$. As in Bolton, Li, Wang, and Yang (2020), deposits are essentially debts with random maturities. A random fraction $\sum_{j \neq i} g_{ij}$ of the newly issued deposits will mature at $t = 1$ while the rest mature at $t = 2$. The random $g_{ij}$ is out of bank $i$’s control as it cannot interfere with its depositors’ payment decisions.

Bank $i$ also receive payment inflows as a result of other banks’ lending. Given bank $j$’s lending amount $y_j$ ($j \neq i$), bank $i$ receives payment inflow equal to $g_{ji} y_j$, where, consistent with the previous definitions, $g_{ji}$ has mean $\mu_{ji}$ and variance $\sigma^2_{ji}$. The correlation between between $g_{ij}$ and $g_{ji}$ is denoted by $\rho_{ij}$. We would expect $\rho_{ij}$ to be negative if economic activities are directional, involving mainly bank $i$’s customers paying $j$’s customers. The correlation $\rho_{ij}$ can also be positive if bank $i$’s customers’ payments to $j$’s customers stimulate economic activities between the two clienteles and result in $j$’s customers making payments to $i$’s customers.\(^\text{17}\)

We define the net payment outflow for bank $i$:

$$x_i = \sum_{j \neq i} g_{ij} y_i - \sum_{j \neq i} g_{ji} y_j$$

Note that payment outflow can also be viewed as depositors’ cash withdrawal (rather electronic transfers to payees’ bank accounts) and their payees’ cash deposits. Different from Diamond and Dybvig (1983) who assume a constant fraction of deposit holders who withdraw at $t = 1$, here the

\(^{17}\)For simplicity, it is assumed that the flow fractions are independent across bank pairs.
withdrawal fraction, $\sum_{j \neq i} g_{ij}$, is random.\textsuperscript{18} Our emphasis on the randomness in $g_{ij}$ is consistent with the findings that payment risk is a critical determinant of bank lending (Li and Li, 2021).

Bank $i$’s costs of covering payment outflow are specified as follows:

$$\tau_1(x_i - m_i) + \frac{\tau_2}{2} (x_i - m_i)^2 + \frac{\kappa}{2} z_i^2,$$

where $\tau_1 > 0$, $\tau_2 > 0$, and $\kappa > 0$.\textsuperscript{(3)}

If $x_i - m_i > 0$ (i.e., the bank does not have enough reserves to cover the outflow), this represents an increasing and convex cost of interbank borrowing. The convexity, as microfounded in Bigio and Sannikov (2019) and Parlour, Rajan, and Walden (2020), captures the impact of interbank market frictions (Afonso and Lagos, 2015).\textsuperscript{19} When $x_i - m_i < 0$, this quadratic form presents an increasing and concave return on interbank lending, and the concavity is again due to the frictions in the interbank market. Finally, since $x_i$, defined in (2), is the net flow, we add an additional term, $\frac{\kappa}{2} z_i^2$ (where the gross outflow, $z_i$, is defined in (1)), to capture the fact that netting may not happen instantaneously, especially in the real-time gross settlement (RTGS) systems adopted by most of the advanced economies. As a result, payment outflow may incur additional costs associated with intraday payment stress (Poole, 1968; Afonso, Kovner, and Schoar, 2011; Ashcraft, McAndrews, and Skeie, 2011; Ihrig, 2019; Copeland, Duffie, and Yang, 2021; d’Avernas and Vandeweyer, 2021; Afonso, Duffie, Rigon, and Shin, 2022; Yang, 2022). Kahn and Roberds (2009) review the earlier literature on payment system.

Payment flows affect both banks and their customers. For bank $i$, payment outflows cause its reserves to decline and, at the same time, its customers’ deposits to decline by the same amount; likewise, payment inflows imply reserve gain for bank $i$ and an increase in deposit holdings of $i$’s customers. The simultaneous effects of payment flows on both banks and their customers is a

\textsuperscript{18}Related, Drechsler, Savov, and Schnabl (2021) emphasize that deposits are long-duration liabilities.

\textsuperscript{19}Banks may borrow from the central bank, but in practice, they are discouraged from utilizing discount window and payment-system overdrafts (Copeland, Duffie, and Yang, 2021).
direct implication of the two-layer design of payment systems where settlement between banks is
done via reserves and settlement between bank customers done via deposits. The impact on bank
customers may in turn affect banks’ lending opportunities and thus ought to be considered.

Consider \( x_i > 0 \), i.e., bank \( i \) and its depositors receive outflows. The depositors now have
less liquidity held in the form of deposits, so their demand for bank loans in the future increases,
which enhances bank \( i \)’s future profitability. The impact on bank \( i \)’s (continuation) value is

\[
\theta_1 x_i + \frac{\theta_2}{2} x_i^2, \quad \text{where } \theta_1 > 0 \text{ and } \theta_2 > 0.
\]

This mechanism relies a certain overlap between depositors and loan clientele, which is common
in practice. In Appendix B, we provide a microfoundation for (4) based on depositors’ liquidity
management problem. The first term, which is positive if \( x_i > 0 \), arises from bank \( i \)’s customers
having less liquidity holdings (deposits) and relying more on future bank credit. The second term
captures the increasing marginal impact: As bank \( i \)’s customers lose liquidity, their marginal value
of liquidity increases, which allows the bank to profit more from credit provision.\(^{20}\) If \( x_i < 0 \), bank
\( i \)’s profits may decline as customers receive payments and hold more liquidity. A greater inflow
(i.e., a more negative \( x_i \)) and a sharper decline of customers’ marginal value of liquidity imply a
lower marginal profits \( (\theta_1 + \theta_2 x_i) \) from lending to meet customers’ future liquidity needs.

Let \( R_i + \epsilon_i \) denote the loan return for bank \( i \), where \( R_i \) is a constant and \( \epsilon_i \) represents a shock
that is realized before bank \( i \) makes its lending decision at \( t = 0 \). The profit shock, \( \epsilon_i \), may originate
from the credit demand side, such as the profitability and scalability of borrowers’ projects and
collateral value. The shock can also arise from the credit supply side and depends on factors such
as bank \( i \)’s loan market power (e.g., Scharfstein and Sunderam, 2016), competition from non-

\(^{20}\)Such response in the marginal value of liquidity arises in static settings (see Appendix B) and dynamic settings
(Riddick and Whited, 2009; Bolton, Chen, and Wang, 2011; Décamps, Mariotti, Rochet, and Villeneuve, 2011).
bank lenders (e.g., Buchak, Matvos, Piskorski, and Seru, 2018b,a; Chernenko, Erel, and Prilmeier, 2022) and regulatory costs of lending (e.g., Blattner, Farinha, and Rebelo, 2019). Under a zero deposit rate, the net interest margin is \( R_i + \varepsilon_i - 1 \). Later we will show that the equilibrium level of bank lending can be solved by applying a network propagation operator to standalone (network-independent) lending that encapsulates the shock, reserves, net interest margin, etc. We will treat the entire standalone lending as a random variable and ignore the internal structure. Therefore, \( \varepsilon_i \) can be any shock to bank \( i \)'s standalone lending, such as a liquidity shock the depletes \( m_i \) before the lending decision or a shock to the net interest margin, and the deposit rate can be non-zero.

Collecting the net interest margin and the quadratic forms (3) and (4), we obtain the expected profits (i.e., bank \( i \)'s objective function):

\[
\max_{y_i} \mathbb{E} \left[ (R_i + \varepsilon_i - 1)y_i - \tau_1 (x_i - m_i) - \frac{\tau_2}{2} (x_i - m_i)^2 - \frac{\kappa}{2} z_i^2 + \theta_1 x_i + \frac{\theta_2}{2} x_i^2 \right].
\]

(5)

We impose the following parameter restriction to ensure the concavity in \( y_i \):

\[
\tau_2 + \kappa > \theta_2.
\]

(6)

Note that \( \varepsilon_i \) is realized before the choice of \( y_i \), so the expectation operator is taken over \( g_{ij} \). The costs of bank \( i \) losing liquidity as a result of payment outflows, \( x_i > 0 \), are offset by its depositors’ increasing marginal value of liquidity (and the associated lending profits) because as bank \( i \) loses liquidity (reserves), its customers lose liquidity (deposits) as well. Similarly, payment inflows, \( x_i < 0 \), lead to increasing and concave profits from reserve surplus that are offset by a decrease in profits from offering credit to depositors as depositors hold more liquidity (deposits) and their credit needs decline.\(^{21}\) Our focus is on a bank’s normal-time operations rather than banking crises.

\(^{21}\)Another cost of payment inflows for banks is related to regulations as pointed out by Bolton, Li, Wang, and Yang
We assume that even when the realized payment flows cause the largest possible loss (which is finite under $g_{ij}, g_{ji} \in [0, 1]$), the bank’s realized profits stay positive.

Before solving $y_i$, we clarify that the bank finances lending with deposits instead of reserves. Deposit issuance only causes a probabilistic reserve drawdown (as some of the borrowers’ payees may be the bank’s own depositors) while lending out reserves causes a direct drawdown. Therefore, as long as the marginal cost of spending reserves is above the deposit rate, the bank prefers financing lending with deposits over reserves. We assume this is the case, in line with the evidence that deposits rates are below the fed funds rate in our sample and other findings (e.g., Rose and Kolari, 1985; Drechsler, Savov, and Schnabl, 2017a; Li and Li, 2021).

### 2.2 Equilibrium on the payment network

We characterize the equilibrium of the network lending game of simultaneous actions. First, we take as given $y_j (j \neq i)$ and solve bank $i$’s optimal choice of credit creation and deposit issuance, $y_i$ (i.e., bank $i$’s optimal response to other banks’ decisions). To simplify the notations, we introduce

---

**References**

Reserve and deposit inflows force banks to expand balance sheets and tighten the supplementary leverage ratio (SLR) regulation imposed on total leverage. Moreover, banks cannot simply lend out reserves to earn higher interest income because, with more deposits (especially the less sticky wholesale deposits), liquidity coverage ratio regulation requires banks to hold more liquid assets. Therefore, payment inflows squeeze banks’ balance-sheet capacities. During the Covid-19 pandemic, banks received massive deposit inflows as a result of policy stimulus and, under the regulatory constraints, banks active seek options to turn down deposit inflows (Moise, 2021, Financial Times).

Moreover, it is assumed that deposits are cheaper sources of financing for lending than issuing bond or equity in line with the literature on money premium that reduces banks’ cost of issuing deposits (e.g., Stein, 2012; DeAngelo and Stulz, 2015; Krishnamurthy and Vissing-Jørgensen, 2015; Nagel, 2016; Begenau, 2020).

Moreover, because borrowers do not have accounts at the central bank, dollar bills have to be redeemed if the bank decides to lend out reserves. This institutional barrier implies that it is more convenient to credit borrowers’ checking accounts with new deposits than to lend out reserves. In practice, credit creation is a debt swap. Bank $i$ obtains the borrowers’ debts (loans) while the borrowers obtain bank $i$’s debts (new deposits). Then the borrowers make payments and the lending bank intermediates between depositors (i.e., the borrowers’ payees) and borrowers. This practice has been adopted in the modern banking system (Gurley and Shaw, 1960; Tobin, 1963; Bianchi and Bigio, 2014; McLeay, Radia, and Thomas, 2014) and throughout the history (Wicksell, 1907; Donaldson, Piacentino, and Thakor, 2018).
the mean of total payment outflows as a fraction of $y_i$:

$$
\bar{\mu}_{-i} \equiv \mathbb{E} \left[ \sum_{j \neq i} g_{ij} \right],
$$

(7)

and the variance of total payment outflows as a fraction of $y_i$:

$$
\sigma^2_{-i} = \text{Var} \left( \sum_{j \neq i} g_{ij} \right).
$$

(8)

We derive the following first-order condition for $y_i$ (derivation details in the appendix):

$$
R_i + \varepsilon_i - 1 = (\tau_1 - \theta_1) \bar{\mu}_{-i} + y_i (\kappa + \tau_2 - \theta_2) (\sigma^2_{-i} + \bar{\mu}^2_{-i}) - \tau_2 \bar{\mu}_{-i} m_i
$$

$$
- (\tau_2 - \theta_2) \sum_{j \neq i} (\bar{\mu}_{-i} \mu_{ji} + \rho_{ij} \sigma_{ij} \sigma_{ji}) y_j.
$$

(9)

The marginal benefit of lending (i.e., the net interest margin on the left side) is equal to the marginal cost that incorporates both the negative and positive effects of payment outflows. The first term on the right side, $(\tau_1 - \theta_1) \bar{\mu}_{-i}$, reflects the negative effect of draining reserves on bank profits and the positive effect of customers losing liquidity and relying more on bank credit in the future. The second term captures the payment-flow risk (i.e., the randomness in $P_{j \neq i} g_{ij}$) associated with one more dollar of lending with the parameter $\kappa$ representing additional cost of gross payment outflows as previously discussed. The third term shows that having more reserves reduces the marginal cost of outflow by reducing the needs for costly reserve borrowing.

