A Theory: supplementary results

A.1 Value function

Let $J_t^i(a,c) : \mathbb{N} \times \mathbb{T} \times \mathbb{R}^2 \to \mathbb{R}$ denote the maximum attainable payoff to a bank of type i that at time $t \in \mathbb{T}$ has reserve balance $a \in \mathbb{R}$ and net credit position $c \in \mathbb{R}$. Then, $J_t^i(a,c)$ satisfies

$$\begin{aligned}
&J_t^i(a,c) \\
&= \mathbb{E}_t \Bigg\{ \mathbb{I}_{\{T-t \le \min[\tau(\beta_i),\tau(\lambda_i)]\}} \left[\int_t^T e^{-r(s-t)} u_i(a)\, ds + e^{-r(T-t)} \left[U_i(a) + e^{-r(\bar{T}-T)} c \right] \right] \\
&+ \mathbb{I}_{\{\tau(\lambda_i) < \min[\tau(\beta_i),T-t]\}} \Bigg[\int_t^{t+\tau(\lambda_i)} e^{-r(s-t)} u_i(a)\, ds \\
&+ e^{-r\tau(\lambda_i)} \sum_{j \in \mathbb{N}} \pi_j \int J^i_{t+\tau(\lambda_i)}(a-z,c)\, dG_{ij}(z) \Bigg] \\
&+ \mathbb{I}_{\{\tau(\beta_i) < \min[\tau(\lambda_i),T-t]\}} \Bigg[\int_t^{t+\tau(\beta_i)} e^{-r(s-t)} u_i(a)\, ds \\
&+ e^{-r\tau(\beta_i)} \sum_{j \in \mathbb{N}} \sigma_j \int J^i_{t+\tau(\beta_i)} \left[a - b^{ij}_{t+\tau(\beta_i)}(a,\tilde{a}), c + R^{ji}_{t+\tau(\beta_i)}(\tilde{a},a) \right] dF^j_{t+\tau(\beta_i)}(\tilde{a}) \Bigg] \Bigg\}, \qquad (12)
\end{aligned}$$

where $\tau(\zeta)$ denotes the exponentially distributed first passage time of the Poisson process with arrival rate ζ,

$$\pi_j \equiv \frac{\lambda_j n_j}{\sum_{i \in \mathbb{N}} \lambda_i n_i}$$
$$\sigma_j \equiv \frac{\beta_j n_j}{\sum_{k \in \mathbb{N}} \beta_k n_k},$$

and

$$\begin{aligned}
&(b_t^{ij}(a,\tilde{a}), R_t^{ji}(\tilde{a},a)) \\
&= \arg \max_{(b,R) \in \mathbb{R}^2} \left[J_t^i(a-b,c+R) - J_t^i(a,c) \right]^{\theta_{ij}} \left[J_t^j(\tilde{a}+b,c-R) - J_t^j(\tilde{a},c) \right]^{\theta_{ji}}. \qquad (13)
\end{aligned}$$

Lemma 1 *The function*

$$J_t^i(a,c) = V_t^i(a) + e^{-r(\bar{T}-t)} c \qquad (14)$$

satisfies (12) if and only if $V_t^i(a)$ *satisfies*

$$V_t^i(a) = \mathbb{E}_t\Bigg\{\mathbb{I}_{\{T-t\leq\min[\tau(\beta_i),\tau(\lambda_i)]\}}\left[\int_t^T e^{-r(s-t)}u_i(a)\,ds + e^{-r(T-t)}U_i(a)\right]$$
$$+\mathbb{I}_{\{\tau(\lambda_i)<\min[\tau(\beta_i),T-t]\}}\left[\int_t^{t+\tau(\lambda_i)} e^{-r(s-t)}u_i(a)\,ds\right.$$
$$\left.+e^{-r\tau(\lambda_i)}\sum_{j\in\mathbb{N}}\pi_j\int V_{t+\tau(\lambda_i)}^i(a-z)\,dG_{ij}(z)\right]$$
$$+\mathbb{I}_{\{\tau(\beta_i)<\min[\tau(\lambda_i),T-t]\}}\left[\int_t^{t+\tau(\beta_i)} e^{-r(s-t)}u_i(a)\,ds\right.$$
$$\left.\left.+e^{-r\tau(\beta_i)}\sum_{j\in\mathbb{N}}\sigma_j\int\left[V_{t+\tau(\beta_i)}^i(a-b_{t+\tau(\beta_i)}^{ij}(a,\tilde{a}))+\bar{R}_{t+\tau(\beta_i)}^{ji}(\tilde{a},a)\right]dF_{t+\tau(\beta_i)}^j(\tilde{a})\right]\right\}, \quad (15)$$

with

$$\bar{R}_{t+\tau(\beta_i)}^{ji}(\tilde{a},a)\equiv e^{-r\left\{\bar{T}-[t+\tau(\beta_i)]\right\}}R_{t+\tau(\beta_i)}^{ji}(\tilde{a},a),$$

and $(b_t^{ij}(a,\tilde{a}), R_t^{ji}(\tilde{a},a))$ *given by (1) and (2).*

Proof. With (14), (13) becomes equivalent to (1) and (2). Substitute (14) into (12) to get

$$V_t^i(a)+e^{-r(\bar{T}-t)}c$$
$$=\mathbb{E}_t\Bigg\{\mathbb{I}_{\{T-t\leq\min[\tau(\beta_i),\tau(\lambda_i)]\}}\left[\int_t^T e^{-r(s-t)}u_i(a)\,ds+e^{-r(T-t)}\left[U_i(a)+e^{-r(\bar{T}-T)}c\right]\right]$$
$$+\mathbb{I}_{\{\tau(\lambda_i)<\min[\tau(\beta_i),T-t]\}}\left[\int_t^{t+\tau(\lambda_i)} e^{-r(s-t)}u_i(a)\,ds\right.$$
$$\left.+e^{-r\tau(\lambda_i)}\sum_{j\in\mathbb{N}}\pi_j\int\left[V_{t+\tau(\lambda_i)}^i(a-z)+e^{-r\{\bar{T}-[t+\tau(\lambda_i)]\}}c\right]dG_{ij}(z)\right]$$
$$+\mathbb{I}_{\{\tau(\beta_i)<\min[\tau(\lambda_i),T-t]\}}\left[\int_t^{t+\tau(\beta_i)} e^{-r(s-t)}u_i(a)\,ds+e^{-r\tau(\beta_i)}\sum_{j\in\mathbb{N}}\sigma_j\int\left[V_{t+\tau(\beta_i)}^i(a-b_{t+\tau(\beta_i)}^{ij}(a,\tilde{a}))\right.\right.$$
$$\left.\left.\left.+e^{-r\{\bar{T}-[t+\tau(\beta_i)]\}}[c+R_{t+\tau(\beta_i)}^{ji}(\tilde{a},a)]\right]dF_{t+\tau(\beta_i)}^j(\tilde{a})\right]\right\},$$

which after cancelling the terms proportional to c, becomes identical to (15). ■

Lemma 2 *The Bellman equation (15) can be written as*

$$\begin{aligned} V_t^i(a) &= \left[1 - e^{-(r+\beta_i+\lambda_i)(T-t)}\right] \frac{u_i(a)}{r+\beta_i+\lambda_i} + e^{-(r+\beta_i+\lambda_i)(T-t)} U_i(a) \\ &+ \lambda_i \int_t^T e^{-(r+\beta_i+\lambda_i)(\tau-t)} \left[\sum_{j\in\mathbb{N}} \pi_j \int V_\tau^i(a-z)\, dG_{ij}(z)\right] d\tau \\ &+ \beta_i \int_t^T e^{-(r+\beta_i+\lambda_i)(\tau-t)} \left[V_\tau^i(a) + \sum_{j\in\mathbb{N}} \sigma_j \theta_{ij} \int \max_{b\in\bar{\mathbb{R}}} S_\tau^{ij}(a,\tilde{a},b)\, dF_\tau^j(\tilde{a})\right] d\tau \end{aligned} \tag{16}$$

or equivalently, as (3) with boundary condition $V_T^i(a) = U_i(a)$.