In the last term on the right side of (9), the network effects can be decomposed into the liquidity externality and hedging externality. The first component, $\bar{\mu}_{-i} \mu_{ji} y_j$, shows that if bank $i$ lends more and incurs the marginal outflow $\bar{\mu}_{-i}$, bank $j$’s lending and its payment flow to $i$ (i.e., $\mu_{ij} y_i$) alleviates $i$’s reserve drain and thus has a greater marginal benefit in reducing $i$’s cost of
lending. We call this term the *liquidity externality of payment network* following Parlour, Rajan, and Walden (2020). Hedging externality is captured by the second component, $\rho_{ij}\sigma_{ij}\sigma_{ji}y_j$. Given bank $j$’s lending, $y_j$, one more dollar of lending by bank $i$ causes itself (and its customers) to receive more inflow if $\rho_{ij}\sigma_{ij}\sigma_{ji}$, the covariance between $g_{ij}$ and $g_{ji}$, is positive, in which case bank $i$’s lending stimulates economic activities that cause $j$’s customers to pay $i$’s customers; if the covariance is negative, the more bank $i$ lends, the more outflow from $i$ to $j$, with the overall impact scaled by $j$’s lending $y_j$. We call this term, $\rho_{ij}\sigma_{ij}\sigma_{ji}y_j$, the hedging externality.\footnote{The network connections arise from risk sharing as in Eisfeldt, Herskovic, Rajan, and Siriwardane (2019).}

Rearranging the first-order condition (9), we solve the optimal $y_i$:

$$y_i = \phi \sum_{j \neq i} w_{ij} y_j + a_i$$

where the network attenuation factor, $\phi$, is given by

$$\phi = \frac{\tau_2 - \theta_2}{\kappa + \tau_2 - \theta_2},$$

and the $ij$-th element of the network adjacency matrix, denoted by $W$, is given by

$$w_{ij} = \frac{\mu_{-i}\mu_{ji} + \rho_{ij}\sigma_{ij}\sigma_{ji}}{\sigma_{-i}^2 + \mu_{-i}^2}.$$

The other terms are collected into $a_i$ ($a = [a_1, \ldots, a_N]$ in vector form):

$$a_i = \frac{R_i + \varepsilon_i - 1 - (\tau_1 - \theta_1)\mu_{-i} + \tau_2\mu_{-i}m_i}{(\kappa + \tau_2 - \theta_2)(\sigma_{-i}^2 + \mu_{-i}^2)}.$$

Note that the denominator in (12) and (13) gives the second moment of total payment outflow as a fraction of deposits (see (7) and (8)). It scales down bank $i$’s lending given bank $j$’s lending ($j \neq i$)
and bank $i$’s characteristics in (13). This negative impact of payment flow risk on bank lending has been documented by Li and Li (2021). This paper focuses on the network externalities.\(^{25}\)

The bilateral network effects depend on the network attenuation factor and the $ij$-th element of the network adjacency matrix:

$$
\phi w_{ij} = \left( \frac{\tau_2 - \theta_2}{\kappa + \tau_2 - \theta_2} \right) \left( \frac{\mu_{-i} \mu_{ji} + \rho_{ij} \sigma_{ij} \sigma_{ji}}{\sigma_{-i}^2 + \sigma_{-i}^2} \right).
$$

(14)

If $\phi w_{ij} > 0$, the pair $\{i, j\}$ feature strategic complementarity in their lending decisions. If $\tau_2 > \theta_2$ (i.e., $\phi > 0$), the benefit of payment inflow from alleviating bank $i$’s reserve drain dominates the cost from reducing future lending opportunities (by having $i$’s customers holding more liquidity). Therefore, when bank $j$ lends more, the expected marginal outflow, $\mu_{ji}$, goes to bank $i$. The liquidity externality is valuable especially when bank $i$’s expected outflow per dollar lent, $\mu_{-i}$, is large. Moreover, strategic complementarity is amplified by the hedging externality if the covariance between $g_{ij}$ and $g_{ji}$, $\rho_{ij} \sigma_{ij} \sigma_{ji}$, is positive, i.e., bank $j$’s lending triggers payment and reserve flows to bank $i$ precisely when bank $i$ loses reserves via payment outflows to $j$. If the covariance is negative, strategic complementarity is dampened and the pair may even flip to strategic substitution.\(^{26}\)

The pair $\{i, j\}$ exhibits strategic substitution in their lending decisions if $\phi w_{ij} < 0$. If $\tau_2 < \theta_2$ (i.e., $\phi < 0$), the cost of payment inflow from reducing future lending opportunities (by increasing bank $i$’s customers’ liquidity holdings) dominates the benefit from alleviating bank $i$’s reserve drain. In this case, bank $i$ is averse to payment inflows and lends less if it expects to receive more inflows from bank $j$. If $\rho_{ij} \sigma_{ij} \sigma_{ji} > 0$ (thus $\phi w_{ij} < 0$), both the liquidity externality and hedging externality point to more payment inflows to bank $i$ if $j$ lends more, so, under bank $i$’s aversion to inflows (i.e., $\phi < 0$), bank $i$ lends less when $j$ lends more; likewise, if bank $i$ lends.

\(^{25}\)The parameter, $\tau_1$, can be interpreted as the cost of reserve borrowing, which negatively affects bank lending in line with the evidence (Jiménez, Ongena, Peydró, and Saurina, 2012, 2014).

\(^{26}\)In our sample, there are only 0.39% of non-zero $w_{ij}$ being negative. 6.47% of all pairs have non-zero $w_{ij}$.
more, bank $j$ expects to receive more inflows and lends less. Therefore, the pair $\{i, j\}$ exhibits strategic substitution. If $\rho_{ij}\sigma_{ij}\sigma_{ji} < 0$, the substitution effects from $\phi < 0$ are dampened.

In our model, the payment network given by (12) describes the ex ante spillover effects in both the first and second moments of payment flows. As previously discussed, the numerator of (12) captures the hedging externality and liquidity externality from the payment network. A bank’s lending decision depends on other banks’ lending decisions because, under the two-layer design of payment system, both the bank and its customers receive liquidity inflows due to the payments of other banks’ borrowers. The linear and quadratic terms in the bank’s objective function imply that both the expected flows and volatilities enter the banks’ decision making.

**Proposition 1** Suppose $|\phi\lambda_{\text{max}}(W)| < 1$, where the function $\lambda_{\text{max}}(\cdot)$ returns the largest eigenvalue. Then, there is a unique interior solution for the Nash equilibrium outcome given by

$$y^*_i = \{M(\phi, W)\}_i a,$$

where $\{\}_i$ is the operator that returns the $i$-th row of its argument, and

$$M(\phi, W) \equiv I + \phi W + \phi^2 W^2 + \phi^3 W^3 + \ldots = \sum_{k=0}^{\infty} \phi^k W^k = (I - \phi W)^{-1},$$

where $I$ is the $N \times N$ identity matrix.

Proposition 1 summarizes the equilibrium solution.\footnote{The sequence in (16) converges under $|\phi\lambda_{\text{max}}(W)| < 1$ (Debreu and Herstein, 1953). The equilibrium definition is akin to that of Calvo-Armengol, Putacchini, and Zenou (2009) who study peer effects in education.} In vector form, we can rewrite (15):

$$y^* = (I - \phi W)^{-1} a.$$
The condition $|\phi \lambda^{\text{max}}(W)| < 1$ states that network externalities must be small enough in order to prevent the feedback triggered by such externalities to escalate without bounds. Note that equation (10), which leads to equilibrium characterization in Proposition (1), is rather robust in that it could be in principle derived from different micro-foundations and in different settings.\footnote{For instance, customers’ payments can be driven by input-output linkages (Carvalho and Tahbaz-Salehi, 2019).}

The matrix $M(\phi, W)$ has an important economic interpretation: it aggregates all direct and indirect links among banks using an attenuation factor, $\phi$, that penalizes, as in Katz (1953), the contribution of links between distant nodes at the rate $\phi^k$, where $k$ is the length of the path between nodes. In the infinite sum in equation (16), the identity matrix captures the (implicit) link of each bank with itself, the second term in the sum captures all the direct links between banks, the third term in the sum captures all the indirect links corresponding to paths of length two, and so on. The elements of $M(\phi, W)$, given by $m_{ij}(\phi, W) \equiv \sum_{k=0}^{+\infty} \phi^k \{W^k\}_{ij}$, aggregates all paths from $j$ to $i$.

The matrix $M(\phi, W)$ contains information about the network centrality of bank.\footnote{This centrality measure takes into account the number of both direct and indirect connections in a network. For more on the Bonacich centrality measure, see Bonacich (1987) and Jackson (2003). For other economic applications, see Ballester, Calvo-Armengol, and Zenou (2006) and Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012). For an excellent review of the literature, see Jackson and Zenou (2012).} Multiplying the rows (columns) of $M(\phi, W)$ by a unit vector of conformable dimensions, we recover the indegree (outdegree) Katz–Bonacich centrality measure. The indegree centrality measure provides the weighted count of the number of ties directed to each node (i.e., inward paths), while the outdegree centrality measure provides the weighted count of ties that each node directs to the other nodes (i.e., outward paths). The $i$-th row of $M(\phi, W)$ captures how bank $i$ loads on the network as whole, while the $i$-th column of $M(\phi, W)$ captures how the network as a whole loads on $i$.

The matrix $M(\phi, W)$ includes the network topology and network attenuation factor $\phi$. Before the lending game starts, shocks to individual banks (attributed to $\varepsilon_i$) are encoded in $a = [a_1, ..., N]$, observed by banks and their peers. We can decompose $a_i$ given by (13) into a time-
invariant term for bank $i$, denoted by $\bar{\alpha}_i$, and a shock specific to bank $i$ (originating from $\varepsilon_i$ in the model setup), denoted by $\nu_i$, that is independent across banks:

$$a_i = \bar{\alpha}_i + \nu_i,$$

(18)

where $\nu_i$ has mean equal to zero and variance $\delta_i^2$. In the next section, we estimate the mean and volatility of $a_i$ directly without using the solution of $a_i$ in (13). This allows us to stay agnostic about the empirical counterpart of $m_i$, the available reserves for bank $i$ to cover payment shocks. As discussed in the introduction, the bank can potentially sell other (unmodelled) assets to obtain reserves or even pledge loans to borrow reserves, so the available reserves do not simply map to the bank’s reserve holdings. Moreover, treating $a_i$ as a “black box” and directly estimating the mean and volatility allow us to broaden the interpretation of shock $\varepsilon_i$. Specifically, $\varepsilon_i$ does not necessarily enter into the net interest margin. It can, for example, hits $m_i$. The estimation results remain the same as long as $\varepsilon_i$ enters the model in any way that brings randomness to $a_i$. Finally, the deposit rate in $a_i$ is not necessarily to be zero. In fact, it can take any number as our estimation treats the entire $a_i$ as a random variable and ignores its internal structure.

In our model, a money multiplier arises. First, reserves (the monetary base) enter into banks’ network-independent lending, i.e., a given by (13). Second, the network amplifies a through $M(\phi, W)$ to the equilibrium amount of loans and deposits, $y$. Our paper provides a theoretical underpinning of the classic concept of money multiplier, and our estimation focuses on characterizing the network amplification effects (i.e., the network propagation operator, $M(\phi, W)$).

We define vectors $\bar{\alpha} = [\bar{\alpha}_1, ..., \bar{\alpha}_N]$ and $\nu = [\nu_1, ..., \nu_N]$. To see clearly how the network
propagates shocks, we rewrite (17) as

\[ y^* = M(\phi, W) \bar{\alpha} + M(\phi, W) \nu. \]

The matrix \( M(\phi, W) \) itself is not enough to determine the systemic importance of a bank. Regardless of \( M(\phi, W) \), i.e., how the shocks are propagated, banks with large shocks (i.e., large \( \delta^2_i \)) have a large influence on other banks’ lending decisions and the aggregate credit supply. The network not only propagates shocks but also amplifies the impact of \( \bar{\alpha} \) on the level of banks’ lending. In Section 3.3, we show how to utilize the equilibrium solution to identify banks that contribute the most to the systemic risk of aggregate credit supply after we discuss the estimation methodology.

**Discussion: Other assets and financing instruments.** We discuss bank liabilities and assets outside of the model. The quadratic cost of reserve drain, parameterized via \( \tau_1 \) and \( \tau_2 \), captures the costs associated with covering reserve deficits, such as borrowing in the federal funds market and utilizing central bank facilities. It also captures the costs of using other means of financing, such as raising new deposits and issuing bond or equity. At \( t = 0 \), lending is financed by deposits rather than bond or equity issuances, so the implicit assumption is that deposits are cheaper sources of financing in line with the literature on money premium.  

Finally, what happens if the bank issues deposits at \( t = 0 \) to increase reserves or other liquid assets that are liquid at \( t = 1 \)? This introduces additional payment flows at \( t = 1 \), so \( y_j \) in (10) is replaced by the sum of \( y_i \) and the additional

---

30Depositors accept a yield on deposits below other financial assets (even on a risk-adjusted basis) for the convenience of payment services (e.g., Stein, 2012; DeAngelo and Stulz, 2015; Krishnamurthy and Vissing-Jørgensen, 2015; Nagel, 2016; Begnau, 2020). The fact that deposits are relatively cheap sources of financing may also be related to banks’ market power (e.g., Drechsler, Savov, and Schnabl, 2017b; Wang, Whited, Wu, and Xiao, 2018). Finally, when a bank extends a loan, it credits the borrower’s account with a matching amount of deposits. Bank lending is a swap of new debts. The bank obtains the borrower’s debt (loan) and issues debt (new deposits) to the borrower. The borrower then uses newly created deposits for payment. This practice of credit and money (deposit) creation has been adopted through the history of banking (Wicksell, 1907; Gurley and Shaw, 1960; Tobin, 1963; Gersbach, 1998; McLeay, Radia, and Thomas, 2014; Bianchi and Bigio, 2014; Donaldson, Piacentino, and Thakor, 2018; Wang, 2019).
deposits raised for acquiring reserves or liquid assets. More reserves or liquid assets increase \( m_i \), and the impact of a higher \( m_i \) is absorbed in \( a_i \). Our results hold under the extended setup.\(^{31}\)

### 2.3 The planner’s solution

The model characterizes not only the shock amplification mechanism through the payment network but also the externalities. Individual banks make their decisions without internalizing the impact on neighbors. We proceed to a formal analysis of the planner’s problem. We consider a planner that equally weights the objective of each bank and chooses loan provision as follows:

\[
\max_{\{y_i\}_{i=1}^N} \mathbb{E} \left[ \sum_{i=1}^N (R_i + \varepsilon_i - 1)y_i - \tau_1 (x_i - m_i) - \frac{\tau_2}{2} (x_i - m_i)^2 - \frac{\kappa}{2} \sigma_i^2 + \theta_1 x_i + \frac{\theta_2}{2} x_i^2 \right].
\] (20)

We do not aim for welfare implications as the planner’s objective only incorporates banks’ profits instead of the total welfare of banks, borrowers, and depositors. The focus is on characterizing network externalities through the wedge between the planner’s solution and market outcome.

The planner’s first order condition for bank \( i \)’s lending amount, \( y_i \), yields:

\[
R_i + \varepsilon_i - 1 = y_i (\kappa + \tau_2 - \theta_2) \left( \sigma_i^2 + \overline{\sigma}_i^2 \right) - \tau_2 \overline{\sigma}_i m_i - (\tau_2 - \theta_2) \sum_{j \neq i} (\overline{\mu}_{ij} + \rho_{ij} \sigma_i \sigma_j) y_j \\
+ y_i (\tau_2 - \theta_2) \overline{\sigma}_i^2 - (\tau_2 - \theta_2) \sum_{j \neq i} (\overline{\mu}_{ij} + \rho_{ij} \sigma_i \sigma_j) y_j \\
+ (\tau_2 - \theta_2) \sum_{j \neq i} \left( \sum_{k \neq j} \mu_{kji} y_k \right) \mu_{ij} - \sum_{j \neq i} \tau_2 m_j \mu_{ij}
\] (21)

The planner’s marginal cost of bank \( i \)’s lending is on the right side of (21). Its first three terms

\(^{31}\)For the existence of an interior solution of deposits issued to increase \( m_i \), we need to impose an increasing and convex cost; otherwise it’s optimal to go infinite because for every one dollar of reserves obtained via deposit financing, the deposits raised will be gone at \( t = 1 \) with a probability smaller than one (the marginal benefit is always positive).
also appear on the right side of the first-order condition (9) in the market equilibrium but the rest differ and reflect the planner’s internalization of the spillover effects of bank $i$’s lending. First, bank $i$’s costs or benefits associated with the expected outflow, $(\tau_1 - \theta_1)\mu_{-i}$ in (9), disappears because, from the planner’s perspective, bank $i$’s expected outflow is the other banks’ expected inflow and thus $i$’s losses are offset by $j$’s gains. Second, the additional term, $y_i(\tau_2 - \theta_2)\sigma_{-i}^2$, reflects the fact that when bank $i$ lends more, it adds payment flow risk not only to itself (via the first term on the right side of (21)) but also to its neighbouring banks. Third, the fifth term, $-(\tau_2 - \theta_2)\sum_{j \neq i}(\mu_{-j} + \rho_{ij}\sigma_{ij})y_j$, captures the liquidity externality and hedging externality of bank $i$’s lending on bank $j$ ($j \neq i$). In particular, the liquidity externality of bank $i$’s marginal lending (through the marginal outflow, $\mu_{ij}$) has a stronger impact on bank $j$ when $j$ expected a large outflow $\mu_{-j}$. The sixth term, $(\tau_2 - \theta_2)\sum_{j \neq i}(\mu_{-j} + \rho_{ij}\sigma_{ij})\mu_{ij}$, shows that if bank $j$ already receives inflows due to bank $k$’s lending ($k \neq j$), the marginal impact of liquidity from bank $i$ (i.e., $\mu_{ij}$) is smaller. Finally, the last term shows that if bank $j$ already has large reserve holdings, the marginal impact of liquidity from bank $i$ is smaller.