Proof. With the bargaining outcomes (1) and (2), (15) can be rewritten as

$$\begin{aligned} V_t^i(a) = \mathbb{E}_t \Bigg\{ & \mathbb{I}_{\{T-t\le\min[\tau(\beta_i),\tau(\lambda_i)]\}} \left[\int_t^T e^{-r(s-t)} u_i(a)\, ds + e^{-r(T-t)} U_i(a)\right] \\ &+ \mathbb{I}_{\{\tau(\lambda_i)<\min[\tau(\beta_i),T-t]\}} \left[\int_t^{t+\tau(\lambda_i)} e^{-r(s-t)} u_i(a)\, ds + e^{-r\tau(\lambda_i)} \sum_{j\in\mathbb{N}} \pi_j \int V_{t+\tau(\lambda_i)}^i(a-z)\, dG_{ij}(z)\right] \\ &+ \mathbb{I}_{\{\tau(\beta_i)<\min[\tau(\lambda_i),T-t]\}} \Bigg[\int_t^{t+\tau(\beta_i)} e^{-r(s-t)} u_i(a)\, ds \\ &+ e^{-r\tau(\beta_i)} \sum_{j\in\mathbb{N}} \sigma_j \int \left[V_{t+\tau(\beta_i)}^i(a) + \theta_{ij} \max_{b\in\bar{\mathbb{R}}} S_{t+\tau(\beta_i)}^{ij}(a,\tilde{a},b)\right] dF_{t+\tau(\beta_i)}^j(\tilde{a})\Bigg]\Bigg\}, \end{aligned}$$

where

$$S_t^{ij}(a,\tilde{a},b) \equiv V_t^i(a-b) + V_t^j(\tilde{a}+b) - V_t^i(a) - V_t^j(\tilde{a}).$$

The first term on the right side of $V_t^i(a)$ can be written as

$$\begin{aligned} &\mathbb{E}_t \left\{\mathbb{I}_{\{T-t\le\min[\tau(\beta_i),\tau(\lambda_i)]\}} \left[\int_t^T e^{-r(s-t)} u_i(a)\, ds + e^{-r(T-t)} U_i(a)\right]\right\} \\ &= e^{-(\beta_i+\lambda_i)(T-t)} \left\{\left[1 - e^{-r(T-t)}\right] \frac{u_i(a)}{r} + e^{-r(T-t)} U_i(a)\right\}. \end{aligned}$$

The second term on the right side of $V_t^i(a)$ can be written as

$$\mathbb{E}_t\left\{\mathbb{I}_{\{\tau(\lambda_i)<\min[\tau(\beta_i),T-t]\}}\left[\int_t^{t+\tau(\lambda_i)} e^{-r(s-t)}u_i(a)\,ds+e^{-r\tau(\lambda_i)}\sum_{j\in\mathbb{N}}\pi_j\int V_{t+\tau(\lambda_i)}^i(a-z)\,dG_{ij}(z)\right]\right\}$$

$$=\frac{\lambda_i}{\beta_i+\lambda_i}\frac{r\left[1-e^{-(\beta_i+\lambda_i)(T-t)}\right]-(\beta_i+\lambda_i)\,e^{-(\beta_i+\lambda_i)(T-t)}\left[1-e^{-r(T-t)}\right]}{r+\beta_i+\lambda_i}\frac{u_i(a)}{r}$$

$$+\int_0^{T-t}\lambda_i e^{-(r+\beta_i+\lambda_i)y}\left[\sum_{j\in\mathbb{N}}\pi_j\int V_{t+y}^i(a-z)\,dG_{ij}(z)\right]dy.$$

The third term on the right side of $V_t^i(a)$ can be written as

$$\begin{aligned}V_t^i(a)=\mathbb{E}_t&\left\{\mathbb{I}_{\{\tau(\beta_i)<\tau(\lambda_i)\}}\mathbb{I}_{\{\tau(\beta_i)<T-t\}}\left[\int_t^{t+\tau(\beta_i)} e^{-r(s-t)}u_i(a)\,ds\right.\right.\\
&\left.\left.+e^{-r\tau(\beta_i)}\sum_{j\in\mathbb{N}}\sigma_j\int\left[V_{t+\tau(\beta_i)}^i(a)+\theta_{ij}\max_{b\in\bar{\mathbb{R}}}S_{t+\tau(\beta_i)}^{ij}(a,\tilde{a},b)\right]dF_t^j(\tilde{a})\right]\right\}\\
&=\frac{\beta_i}{\beta_i+\lambda_i}\frac{r\left[1-e^{-(\beta_i+\lambda_i)(T-t)}\right]-(\beta_i+\lambda_i)\,e^{-(\beta_i+\lambda_i)(T-t)}\left[1-e^{-r(T-t)}\right]}{r+\beta_i+\lambda_i}\frac{u_i(a)}{r}\\
&+\int_0^{T-t}\beta_i e^{-(r+\beta_i+\lambda_i)z}\left[\sum_{j\in\mathbb{N}}\sigma_j\int\left[V_{t+z}^i(a)+\theta_{ij}\max_{b\in\bar{\mathbb{R}}}S_{t+z}^{ij}(a,\tilde{a},b)\right]dF_{t+z}^j(\tilde{a})\right]dz.\end{aligned}$$

Thus, we can write

$$\begin{aligned}V_t^i(a)=&\left[1-e^{-(r+\beta_i+\lambda_i)(T-t)}\right]\frac{u_i(a)}{r+\beta_i+\lambda_i}+e^{-(r+\beta_i+\lambda_i)(T-t)}U_i(a)\\
&+\lambda_i\int_0^{T-t}e^{-(r+\beta_i+\lambda_i)y}\left[\sum_{j\in\mathbb{N}}\pi_j\int V_{t+y}^i(a-s)\,dG_{ij}(s)\right]dy\\
&+\beta_i\int_0^{T-t}e^{-(r+\beta_i+\lambda_i)z}\left[\sum_{j\in\mathbb{N}}\sigma_j\int\left[V_{t+z}^i(a)+\theta_{ij}\max_{b\in\bar{\mathbb{R}}}S_{t+z}^{ij}(a,\tilde{a},b)\right]dF_{t+z}^j(\tilde{a})\right]dz.\end{aligned}$$

With a change of variables in the integrals with respect to time,

$$\begin{aligned}V_t^i(a)=&\left[1-e^{-(r+\beta_i+\lambda_i)(T-t)}\right]\frac{u_i(a)}{r+\beta_i+\lambda_i}+e^{-(r+\beta_i+\lambda_i)(T-t)}U_i(a)\\
&+\lambda_i\int_t^{T}e^{-(r+\beta_i+\lambda_i)(\tau-t)}\left[\sum_{j\in\mathbb{N}}\pi_j\int V_{\tau}^i(a-z)\,dG_{ij}(z)\right]d\tau\\
&+\beta_i\int_t^{T}e^{-(r+\beta_i+\lambda_i)(\tau-t)}\left[V_\tau^i(a)+\sum_{j\in\mathbb{N}}\sigma_j\theta_{ij}\int\max_{b\in\bar{\mathbb{R}}}S_{\tau}^{ij}(a,\tilde{a},b)\,dF_{\tau}^j(\tilde{a})\right]d\tau.\end{aligned}\tag{17}$$

To obtain (3), simply differentiate (16) with respect to t. ■

A.2 Extension: regulatory borrowing costs

In this section we generalize the theory to allow for proportional borrowing costs to proxy for the effects of regulatory constraints that affect banks' incentives to buy fed funds. Let

$$\Gamma_t^i(a,b,R) \equiv V_t^i(a-b) - V_t^i(a) + \left[1+\mathbb{I}_{\{b<0\}}\kappa_i\right] e^{-r(\bar{T}-t)}R \tag{18}$$

denote payoff of a bank of type $i \in \mathbb{N}$, with pre-trade balance a, that at time t sells a loan of size b in exchange for a repayment of size R delivered at time $\bar{T}$, with $\kappa_i \in \mathbb{R}_+$. Intuitively, if $b, R \in \mathbb{R}_+$, then the bank is "selling fed funds" (i.e., lending) and the gain from trade is as in Section 2.2. Conversely, if $b, R \in \mathbb{R}_-$, then the bank is "buying fed funds" (i.e., *borrowing*), and κ_i captures the effects of policies that increase the shadow cost of the bank's liabilities. In all our calibrations we set κ_{GSE} large enough to make our theory consistent with the fact that the business model of a GSE consists of lending, but not borrowing in the fed funds market. In our 2019 calibration we use κ_i for $i \in \{F, M, S\}$ to capture the effects of the prudential liquidty regulations discussed in Appendix B (Section B.2). With borrowing costs, the bargaining outcome at time t between two banks of type i and j, with respective balances a and $\tilde{a}$, denoted $(b_t^{ij}(a,\tilde{a}), R_t^{ji}(\tilde{a},a))$, is the solution to

$$\max_{(b,R)\in\bar{\mathbb{R}}\times\mathbb{R}} \Gamma_t^i(a,b,R)^{\theta_{ij}}\Gamma_t^j(\tilde{a},-b,-R)^{\theta_{ji}}. \tag{19}$$

The correspondig first-order condition with respect to R is

$$\theta_{ij}\left[1+\mathbb{I}_{\{b<0\}}\kappa_i\right]\Gamma_t^j(\tilde{a},-b,-R) = \theta_{ji}\left[1+\mathbb{I}_{\{0<b\}}\kappa_j\right]\Gamma_t^i(a,b,R),$$

which implies $R_t^{ji}(\tilde{a},a)$ is given by

$$\begin{aligned} e^{-r(\bar{T}-t)}R_t^{ji}(\tilde{a},a) &= \frac{\theta_{ij}}{1+\mathbb{I}_{\{0<b_t^{ij}(a,\tilde{a})\}}\kappa_j}[V_t^j(\tilde{a}+b_t^{ij}(a,\tilde{a})) - V_t^j(\tilde{a})] \\ &\quad + \frac{\theta_{ji}}{1+\mathbb{I}_{\{b_t^{ij}(a,\tilde{a})<0\}}\kappa_i}[V_t^i(a) - V_t^i(a-b_t^{ij}(a,\tilde{a}))], \end{aligned} \tag{20}$$

and

$$b_t^{ij}(a,\tilde{a}) \in \arg\max_{b\in\bar{\mathbb{R}}} \hat{S}_t^{ij}(a,\tilde{a},b), \tag{21}$$

where

$$\hat{S}_t^{ij}(a,\tilde{a},b) \equiv \hat{\Gamma}_t^{ij}(a,\tilde{a},b)^{\theta_{ij}}\hat{\Gamma}_t^{ji}(\tilde{a},a,-b)^{\theta_{ji}},$$

with

$$\hat{\Gamma}_t^{ij}(a,\tilde{a},b) \equiv \theta_{ij}\left\{S_t^{ij}(a,\tilde{a},b) - \frac{\mathbb{I}_{\{0<b\}}\kappa_j - \mathbb{I}_{\{b<0\}}\kappa_i}{1+\mathbb{I}_{\{0<b\}}\kappa_j}[V_t^j(\tilde{a}+b) - V_t^j(\tilde{a})]\right\}$$

$$\hat{\Gamma}_t^{ji}(\tilde{a},a,-b) \equiv \theta_{ji}\left\{S_t^{ij}(a,\tilde{a},b) - \frac{\mathbb{I}_{\{0<b\}}\kappa_j - \mathbb{I}_{\{b<0\}}\kappa_i}{1+\mathbb{I}_{\{b<0\}}\kappa_i}\left[V_t^i(a) - V_t^i(a-b)\right]\right\}.$$