Rearranging the planner’s first-order condition (21), we solve the optimal $y_i$:

$$y_i = \tilde{\phi}_i \sum_{j \neq i} \tilde{w}_{ij}y_j - \tilde{\phi}_i \sum_{j \neq i} \mu_{ij} \left( \sum_{k \neq j} \mu_{kj}y_k \right) + \tilde{a}_i$$

(22)

where the network attenuation factor for bank $i$, $\tilde{\phi}_i$, is given by,

$$\tilde{\phi}_i = \frac{(\tau_2 - \theta_2)(\sigma_{-i}^2 + \mu_{-i}^2)}{(\kappa + \tau_2 - \theta_2)(\sigma_{-i}^2 + \mu_{-i}^2)} = \frac{1}{1 + \frac{\sigma_{-i}^2}{\sigma_{-i}^2 + \mu_{-i}^2}}$$

(23)
and the $ij$-th element of the network adjacency matrix, denoted by $\tilde{W}$, is given by

$$\tilde{w}_{ij} = \frac{\mu_{-i} \mu_{ji} + 2 \rho_{ij} \sigma_{ij} \sigma_{ji} + \mu_{-j} \mu_{ij}}{\sigma_{-i}^2 + \mu_{-i}^2}. \quad (24)$$

The other terms are collected into $\tilde{a}_i$ ($\tilde{\alpha} = [\tilde{a}_1, ..., \tilde{a}_N]$ in vector form):

$$\tilde{a}_i \equiv \frac{\epsilon_i + R_i - 1 + \tau_2 \mu_{-i} m_i - \sum_{j \neq i} \tau_2 m_j \mu_{ij}}{(\kappa + \tau_2 - \theta_2) (\sigma_{-i}^2 + \mu_{-i}^2) + (\tau_2 - \theta_2) \sigma_{-i}^2}. \quad (25)$$

Throughout this paper, “$\sim$” differentiates the variable in the planner’s solution from its counterpart in the decentralized equilibrium. The planner’s network attenuation factor differs from $\phi$ in (11) and is bank $i$-specific due to the additional term, $(\tau_2 - \theta_2) \sigma_{-i}^2$, in the denominator that reflects the payment risk spillover effect of bank $i$’s lending. This additional term scales down bank $i$’s lending and also appears in the denominator of $\tilde{a}_i$ in (25). Different from the decentralized counterpart in (13), the numerator of $\tilde{a}_i$ no longer has the expected outflow (which, from the planner’s perspective, is offset by other banks’ inflow) but it has an additional term $\sum_{j \neq i} \tau_2 m_j \mu_{ij}$ because the liquidity externality of bank $i$’s lending is less valuable when bank $j$ ($j \neq i$) already hold large reserves. Finally, the $ij$-th element of adjacency matrix in (24) differs from its decentralized counterpart in (12) by incorporating the hedging and liquidity externalities of bank $i$’s lending.

Let $\tilde{\Phi}$ denote the diagonal matrix with the $i$-th diagonal element equal to $\tilde{\phi}_i$ and $U$ denote the matrix with the $ij$-th element equal to $\mu_{ij}$. We rewrite the planner’s solution (22) in vector form:

$$y^* = \tilde{\Phi} \tilde{W} y - \tilde{\Phi} U U^\top y + \tilde{\alpha} \quad (26)$$

and in closed-form,

$$y^* = \left( I - \tilde{\Phi} \tilde{W} + \tilde{\Phi} U U^\top \right)^{-1} \tilde{\alpha}. \quad (27)$$
Proposition 2 Suppose $\lambda^{\max} \left( \Phi \tilde{W} + \tilde{\Phi} U U^\top \right) < 1$, where the function $\lambda^{\max}(\cdot)$ returns the largest eigenvalue. Then, the planner's optimal solution is uniquely defined and given by (27).

3 Empirical Methodology

3.1 Data and the empirical specification

We use confidential transaction-level data from Fedwire Funds Service (“Fedwire”) that span from 2010 to 2020. Fedwire is a real-time gross settlement (RTGS) used to electronically settle U.S. dollar payments among member institutions (including more than two thousand banks). The system processes trillions of dollars daily. In 2020, the average weekly transaction value exceeded the U.S. annual GDP. Fedwire accounts for roughly two thirds of the transaction volume in the U.S. The majority of the rest of transactions are mainly settled through Clearing House Interbank Payments System (CHIPS) of 43 members, which, unlike Fedwire, allows netting (potentially at the expense of inducing greater counterparty risks) and therefore does not fit our setting where payments are settled on gross terms without counterparty risks. In Appendix A, we provide more details on the structure of U.S. payment system. Bech and Hobijn (2007) provide an overview on the adoption of real-time gross settlement (RTGS) across countries.

The Federal Reserve maintains accounts for both senders and receivers and settles individual transactions immediately without netting. For each transaction, the Fedwire data provide information on the time and date of the transaction, identities of sender and receiver, payment amount, and transaction type. We focus on transactions instructed by customers, which are out of the banks’ control as in our theoretical model. In particular, we exclude bank-scheduled transfers and banks’
purchases and sales of federal funds. Customer-initiated transactions make up about 85% of trans-
actions (in terms of number of transactions). We obtain data on bank balance sheets and income
statements from U.S. Call Report. We merge the Fedwire data with the Call Report data using Fed-
eral Reserve’s internal identity system. Our merged sample covers 83% of banks in Call Report
(in terms of total assets). We provide the summary statistics in Table D.1 in the appendix.

We set up our empirical specification following the solution of $y_i$ in (10). Our estimation
is based on a quarterly sample. To maintain the standard econometric assumptions of stationarity
and ergodicity of data generating processes (Hayashi, 2000), we use banks’ quarterly loan growth
rates instead of loan amounts. Therefore, we divide both sides of (10) by the loan amount at $t - 1$
to obtain the loan growth rate of bank $i$ at $t$, denoted by $n_{i,t}$

$$n_{i,t} = \frac{y_{i,t}}{y_{i,t-1}} = \phi \sum_{j \neq i} w_{ij} \frac{y_{j,t}}{y_{i,t-1}} + \frac{a_{i,t}}{y_{i,t-1}}. \tag{28}$$

To simplify the notation, we use $a'_{i,t}$ to denote $a_{i,t}/y_{i,t-1}$. For the decomposition in (18), we have

$$a'_{i,t} = \hat{\alpha}'_i + \nu'_{i,t}, \tag{29}$$

where, $\hat{\alpha}'_i = \tilde{\alpha}_i/y_{i,t-1}$, and the shock, $\nu'_{i,t}$, has a zero mean and a conditional variance $\delta'_i^2 (\delta'_i = \delta_i/y_{i,t-1})$. In our quasi-MLE estimation, the parameters enter the probability density of $n_{i,t}$ conditional on $y_{i,t-1}$, and the joint likelihood is the product of conditional probability densities.

Next, we substitute bank $j$’s loan growth rate, $n_{j,t} = \frac{y_{j,t}}{y_{j,t-1}}$ in (28) to obtain:

$$n_{i,t} = \phi \sum_{j \neq i} w'_{ij} n_{j,t} + \hat{\alpha}'_i + \nu'_{i,t}, \tag{30}$$
where the loan amount-adjusted adjacency matrix, denoted by $W'$, has the $ij$-th element given by

$$w'_{ij} \equiv w_{ij} \frac{y_{j,t-1}}{y_{i,t-1}}. \quad (31)$$

To obtain $w'_{ij}$ for quarter $t$, we calculate $w_{ij}$ following the definition (12) and adjust it by the lagged loan amounts of bank $i$ and $j$ as in (31). The statistics in $w_{ij}$, $\mu_{ij}$, $\mu_{ji}$, $\rho_{ij}$, $\sigma_{ij}$, $\sigma_{ji}$, are, respectively, the mean of daily observations of $g_{ij}$ in quarter $t - 1$, the mean of daily observations of $g_{ji}$ in quarter $t - 1$, the correlation between the daily observations of $g_{ij}$ and $g_{ij}$ in quarter $t - 1$, the standard deviation of daily observations of $g_{ij}$ in quarter $t - 1$, and the standard deviation of daily observations of $g_{ji}$ in quarter $t - 1$. Following the theoretical definitions, we scale the payment flows by bank $i$’s deposit stock at the beginning of the quarter to obtain $g_{ij}$. These payment statistics are from the lagged quarter to maintain predeterminancy, and they will form the banks’ belief over
payment flows in quarter \( t \). Our results are robust to extending the calculation window from one quarter to eight lagged quarters. In figure 1, we report the frequency distribution of \( g_{ij} \).

A key target of our estimation is the parameter \( \phi \). An estimate of \( \phi \) that is statistically significant from zero suggests that the network as a whole has a significant impact on bank lending. And, together with the network adjacency matrix, \( W' \), the parameter \( \phi \) determines whether bank lending decisions are strategic complements or substitutes. Instead of directly estimating the equilibrium condition (30) using observations of loan growth rates, we recognize that empirically, a bank’s lending decision depends on bank characteristics and macroeconomic variables outside of our theoretical model. Specifically, our empirical model of loan growth rate has two components, \( q_{i,t} \) that is outside of model of liquidity percolation in the payment system, and \( n_{i,t} \), which is the component dependent on the payment network and modelled in Section 2.

In data, we only observe \( l_{i,t} = q_{i,t} + n_{i,t} \), not \( q_{i,t} \) and \( n_{i,t} \) separately. However, by observing bank characteristics (denoted by \( x_{i,t}^{c} \)) and macroeconomic variables (denoted by \( x_{t}^{p} \)) that drive \( q_{i,t} \), we are able to estimate the network attenuation factor, \( \phi \), effectively using the residuals of \( l_{i,t} \). In our estimation, the bank characteristics include the logarithm of total assets, the ratio of liquid securities (reserves and available-for-trade securities) to total assets, the ratio of equity capital to total assets, the ratio of deposits to total assets, the ratio of loans to total assets, the return on assets, and the macroeconomic variables (from FRED) include the change in effective federal funds rate (EFFR), real GDP growth, inflation, stock market return, and housing price growth.\(^{32}\) All control variables are lagged by one quarter for predeterminancy. We also include the constant as a control variable. We provide the summary statistics in Table D.1 in the appendix.

\(^{32}\)The stock market return is the quarterly change of the Wilshire 5000 Total Market Index (a market-capitalization-weighted index of the market value of all American-stocks actively traded in the United States). The housing price growth is the quarterly change of the S&P/Case-Shiller U.S. National Home Price Index.
In sum, our empirical model of the observed loan growth rate is

\[ l_{i,t} = \sum_{m=1}^{M} \beta_{m}^{\text{bank}} x_{i,t}^{m} + \sum_{p=1}^{P} \beta_{p}^{\text{macro}} x_{t}^{p} + n_{i,t}, \]  

(32)

where, according to (30), we have that

\[ n_{i,t} = \phi \sum_{j \neq i} w_{ij}^{t} n_{j,t} + \bar{\alpha}^{t} + \nu_{i,t}^{t} \sim N \left( 0, \delta_{i}^{t2} \right). \]  

(33)

Equation (32) and (33) together constitute a spatial error model (SEM) (e.g., Anselin, 1988; Elhorst, 2010). Such models allow the joint estimation of \( \beta \) coefficients in the observational equation (32), and \( \bar{\alpha}^{t}, \delta_{i}^{t2}, \) and \( \phi \) in the error (or residual) equation (33). Therefore, even though the econometrician does not observe \( n_{i,t} \) directly, the parameters of the network game can still be recovered.

We can rewrite the system of (32) and (33) in vector form:

\[ \ell_{t} = X_{t} \beta + n_{t}, \]  

(34)

and

\[ n_{t} = \phi W^{t} n_{t} + \bar{\alpha}^{t} + \nu_{t}^{t}. \]  

(35)

Following Proposition 1, we require that \(|\phi \lambda_{\text{max}}(W^{t})| < 1\), where the function \( \lambda_{\text{max}}(\cdot) \) returns the largest eigenvalue. Under this restriction, we have

\[ n_{t} = \left( I - \phi W^{t} \right)^{-1} (\bar{\alpha}^{t} + \nu_{t}^{t}). \]  

(36)

Bank characteristics and macroeconomic variables absorb part of the variation in loan growth rates and only leave the residual variation for identifying the network effect, \( \phi \), and the other
parameters of the network game. This is a conservative approach because any peer effects (or comovement) related to these bank characteristics or common loadings on macroeconomic factors are controlled for, and we only use the residual variations to estimate the parameters of the network lending game. Given the strong heterogeneity in bank sizes, \( w'_{ij} = w_{ij}y_{j,t-1}/y_{i,t-1} \), can be large if bank \( i \) is much smaller than bank \( j \), which then implies that for small banks, the network-dependent component, \( n_{i,t} \), mechanically accounts for a large share of loan growth relative to \( q_{i,t} \) (the component determined by bank characteristics and macroeconomic variables). Our model does not address the relative importance of \( n_{i,t} \) and \( q_{i,t} \) in driving loan growth. We only use the bank characteristics and macroeconomic variables as control variables to absorb loan growth variations from the existing literature. We normalize \( W' \) to be right-stochastic (i.e., to have all row sums equal to one or \( W'1 = 1 \)) by dividing \( w'_{ij} \) by the \( i \)-th row sum so that the relative contributions of \( n_{i,t} \) and \( q_{i,t} \) are not mechanically driven by bank sizes. Moreover, normalizing \( W' \) also prevents the estimation of \( \phi \) from being disproportionately influenced by the small banks’ loan growth.

We estimate the parameters \( \phi, \bar{\alpha}', \delta', \) and \( \beta \) by maximizing the following joint likelihood that is derived by equations (34) and (35):

\[
-\frac{T}{2} \ln \left( (2\pi)^N |\Delta'| \right) - \frac{1}{2} \sum_{t=1}^{T} \left[ (\mathbf{I} - \phi \mathbf{W}') (\ell_t - X_t \beta) - \bar{\alpha}' \right]^\top \Delta'^{-1} \left[ (\mathbf{I} - \phi \mathbf{W}') (\ell_t - X_t \beta) - \bar{\alpha}' \right],
\]

where \( N \) is the number of banks, \( T \) is the total number of quarters, \( \Delta' \) is a diagonal matrix with the \( i \)-th diagonal element equal to \( \delta'^2_i \), and \( |\Delta'| \) is the determinant of \( \Delta' \). When the shocks \( \nu_t' \) are normally distributed, the estimator is the maximum likelihood estimator (MLE) and has the textbook properties of consistency and asymptotic normality. When the shocks are not normally distributed, the estimator is quasi-MLE. Because the score of the normal log-likelihood has the martingale difference property when the first two conditional moments are correctly specified, the
quasi-MLE is consistent and has a limiting normal distribution (Bollerslev and Wooldridge, 1992).