In summary, the bargaining solution, $(b_t^{ij}(a,\tilde{a}), R_t^{ji}(\tilde{a},a))$, is given by (21) and (20), and the value function $V_t^i(a)$ now satisfies

$$rV_t^i(a) - \dot{V}_t^i(a) = u_i(a) + \lambda_i \sum_{j\in\mathbb{N}} \pi_j \int \left[V_t^i(a-z) - V_t^i(a)\right] dG_t^{ij}(z)$$
$$+ \beta_i \sum_{j\in\mathbb{N}} \sigma_j \int \Gamma_t^i(a, b_t^{ij}(a,\tilde{a}), R_t^{ji}(\tilde{a},a)) dF_t^j(\tilde{a}), \tag{22}$$

with Γ_t^i as defined in (18). Notice that (21), (20), and (22) generalize (1), (2), and (3), respectively (and the former reduce to the latter if $\kappa_i = 0$ for all $i \in \mathbb{N}$).

B Institutional background and regulation

In this section we review three financial regulations that affect banks' incentives to borrow and lend in the fed funds market. Two of them directly increase a bank's shadow value of holding reserves by imposing regulatory balance-sheet constraints that can be satisfied with reserve balances (traditional reserve requirements, discussed in Section B.1, and the *Liquidity Coverage Ratio*, discussed in Section B.2.1). The third, is a leverage constraint that increases a bank's shadow cost of all borrowing, including fed funds purchases (the *Supplementary Leverage Ratio*, discussed in Section B.2.2).

B.1 Traditional reserve requirements (Regulation D)

Reserve requirements have been a part of the financial landscape in the United States since before the Federal Reserve Act of 1913 that created the system of Reserve Banks.[98] Regulation D ("Reserve Requirements for Depository Institutions") is the Federal Reserve regulation

[98] Reserve requirements at the national level were first established with the passage of the National Bank Act in 1863. In the original Federal Reserve Act of 1913, for example, banks were required to hold in reserve different percentages of their demand deposits, depending on whether they were classified as *central reserve city banks* (18 percent), *reserve city banks* (15 percent), or *country banks* (12 percent). See Feinman (1993) for more background and references on the history of reserve requirements in the United States.

that stipulates reserve requirements for depository institutions (i.e., commercial banks, savings banks, thrift institutions, credit unions, and agencies and branches of foreign banks located in the United States).

Until March 2020, Regulation D required depository institutions to keep a minimum amount of reserves against their transaction accounts (such as demand deposits).[99] This reserve requirement was 0%, 3%, or 10% of transaction account deposits depending on the size of the bank's reservable liabilities.[100] Institutions had to satisfy reserve requirements by holding cash in their vaults or as a balance in the institution's account at the Federal Reserve Bank in the Federal Reserve District in which the institution is located (either an account of the institution or an account of the institution's Federal Reserve pass-through correspondent).

Reserve requirements were calculated based on a bank's deposit accounts during *computation periods* that depended on the frequency (either weekly or quarterly) with which an institution files an FR 2900 report.[101] Each reserve computation period was used to calculate the reserve requirement that a bank had to satisfy on a lagged basis, i.e., during a 14-day *(reserve) maintenance period* in the future.

For institutions that file the FR 2900 report weekly, a *(FR 2900) reporting period* is one week long, covering the seven consecutive calendar days beginning on a Tuesday and ending on the following Monday. The *computation period* for weekly reporters consisted of two *reporting periods*, i.e., 14 consecutive days beginning on a Tuesday and ending on the second Monday thereafter. A *maintenance period* consisted of 14 consecutive days beginning on a Thursday and ending on the second Wednesday thereafter. Each reserve computation period was used to calculate the reserve requirement that a bank had to satisfy on a lagged basis: The reserve balance requirement that had to be satisfied during a maintenance period was based on the average level of net transaction accounts and vault cash held during the computation period that had ended 17 days earlier.[102]

Federal Reserve Banks were authorized to assess charges for deficiencies at a rate of 1 percentage point per year above the primary credit rate in effect for borrowings from the Federal Reserve Bank on the first day of the calendar month in which the deficiencies occurred. Charges were assessed on the basis of daily average deficiencies during each maintenance period.

[99] There was an explicit exemption from Regulation D for bank obligations in nondeposit form to another bank, which included "federal funds purchased".

[100] The Federal Reserve Board reduced all reserve requirement ratios to 0% effective March 26, 2020.

[101] This report collects information on select deposits and vault cash from depository institutions.

[102] See Federal Reserve Board (2019a) for details.

B.2 Post-GFC regulation

In the years following the Great Financial Crisis (GFC), the Federal Reserve Board (FRB), the Federal Deposit Insurance Corporation (FCC), and the Office of the Comptroller of the Currency (OCC) implemented versions of two regulations agreed to by the Basel Committee on Banking Supervision (BCBS), and consistent with the Dodd-Frank Wall Street Reform and Consumer Protection Act: The *Liquidity Coverage Ratio* (LCR), a prudential liquidty standard, and the *Supplementary Leverage Ratio* (SLR), a prudential leverage standard. Both affect banks' payoffs from trading in the fed funds market. We discuss each in turn.

B.2.1 Liquidity Coverage Ratio (LCR)

The first objective of the Basel III accord agreed upon by the members of the Basel Committee on Banking Supervision (BCBS) is to promote the short-term resilience of the liquidity risk profile of banks. The BCBS developed the LCR to achieve this objective.[103] Specifically, the LCR is designed to ensure that a bank maintains an adequate level of unencumbered, *High Quality Liquid Assets* (HQLA) that can be converted into cash to meet its liquidity needs for a 30-calendar-day time horizon under a liquidity stress scenario specified by supervisors.

The LCR is defined as

$$LCR \equiv \frac{H}{L}, \tag{23}$$

where H denotes HQLA, and L is a measure of total net cash outflows in a 30-day standardized stress scenario. The HQLA consist of Level 1 assets and Level 2 assets. Level 1 assets, which are not subject to haircuts or quantitative caps, include reserves in excess of Regulation D held at a Federal Reserve Bank, as well as securities issued or guaranteed by the U.S. Treasury. Level 2 assets are subject to prescribed haircuts and are capped at no more than 40% of a banking organization's total HQLA.[104] For our purposes, we can think of H as consisting of two components: (i) reserves, denoted Q_0, minus Regulation D required reserves, denoted R_D;

[103]See Basel Committee on Banking Supervision (2010) for more details on the rationale for the regulation.

[104]Level 2 assets are further divided into Level 2A and Level 2B assets. Level 2A assets, which are subject to a 15% haircut, include claims on or guaranteed by a U.S. government-sponsored enterprise (GSE) such as Fannie Mae and Freddie Mac. Level 2B assets, which are subject to a 50% haircut and are capped at no more than 15% of a banking organization's total HQLA, include certain corporate debt securities issued by non-financial companies, and certain publicly traded common equities issued by non-financial companies that are included in the Russell 1000 Index or a foreign equivalent index for shares held in foreign jurisdictions.

and (ii) the value (net of haircut) of all other assets that qualify as HQLA, denoted A, i.e.,

$$H = A + M,$$

where

$$M \equiv \max(Q_1, 0) \tag{24}$$

and $Q_1 \equiv Q_0 - R_D$ denotes the quantity of reserves in excess of the Regulation D requirement. The "max" in (24) reflects that only reserves in excess of Regulation D qualify as HQLA.

Banks report H and L, and these reports are publicly available at a quarterly frequency.[105] Given H, since we have independent information on Q_0 and R_D (and therefore M), we can infer A. The LCR regulation requires

$$1 \leq LCR \tag{25}$$

daily, or monthly, depending on the size and other characteristics of the bank.[106] For our purposes, the key implication of the policy constraint (25) is that it may cause a bank to treat certain holdings of HQLA as required to comply with the LCR regulation. By this we mean that the LCR constraint may cause the bank to inpute an additional shadow cost of reducing its holdings of HQLA on a typical day—including reserve balances. In the specific case of reserve balances, the bank may impute an additional shadow cost of selling fed funds, since this may drive the bank's reserves (net of the Regulation D requirement) below the level of reserves that the bank routinely allocates to comply with the LCR regulation. Thus, in practice, banks may regard some of the reserves in excess of the Regulation D requirement as being "required" to satisfy the LCR constraint. The fact that the LCR regulation allows for substitutability among the HQLA in the numerator of the left side of (25) presents us with an identification challenge when trying to estimate the share of a bank's reserve balances in excess of the Regulation D requirement that the bank treats as "required" to satisfy the LCR constraint. Next, we formalize this identification problem, and describe how we address it.

For each bank, we observe H, M, and A. We want to express M as the sum of a component, $\hat{M}^R$, that represents the quantity of reserves (in excess of the Regulation D requirement) that

[105] E.g., from the S&P Global Capital IQ database. See Appendix D (Section D.1.3) for details.

[106] Relatively large institutions regulated by the FRB must calculate and maintain a liquidity coverage ratio that is equal to or greater than 1 on each business day (or, in the case of a smaller FRB-regulated institutions, on the last business day of the applicable month). The LCR rule is codified at 12 CFR part 50 (OCC), 12 CFR part 249 (FRB), and 12 CFR part 329 (FDIC).

the bank relies on to comply with the LCR regulation, and a component, $\hat{M}^E$, that represents reserves in excess of the Regulation D *and* the LCR requirements. Similarly, a bank may hold HQLA (other than reserves) in excess of what would be necessary to meet the LCR requirement for reasons other than having to comply with the LCR regulation, so we can also decompose A into two (unobserved) components: $\hat{A}^R$, which represents the value of HQLA (other than reserves in excess of the Regulation D requirement) that the bank regards as being necessary to comply with the LCR regulation, and $\hat{A}^E$, which represents the value of HQLA (other than reserves in excess of the Regulation D requirement) that the bank regards as being in excess of what is required to meet the LCR regulation. In summary, $\{\hat{M}^j, \hat{A}^j\}_{j\in\{R,E\}}$ satisfy:

$$M = \hat{M}^R + \hat{M}^E \tag{26}$$

$$A = \hat{A}^R + \hat{A}^E \tag{27}$$

$$\hat{A}^R + \hat{M}^R \leq L, \text{ with "=" if } L \leq A + M \tag{28}$$

$$\hat{A}^E + \hat{M}^E = 0, \text{ if } A + M < L \tag{29}$$

$$\hat{M}^j, \hat{A}^j \in \mathbb{R}_+ \text{ for } j \in \{R, E\}. \tag{30}$$

We are interested in using the policy constraint (25) along with data on M, A, and L, and (26)-(30), to estimate bank-level bounds for $\hat{M}^R$.

There are three special cases in which the constraint (25) together with knowledge of M, A, and L, and the definitions (26)-(30) are sufficient to identify $\hat{M}^R$ and $\hat{A}^R$. First, if a bank has $LCR \leq 1$ (i.e., if it is not complying with the LCR regulation in a given sample period), then the bank is clearly holding no excess HQLA of any type, so $\hat{M}^R = M$, $\hat{A}^R = A$, and $\hat{M}^E = \hat{A}^E = 0$, as implied by (26), (27), (29), and (30). Second, if $LCR \geq 1$ and $Q_1 \leq 0$, then $M = 0$, so the LCR requirement, L, is being satisfied exclusively with HQLA other than reserves, i.e., $\hat{A}^R = L$ and $\hat{A}^E = A - L$, with $\hat{M}^R = \hat{M}^E = 0$, as implied by (26), (27), (28), and (30). Third, if $LCR \geq 1$ and $A = 0$, then the LCR requirement, L, is being satisfied exclusively with reserves, M, i.e., $\hat{M}^R = L$ and $\hat{M}^E = M - L$, with $\hat{A}^R = \hat{A}^E = 0$, as implied by (26), (27), (28), and (30).

In practice, most banks satisfy the LCR constraint (25) with $\min(M, A) \geq 0$, and for such banks it is not obvious how to decompose the level of *required* HQLA, i.e., L, into the two unobserved components, $\hat{M}^R$ and $\hat{A}^R$. However, notice that conditions (26)-(30) imply $\hat{M}^R$

must satisfy the following bounds:

$$\hat{M}^R \begin{cases} = M & \text{if } A+M<L \\ \in [\max(0, L-A), \min(L, M)] & \text{if } L \leq A+M. \end{cases} \tag{31}$$

We can write (31) as

$$\hat{M}^R = \begin{cases} M & \text{if } A+M<L \\ \rho \min(L, M) + (1-\rho)\max(0, L-A) & \text{if } L \leq A+M, \end{cases} \tag{32}$$

for some $\rho \in [0,1]$. For a given ρ, (26)-(30), and (32) imply

$$\hat{A}^R = \begin{cases} A & \text{if } A+M<L \\ (1-\rho)\min(L, A) + \rho \max(0, L-M) & \text{if } L \leq A+M, \end{cases}$$

and given $\hat{M}^R$ and $\hat{A}^R$, $\hat{M}^E$ and $\hat{A}^E$ are implied by (26) and (27).

The parameter $\rho \in [0,1]$ represents the bank's (unobserved) preference for satisfying the LCR requirement, L, with reserves (rather than with other HQLA). For example, if $\rho = 1$, the bank has a strong preference for satisfying the LCR with reserves, and this will reduce the bank's willingness to lend reserves in the fed funds market. If $\rho = 0$, the bank has a strong preference for satisfying the LCR with HQLA *other* than reserves, and will be less constrained by its reserve balance when trading in the fed funds market.

According to elementary theory, the quantity of reserves in excess of regulatory reserve requirements is a key determinant of a bank's "fundamental" incentive to borrow and lend in the fed funds market. For example, a bank whose reserve balance is lower than the minimum regulatory requirement, has a fundamental incentive to borrow (at a rate no larger than the shadow cost of violating the regulatory requirement). Conversely, a bank whose reserve balance is higher than the regulatory requirement, would have, all else equal, an incentive to lend (e.g., to banks with negative excess reserves, at a rate between the lender's and the borrower's respective shadow prices of reserves). For this reason, it is important to impute an accurate notion of "excess reserves" in any empirical implementation of a theory of interbank loans.

The traditional definition of "excess reserves", which only subtracts the Regulation D requirement from the bank's reserve balance is not an adequate notion of excess reserves for institutions that must comply with the LCR regulation.[107] In our empirical and quantitative work we use a more comprehensive notion of "required reserves" that includes not only the level

[107]The LCR regulation applies to bank holding companies (BHCs) and savings and loans holding (SLHCs) with at least \$50 bn in total consolidated assets.

of reserves that a bank is required to hold to comply with Regulation D, but also the level of reserves that the bank holds toward meeting the LCR requirement. Specifically, our benchmark definition of "excess reserves" for any bank that is subject to, and satisfies the LCR constraint (25), is $Q_2 \equiv Q_1 - R_L$, where $R_L \equiv \max(0, L - A)$. In other words, to construct our preferred notion of excess reserves, we start from the traditional notion of reserves in excess of the Regulation D requirement, Q_1, and subtract the minimum level of reserves needed to comply with the LCR requirement, i.e., R_L.[108] Notice that our measure of excess reserves coincides with the traditional measure for a bank that has enough HQLA other than reserves to meet the LCR requirement, i.e., $Q_1 - Q_2 = R_L = 0$ if $L \leq A$. But our measure of excess reserves is lower than the traditional measure for a bank whose holdings of HQLA other than reserves are insufficient to meet the LCR requirement, i.e., if $A < L$, then $0 < Q_1 - Q_2 = R_L = L - A$.

B.2.2 Supplementary Leverage Ratio (SLR)

The SLR is the U.S. banking agencies' implementation of the "Basel III Tier 1 Leverage Ratio", which is defined as

$$SLR \equiv \frac{\text{Tier 1 Capital}}{\text{Total Leverage Exposure}}. \tag{33}$$

The numerator (defined in U.S. Basel III) includes common stock and retained earnings. The denominator is a comprehensive measure of assets, composed of four elements: (1) on-balance sheet assets, (2) derivative exposures, (3) repo-style transaction exposures, and (4) other off-balance sheet exposures. The SLR regulation requires a bank to maintain an SLR above a threshold; specifically, either $SLR \geq 0.03$, or $SLR \geq 0.05$.[109]

B.2.3 Resolution Planning

In the aftermath of the GFC, regulatory authorities started requiring large "systemically important" financial institutions (e.g., BHCs with total consolidated assets of \$50 bn or more) to periodically submit a resolution plan (also known as "living will") to the Federal Reserve

[108] From (32), we see that R_L is the same as $\hat{M}^R$ when $\rho = 0$ (in the empirically relevant case with $L \leq A + M$). In this sense, our preferred notion of excess reserves selects the largest level of excess reserves that is consistent with the LCR constraint, (25).

[109] The threshold equals 3% for *advanced approaches firms*, which include state banks, savings associations, bank holding companies (BHCs), and saving and loan holding companies (SLHCs) with more than \$250 bn in total consolidated assets, or more than \$10 bn of on-balance sheet foreign exposures. The threshold equals 5% for the 8 US bank-holding companies that have been identified by the Financial Stability Board as global systemically important banks (and their U.S. insured depository institution subsidiaries).

and the Federal Deposit Insurance Corporation. A resolution plan describes in some detail the company's strategy for rapid and orderly resolution in the event of material financial distress or failure of the company.

B.2.4 Effects of LCR and SLR regulation on fed funds trading incentives

In this section we discuss the effects of the LCR and SLR reguation on banks' incentives to borrow and lend in the fed funds market.