We follow Bollerslev and Wooldridge (1992) to calculate the asymptotic standard errors that are robust to the non-normality of shocks.

### 3.2 Parameter identification

To fix intuition about how the key network parameter, $\phi$, is identified from the data, it is useful to consider a simplified version of the model in equations (32) and (33). Our analysis follows (Denbee, Julliard, Li, and Yuan, 2021). Let $L_t \in \mathbb{R}^N$ denote the vector containing loan growth rates of individual banks at quarter $t$, and to simplify exposition let us disregard the fixed effects, $\bar{\alpha}_i'$, in equation (33) and assume that the network matrix has constant weights $W'$. The model given by (34) and (35) can be rewritten in vector form:

$$
\ell_t = X_t\beta + n_t, \quad n_t \sim N(0_N, \Omega),
$$

where $0_N$ denotes a $N$-dimensional vector of zeros, $\Omega = M\Delta'M^\top$ with $M = (I - \phi W')^{-1}$, $\Delta'$ is a diagonal matrix with elements given by $\left\{\delta_{i}^2\right\}_{i=1}^N$. In deriving the covariance $\Omega$, we used equation (33), i.e., that in equilibrium we can rewrite $n_t$ (having, for now, removed $\bar{\alpha}_i$) as $n_t = (I - \phi W')^{-1}\nu'_t$, where $\nu'_t$ has a distribution with zero mean and a diagonal covariance matrix $\Delta'$.

The reduced form specification in (37) has the same structure and properties as the Seemingly Unrelated Regressions (SUR, see e.g. Zellner (1962)). Hence, one can consistently estimate the mean equation parameters, $\beta$, (e.g., via linear projections), and use the fitted residuals to construct a consistent estimator of covariance matrix $\Omega$. Note that if we knew the parameters $\phi$ and $\left\{\delta_{i}^2\right\}_{i=1}^N$ we could actually premultiply the specification in equation (37) by the Cholesky decomposition of $\Omega^{-1}$, obtaining a transformed system with spherical errors, and therefore gaining efficiency of
the estimates – e.g., we could do the canonical GLS transformation. For this reason, rather than employing a two-step procedure, we jointly estimate the mean equation and covariance parameters by maximizing the quasi-maximum likelihood function.

The key question is whether we can recover the structural parameters \( \phi \) and \( \left\{ \sigma^2_i \right\}_{i=1}^N \). Being symmetric, the estimated \( \hat{\Omega} \) gives \( N(N+1)/2 \) equations, while we have to recover \( N + 1 \) parameters in \( M\Delta'M^T \). Therefore, as long as \( \Omega \) is full-rank, the system is over-identified if we have three or more banks (with linearly independent links). In a nutshell, the identification of this spatial error formulation works as that of structural vector autoregressions (Sims and Zha, 1999) where the contemporaneous propagation of shocks among dependent variables (captured by \( \phi \) in our setting) can be recovered from the reduced-form covariance structure.\(^{33}\) Note that what allows the identification of \( \phi \) and \( \left\{ \sigma^2_i \right\}_{i=1}^N \) are exactly the following two properties: (1) the observed loan growth rate, \( l_{i,t} \), can be decomposed into \( q_{i,t} \), driven by the control variables \( X_t \), and \( n_{i,t} \), the component dependent on the payments; (2) Proposition 1 states how the network component \( n_{i,t} \) depends on the structural shocks in equilibrium. The first restriction defines the mean equation in (37), allowing us to recover \( n_{i,t} \) as residuals.\(^{34}\) The second restriction imposes a structure on the covariance matrix of \( n_{i,t} \), allowing us to recover \( \phi \) and \( \left\{ \sigma^2_i \right\}_{i=1}^N \).

To sharpen the intuition, let us consider a system of three banks and the simplest network, a

---

\(^{33}\)For an extensive discussion of estimation and identification of spatial models see Anselin (1988), and chapter 8 in particular for the Spatial Error Model.

\(^{34}\)Ideally, if we were to observe \( q_{i,t} \) and \( n_{i,t} \) separately, we could estimate \( \phi \) and \( \left\{ \sigma^2_i \right\}_{i=1}^N \) only using the data on \( n_{i,t} \). But as econometricians we only observe \( l_{i,t} = q_{i,t} + n_{i,t} \) and the control variables that drive \( q_{i,t} \), so we estimate \( \phi \) and \( \left\{ \sigma^2_i \right\}_{i=1}^N \) and the control variables’ coefficients jointly.
chain: Bank 1 borrows from Bank 2, and 2 from 3, so

\[
W' = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0
\end{bmatrix}, \quad \text{and } M\Delta'M^\top = \begin{bmatrix}
\delta_{1}^{\prime 2} + \phi^2\delta_{2}^{\prime 2} + \phi^4\delta_{3}^{\prime 2} & \phi\delta_{2}^{\prime 2} + \phi^3\delta_{3}^{\prime 2} & \phi^2\delta_{3}^{\prime 2} \\
\phi\delta_{2}^{\prime 2} + \phi^3\delta_{3}^{\prime 2} & \delta_{2}^{\prime 2} + \phi^2\delta_{3}^{\prime 2} & \phi\delta_{3}^{\prime 2} \\
\phi^2\delta_{2}^{\prime 2} & \phi\delta_{3}^{\prime 2} & \delta_{3}^{\prime 2}
\end{bmatrix}.
\]

The volatility of \(n_1\) is \(\delta_{1}^{\prime 2} + \phi^2\delta_{2}^{\prime 2} + \phi^4\delta_{3}^{\prime 2}\). The first term is the volatility of Bank 1’s structural shock, \(\nu_{1}'\). The second term is the volatility of Bank 2’s structural shock transmitted by one step to Bank 1, i.e., \(\phi\nu_{2}\), and the third term reflects Bank 3’s shock transmitted by two steps (via Bank 2) to Bank 1, i.e., \(\phi^2\nu_{3}\). By the same logic, the volatility of \(n_2\) is \(\delta_{2}^{\prime 2} + \phi^2\delta_{3}^{\prime 2}\), capturing Bank 2’s exposure to its own shock and Bank 3’s shock, while Bank 3 only loads on its own shock. The covariance between \(z_1\) and \(z_2\) is \(\phi\delta_{2}^{\prime 2} + \phi^3\delta_{3}^{\prime 2}\), reflecting Bank 1’s and 2’s exposure to Bank 2’s and 3’s shocks. The covariance between \(z_2\) and \(z_3\) is \(\phi\delta_{3}^{\prime 2}\) as it only arises from the one-step transmission of Bank 3’s shock to Bank 2, i.e., \(\phi z_3\). Such covariances are due to network connections, and their estimates identify the network effect parameter, \(\phi\). Given \(\delta_{3}^{\prime 2} = \{\Omega\}_{3,3}\), we can solve for \(\phi\) using either the covariance between \(n_1\) and \(n_3\), i.e., \(\{\Omega\}_{1,3} = \phi^2\delta_{3}^{\prime 2}\), or the covariance between \(n_2\) and \(n_3\), i.e., \(\{\Omega\}_{2,3} = \phi\delta_{3}^{\prime 2}\), so the system is clearly over-identified. Moreover, given the estimates of \(\delta_{3}^{\prime 2}\) and \(\phi\), either the volatility of \(n_2\), i.e., \(\{\Omega\}_{2,2} = \delta_{2}^{\prime 2} + \phi^2\delta_{3}^{\prime 2}\), or the covariance between \(n_1\) and \(n_2\), i.e., \(\{\Omega\}_{1,2} = \phi\delta_{2}^{\prime 2} + \phi^3\delta_{3}^{\prime 2}\), give a solution for \(\delta_{2}^{\prime 2}\). Finally, given \(\phi\), \(\delta_{2}^{\prime 2}\), and \(\delta_{3}^{\prime 2}\), \(\{\Omega\}_{1,1}\) pins down \(\delta_{1}^{\prime 2}\).

A key identifying assumption is that the structural shocks, \(\nu_i'\), are independent across banks, and thus, after controlling for the observed bank characteristics and macro variables, the residuals’ (i.e., \(n_i'\)’s) correlations only arise from the network linkages. Therefore, the impact of network, \(\phi\), is identified by such correlations. Accordingly, in the estimation, we saturate the mean equation by controlling for a rich set of bank characteristics, so the residual correlations are driven by the
network linkages instead of missing variables that induce comovement among banks’ decisions.\footnote{This identification argument is not affected by time variation in G as long as we have a well-defined unconditional variance. This case is analogous to the one of S-VARs with time-varying volatility as, e.g., in Primiceri (2005).}

### 3.3 Systemic risk

In our model, shocks are realized before banks’ lending decisions and, after banks choose the loan amounts, the shocks are propagated through the payment network. The system given by equations (34) and (35) highlights the propagation mechanism: A shock to bank $j$ is transmitted to bank $i$ through $\phi w_{ij,t}$, so if $\phi w_{ij,t} > 0$ (strategic complementarity), the network amplifies shocks, and if $\phi w_{ij,t} < 0$ (strategic substitution), the network buffers shocks. Given the realized shocks, $\nu'_t = [\nu'_{1,t}, ..., \nu'_{n,t}]^\top$, the ultimate impact of shocks to all banks is given by the following vector

$$
\epsilon_t = (I - \phi W')^{-1} \nu'_t = M(\phi, W') \nu'_t,
$$

where the matrix $M(\phi, W')$ records the routes that propagate the shocks:

$$
M(\phi, W') \equiv I + \phi W' + \phi^2 W'^2 + \phi^3 W'^3 + ... = \sum_{k=0}^{\infty} \phi^k W'^k = (I - \phi W')^{-1},
$$

where the first term captures direct effects of shocks, the second is the sum of direct outbound links, the third element is the sum of second-order links, and so on.

Consider unitary shocks to all banks. $W'$ being right-stochastic (i.e., $W'1 = 1$) implies

$$
\epsilon_t = (I - \phi W')^{-1} 1 = M(\phi, W') 1 = I + \phi W'1 + \phi^2 W'^2 1 + \phi^3 W'^3 1 + ... = \frac{1}{1 - \phi} 1.
$$

Therefore, the network attenuation factor, $\phi$, can serve as a proxy for the strength of network...
amplification mechanism. In the following, we define the network multiplier.

**Definition 1 (Network Multiplier)** The network multiplier is defined as $\frac{1}{1-\phi}$.

Given the estimates of $\phi$, $\bar{\alpha}'$, and $\delta'^2$, we use our structural model to identify systemically important banks. A bank is systemically important if its shock has a disproportionately large impact on the aggregate credit supply. We call such bank the **volatility key bank** as our approach provides a decomposition of credit-supply volatility into different banks’ contributions.

Let $N_t$ denote the network-dependent component of aggregate credit supply. Note that our estimation uses the loan growth rates rather than the loan amounts, so, the link between $N_t$ and the network-dependent component of loan growth rate is given by

$$N_t = \sum_{i=1}^{N} y_{i,t-1} n_{i,t} = y_{t-1}^T n_t .$$  \hspace{1cm} (41)

Substituting in the solution of $n_t$ in (36), we obtain

$$N_t = y_{t-1}^T (I - \phi W')^{-1} (\bar{\alpha}' + \nu'_t) = y_{t-1}^T M (\phi, W') (\bar{\alpha}' + \nu'_t) .$$ \hspace{1cm} (42)

Before the shocks are realized, we calculate the conditional mean of $N_t$,

$$\mathbb{E}_{t-1}[N_t] = y_{t-1}^T M (\phi, W') \bar{\alpha}' ,$$  \hspace{1cm} (43)

and the conditional variance of $N_t$,

$$\text{Var}_{t-1} (N_t) = y_{t-1}^T M (\phi, W') \Delta' M (\phi, W')^T y_{t-1} ,$$ \hspace{1cm} (44)

where $\Delta'$ is the covariance matrix of $\nu'_t$, a diagonal matrix whose $i$-th diagonal element is $\delta'^2_i$. The
conditional mean and variance of aggregate credit supply characterize in expectation the strength of the payment network in generating bank credit provision and propagating shocks. Next, we define the volatility key bank through the network impulse response function.

**Definition 2 (Network Impulse Response Function and Volatility Key Bank)** The impulse response of aggregate credit supply to a one standard-deviation shock to a bank $i$ is given by

$$NIRF_{i,t-1} (\phi, \delta_i', W') \equiv \frac{\partial N_t}{\partial \nu_{i,t}} \delta_i' = y_{t-1}^\top \{M(\phi, W')\}_{i} \delta_i'$$

where the operator $\{\}_{i}$ returns the $i$-th column of its argument. The volatility key bank, given by

$$i^*_t = \arg \max_{i \in \{1, \ldots, N\}} NIRF_{i,t-1} (\phi, \delta_i', W'),$$

is the one that contributes the most to the conditional volatility of aggregate credit growth.

A bank’s NIRF records the impact of its shock on the aggregate credit supply. It depends on the network attenuation factor, $\phi$, the network topology given by $W'$, and the size of the bank’s shock, $\delta_i'$. Our estimation method allows us to identify both $\phi$ and $\delta_i'$. Next, we show that NIRFs measure banks’ contributions to the conditional volatility of aggregate credit supply and thus identifies the volatility key bank by providing a clear ranking of each bank’s volatility contribution.

**Proposition 3 (Credit-Supply Volatility Decomposition)** The network impulse response functions (NIRFs) decompose the conditional volatility of aggregate credit supply:

$$\text{Var}_{t-1} (N_t) = \text{vec} \left( \{NIRF_{i,t-1} (\phi, \delta_i', W')\}_{i=1}^N \right)^\top \text{vec} \left( \{NIRF_{i,t-1} (\phi, \delta_i', W')\}_{i=1}^N \right),$$

where “$\text{vec}$” is the vectorization operator.
3.4 Comparing the planner’s solution and market equilibrium

We compare the conditional expectation and conditional volatility of aggregate credit supply from the market equilibrium and those from the planner’s solution. First, we show how to utilize the parameter estimates in calculating the planner’s solution. Following Section 3.1, we define

\[ e_{wij} = e_{wij}' y_j,t - 1 \]

(48)

where \( e_{wij} \) is defined in (24), and

\[ \mu_{ij} = \mu_{ij}' y_j,t - 1 \]

(49)

The network-dependent component in the planner’s solution (50) can be written as

\[ n_i,t = \phi_i \sum_{j \neq i} w_{ij} n_j,t - \phi_i \sum_{j \neq i} \mu_{ij} \left( \sum_{k \neq j} \mu_{kj} n_k,t \right) + a_i,t' \]

(50)

where \( \phi_i = \left( \frac{1}{\phi^2 + \sigma_i^2 + \sigma_i'^2} \right)^{-1} \) is defined (23) and \( a_i,t' \equiv a_{i,t}/y_{i,t-1} \) (\( a_{i,t} \) given by (25)). Following Section 2.3, let \( \hat{W}' \) and \( U' \) denote the matrices whose the \( ij \)-th elements are equal to \( w_{ij}' \) and \( \mu_{ij}' \), respectively. Let \( \hat{a}_i \) denote the vector for \( a_i,t' \), \( i = 1, \ldots, N \). And, let \( \hat{\Phi} \) denote the diagonal matrix with the \( i \)-th diagonal element equal to \( \phi_i \). In vector form, we have:

\[ \hat{n}_t = \hat{\Phi} \left( \hat{W}' - UU'^T \right) \hat{n}_t + \hat{a}_t' \]

(51)

The planner’s choice of individual banks’ lending can be solved as follows:

\[ \hat{n}_t = \hat{M} \left( \hat{\Phi}, \hat{W}', U, U' \right) \hat{a}_t' \]

(52)
where we define
\[
\tilde{M} \left( \bar{\Phi}, \tilde{W}, U, U' \right) \equiv \left( I - \bar{\Phi} \tilde{W}' + \bar{\Phi} U U'^\top \right)^{-1} .
\] (53)

Next, we explain how to calculate the planner’s solution with payment data and parameters from our estimation of market equilibrium. Following the calculation of \( w_{ij} \) of the market equilibrium in Section 3.1, we calculate \( \tilde{w}_{ij} \) following the definition (24) using the statistics of payment flows in quarter \( t - 1 \) and obtain \( \mu_{ij} \) by calculating the average of \( g_{ij} \) in quarter \( t - 1 \). \( \mu'_{ij} \) is calculated following (49). Following Section 2.3, we normalize \( \tilde{W}' - U U'^\top \) to be right-stochastic.