First, we consider the effect of LCR regulation on banks' incentives to borrow and lend in the fed funds market. Reserves appear (with weight =1) in the numerator of the LCR in (23), and overnight fed fund liabilities appear in the denominator (also with weight =1). Consider a bank that borrows ℓ in the fed funds market. The LCR before the trade is $\frac{H}{L}$ and after the trade it is $\frac{H+\ell}{L+\ell}$. Since

$$\frac{\partial}{\partial \ell}\left(\frac{H+\ell}{L+\ell}\right)=\frac{L-H}{(L+\ell)^2},$$

it follows that the trade does not affect the LCR if the bank is satisfying it exactly pre-trade (i.e., if $LCR \equiv \frac{H}{L}=1$), increases the LCR if the borrowing bank is below the LCR target pre trade (i.e., if $LCR \equiv \frac{H}{L}<1$), and decreases the LCR if the borrowing bank is above the LCR target pre trade (i.e., if $LCR \equiv \frac{H}{L}>1$). For a bank that lends ℓ in the fed funds market, the LCR before the trade is $\frac{H}{L}$ and after the trade it is $\frac{H-\ell}{L}$.[110] Hence, selling fed funds unambiguously reduces the LCR. To summarize, LCR regulation increases the shadow cost of selling fed funds (because lending reserves tightens the LCR constraints of lenders), and increases the shadow cost of borrowing for banks whose LCR constraints are slack at the time of the trade (because borrowing reserves tightens the LCR constraints of such banks).

Second, we consider the effect of SLR regulation on banks' incentives to borrow and lend in the fed funds market. Let $\mathcal{A}$ denote assets, $\mathcal{L}$ denote liabilities, and $\mathcal{C} \equiv \mathcal{A}-\mathcal{L}$ denote capital. Then, we can write (33) as

$$SLR \equiv \frac{\mathcal{C}}{\mathcal{A}}=\frac{\mathcal{A}-\mathcal{L}}{\mathcal{A}}. \tag{34}$$

Notice that *lending* in the fed funds market does not change the SLR because the bank that acts as a lender is just exchanging reserves for an overnight credit of reserves, which leaves both $\mathcal{L}$ and $\mathcal{A}$ unchanged. However, *borrowing* in the fed funds market reduces the SLR, since

[110] The quantity of reserves sold, ℓ, is subtracted from the HQLA of the lender, but the corresponding fed funds credit is not added to the total of HQLA of the lender because fed funds not qualify as a HQLA.

borrowing ℓ dollars worth of reserves increases liabilities from $\mathcal{L}$ to $\mathcal{L}+\ell$, and increases assets from $\mathcal{A}$ to $\mathcal{A}+\ell$, and therefore the SLR is *reduced* from $\frac{\mathcal{A}-\mathcal{L}}{\mathcal{A}}$ to $\frac{\mathcal{A}-\mathcal{L}}{\mathcal{A}+\ell}$. To summarize, SLR regulation has no effect on the shadow cost of lending fed funds (because lending reserves does not alter the SLR constraint), but increases the shadow cost of buying fed funds (because borrowing reserves tightens the SLR constraint of solvent banks).

C Computation

In this section we discuss computational issues. Section C.1 outlines the solution algorithm. Section C.2 explains how we compute, in the quantitative theory, the statistics that we compare with their empirical counterparts.

C.1 Solution algorithm

The steps we use to solve for the equilibrium of the model are as follows.

Step 0: Set grids. We think of the time interval $[0,T]$ as corresponding to a trading day in the fed-funds market, which consists of 9.5 hours (from 9.00 AM to 5.30 PM). We divide the interval $[0,T]$ into N_T+1 periods, denoted $t=0,1,\ldots,N_T$, and set $N_T=799$. As we have 800 periods, each period represents approximately 42 seconds (i.e., $\frac{9.5\times 60\times 60}{800}=42.75$ seconds).

For each bank type $i\in\mathbb{N}$, we construct an equally spaced grid for reserve balances, $\mathbb{A}^i=\left\{a_1^i,a_2^i,\ldots,a_{N_a}^i\right\}$, with $N_a=150$. We interpret each unit of reserves in the model as corresponding to \$10 bn in the data. For the benchmark years 2017 and 2019, we set a_1^i and $a_{N_a}^i$ equal to the 0.5^{th} and 99.5^{th} percentiles of the kernel estimate of the beginning-of-day distributions, respectively (see Section 3.3). We use the interpolation procedure explained in Section 3.6 to construct grids whenever we change the total quantity of balances. In all cases we add 5 additional points to the grid, $\{-0.2,-0.1,\ 0,\ 0.1,\ 0.2\}$.[111]

For each pair of bank types, $i,j\in\mathbb{N}$, we construct a grid for payment sizes, $\mathbb{Z}^{ij}=\left\{z_1^{ij},z_2^{ij},\ldots,z_{N_z}^{ij}\right\}$, with $N_z=35$. The probability mass function for payment sizes, $\{G_{ij}(z)\}_{z\in\mathbb{Z}^{ij}}$, is constructed with the procedure described in Section 3.2.

[111] We add these grid points because the value functions are numerically close to having a kink around $a=0$ towards the end of the trading day (i.e., as t gets closer to N_T).

Step 1: Guess the distribution of balances. For each $a \in \mathbb{A}^i$, let $f_t^i(a)$ be the fraction of banks of type $i \in \mathbb{N}$ that hold a quantity of reserves equal to a at the beginning of period t, with $\sum_{a\in\mathbb{A}^i} f_t^i(a) = 1$. The beginning-of-day distribution, $f_0^i(\cdot)$, is given since $F_0^i(a) \equiv \sum_{x\in\mathbb{A}^i} f_0^i(x)\mathbb{I}_{\{x\leq a\}}$ is estimated from the data with the procedure described in Section 3.3. Guess the distributions $\{f_t^i(a)\}_{a\in\mathbb{A}^i,i\in\mathbb{N}}$ for each $t \in \{1, 2, \ldots, N_T\}$.

Step 2: Compute the value function. Since for each $i \in \mathbb{N}$ and $a \in \mathbb{A}^i$ we have the terminal condition, $V_{N_T}^i(a) = U_i(a)$, where $U_i(\cdot)$ is the exogenous end-of-day payoff function, we can then solve backwards for the value function, $\{V_t^i(a)\}_{a\in\mathbb{A}^i,i\in\mathbb{N},t\in\{0,\ldots,N_T-1\}}$. Each of these backward iterations between period $t \in \{N_T, \ldots, 1\}$ and period $t-1$ consists of two steps. In the first step, for each pair of bank types $i, j \in \mathbb{N}$, we compute the bargaining outcomes, $\left\{b_t^{ij}(a^i, a^j), R_t^{ji}(a^j, a^i)\right\}_{a^i\in\mathbb{A}^i,a^j\in\mathbb{A}^j}$, taking $\{V_t^i(a)\}_{a\in\mathbb{A}^i,i\in\mathbb{N}}$ as given. In the second step we solve for the value function backwards, i.e., we solve for $\{V_{t-1}^i(a)\}_{a\in\mathbb{A}^i,i\in\mathbb{N}}$ given the one-period-ahead bargaining outcomes and values, i.e., $\left\{b_t^{ij}(a^i, a^j), R_t^{ji}(a^j, a^i), V_t^i(a^i)\right\}_{a^i\in\mathbb{A}^i,a^j\in\mathbb{A}^j,(i,j)\in\mathbb{N}^2}$. Next, we explain these two steps in detail.

Step 2.1: *Solve for $b_t^{ij}(\cdot,\cdot)$ and $R_t^{ji}(\cdot,\cdot)$.* Given the values $\{V_t^i(\cdot)\}_{i\in\mathbb{N}}$, we compute the bargaining outcome for the loan size, $b_t^{ij}(a, \tilde{a})$, as in (21), which can be written as:

$$b_t^{ij}(a, \tilde{a}) = \arg\max_b \left\{ \frac{1}{1 + \mathbb{I}_{\{b<0\}}\kappa_i} V_t^i(a-b) + \frac{1}{1 + \mathbb{I}_{\{0<b\}}\kappa_j} V_t^j(\tilde{a}+b) - \epsilon|b| \right\}, \tag{35}$$

where $\epsilon = 1\text{e-}9$ is a small trading cost introduced to rule out loans with negligible gains from trade. Since a unit of reserves in the model corresponds to \$10 bn in the data, the value of ϵ implies a trading cost of \$10 for a loan of size \$1 bn. We use a Golden search routine to solve for $b_t^{ij}(a, \tilde{a})$ in (35), for each $a \in \mathbb{A}^i$, $\tilde{a} \in \mathbb{A}^j$, and $(i, j) \in \mathbb{N}\times\mathbb{N}$. Given the bargained loan sizes, $\{b_t^{ij}(a^i, a^j)\}_{a^i\in\mathbb{A}^i,a^j\in\mathbb{A}^j,(i,j)\in\mathbb{N}^2}$, we can compute the associated repayments, $R_t^{ji}(a^j, a^i)$, as in (20), and the gain from trade to the bank of type i and balance a^i, i.e., $\Gamma_t^i(a, b_t^{ij}(a^i, a^j), R_t^{ji}(a^j, a^i))$, as in (18).

Step 2.2: *Solve for $V_t^i(a)$ backwards.* We divide each period into two stages. Random payments between pairs of banks take place in the first stage. Trade between pairs of banks takes place in the second stage. The first stage is divided further into N_S subperiods, each indexed by $s \in \{1, 2, \ldots, N_S\}$ with $N_S = 42$, so each of these subperiods corresponds to 1 second (since each full model period corresponds to approximately 42 seconds). We solve for the value function within each model period backwards: we start by solving for the value of trade decisions in the second stage, and then integrate the value of the payment shocks in the 42 subperiods of the first stage.