We calculate \( \bar{\phi}_i \) using the estimate of \( \phi \) and the payment statistics, \( \bar{\sigma}^2_i \) and \( \bar{\mu}_i \) (see Section 3.1). To compute the mean and standard deviation of \( \bar{a}'_{i,t} \), we solve the connection between \( \bar{a}'_{i,t} \) in the planner’s solution and \( a'_{i,t} \) in (29) of the market equilibrium:
\[
\bar{a}'_{i,t} = \frac{a'_{i,t}}{y_{i,t-1}} = b'_{i,t} + \frac{\bar{\phi}_i}{\phi} a'_{i,t} ,
\] (54)

where,
\[
b'_{i,t} = \frac{\bar{\phi}_i}{\left( \bar{\sigma}^2_i + \bar{\mu}_i \right)} \left[ \left( \frac{\tau_1 - \theta_1}{\tau_2 - \theta_2} \right) \frac{\bar{\mu}_i}{y_{i,t-1}} - \left( \frac{\tau_2}{\tau_2 - \theta_2} \right) \sum_{j \neq i} \mu_{ij} \frac{m_j}{y_{j,t-1}} \right].
\] (55)

We rewrite the planner’s solution (52) in vector form:
\[
\tilde{n}_t = \tilde{M} \left( \bar{\Phi}, \tilde{W}, U, U' \right) b_{t-1}' + \tilde{M} \left( \bar{\Phi}, \tilde{W}', U, U' \right) \frac{1}{\phi} \bar{\phi} a'_t .
\] (56)

The network-dependent component of aggregate credit supply in the planner’s solution is
\[
\tilde{N}_t = \sum_{i=1}^N y_{i,t-1} \tilde{n}_{i,t} = y_{t-1}' \tilde{n}_t .
\] (57)

After obtaining the estimates of \( \phi \), \( \bar{a}'_{i,t} \) (the mean of \( a'_{i,t} \)) and \( \bar{\sigma}^2_i \) (the volatility of \( a'_{i,t} \)), we compute
the mean and volatility of second term in $\bar{a}_{i,t}$ and thus obtain the conditional mean and volatility of the second term in $\bar{n}_t$. Because the first term in $\bar{n}_t$ (i.e., $\bar{M}(\bar{\Phi}, \bar{W}', U, U')b'_{t-1}$) does not contribute to the conditional volatility, we can solve the conditional volatility of the planner’s solution of $\bar{N}_t$:

$$Var_{t-1}\left[\bar{N}_t\right] = \frac{1}{\phi^2}y_{t-1}^\top \bar{M} \left(\bar{\Phi}, \bar{W}', U, U'\right) \bar{\Phi}^2 \Delta' \bar{M} \left(\bar{\Phi}, \bar{W}', U, U'\right)^\top y_{t-1},$$  \hspace{1cm} (58)

where, as previously defined, $\Delta'$ is a diagonal matrix with the $i$-th diagonal element equal to $\delta_i'^2$.

The calculation of the conditional mean of $\bar{N}_t$,

$$E_{t-1}\left[\bar{N}_t\right] = y_{t-1}^\top \bar{M} \left(\bar{\Phi}, \bar{W}', U, U'\right) b'_{t-1} + y_{t-1}^\top \bar{M} \left(\bar{\Phi}, \bar{W}', U, U'\right) \frac{1}{\phi} \bar{\Phi} \bar{\alpha}',$$  \hspace{1cm} (59)

requires the first term in $\bar{a}_{i,t}$, and the first term in $\bar{a}_{i,t}$ depends on the parameters, $\tau_1$, $\tau_2$, $\theta_1$, and $\theta_2$ that cannot be separately identified in our estimation (as we only estimate $\phi = \frac{\tau_2 - \theta_2}{\kappa + \tau_2 - \theta_2}$ defined in (11)). Therefore, when comparing the conditional mean of $N_t$ of the market equilibrium and the conditional mean of $\bar{N}_t$ of the planner’s solution, we focus on the second component of $E_{t-1}[^{\bar{N}}_t]$ that can be computed from our parameter estimates. Moreover, the second component, $y_{t-1}^\top \bar{M} \left(\bar{\Phi}, \bar{W}', U, U'\right) \frac{1}{\phi} \bar{\Phi} \bar{\alpha}'$, is more comparable to the market-equilibrium counterpart, $E_{t-1}[^{N}_t] = y_{t-1}^\top M \left(\phi, W'\right) \bar{\alpha}'$ in (43) because the only differences are in the network propagation (i.e., $\bar{M}(\bar{\Phi}, \bar{W}', U, U')$ vs. $M(\phi, W')$) and the deviations of $\bar{\phi}_i$ from $\phi$ (captured by $\frac{1}{\phi} \bar{\Phi}$).
Table 1: Network multiplier. The table reports the estimate of $\phi$ in the system of equations (32) and (33). The $t$-statistics are calculated with quasi-MLE robust standard errors and are reported in parentheses under the estimated coefficients. The network multiplier, $1/(1 - \hat{\phi})$, is reported in the second line, and the $R^2$ in the third line is the fraction of variation explained by the control variables (i.e., the bank characteristics and macroeconomic variables).

<table>
<thead>
<tr>
<th>Number of Banks</th>
<th>500</th>
<th>500 (Not winsorized)</th>
<th>300</th>
<th>400</th>
<th>600</th>
<th>700</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\phi}$</td>
<td>0.1452</td>
<td>0.1377 (3.44)</td>
<td>0.1499 (3.16)</td>
<td>0.1562 (3.42)</td>
<td>0.1396 (3.17)</td>
<td>0.1373 (3.06)</td>
</tr>
<tr>
<td>$1/(1 - \hat{\phi})$</td>
<td>1.1698</td>
<td>1.1597</td>
<td>1.1764</td>
<td>1.1852</td>
<td>1.1623</td>
<td>1.1591</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.1139</td>
<td>0.1138</td>
<td>0.1205</td>
<td>0.1150</td>
<td>0.1183</td>
<td>0.1167</td>
</tr>
</tbody>
</table>

4 Estimation Results

4.1 The network multiplier

In this section, we present our empirical results. Table 1 reports the estimate of the key parameter $\phi$, the network attenuation factor and the implied network multiplier. Our estimation is done on different subsamples of banks ranked by the size of their deposit liabilities. The main specification includes the top 500 banks, and the results are reported in the first column. In the second column, we show that the results are similar without winsorizing $g_{ij}$ at 0.5% for the calculation of the payment-flow statistics (such as $\mu_{ij}$, $\sigma_{ij}$, and $\rho_{ij}$). In the last four columns, we report the results based on top 300, 400, 600, and 700 banks and show that the results are similar.

A key finding is that $\phi$ is positive and the network multiplier is greater than one. As discussed in Section 3.1, under $\phi > 0$ or $1/(1 - \phi) > 1$, the network amplifies unitary shocks to all banks by the amount of $1/(1 - \phi) - 1$. For example, an estimate of $\phi$ equal to 0.1452 (and a network multiplier equal to 1.1698) implies that the network amplifies the shocks by around 17%. The finding of a stable estimate of $\phi$ across different numbers of banks shows robust network effects.
that are not drive by a (core) subset of banks of large sizes.\textsuperscript{36}

The finding of $\phi > 0$ also suggests that the bank liquidity management channel dominates the customer liquidity management channel. As previously discussed in Section 2, the key feature of the two-layer payment system is that when payment outflows happen, a bank experiences reserve outflows and its depositors experience deposit outflows. The former implies a cost on the bank, while the latter implies an increase in the customers’ marginal value of liquidity and future lending opportunities for the bank. When $\phi > 0$, which implies $\tau_2 > \theta_2$, the bank’s marginal cost of losing liquidity dominates the marginal benefit of having more lending opportunities in the future. Moreover, as discussed in Section 2.2, the sign of $\phi w'_{ij}$ determines whether banks’ lending decisions are strategic complements or substitutes. In our sample, there are only 0.39\% of non-zero $w'_{ij}$ being negative.\textsuperscript{37} Therefore, a positive estimate of $\phi$ indicates strategic complementarity.

We have hundreds of banks (i.e., hundreds of $\bar{\alpha}'_i$ and $\delta'_i$) in each sample, and the samples differ in the number of $\bar{\alpha}'_i$ and $\delta'_i$, so it is more convenient to compare the estimation of $\bar{\alpha}'_i$ and $\delta'_i$ through the frequency distribution in Figure 2. The figure shows that across subsamples, the distributions of these parameters are fairly consistent, which again suggests the robustness of equilibrium characteristics of the network lending game to the selection of subsamples of banks ranked by deposit sizes. We report the estimates of control variable coefficients in Table D.2 and show that these estimates are statistically close in Figure D.1 in the appendix.

In the following, our analysis is based on the sample of top 500 banks. We analyze the impact of network externalities on aggregate credit supply. Equation (39) shows that under $\phi > 0$, each round of network propagation amplifies banks’ responses in loan growth to their own and other banks’ expected levels ($\bar{\alpha}'_i$) and shocks ($\nu'_{i,t}$). Therefore, aggregate credit supply depends on both

\textsuperscript{36}The network adjacency matrix, $W'$, is independently constructed for each subsample with only banks in the subsample as nodes on the network.

\textsuperscript{37}Among all the potential pairs, there are 6.47\% have non-zero $w'_{ij}$. 

44
Figure 2: The estimates of $\bar{\alpha}'_i$ and $\delta'_i$. This figure reports the frequency distribution of the estimates of $\bar{\alpha}'_i$ (Panel A) and $\delta'_i$ (Panel B) across different samples of banks ranked by the size of deposit liabilities.

the expected levels and shocks of individual banks, i.e., the standalone (network-independent) loan growth, but more importantly, the network, $W'$, and the network attenuation factor, $\phi$.

In Figure 3, we decompose the mean (Panel A) and volatility (Panel B) of aggregate credit supply conditional on the previous period’s bank lending ($y_{i,t-1}$) equal to the sample average. In both Panel A and Panel B, the first column shows the standalone (network-independent) value and each subsequent column corresponds to the cumulative effect after each round of network propagation. For the network adjacency matrix, $W'$, we use the average across the 44 quarters in our sample. For both conditional mean and volatility, the second and third columns correspond respectively to the direct network linkages and the first layer of indirect network linkages. Both direct and indirect linkages have significant influence on the equilibrium level of aggregate credit supply. Linkages that are more than two steps away are relatively less important. The key to this feature is the value of $\phi$. The smaller $\phi$ is, the weaker effects of distant network linkages, because
Figure 3: **Network propagation and aggregate credit supply.** This figure reports the mean (Panel A) and volatility (Panel B) of aggregate credit supply conditional on the outstanding loan amounts of the previous period (i.e., \( \{y_{i,t-1}\}_{t=1}^{N} \)). In both panels, the statistics are decomposed into each round of network propagation. We show results based on our data network and a counterfactual network where all banks are equally connected (i.e., \( w'_{ij} = 1/(N-1) \)).

As shown in (39), \( \phi \) determines the discount factor for network linkages.

In Figure 3, we also explore the importance of network topology in determining the network propagation mechanism. In the counterfactual network, which we call the uniform network, banks are equally connected (i.e., \( w'_{ij} = 1/(N-1) \)). In Panel A, relative to the hypothetical uniform network, the data network generates a lower expected level of aggregate credit supply, and in each round of network propagation, the cumulative effects of the hypothetical network are dominated those of the uniform network. In Panel B, relative to the uniform network, the data network generates a higher volatility of aggregate credit supply. Note that in both panels, the first columns under the two networks have the same value because they represent the standalone values without network propagation. The divergence happens starting the first round of network propagation. While
Figure 4: **Network topology.** This figure compares the data networks given by the average adjacency matrix $W'$ in our sample and the hypothetical uniform network. The size of node $i$ is proportional to $\delta'_i$ (the volatility of structural shock to loan growth). We apply the algorithm in Fruchterman and Reingold (1991): Linked nodes should be close and notes should be distributed widely for visibility.

Both networks generated a similar expected level, the volatility difference is large in magnitude. In our sample of top 500 banks, the average of aggregate bank lending is $6.4$ trillion. We calculate the annualized standard deviation by multiplying the quarterly value of $54$ billions per quarter in Panel B of Figure 3 by 4. Therefore, the annualized volatility generated by the payment network is $54 \times 4 / 6400 = 3.4\%$. In contrast, the counterfactual network of equally connected banks generates an annualized volatility of $2.8\%$ (implied by $45$ billions in Panel B of Figure 3).

In Figure 5, we compare the data network given by the average adjacency matrix $W''$ and the uniform network. The size of node $i$ is proportional to $\delta'_i$ (the volatility of bank-specific shock to loan growth). The most connected nodes are placed at the center while the least connected at the periphery (Fruchterman and Reingold, 1991). The distribution of edges (linkages) of the data network is much more uneven, suggesting less heterogeneity in banks’ network positions.
Figure 5: Network and fat tails. We simulate 10,000 times a vector of 500 i.i.d. standard normal shocks and, for each simulation, we calculate the average of the shocks and the averages of shocks (denoted by $\nu$) amplified by two networks, i.e., the averages of vector $(\mathbf{I} - \phi \mathbf{G})^{-1} \nu$ where the two networks are $\mathbf{G} = \mathbf{W}'$ (the data network) and the uniform network. The figure reports the frequency distribution of the averages of 10,000 simulations.

The topology of payment network directly affects the aggregation of bank-level (granular) shocks. As emphasized by Acemoglu, Carvalho, Ozdaglar, and Tahbaz-Salehi (2012), network propagation may cause the law of large numbers to fail on the aggregation of idiosyncratic shocks as the number of nodes goes to infinity. While we cannot examine the asymptotic behavior as our sample contains a finite number (500) of banks, we show in Figure 5 that the data network generates fatter tails than the uniform network. Specifically, we simulate 10,000 times a vector of 500 i.i.d. standard normal shocks. For each simulation, we calculate the simple average (which has a standard deviation of $\sqrt{\frac{1}{500}} = 0.045$) and the average of shocks (denoted by $\nu$) amplified by the two networks, i.e., the averages of vector $(\mathbf{I} - \phi \mathbf{G})^{-1} \nu$ where the two networks are $\mathbf{G} = \mathbf{W}'$ (the data network) and the uniform network. The figure reports the frequency distribution of the and shows fatter tails from the shock propagation of the payment network.
Figure 6: **Bank shock size and NIRF.** In this figure, we plot the size of bank-specific shock, $\delta_i'$, and network impulse response function (NIRF) for the five hundred banks in our sample.

### 4.2 Volatility key bank

We define volatility key bank in (45) as the bank with the highest network impulse response function (NIRF) and, in (47), we show that the volatility of (network-dependent component of) aggregate credit supply conditional on the lending distribution in the previous period (i.e., $\{y_{i,t-1}\}_{i=1}^{N}$) can be decomposed into individual banks’ NIRFs. Therefore, ranking banks by their NIRFs is equivalent to ranking banks by their contributions to credit-supply volatility. Next, we analyze how banks’ positions in the network given by the adjacency matrix, $W'$, and the sizes of their structural shocks, $\{\delta_i'^2\}_{i=1}^{N}$ determine their NIRFs. As shown in Figure 5, banks differ significantly in both aspects. Therefore, we expect to see strong cross-section heterogeneity in NIRFs.

In Figure 6, we plot the loan amount implied by the size of bank-specific shock to growth rate (i.e., $y_{i,t-1}\delta_i'$), and network impulse response function (NIRF) for the top five hundred banks by deposit size. For both quantities, we set the loan amounts from the previous period, $y_{i,t-1}$, to the
When \( y_{i,t-1} \delta'_i \) and NIRF are close for a bank, the payment network does not have a significant effect on the bank’s contribution to the volatility of aggregate credit supply. In other words, what the bank contributes is close in magnitude to the size of its own shock. In contrast, when NIRF and \( y_{i,t-1} \delta'_i \) are very different for a bank, the bank’s position in payment network significantly affects its contribution to the volatility of aggregate credit supply. In Figure 6, we see the wedge between NIRF and \( y_{i,t-1} \delta'_i \) is particularly large for a handful of banks. This finding suggests that the payment network amplifies the shocks to a relatively small number of banks and therefore generates heterogeneity in banks’ contribution to the volatility of aggregate credit supply that is beyond the heterogeneity from banks’ difference in the size of their shocks \( \delta'_i \).

Beyond the implications on aggregate credit supply, our finding in Figure 6 also sheds light on how payment network externalities affect the cross-sectional distribution of credit-supply volatility. The volatilities of individual banks’ lending are main sources of uncertainty in the funding environment of bank-dependent firms and households. When the payment network amplifies volatilities for certain banks and dampen volatilities for others, the ultimate impact on the real economy depends on whether borrowers are able to smooth out volatilities by switching between different lenders. Frictions that limit borrowers’ mobility transmit credit-supply volatilities to bank-financed investment of firms and households’ purchases of services, goods, and real estate.\(^{38}\)

In Panel A of Figure 7, we take the ratio of a bank’s network impulse response function (NIRF) to its average loan amount in our sample. We rank banks by their NIRF and plot the ratio for each bank. Note that a bank’s NIRF is comparable in magnitude to its loan value. As shown in the definition (45), NIRF is given by the product between a bank’s lending in the previous period and its equilibrium growth loan growth rate given the realized shock equal to the standard deviation \( \delta'_i \). If bank size is an adequate proxy for a bank’s systemic importance, we would expect a relatively

---

\(^{38}\)Ongena and Smith (2001) empirically characterize firms with greater mobility in lending relationships.
Figure 7: **Network topology and bank NIRF.** In Panel A, we take the ratio of a bank’s network impulse response function (NIRF) and the bank’s average loan amount in our sample. The flat line is drawn from the average NIRF divided by the cross-section average of banks’ average loan amount in our sample. In Panel B, we take the ratio of a bank’s NIRF to the counterfactual NIRF implied by a uniform network, where all banks are equally connected (i.e., \( w_{ij}' = 1/(N - 1) \)). The flat line is drawn from the average NIRF dividend by the average NIRF implied by the uniform network. When calculating both NIRFs, we use the same estimates of parameters of the lending game. In both panels, we rank banks by their NIRFs and plot the ratio for each bank.

flat line. In contrast, the figure shows strong heterogeneity. Scaled by the size of lending, banks differ significantly in their contributions to the credit-supply volatility. In other words, larger banks are not necessarily more important in the sense of generating systemic risk in the credit supply.

To further investigate on the impact of network topology on banks’ contributions to credit-supply volatility, we take the ratio of a bank’s NIRF to the counterfactual NIRF implied by a uniform network, where all banks are equally connected (i.e., \( w_{ij}' = 1/(N - 1) \)). If the ratio is close to one, the topology of payment network does not affect the bank’s contribution to credit-supply volatility relative to an equally connected network. If the ratio is greater (smaller) than one, the payment network has an amplification (dampening) effect. In Panel B of Figure 7, we rank banks by their NIRFs and plot the ratio for each bank. Except for less than fifty banks having a ratio...
greater than one, the network actually has a buffering effect, relative to a uniform network, when it comes to the propagation of individual banks’ shocks to the aggregate credit supply. However, for banks with the ratio greater than one, the amplification effect is significant. As discussed in Section 4.1, strategic complementarity under $\phi > 0$ generates a shock amplification mechanism. Our analysis in Figure 6 and 7 shows that the amplification works through a small subset of banks.

As shown in (47), the volatility of aggregate credit supply can be decomposed into individual banks’ NIRFs. In Figure 8, we rank banks by their NIRFs and, starting from the bank with the highest NIRF, we accumulate banks’ contribution to the conditional volatility of aggregate credit supply (conditional on the lending distribution of previous period, i.e., $\{y_{i,t-1}\}_{i=1}^N$, being equal to the sample-average lending distribution). The cumulative volatility is divided by the total conditional volatility of the network-dependent component of aggregate credit supply given by (44). The curve ends at 100% because after
4.3 Comparing the planner’s solution and market equilibrium

We apply the framework in Section 3.4 to compare the market equilibrium and the planner’s solution. The planner maximizes the total profits of all banks, internalizing the liquidity externality
and hedging externality through the payment network. In Panel A of Figure 9, we decompose the expected aggregate credit supply (conditional on previous loan amounts, i.e., \( \{y_{i,t-1}\}_{i=1}^{N} \)) into rounds of network propagation. The first column in both cases is generated by the loan growth rate independent from any network effects (i.e., \( \bar{\alpha}'_i \) for the market equilibrium and \( \hat{\phi}_i \alpha'_i / \phi \) in the planner’s solution). The second column adds to the first column the impact of direct network linkages, and the third column adds to the second column the impact of first-degree indirect linkages. The planner’s solution differs from the market equilibrium by internalizing the spillover effects of banks’ lending decisions. Once the network effects are activated (i.e., starting from the second column), the planner’s solution features a higher expected level of credit supply. The wedge is stable across rounds of network propagation, suggesting that the main difference between the planner’s solution and market equilibrium is due to the direct network linkages.

In Panel B of Figure 9, we decompose the volatility of aggregate credit supply (conditional on \( \{y_{i,t-1}\}_{i=1}^{N} \)) into rounds of network propagation. By internalizing the spillover effects of individual banks’ lending decisions, the planner responds to the shocks to individual banks differently from the market equilibrium, so the planner’s aggregate credit supply features a volatility that is around 10% below that of the market equilibrium. Overall, the planner’s solution features a risk-return trade-off that is superior to that implied by the market equilibrium. In other words, payment network externalities induce a lower expected level of credit supply and higher volatility.

In Figure 10, we compare the planner’s solution and market equilibrium through the distribution of lending volatility and expected level across banks. Many borrowers rely on relationship lending. Therefore, the distribution of credit across banks affects the real economy. In Panel A of Figure 10, we plot the histogram of banks’ volatilities of banks’ lending given by the market equilibrium condition (36). Using the planner’s solution (52), we also calculate the volatility of banks’ lending implied by the planner’s solution. The volatility distribution of the market equilibrium is
Figure 10: **NIRF and expected network lending distribution: Market equilibrium vs. planner’s solution.** In Panel A, we plot the histogram of banks’ NIRFs obtained from the market equilibrium and planner’s solution. In Panel B, we plot banks’ expected lending in the network game from the market equilibrium and planner’s solution.

The distribution of NIRFs is tilted to the right relative to the planner’s distribution, suggesting more volatile credit supply at bank level. A borrower can switch from a bank with a higher lending volatility to a more stable lender can benefit from having a more stable credit supply condition.

In Panel B of Figure 10, we calculate the expected levels of lending for individual banks using the market equilibrium condition (36) and the planner’s solution (52) and plot the histogram for both cases. Note that, as discussed in Section 3.1, the constant among control variables absorbs the average lending, so the estimates of $\bar{\alpha}_t$ can potentially be negative. The distribution of expected lending in the market equilibrium exhibits wider dispersion than that of the planner’s solution. This finding suggests that payment network externalities generate a greater cross-sectional dispersion of bank lending and thus makes any frictions limiting borrowers’ mobility more costly.

In Figure 11, we present the rolling estimation results. We conduct rolling estimation with
Figure 11: Rolling estimation: market equilibrium vs. the planner’s solution. In this figure, we report the rolling estimation results with each rolling window containing twenty two quarters (i.e., half of the total forty four quarters in our sample). We report the estimate of network multiplier in Panel A together with the confidence band of two standard errors. In Panel B and C, we compare respectively the volatility and expectation of aggregate credit supply implied by the loan growth rates in the market equilibrium and planner’s solution (conditional on previous lending amounts, \( \{ y_{i,t-1} \}_{i=1}^{N} \) where \( y_{i,t-1} \) is set to the full-sample average). In Panel B, we also plot the sum of banks’ network-independent volatilities conditional on previous loan amounts (i.e., \( \{ y_{i,t-1} \delta_{i}^{t} \}_{i=1}^{N} \)). In Panel C, we also plot the sum of banks’ network-independent expected lending conditional on previous loan amounts (i.e., \( \{ y_{i,t-1} \bar{\alpha}_{i}^{t} \}_{i=1}^{N} \)).

each rolling window containing twenty two quarters (i.e., half of the total forty four quarters in our sample). In Panel A of Figure 11, we report the estimate of the network multiplier and the confidence interval of two standard errors from the method of Bollerslev and Wooldridge (1992) that is robust to non-normality of shocks in quasi-MLE. The estimate is plotted against the last quarter of the rolling sample. The multiplier demonstrates significant variation over time. During the Covid-19 pandemic, banks experience larger shocks and greater heterogeneity in shock exposure, so our estimate of \( \phi \) contains more noise and has a wider standard-error band.

Next, we compare the volatility and expectation of aggregate credit supply implied by the loan growth rates in the market equilibrium and planner’s solution (conditional on previous lending
amounts, \( \{y_{i,t-1}\}_{i=1}^{N} \) where \( y_{i,t-1} \) is set to the full-sample average). The dynamics of wedge between the two equilibria follow the dynamics of network multiplier. When \( \phi \) is higher, the network externalities are stronger, which then implies a greater difference between the two equilibria.

In Panel B of Figure 11, we show that during the period of low \( \phi \) (the rolling windows ending between 2018 and 2019), the conditional volatility of planner’s credit supply is close to the simple sum of banks’ volatilities independent of network effects (i.e., \( \{y_{i,t-1}\delta_{i}'\}_{i=1}^{N} \)). During this period, payment network externalities amplify individual banks’ shocks so market equilibrium generates a higher volatility of aggregate credit supply than the sum of banks’ network-independent volatilities. The volatility wedge can be as high as $8 billions per quarter (i.e., annualized volatility of \( 8 \times 4/6400 = 0.5\% \) given the average aggregate bank credit of $6.4 trillions in our sample).

In Panel C of Figure 11, we plot the conditional expectation. Both the market equilibrium and planner’s solution feature a higher level of credit supply than what is implied by the simple sum of banks’ network-independent credit provision. Therefore, the payment network has a overall positive effect on amplifying the aggregate credit supply through the circulation of liquidity among banks. Across different time periods, the wedge between the market equilibrium and planner’s solution is larger when the estimate of \( \phi \) is larger in Panel A. Over time, both the conditional volatility (in Panel B) and expectation (in Panel A) of planner’s credit supply exhibits much smaller variations than those of the market equilibrium, suggesting that payment network externalities generate significant uncertainty in the credit conditions for the real economy.

5 Conclusion

We provide the first evidence on how payment-flow topology affects the supply of bank credit. The payment network generates strategic complementarity in banks’ lending decisions and amplifies
shocks to individual banks. Our analysis reveals a subset of systemically important banks that have a large influence on the level and fluctuation of aggregate credit supply due to their special positions in the payment network. We quantify the network externalities and show that policy interventions targeted at such externalities may improve the risk-return profile of credit supply.

Our framework offers a new theoretical underpinning of the concept of money multiplier. In our model, banks finance lending with deposits and hold reserves to cover payment outflows under real-time gross settlement (RTGS), creating a natural link between the monetary base and the creation of credit and deposits. Importantly, liquidity percolation through the payment network generates interconnectedness in banks’ liquidity conditions. Therefore, the money multiplier in equilibrium depends on the topology of payment flows.

Our paper offers policy guidance in the rapidly growing space of digital payment. Technology-driven entrants rewire the payment flows, and central banks around the world actively research on the implications of central bank digital currency (CBDC). The current discussion on payment system reforms focuses on operational efficiency and technological vulnerabilities. Our paper broadens the attention to implications on credit supply and provides an equilibrium-based empirical framework to quantify the impact of payment-network changes on credit conditions.
References


Copeland, A., D. Duffie, and Y. Yang (2021). Reserves were not so ample after all. Staff Reports 974, Federal Reserve Bank of New York.


64


65


Appendix: Background Information on Payment Systems

The Fedwire Funds Service is the primary payment system in U.S. for large-value domestic and international USD payments. It is a real-time gross settlement system that enables participants to initiate funds transfer that are immediate, final, and irrevocable once processed. The service is operated by the Federal Reserve Banks. Financial institutions that hold an account with a Federal Reserve Bank are eligible to participate in the service and electronically transfer funds between each other. Such institutions include Federal Reserve member banks, nonmember depository institutions, and certain other institutions, such as U.S. branches and agencies of foreign banks.

Participants originate funds transfers by instructing a Federal Reserve Bank to debit funds from its own account and credit funds to the account of another participant. To make transfers, the following information is submitted to the Federal Reserve: the receiving bank’s routing number, account number, name and dollar amount being transferred. Each transaction is processed individually and settled upon receipt. Wire transfers sent via Fedwire are completed the same business day, with many being completed instantly. Participants may originate funds transfers online, by initiating a secure electronic message, or offline, via telephone procedures.

Participants of Fedwire Funds Service can use it to send or receive payments for their own accounts or on behalf of corporate or individual clients. In the paper, we focus on Fedwire fund transfers made on behalf of banks’ corporate or individual clients, which make up about 80% of total transactions in terms of transaction number.