Let $\hat{V}_t^i(a)$ be the value of a bank of type $i \in \mathbb{N}$ with balance $a \in \mathbb{A}^i$ at the beginning of the second stage of period t. This value satisfies

$$(1+\Delta r)\hat{V}_t^i(a) = \Delta u_i(a) + \Delta\beta_i \sum_{j\in\mathbb{N}} \sigma_j \sum_{\tilde{a}\in\mathbb{A}_j} \bar{\Gamma}_{t+1}^{ij}(a,\tilde{a}) f_{t+1}^j(\tilde{a}) + V_{t+1}^i(a), \tag{36}$$

where $\bar{\Gamma}_t^{ij}(a,\tilde{a}) \equiv \Gamma_t^i(a, b_t^{ij}(a,\tilde{a}), R_t^{ji}(\tilde{a},a))$, and $\Delta = 1/\left[N_S(N_T+1)\right]$ is the size of the time interval (including all trade and payment periods in the day). Let $\tilde{V}_{t,s}^i(a)$ be the value of a bank of type $i \in \mathbb{N}$ with balance $a \in \mathbb{A}^i$ at the beginning of subperiod s of the first stage of period t. This value satisfies

$$(1+\Delta r)\tilde{V}_{t,s}^i(a) = \Delta u_i(a) + \Delta\lambda_i \sum_{j\in\mathbb{N}} \pi_j \sum_{z\in\mathbb{Z}^{ij}} \left[V_t^i(a-z) - V_t^i(a)\right] G_{ij}(z) + \tilde{V}_{t,s+1}^i(a), \tag{37}$$

for $s = 1, \ldots, N_S$, with boundary conditions $\tilde{V}_{t,N_S+1}^i(a) = \hat{V}_t^i(a)$, and $\tilde{V}_{t,1}^i(a) = V_t^i(a)$. Equations (36) and (37) are the discrete-time approximations to the Bellman equation (22).

We solve (36)-(37) backwards, as follows. Given $V_{t+1}^i(\cdot)$ (recall the terminal condition $V_{N_T}^i(\cdot) = U_i(\cdot)$), we compute $\bar{\Gamma}_{t+1}^{ij}(\cdot,\cdot)$, and given our guess of $\{f_{t+1}^i(\cdot)\}_{i\in\mathbb{N}}$, we compute $\hat{V}_t^i(\cdot)$ using (36). We then compute $\{\tilde{V}_{t,s}^i(\cdot)\}_{s\in\{1,2,\ldots,N_S\}}$ using (37) and the terminal condition $\tilde{V}_{t,N_S+1}^i(\cdot) = \hat{V}_t^i(\cdot)$ by iterating backwards, which delivers $V_t^i(\cdot) = \tilde{V}_{t,1}^i(\cdot)$.

Step 3: Compute the implied distribution of balances Given the negotiated loan sizes, $b_t^{ij}(\cdot,\cdot)$ and the distribution of random payments, we can solve for the distribution of balances forward from an initial condition, $f_0^i(a)$. As in **step 2**, we need to compute the evolution of balances for the two within-period stages (the payments stage, and the trading stage). Since we are solving for the distribution of reserves forward, we start with the first stage and then move to the second stage.

Let $\tilde{f}_{t,s}^{i,\text{new}}(a_m)$ be the fraction of banks of type $i \in \mathbb{N}$ with balance $a_m \in \mathbb{A}^i$, at the beginning of subperiod s of the first stage of period t. We use the superscript "new" to emphasize that this is the new distribution implied by the bargaining outcomes in **step 2** (rather than the distribution that was used to derive those outcomes). Then,

$$\tilde{f}_{t,s}^{i,\text{new}}(a_m) = (1-\Delta\lambda_i)\tilde{f}_{t,s-1}^{i,\text{new}}(a_m) + \Delta\lambda_i \sum_{j\in\mathbb{N}} \sum_{a\in\mathbb{A}^i} \sum_{z\in\mathbb{Z}^{ij}} \pi_j \mathbb{L}\left(a_m, a-z\right) G_{ij}(z) \tilde{f}_{t,s-1}^{i,\text{new}}(a) \tag{38}$$

for $s = 1, \dots, N_S$, with initial condition $\tilde{f}^{i,\text{new}}_{t,0}(a_m) = f^{i,\text{new}}_t(a_m)$, and where

$$\mathbb{L}(a_m, x) \equiv \mathbb{I}_{\{x \in (a_{m-1},\ a_m]\}} \frac{x - a_{m-1}}{a_m - a_{m-1}}$$

implements a linear interpolation. The recursion (38) is initialized with the exogenous time-0 distribution of balances, i.e., $f^{i,\text{new}}_0(a_m) = f^i_0(a_m)$.

Let $\hat{f}^{i,\text{new}}_t(a_m)$ be the fraction of banks of type $i \in \mathbb{N}$ with balance $a_m \in \mathbb{A}^i$ after the trades in the second stage of period t; it is given by

$$\begin{aligned}\hat{f}^{i,\text{new}}_t(a_m) &= (1 - \Delta\beta_i)\tilde{f}^i_{t,N_S}(a_m) \\ &\quad + \Delta\beta_i \sum_{j \in \mathbb{N}} \sum_{a \in \mathbb{A}^i} \sum_{\tilde{a} \in \mathbb{A}^j} \sigma_j \mathbb{L}\left(a_m, a - b^{ij}_t(\hat{a}, \tilde{a})\right) \tilde{f}^{i,\text{new}}_{t,N_S}(a) f^{j,\text{new}}_{t,N_S}(\tilde{a}).\end{aligned}$$

Having solved for $\hat{f}^{i,\text{new}}_t(\cdot)$, set $f^{i,\text{new}}_{t+1}(\cdot) = \tilde{f}^{i,\text{new}}_{t+1,0}(\cdot) = \hat{f}^{i,\text{new}}_t(\cdot)$, and move to next period.

Step 4: Check for convergence. We use two criteria for convergence.

Criterion 1. We determine that the algorithm has converged if the probability distribution in **step 1** is close enough to the probability distribution obtained after **step 4**. Specifically, we consider the algorithm has converged if $\mathcal{E}(f) \equiv \max_{a,i,t}|f^i_t(a) - f^{i,\text{new}}_t(a)| <$ 1e-4.

Criterion 2. We determine that the algorithm has converged if some key theoretical moments have stabilized across iterations. In particular, we look at convergence in the distribution of interest rates and measures of trading activity.[112] Specifically, let ρ^p_t denote the p-percentile of the (volume weighted) distribution of interest rates at time t. Every 10 iterations of the algorithm, we compute the rate percentiles ρ^p_t for $p \in \{0.05, 0.10, 0.30, 0.50, 0.70, 0.90, 0.95\}$, and then compute the error $\mathcal{E}(\rho) \equiv \max_{p,t}|\rho^p_t - \rho^{p,\text{new}}_t|$. Every 10 iterations, we also compute: the effective fed-funds rate (EFFR), the participation for each bank, $\mathcal{P}_i$, and the reallocation for each bank, $\mathcal{R}_i$. We consider the algorithm has converged if after 10 iterations, we have: (i) $\mathcal{E}(f) <$ 1e-3, (ii) $\mathcal{E}(\rho) <$ 1e-4, and (iii) the errors for the EFFR, $\mathcal{P}_i$, and $\mathcal{R}_i$ all below 1e-4. For all these error computations, we check errors comparing results 10 iterations apart (e.g.: the EFFR this iteration compared with the EFFR 10 iterations ago), which ensure that results are stable across algorithm iterations.

The reason we sometimes use **Criterion 2** is that, despite our using the trading cost ϵ in equation (35), we sometimes observe loans that entail very small gains from trade, but still

[112] See Section C.2 for details on computation of theoretical moments.

affect the distributions $\{f_t^i(\cdot)\}$. These small-surplus trades may keep the error $\mathcal{E}(F)$ above the convergence tolerance, but have no significant effect on the distribution of rates nor in the relevant measures of trading activity. To ensure the algorithm has stabilized, we only start implementing **Criterion 2** after 25 iterations, and check errors $\mathcal{E}(\rho)$, EFFR, $\mathcal{P}_i$, and $\mathcal{R}_i$ every 10 iterations. We have found that using **Criterion 1** exclusively has no significant effect on the main results, but it typically takes longer for the algorithm to converge.

C.2 Computation of theoretical moments

Many of the statistics that we compute from model output are volume-weighted, which is the standard way many official statistics are computed (e.g., the EFFR). In this section we provide more details on how to perform these calculations in the theory.

Let $\omega_t^{ij}(a,\tilde{a})$ be the share of loans between banks type $i \in \mathbb{N}$ and $j \in \mathbb{N}$ with balances $a \in \mathbb{A}^i$ and $\tilde{a} \in \mathbb{A}_j$ at time t, relative to the total volume of loans in the trading-day, υ. That is,

$$\omega_t^{ij}(a,\tilde{a}) = \frac{\tilde{\upsilon}_t^{ij}(a,\tilde{a})}{\upsilon}$$

where

$$\tilde{\upsilon}_t^{ij}(a,\tilde{a}) = (\Delta\beta_i)\, n_i \sigma_j F_t^i(a) F_t^i(\tilde{a}) |b_t^{ij}(a,\tilde{a})|,$$

and

$$\upsilon = \sum_t \sum_{i,j\in\mathbb{N}} \sum_{(a,\tilde{a})\in\mathbb{A}^i\times\mathbb{A}^j} \tilde{\upsilon}_t^{ij}(a,\tilde{a})$$

is the total volume of loans in the trading day.