The Fedwire Funds Service business day begins at 9:00 p.m. eastern standard time (EST) on the preceding calendar day and ends at 7:00 p.m. EST, Monday through Friday, excluding designated holidays. For example, the Fedwire Funds Service opens for Monday at 9:00 p.m. on the preceding Sunday. The deadline for initiating transfers for the benefit of a third party (such as a bank’s customer) is 6:00 p.m. EST each business day and 7:00 p.m. EST for banks own transactions. Under certain circumstances, Fedwire Funds Service operating hours may be extended by the Federal Reserve Banks.

To facilitate the smooth operation of the Fedwire Funds Service, the Federal Reserve Banks offer intraday credit, in the form of daylight overdrafts, to financially healthy Fedwire participants with regular access to the discount window. Many Fedwire Funds Service participants use daylight credit to facilitate payments throughout the operating day. Nevertheless, the Federal Reserve Policy on Payment System Risk prescribes daylight credit limits, which can constrain some Fedwire
Funds Service participants’ payment operations. Each participant is aware of these constraints and is responsible for managing its account throughout the day.

The usage of Fedwire Funds Service grows over our sample period from 2010 to 2020, with total number of transfers and transaction dollar value increasing by 47% and 38%, respectively. In 2020, approximately 5,000 participants initiate funds transfers over the Fedwire Funds Service, and the Fedwire Funds Service processed an average daily volume of 727,313 payments, with an average daily value of approximately $3.3 trillion. The distribution of these payments is highly skewed, with a median value of $24,500 and an average value of approximately $4.6 million. In particular, only about 7% of Fedwire fund transfers are for more than $1 million.

The other important interbank payment system in U.S. is the Clearing House Interbank Payments System (CHIPS), which is a private clearing house for large-value transactions between banks. In 2020, CHIPS processed an average daily volume of 462,798 payments, with an average daily value of approximately $1.7 trillion, about half of the daily value processed by Fedwire. There are three key differences between CHIPS and Fedwire Funds Service. First, CHIPS is privately owned by The Clearing House Payments Company LLC, while Fedwire is operated by the Federal Reserve. Second, CHIPS has less than 50 member participants as of 2020, compared with thousands of banking institutions making and receiving funds via Fedwire. Third, CHIPS is not a real-time gross settlement (RTGS) system like Fedwire, but a netting engine that uses bilateral and multi-lateral netting to consolidates pending payments into single transactions. The netting mechanism significantly reduces the impact of payment flows on banks’ decision making (and therefore our sample focuses on the RTGS, Fedwire) but exposes banks to potential counterparty risks.

---

39 Data source: www.frbservices.org. Federal Reserve also operates two smaller payment systems, National Settlement Service (NSS) with an average daily settlement value of $93 billions in 2020 (source: www.frbservices.org), and FedACH with an average daily settlement value of $122.8 billion in 2020 (source: www.federalreserve.gov).

40 Data source: https://www.theclearinghouse.org
B Appendix: Bank Customer Liquidity Management

In this section, we microfound the component, \( \theta_1 x_i + \frac{\theta_2}{2} x_i^2 \), of bank \( i \)'s objective function by modelling the liquidity management problem of bank \( i \)'s customers.

In aggregate, bank \( i \)'s customers lose liquidity \( x_i \), which is equal to the payment outflow to other banks' customers. To cover the liquidity shortfall, bank \( i \)'s customers may borrow from bank \( i \), for example, in the form of lines of credit.\(^\text{41}\) Consider a unit mass of customers and the evenly distributed loss of liquidity (i.e., each customer’s loss of liquidity is equal to \( x_i \)). A representative customer chooses \( c \), the amount of liquidity obtained from bank \( i \) (for example, the size of lines of credit). Bank \( i \) charges a proportional price \( P_c \). The customer’s problem is given by

\[
\max_c \xi_1 \left[ c - x_i - \frac{1}{2\xi_2} (c - x_i)^2 \right] - c P_c,
\]

where the parameter \( \xi_1 (> 0) \) captures the overall demand for liquidity and the parameter \( \xi_2 (> 0) \) captures the decreasing return to liquidity. A key economic force is that a higher \( x_i \) increases the marginal benefit of \( c \). In other words, when bank \( i \)'s customers lose liquidity through payment outflows to other banks’ customers, they rely more on bank \( i \) for liquidity provision.

From the customer’s first order condition for \( c \),

\[
\xi_1 - \frac{\xi_1}{\xi_2} (c - x_i) = P_c,
\]

we solve the optimal \( c \):

\[
c = \xi_2 \left( 1 - \frac{P_c}{\xi_1} \right) + x_i.
\]

The customer’s liquidity demand is stronger following a greater payment outflow, \( x_i \) and when the marginal value of liquidity declines slower (i.e., under a greater value of \( \xi_2 \)). A higher value of \( \xi_1 \) or a lower price \( P_c \) also increase \( c \). Under the homogeneity of bank \( i \)'s customers, equation (B.3) is also the aggregate liquidity demand for the unit mass of bank \( i \)'s customers.

Bank \( i \) sets the price \( P_c \) to maximize its profits from liquidity provision:

\[
\max_{P_c} \xi_2 \left( 1 - \frac{P_c}{\xi_1} \right) + x_i \right] P_c.
\]

\(^{41}\)Empirically, cash and lines of credit are substitutes (Lins, Servaes, and Tufano, 2010).
Here we assume relationship banking so bank i’s customers cannot obtain liquidity elsewhere. This translates into bank i’s market power and monopolistic profits. From the first-order condition for \( P_c \),

\[
- \frac{\xi_2}{\xi_1} P_c + \xi_2 \left( 1 - \frac{P_c}{\xi_1} \right) + x_i = 0, \tag{B.5}
\]

we solve the optimal \( P_c \):

\[
P_c = \frac{\xi_1}{2} \left( 1 + \frac{x_i}{\xi_2} \right). \tag{B.6}
\]

Substituting the optimal \( P_c \) into bank i’s profits, we obtain the maximized profits:

\[
\frac{\xi_1 \xi_2}{4} \left( 1 + \frac{x_i}{\xi_2} \right)^2 = \frac{\xi_1 \xi_2}{4} + \frac{\xi_1}{2} x_i + \frac{\xi_1}{4 \xi_2} x_i^2, \tag{B.7}
\]

which corresponds to the component, \( \theta_1 x_i + \frac{\theta_2}{2} x_i^2 \), of bank i’s objective function in the main text with

\[
\theta_1 = \frac{\xi_1}{2}, \text{ and, } \theta_2 = \frac{\xi_1}{2 \xi_2}. \tag{B.8}
\]

The constant \( \frac{\xi_1 \xi_2}{4} \) is omitted in bank i’s objective function in the main text.
Appendix: Derivation Details

C.1 Solving the equilibrium

Let $\phi$ denote the correlation (not negative of correlation). We have

$$
\mathbb{E} \left[ (x_i - m_i)^2 \right] = \text{Var}(x_i) + \mathbb{E} [x_i - m_i]^2
$$

(C.1)

$$
\begin{align*}
\mathbb{E} \left[ (x_i - m_i)^2 \right] &= \text{Var} \left( \sum_{j \neq i} g_{ij} y_i - \sum_{j \neq i} g_{ji} y_j \right) + \mathbb{E} \left[ \sum_{j \neq i} g_{ij} y_i - \sum_{j \neq i} g_{ji} y_j - m_i \right]^2 \\
&= \sum_{j \neq i} \text{Var} \left( g_{ij} y_i - g_{ji} y_j \right) + \left( \sum_{j \neq i} \mu_{ij} y_i - \sum_{j \neq i} \mu_{ji} y_j - m_i \right)^2 \\
&= \sum_{j \neq i} \left( y_i^2 \sigma_{ij}^2 + y_j^2 \sigma_{ji}^2 - 2y_i y_j \sigma_{ij} \sigma_{ji} \rho_{ij} \right) + \left( \sum_{j \neq i} \mu_{ij} y_i - \sum_{j \neq i} \mu_{ji} y_j - m_i \right)^2,
\end{align*}
$$

(C.2)

$$
\mathbb{E} \left[ x_i^2 \right] = \text{Var}(x_i) + \mathbb{E} [x_i]^2 = \text{Var}(x_i) + \mathbb{E} \left[ \sum_{j \neq i} g_{ij} y_i - \sum_{j \neq i} g_{ji} y_j \right]^2
$$

$$
\begin{align*}
\mathbb{E} \left[ x_i^2 \right] &= \sum_{j \neq i} \text{Var} \left( g_{ij} y_i - g_{ji} y_j \right) + \left( \sum_{j \neq i} \mu_{ij} y_i - \sum_{j \neq i} \mu_{ji} y_j \right)^2 \\
&= \sum_{j \neq i} \left( y_i^2 \sigma_{ij}^2 + y_j^2 \sigma_{ji}^2 - 2y_i y_j \sigma_{ij} \sigma_{ji} \rho_{ij} \right) + \left( \sum_{j \neq i} \mu_{ij} y_i - \sum_{j \neq i} \mu_{ji} y_j \right)^2,
\end{align*}
$$

(C.2)

$$
\mathbb{E} \left[ z_i^2 \right] = \text{Var}(z_i) + \mathbb{E} [z_i]^2 = \text{Var}(z_i) + \mathbb{E} \left[ \sum_{j \neq i} g_{ij} y_i \right]^2
$$

$$
\begin{align*}
\mathbb{E} \left[ z_i^2 \right] &= \sum_{j \neq i} \text{Var} \left( g_{ij} y_i \right) + \left( \sum_{j \neq i} \mu_{ij} y_i \right)^2 \\
&= \sum_{j \neq i} \left( y_i^2 \sigma_{ij}^2 + \mu_{ij}^2 \right) = y_i^2 \left( \sigma_{-i}^2 + \mu_{-i}^2 \right),
\end{align*}
$$

(C.3)

where, to simplify the notations, we define

$$
\overline{\mu}_{-i} \equiv \sum_{j \neq i} \mu_{ij}
$$

(C.4)
and
\[
\sigma_{-i}^2 = \sum_{j \neq i} \sigma_{ij}^2 = \text{Var} \left( \sum_{j \neq i} g_{ij} \right) = \mathbb{E} \left[ \left( \sum_{j \neq i} g_{ij} \right)^2 \right] - \left( \mathbb{E} \left[ \sum_{j \neq i} g_{ij} \right] \right)^2
\]  \hspace{1cm} (C.5)

where the second equality is based on the fact that \(g_{ij}\) is independent across \(j\) (pairs).

To solve the first-order condition for \(y_i\), we use
\[
\frac{\partial \mathbb{E} \left[ x_i - m_i \right]}{\partial y_i} = \sum_{j \neq i} \mu_{ij} = \mu_{-i},
\]  \hspace{1cm} (C.6)
\[
\frac{\partial \mathbb{E} \left[ x_i \right]}{\partial y_i} = \sum_{j \neq i} \mu_{ij} = \mu_{-i},
\]  \hspace{1cm} (C.7)
\[
\frac{\partial \mathbb{E} \left[ (x_i - m_i)^2 \right]}{\partial y_i} = 2 \sum_{j \neq i} \left( y_i \sigma_{ij}^2 - \rho_{ij} \sigma_{ij} \sigma_{ji} y_j \right) + 2 \left( \sum_{j \neq i} \mu_{ij} y_i - \sum_{j \neq i} \mu_{ji} y_j - m_i \right) \left( \sum_{j \neq i} \mu_{ij} \right)
\]  \hspace{1cm} (C.8)
\[
\frac{\partial \mathbb{E} \left[ x_i^2 \right]}{\partial y_i} = 2 \sum_{j \neq i} \left( y_i \sigma_{ij}^2 - \rho_{ij} \sigma_{ij} \sigma_{ji} y_j \right) + 2 \left( \sum_{j \neq i} \mu_{ij} y_i - \sum_{j \neq i} \mu_{ji} y_j \right) \left( \sum_{j \neq i} \mu_{ij} \right)
\]  \hspace{1cm} (C.9)
\[
\frac{\partial \mathbb{E} \left[ z_i^2 \right]}{\partial y_i} = 2 y_i (\sigma_{-i}^2 + \mu_{-i}^2)
\]  \hspace{1cm} (C.10)

The first-order condition for \(y_i\):
\[
0 = \varepsilon_i + R - 1 - \tau_1 \mu_{-i} + \theta_1 \mu_{-i} - y_i \kappa (\sigma_{-i}^2 + \mu_{-i}^2)
\]  
\[
- \tau_2 \left[ \sum_{j \neq i} \left( y_i \sigma_{ij}^2 - \rho_{ij} \sigma_{ij} \sigma_{ji} y_j \right) + \left( y_i \mu_{-i} - \sum_{j \neq i} \mu_{ji} y_j - m_i \right) \mu_{-i} \right]
\]  
\[
+ \theta_2 \left[ \sum_{j \neq i} \left( y_i \sigma_{ij}^2 - \rho_{ij} \sigma_{ij} \sigma_{ji} y_j \right) + \left( y_i \mu_{-i} - \sum_{j \neq i} \mu_{ji} y_j \right) \mu_{-i} \right]
\]
which can be further simplified to
\[
0 = \varepsilon_i + R - 1 - (\tau_1 - \theta_1)\mu_{-i} + \tau_2\mu_{-i}m - y_i (\kappa + \tau_2 - \theta_2) (\sigma^2_{-i} + \mu^2_{-i})
+ (\tau_2 - \theta_2) \sum_{j \neq i} (\rho_{ij}\sigma_{ij}\sigma_{ji} + \mu_{-i}\mu_{ji}) y_j
\] (C.11)

From this condition, we solve the optimal \(y_i\).

### C.2 Solving the planner’s solution

To solve the planner’s solution, we calculate the following derivatives:

\[
\frac{\partial \mathbb{E} [(x_j - m_j)^2]}{\partial y_i} = \frac{\partial}{\partial y_i} \left\{ \sum_{k \neq j} \left( y_j^2\sigma^2_{jk} + y_k^2\sigma^2_{kj} - 2y_jy_k\sigma_{jk}\sigma_{kj}\rho_{jk} \right) + \left( \sum_{k \neq j} \mu_{jk}y_j - \sum_{k \neq j} \mu_{kj}y_k - m_j \right)^2 \right\}
\]

\[
= 2 \left( y_i\sigma^2_{ij} - \rho_{ij}\sigma_{ij}\sigma_{ji}y_j \right) - 2 \left( \sum_{k \neq j} \mu_{jk}y_j - \sum_{k \neq j} \mu_{kj}y_k - m_j \right) \mu_{ij}
\]

\[
= 2 \left( y_i\sigma^2_{ij} - \rho_{ij}\sigma_{ij}\sigma_{ji}y_j \right) - 2 \left( y_j\mu_{-j} - \sum_{k \neq j} \mu_{kj}y_k - m_j \right) \mu_{ij}
\] (C.12)

\[
\frac{\partial \mathbb{E} [x_j^2]}{\partial y_i} = \frac{\partial}{\partial y_i} \left\{ \sum_{k \neq j} \left( y_j^2\sigma^2_{jk} + y_k^2\sigma^2_{kj} - 2y_jy_k\sigma_{jk}\sigma_{kj}\rho_{jk} \right) + \left( \sum_{k \neq j} \mu_{jk}y_j - \sum_{k \neq j} \mu_{kj}y_k \right)^2 \right\}
\]

\[
= 2 \left( y_i\sigma^2_{ij} - \rho_{ij}\sigma_{ij}\sigma_{ji}y_j \right) - 2 \left( \sum_{k \neq j} \mu_{jk}y_j - \sum_{k \neq j} \mu_{kj}y_k \right) \mu_{ij}
\]

\[
= 2 \left( y_i\sigma^2_{ij} - \rho_{ij}\sigma_{ij}\sigma_{ji}y_j \right) - 2 \left( y_j\mu_{-j} - \sum_{k \neq j} \mu_{kj}y_k \right) \mu_{ij}
\] (C.13)

\[
\frac{\partial \mathbb{E} [z_j^2]}{\partial y_i} = 0
\] (C.14)
The first-order condition for $y_i$:

$$0 = \varepsilon_i + R - 1 - \tau_1 \mu_{-i} + \theta_1 \mu_{-i} - y_i \kappa (\sigma_{-i}^2 + \bar{\mu}_{-i}^2)$$

$$- \tau_2 \left[ \sum_{j \neq i} (y_i \sigma_{ij}^2 - \rho_{ij} \sigma_{ij} \sigma_{ji} y_j) + \left( y_i \bar{\mu}_{-i} - \sum_{j \neq i} \mu_{ji} y_j - m_i \right) \bar{\mu}_{-i} \right]$$

$$+ \theta_2 \left[ \sum_{j \neq i} (y_i \sigma_{ij}^2 - \rho_{ij} \sigma_{ij} \sigma_{ji} y_j) + \left( y_i \bar{\mu}_{-i} - \sum_{j \neq i} \mu_{ji} y_j \right) \bar{\mu}_{-i} \right]$$

$$+ \sum_{j \neq i} (\tau_1 - \theta_1) \mu_{ij} - (\tau_2 - \theta_2) (y_i \sigma_{ij}^2 - \rho_{ij} \sigma_{ij} \sigma_{ji} y_j) + (\tau_2 - \theta_2) \left( y_j \bar{\mu}_{-j} - \sum_{k \neq j} \mu_{kj} y_k \right) \mu_{ij} - \tau_2 m_j \mu_{ij}$$

$$= \varepsilon_i + R - 1 - (\tau_1 - \theta_1) \bar{\mu}_{-i} + \tau_2 \bar{\mu}_{-i} m - y_i (\kappa + \tau_2 - \theta_2) (\sigma_{-i}^2 + \bar{\mu}_{-i}^2)$$

$$+ (\tau_2 - \theta_2) \sum_{j \neq i} (\rho_{ij} \sigma_{ij} \sigma_{ji} + \bar{\mu}_{-i} \mu_{ji}) y_j$$

$$+ (\tau_1 - \theta_1) \bar{\mu}_{-i} - (\tau_2 - \theta_2) y_i \sigma_{-i}^2 + (\tau_2 - \theta_2) \sum_{j \neq i} (\rho_{ij} \sigma_{ij} \sigma_{ji} + \bar{\mu}_{-j} \mu_{ij}) y_j$$

$$- (\tau_2 - \theta_2) \sum_{j \neq i} \left( \sum_{k \neq j} \mu_{kj} y_k \right) \mu_{ij} - \sum_{j \neq i} \tau_2 m_j \mu_{ij}$$

$$= \varepsilon_i + R - 1 + \tau_2 \bar{\mu}_{-i} m - y_i (\kappa + 2 \tau_2 - 2 \theta_2) \sigma_{-i}^2 - y_i (\kappa + \tau_2 - \theta_2) \bar{\mu}_{-i}^2$$

$$+ (\tau_2 - \theta_2) \sum_{j \neq i} (2 \rho_{ij} \sigma_{ij} \sigma_{ji} + \bar{\mu}_{-i} \mu_{ji} + \bar{\mu}_{-j} \mu_{ij}) y_j - (\tau_2 - \theta_2) \sum_{j \neq i} \mu_{ij} \left( \sum_{k \neq j} \mu_{kj} y_k \right) - \sum_{j \neq i} \tau_2 m_j \mu_{ij}$$

From this condition, we solve the planner’s choice of optimal $y_i$. 
### Appendix: Additional Tables and Figures

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Mean</th>
<th>S.D.</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quarterly loan growth rate</td>
<td>22000</td>
<td>0.0230</td>
<td>0.0550</td>
<td>-0.0016</td>
<td>0.0143</td>
<td>0.0341</td>
</tr>
<tr>
<td><strong>Bank Characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(Asset) (unit: log(USD '000))</td>
<td>22000</td>
<td>15.13</td>
<td>1.41</td>
<td>14.15</td>
<td>14.69</td>
<td>15.72</td>
</tr>
<tr>
<td>Liquid Assets/Total Assets</td>
<td>22000</td>
<td>0.18</td>
<td>0.12</td>
<td>0.10</td>
<td>0.16</td>
<td>0.24</td>
</tr>
<tr>
<td>Capital/Total Assets</td>
<td>22000</td>
<td>0.11</td>
<td>0.03</td>
<td>0.09</td>
<td>0.10</td>
<td>0.12</td>
</tr>
<tr>
<td>Deposits/Total Assets</td>
<td>22000</td>
<td>0.68</td>
<td>0.12</td>
<td>0.63</td>
<td>0.70</td>
<td>0.75</td>
</tr>
<tr>
<td>Return on asset</td>
<td>22000</td>
<td>0.0026</td>
<td>0.0025</td>
<td>0.0018</td>
<td>0.0025</td>
<td>0.0033</td>
</tr>
<tr>
<td>Loans/Total Assets</td>
<td>22000</td>
<td>0.67</td>
<td>0.15</td>
<td>0.60</td>
<td>0.70</td>
<td>0.77</td>
</tr>
<tr>
<td><strong>Macroeconomic Variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effective Fed Funds Rate change (%)</td>
<td>22000</td>
<td>-0.0007</td>
<td>0.2361</td>
<td>-0.0101</td>
<td>0.0119</td>
<td>0.0521</td>
</tr>
<tr>
<td>GDP growth (%)</td>
<td>22000</td>
<td>0.51</td>
<td>3.09</td>
<td>-2.59</td>
<td>1.43</td>
<td>2.29</td>
</tr>
<tr>
<td>Inflation (%)</td>
<td>22000</td>
<td>0.43</td>
<td>0.66</td>
<td>0.11</td>
<td>0.46</td>
<td>0.82</td>
</tr>
<tr>
<td>Stock market return (%)</td>
<td>22000</td>
<td>3.68</td>
<td>8.06</td>
<td>0.51</td>
<td>4.52</td>
<td>7.97</td>
</tr>
<tr>
<td>Housing price growth (%)</td>
<td>22000</td>
<td>1.13</td>
<td>1.87</td>
<td>0.14</td>
<td>1.15</td>
<td>2.29</td>
</tr>
<tr>
<td><strong>Cross-Section Payment Statistics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average net daily payment flow/Deposits (%)</td>
<td>500</td>
<td>0.01</td>
<td>0.91</td>
<td>-0.14</td>
<td>0.01</td>
<td>0.17</td>
</tr>
<tr>
<td>s.d. of net daily payment flow/Deposits (%)</td>
<td>500</td>
<td>0.97</td>
<td>0.83</td>
<td>0.48</td>
<td>0.74</td>
<td>1.14</td>
</tr>
<tr>
<td>Average gross daily outflow/Deposits (%)</td>
<td>500</td>
<td>1.82</td>
<td>4.31</td>
<td>0.35</td>
<td>0.74</td>
<td>1.47</td>
</tr>
<tr>
<td>s.d. of gross daily outflow/Deposits (%)</td>
<td>500</td>
<td>1.11</td>
<td>1.34</td>
<td>0.41</td>
<td>0.70</td>
<td>1.18</td>
</tr>
</tbody>
</table>

Table D.1: **Summary Statistics.** The table reports the number of observations, mean, standard deviation, and percentiles of variables in our sample. Our sample contains 500 banks and 44 quarters from 2010 to 2020. We calculate $\mu_{ij}$ ($\sigma_{ij}$) as the within-quarter average (standard deviation) of daily payment outflows from bank $i$ to bank $j$ divided by bank $i$’s deposits at the beginning of the quarter. Therefore, $\sum_{j \neq i} \mu_{ij}$ is the average daily payment outflow as a fraction of deposits for bank $i$ within a quarter and $\sum_{j \neq i} \sigma_{ij}^2$ measures the payment-flow risk for bank $i$. 
The abbreviation, EFFR, is for effective fund funds rate.

Table D.2: **Control Variable Coefficients.** The table reports the estimates of control variable coefficients across samples of different sizes that contain banks ranked by the size of their deposits. The t-stats are in the parentheses. The abbreviation, EFFR, is for effective fund funds rate.

<table>
<thead>
<tr>
<th>Number of Banks:</th>
<th>500</th>
<th>500</th>
<th>300</th>
<th>400</th>
<th>600</th>
<th>700</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Not winsorized)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Constant</strong></td>
<td>0.0897</td>
<td>0.0863</td>
<td>0.1449</td>
<td>0.0973</td>
<td>0.0765</td>
<td>0.0609</td>
</tr>
<tr>
<td></td>
<td>(0.90)</td>
<td>(0.86)</td>
<td>(1.33)</td>
<td>(0.96)</td>
<td>(0.76)</td>
<td>(0.60)</td>
</tr>
<tr>
<td><strong>Bank Characteristics:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(Asset)</td>
<td>−0.0039</td>
<td>−0.0038</td>
<td>−0.0067</td>
<td>−0.0046</td>
<td>−0.0042</td>
<td>−0.0032</td>
</tr>
<tr>
<td></td>
<td>(−0.80)</td>
<td>(−0.76)</td>
<td>(−1.37)</td>
<td>(−0.92)</td>
<td>(−0.81)</td>
<td>(−0.62)</td>
</tr>
<tr>
<td>Liquid Assets/Assets</td>
<td>0.0144*</td>
<td>0.0142*</td>
<td>−0.0011</td>
<td>0.0008</td>
<td>0.0200***</td>
<td>0.0252***</td>
</tr>
<tr>
<td></td>
<td>(1.77)</td>
<td>(1.74)</td>
<td>(−0.12)</td>
<td>(1.03)</td>
<td>(2.58)</td>
<td>(3.52)</td>
</tr>
<tr>
<td>Capital/Assets</td>
<td>0.0931***</td>
<td>0.0941***</td>
<td>0.1086***</td>
<td>0.0971***</td>
<td>0.0607</td>
<td>0.0531</td>
</tr>
<tr>
<td></td>
<td>(3.28)</td>
<td>(3.28)</td>
<td>(4.63)</td>
<td>(3.58)</td>
<td>(1.51)</td>
<td>(1.32)</td>
</tr>
<tr>
<td>Deposits/Assets</td>
<td>−0.0108**</td>
<td>−0.0104**</td>
<td>−0.0080</td>
<td>−0.0079</td>
<td>−0.0111***</td>
<td>−0.0097**</td>
</tr>
<tr>
<td></td>
<td>(−2.27)</td>
<td>(−2.22)</td>
<td>(−1.47)</td>
<td>(−1.51)</td>
<td>(−2.65)</td>
<td>(−2.34)</td>
</tr>
<tr>
<td>Return on asset</td>
<td>1.2726***</td>
<td>1.2789***</td>
<td>1.3469***</td>
<td>1.3646***</td>
<td>1.2911***</td>
<td>1.3236***</td>
</tr>
<tr>
<td></td>
<td>(4.19)</td>
<td>(4.17)</td>
<td>(3.54)</td>
<td>(4.00)</td>
<td>(4.55)</td>
<td>(4.67)</td>
</tr>
<tr>
<td>Loans/Assets</td>
<td>−0.0296***</td>
<td>−0.0294***</td>
<td>−0.0386***</td>
<td>−0.0331***</td>
<td>−0.0241***</td>
<td>−0.0172***</td>
</tr>
<tr>
<td></td>
<td>(−2.96)</td>
<td>(−2.90)</td>
<td>(−4.14)</td>
<td>(−3.47)</td>
<td>(−2.32)</td>
<td>(−1.64)</td>
</tr>
<tr>
<td><strong>Macro. Variables:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EFFR change (%)</td>
<td>−0.0111</td>
<td>−0.0111</td>
<td>−0.0102</td>
<td>−0.0108</td>
<td>−0.0125</td>
<td>−0.0123</td>
</tr>
<tr>
<td></td>
<td>(−1.30)</td>
<td>(−1.31)</td>
<td>(−1.41)</td>
<td>(−1.27)</td>
<td>(−1.37)</td>
<td>(−1.31)</td>
</tr>
<tr>
<td>GDP growth (%)</td>
<td>−0.0007</td>
<td>−0.0007</td>
<td>−0.0005</td>
<td>−0.0007</td>
<td>−0.0009</td>
<td>−0.0009</td>
</tr>
<tr>
<td></td>
<td>(−1.00)</td>
<td>(−0.98)</td>
<td>(−0.85)</td>
<td>(−0.97)</td>
<td>(−1.16)</td>
<td>(−1.15)</td>
</tr>
<tr>
<td>Inflation (%)</td>
<td>0.0032</td>
<td>0.0032</td>
<td>0.0032</td>
<td>0.0029</td>
<td>0.0036</td>
<td>0.0036</td>
</tr>
<tr>
<td></td>
<td>(1.11)</td>
<td>(1.13)</td>
<td>(1.13)</td>
<td>(1.03)</td>
<td>(1.23)</td>
<td>(1.23)</td>
</tr>
<tr>
<td>Stock return (%)</td>
<td>−0.0009**</td>
<td>−0.0009**</td>
<td>−0.0008**</td>
<td>−0.0009**</td>
<td>−0.0009**</td>
<td>−0.0009**</td>
</tr>
<tr>
<td></td>
<td>(−2.28)</td>
<td>(−2.30)</td>
<td>(−2.50)</td>
<td>(−2.26)</td>
<td>(−2.22)</td>
<td>(−2.17)</td>
</tr>
<tr>
<td>Housing price growth (%)</td>
<td>0.0022**</td>
<td>0.0022**</td>
<td>0.0018**</td>
<td>0.0020**</td>
<td>0.0024**</td>
<td>0.0025***</td>
</tr>
<tr>
<td></td>
<td>(2.52)</td>
<td>(2.50)</td>
<td>(2.19)</td>
<td>(2.25)</td>
<td>(2.56)</td>
<td>(2.64)</td>
</tr>
</tbody>
</table>

(* p<0.10 ** p<0.05 *** p<0.01)
Figure D.1: **Control variable coefficients across samples.** This figure reports the ratio of an estimate from an alternative sample to the estimate from our main sample of the top 500 banks by deposit size. A ratio around one shows the two estimates are close. We plot the 95% confidence interval of each estimate from our main sample scaled by the estimate so the mid-point is equal to one.
Figure D.2: **Eigenvalues of network adjacency matrix.** In this figure, we plot the absolute values of five largest eigenvalues of $W'$. $W'$ for quarter $t$ is calculated from payment data from quarter $t - 1$ (see Section 2.1).