The EFFR is the volume-weighted mean of all daily traded rates, i.e.,

$$\text{EFFR} = \sum_t \sum_{i,j\in\mathbb{N}} \sum_{(a,\tilde{a})\in\mathbb{A}^i\times\mathbb{A}^j} \omega_t^{ij}(a,\tilde{a}) \rho_t^{ij}(a,\tilde{a}). \tag{39}$$

Let υ_i^e and υ_i^r denote the values of all the loans that were extended and received, respectively, throughout the trading day by all banks of type $i \in \mathbb{N}$, i.e.,

$$\begin{aligned}
\upsilon_i^e &= \sum_t \sum_{i,j\in\mathbb{N}} \sum_{(a,\tilde{a})\in\mathbb{A}^i\times\mathbb{A}^j} \omega_t^{ij}(a,\tilde{a}) b_t^{ij}(a,\tilde{a}) \mathbb{I}_{\left\{b_t^{ij}(a,\tilde{a})>0\right\}} \\
\upsilon_i^r &= \sum_t \sum_{i,j\in\mathbb{N}} \sum_{(a,\tilde{a})\in\mathbb{A}^i\times\mathbb{A}^j} \omega_t^{ij}(a,\tilde{a}) b_t^{ij}(a,\tilde{a}) \mathbb{I}_{\left\{b_t^{ij}(a,\tilde{a})<0\right\}}.
\end{aligned}$$

The participation and reallocation measures are $\mathcal{P}_i = (\upsilon_i^e + \upsilon_i^r)/\upsilon$, and $\mathcal{R}_i = (\upsilon_i^e - \upsilon_i^r)/(\upsilon_i^e + \upsilon_i^r)$, respectively.

D Data

This appendix discusses the data we use in the paper. Section D.1 describes the data sources, how we merged them, and our sample selection procedure. Section D.2 describes our own calculations of statistics that we used in the empirical and quantiative sections of the paper. Section D.3 gives a detailed account of the market events of September 13–20, 2019.

D.1 Reserve balances, transfers, and regulatory requirements

We used three databases for bank-level data on: (1) reserve balances and Regulation-D requirements, (2) high-frequency reserve transfers, and (3) Liquidity Coverage Ratio (LCR) requirements. We discuss each below.

D.1.1 Reserve balances and Regulation D

Bank-level end-of day balances at daily frequency were provided by the Monetary Policy Operations and Analysis (MPOA) section of the Monetary Affairs Division at the Federal Reserve Board of Governors. MPOA also supplied us the bank-level data on Regulation-D reserve requirements for each two-week maintenance period. Reserve balances and Regulation-D requirements are reported at the level of the bank holding company (and we used the bank holding company as the relevant unit of observation throughout). MPOA reports end-of-day balances as of 6:30 pm EST. We imputed next-day beginning-of-day balances as of 9:00 am EST with the procedure explained in Section D.2.1.

D.1.2 Reserve tranfers

Fedwire Funds Services (*Fedwire*) is an electronic large-value real-time gross settlement system operated by the Federal Reserve Banks. Fedwire participants include commercial banks, savings banks, thrift institutions, credit unions, agencies and branches of foreign banks in the United States, government securities dealers, government agencies such as federal or state governments, and Government Sponsored Enterprises (GSEs, e.g., Freddie Mac, Fannie Mae, and Federal Home Loan Banks). These institutions hold reserve balances in accounts at the Federal Reserve, and use Fedwire to transfer reserves to other participants, e.g., to settle payments, or to lend or repay loans of reserve balances.

Every Fedwire participant is identified by a *Fedwire account number*. Whenever an institution uses multiple Fedwire account numbers, we followed the guidelines from the Reserve Bank

Operations and Payment Systems Division at the Federal Reserve Board for linking those Fedwire account numbers to a single *bank ID*. Whenever institutions with different bank IDs belong to the same bank-holding company, we aggregated them into a single entity (since regulations, e.g., reserve requirements, LCR and SLR requirements, and interest-on-reserves calculations, etc., typically apply at the level of the bank-holding-company level). In a few instances, a bank ID could not be matched to a bank-holding company. Those account numbers were excluded from the sample. We also excluded any bank ID that did not have any fed funds trading activity in a given year. Our sample consists of 754 Fedwire participants for the year 2006, 404 for the year 2014, 395 for the year 2017, and 412 for the year 2019.

Having mapped Fedwire account numbers to bank-holding companies, we assigned the identity of each Fedwire sender or receiver to a bank holding company. We used the output of the Furfine algorithm to identify the set of overnight loans from the universe of Fedwire transfers, and treated the remaining transfers as *payments* unrelated to overnight borrowing and lending. All individual payments with value lower than $10,000 between a pair of banks during a trading day are consolidated into a single payment.

D.1.3 Liquidity Coverage Ratio

LCR regulation requires a bank to maintain (typically on a daily basis) a quantity of *High Quality Liquid Assets* (HQLA) at least as large as a measure of total net cash outflows in a 30-day standardized stress scenario. If we let $\mathrm{H}_m(d)$ denote the quantity of qualifying HQLA held by bank m in a trading period d, and $\mathrm{L}_m(d)$ denote the corresponding measure of outflows in the stress scenario, the LCR regulation requires $\mathrm{L}_m(d) \leq \mathrm{H}_m(d)$.[113]

Both, $\mathrm{L}_m(d)$ and $\mathrm{H}_m(d)$ are made public by each bank at a quarterly frequency. We obtain data on the ratio of these quantities from S&P Global Capital IQ database.[114] We used SNL Classic Data and run a Companies (Classic) screener to search for our data. We extracted quarterly LCR (LIQUIDITY_COV_RATIO) data from 1990Q1 to 2021Q2. For some banks, LCR data were missing in some quarters. For these cases, we obtained the LCR data directly

[113] Appendix B (Section B.2.1) describes the LCR regulation in greater detail.
[114] The S&P database can be accessed at: https://www.spglobal.com/marketintelligence/en/.

from the bank's website.[115]

We merged our balances data from MPOA (described in Section D.1.1) with the S&P LCR data using the Replication Server System Database (RSSD) ID. (The balances data from MPOA contains the RSSD of each bank holding company.) We then created a manual cross-walk to match RSSDs to parent company names in the S&P database, using the National Information Center repository from the Federal Financial Institutions Examination Council (https://www.ffiec.gov/NPW). We always matched RSSDs to the parent bank holding company to which the LCR regulation applies. In general, this procedure implies matching the RSSDs to the highest level parent company in the corporate structure, except for cases in which the parent company is a sovereign government and the LCR constraint applies to the second highest parent company level.

D.2 Empirical computations

D.2.1 Balances: beginning-of-day imputation

This section provides further details about the construction of beginning-of-day (BOD) balances that we discussed in Section 3.3. The BOD balances used in the paper were obtained from the following three-step procedure for each bank:

- Step 1. We started with the end-of-day (EOD) balance for trading day $d-1$ obtained from MPOA, and calculated a "basic" measure of the BOD balance for trading day d, by adding (subtracting) the repayments received (sent) corresponding to loans extended (received) during trading day d.

- Step 2. From the "basic" measure of BOD balance calculated in Step 1, we calculated an "adjusted" measure of BOD balance by subtracting the quantity of required reserves, i.e., the minimum level of reserves that the bank must hold during the maintenance period in order to comply with Regulation D and the minimum LCR requirement.

[115] This was the case for the following three banks:

- Credit Agricole Group (https://www.credit-agricole.com/en/pdfPreview/186985)
- DNB ASA (https://ir.dnb.no/capital-framework)
- State Street Corporation (https://investors.statestreet.com/filings-and-reports/u-s-liquidity-coverage-ratio-disclosures/default.aspx).

- Step 3. From the "adjusted" measure of BOD balance calculated in Step 2, we calculated a measure of "unencumbered" BOD reserve balance for trading day d, by the netting predictable payments that take place during trading day d.

Next, we discuss each step in more detail.

Step 1: Netting repayments of previous-day loans. For each bank holding company m in our sample, we obtained the EOD balance as of 6:30 pm EST of day d from MPOA (see Section D.1.1), which we denote $\mathrm{a}_m^{\mathrm{eod}}(d)$. For each bank m, we used the output of the Furfine algorithm to compute the repayments to be sent and received on day d corresponding to loans originated during day $d-1$. Let $\mathrm{receive}_m(d)$ and $\mathrm{send}_m(d)$ denote the amounts of reserves that bank m will receive or send, respectively, on day d, and define the net repayment corresponding to loans originated during day $d-1$, as $\mathrm{net}_m(d) \equiv \mathrm{receive}_m(d) - \mathrm{send}_m(d)$. We then computed $\mathrm{a}_m(d) = \mathrm{a}_m^{\mathrm{eod}}(d) + \mathrm{net}_m(d)$, which is our "basic" measure of BOD balance for bank m on day d. Finally, we computed the BOD "basic" balance for the maintenance period h as the average of $\mathrm{a}_m(d)$ for days $d \in h$: $\mathrm{a}_m(h) = \frac{1}{N_h} \sum_{d \in h} \mathrm{a}_m(d)$, where N_h is the number of trading days in a maintenance period h.

As mentioned in Section 3.2 (footnote 24), for the purpose of calculating the "basic" BOD balance, we treated GSEs differently than banks. In the case of a GSE, we did not only net out the repayments corresponding to loans issued on day $d-1$ (i.e., $\mathrm{net}_m(d)$), but *all transfers* sent or received during trading day d—involving *any* counterparty, not only those that meet the sample selection criteria described Section D.1.2. The rationale for netting all transfers that will occur during day d to obtain the GSE's "basic" BOD balance for day d is that a GSE's business model generates very predictable cashflows, so through the lens of our theory, we regard the GSE as being able to predict all its intraday Fedwire transfers at the beginning of the trading day.

Step 2: Subtracting reserve requirements. For each bank m in maintenance period h, we computed "adjusted" (excess) reserves as $\mathrm{x}_m(h) \equiv \mathrm{a}_m(h) - \underline{\mathrm{a}}_m^D(h) - \underline{\mathrm{a}}_m^L(h)$, where $\underline{\mathrm{a}}_m^D(h)$ and $\underline{\mathrm{a}}_m^L(h)$ denote the Regulation D and LCR reserve requirements, respectively. The bank-level Regulation-D requirement, $\underline{\mathrm{a}}_m^D(h)$, was provided by MPOA. The reserve requirement implied by the LCR regulation is less straightforward, as we discuss next.

As explained in Appendix B (Section B.2.1), the LCR regulation requires a bank to maintain (on a daily or monthly basis) a quantity of *High Quality Liquid Assets* (HQLA) at least as large as a measure of total net cash outflows in a 30-day standardized stress scenario. Specifically, if we let $\mathrm{H}_m(d)$ denote the quantity of qualifying HQLA held by bank m in a trading period d (a day or a month, depending on the type of institution, see footnote 106) and $\mathrm{L}_m(d)$ denote the corresponding measure of outflows in the stress scenario, the LCR regulation requires $\mathrm{L}_m(d) \leq \mathrm{H}_m(d)$. The set of qualifying HQLA includes reserves in excess of Regulation D, as well as securities issued or guaranteed by the U.S. Treasury (and also other securities, but subject to caps and haircuts). The fact that the LCR regulation allows banks to meet the requirement with assets other than reserves presents a challenge when trying to identify the quantity of reserves that bank m treats as "required" to satisfy the LCR constraint in period d, i.e., $\underline{\mathrm{a}}_m^L(d)$. Our strategy to tackle this identification problem is to set $\underline{\mathrm{a}}_m^L(d) = \max(0, \mathrm{L}_m(d) - \mathrm{A}_m(d))$, where $\mathrm{A}_m(d) \equiv \mathrm{H}_m(d) - \max\left(0, \mathrm{a}_m(d) - \underline{\mathrm{a}}_m^D(d)\right)$ is the quantity of qualifying HQLA in excess of (i.e., other than) reserves net of the Regulation D requirement.[116] Our proposed measure of excess reserves, $\mathrm{x}_m(h) \equiv \mathrm{a}_m(h) - \underline{\mathrm{a}}_m^D(h) - \underline{\mathrm{a}}_m^L(h)$, selects the largest level of excess reserves net of the Regulation D requirement that is consistent with the LCR constraint.

For banks that are not subject to LCR regulation (such as banks with assets below \$50 bn in our sample period), we set $\underline{\mathrm{a}}_m^L(h) = 0$. Since GSEs are not subject to Regulation D or LCR regulation, we set $\underline{\mathrm{a}}_m^D(h) = \underline{\mathrm{a}}_m^L(h) = 0$ for $m \in \mathbb{B}_{GSE}$. Finally, since we only have quarterly LCR observations (see Section D.1.3), we imputed the same LCR-induced reserve requirement for all maintenance periods within the quarter.

Step 3: Netting predictable payments. To go from the bank-level "adjusted" measure of BOD balance calculated in Step 2, to the bank-level measure of "unencumbered" BOD reserve balance for period h, we netted (at the individual bank level) all *predictable payments* that take place during period h, as explained in Section 3.3.

D.2.2 Network statistics

In this section we describe the calculations of the network statistics reported in Figure 3.[117] We begin by introducing some notation. Let υ_{md}^e denote the dollar value of all loans extended, and υ_{md}^r denote the dollar value of loans received, by bank m on day d. Let υ_{mh}^e and υ_{mh}^r

[116]See Section B.2.1 in Appendix B for a more detailed explanation of our strategy to identify the quantity of required reserves induced by the LCR regulation.

[117]The theoretical counterparts of these computations are discussed in Appendix C (Section C.2).

denote the dollar values of loans extended and received, respectively, during the maintenance period h, i.e., $\upsilon^e_{mh} = \sum_{d \in h} \upsilon^e_{md}$ and $\upsilon^r_{mh} = \sum_{d \in h} \upsilon^r_{md}$. Finally, let $\upsilon_h = \sum_m \upsilon^e_{mh}$ denote the total dollar value of loans extended in maintenance period h. We compute the participation and reallocation values by *bank type*, $i \in \{F, M, S, GSE\}$, as follows.

Participation rate by bank type. We computed the participation rate for bank type $i \in \{F, M, S, GSE\}$ during maintenance period h as $\mathcal{P}_{ih} = \sum_{m \in i} \frac{\upsilon^e_{mh} + \upsilon^r_{mh}}{\upsilon_h}$. We then computed the participation rate of bank type $i \in \{F, M, S, GSE\}$ in a given year as $\mathcal{P}_i = \frac{1}{N_h} \sum_h \mathcal{P}_{ih}$, where N_h denotes the number of maintenance periods in the year.

Reallocation index by bank type. We computed the reallocation index for bank type $i \in \{F, M, S, GSE\}$ during maintenance period h as $\mathcal{R}_{ih} = \frac{\sum_{m \in i} \upsilon^e_{mh} - \sum_{m \in i} \upsilon^r_{mh}}{\sum_{m \in i} \upsilon^e_{mh} + \sum_{m \in i} \upsilon^r_{mh}}$. We then computed the reallocation index of group i in a given year as $\mathcal{R}_i = \frac{1}{N_h} \sum_h \mathcal{R}_{ih}$, where N_h denotes the number of maintenance periods in the year.

As explained in Section 3.1, the arrows from one node to another in Figure 3 represent loans extended from banks of that type to the other. The arrow width is proportional to the volume of trade between the bank types connected by the arrow. The node size is proportional to the volume of trade between banks of a given type. The arrow widths and node sizes are defined relative to the trades within a year, so they are not comparable across years. The colors of the arrows and nodes are: light blue, dark blue, light red, or dark red, depending on whether the volume-weighted average interest rate on the loans between the two types of banks, expressed as a spread over the EFFR, falls in the first, second, third, or fourth quartile, respectively.

D.2.3 Kernel density estimations

We use Gaussian kernel densities to estimate the distributions of payment shocks, beginning-of-day reserve balances, and aggregate reserve-draining shocks. For the distributions of payment shocks, and the distribution of reserve-draning shocks, we set the smoothing parameter, h, using a standard "rule of thumb", namely $h = 0.9 \min\left(\hat{\sigma}, \frac{\mathrm{IQR}}{1.34}\right) \mathrm{n}^{-1/5}$, where n, $\hat{\sigma}$, and IQR denote the number of observations, standard deviation, and interquartile range of the sample, respectively. For the distributions of beginning-of-day reserves we use the [iterative] methodology described in Botev et al. (2010) to set the smoothing parameter (since they may be multimodal, as seen in Figures 6-9).

D.2.4 Reduced-form estimation of the reserve demand (11)

As in Section 6.2, let s_t denote the EFFR-IOR spread on day t, and Q_t denote the aggregate quantity of reserves at the end of day t. We estimated equation (11) using a nonlinear least-squares procedure. For each sample period, we estimated the vector of parameters, $\nu \equiv \{\underline{s}, \tilde{s}, \xi, Q_0\}$, with $\tilde{s} \equiv \overline{s} - \underline{s}$, to solve

$$\Xi = \min_{\nu} \left\{ \sum_t (s_t - D(Q_t))^2 \right\} \text{ s.t. } 0 \leq \tilde{s}, 0 \leq \xi, \tag{40}$$

where $D(\cdot)$ is as defined in equation (11). We found the solution to (40) by following two steps. In the first step, we did a thorough grid search: we set equally spaced grids for each parameter in ν, computed the hypercube combining all these grids, and then evaluated Ξ for each entry in this hypercube.[118] Let ν_{grid} be the vector or parameters that delivered the lowest value of Ξ. In the second step, we used a Nelder-Mead optimization starting from ν_{grid}.

D.2.5 A mapping between reserves of all banks and reserves of active banks

Let Q_t^D denote the quantity of *total reserves* on day t in the sample of all banks in the data (e.g., the quantity of reserves shown in Figure 16). Let Q_t^M denote the quantity of *active excess reserves* on day t that we use to calibrate our model to the year 2019 (and in the interpolation procedure described in Section 3.6, which also uses the year 2017 as an endpoint).[119] Let $\mathbb{T}$ denote a subset of trading days and let T be the cardinality of this set. For any sample $\left\{Q_t^D\right\}_{t\in\mathbb{T}}$, define $\bar{Q}_{\mathbb{T}}^D \equiv \frac{1}{T}\sum_{t\in\mathbb{T}} Q_t^D$. Similarly, for any sample $\left\{Q_t^M\right\}_{t\in\mathbb{T}}$, define $\bar{Q}_{\mathbb{T}}^M \equiv \frac{1}{T}\sum_{t\in\mathbb{T}} Q_t^M$.

Our model output, e.g., the aggregate demand for reserves, is computed using a quantity of reserves $Q \in \mathbb{R}$ constructed with the interpolation procedure described in Section 3.6, which uses $\bar{Q}_{2017}^M$ and $\bar{Q}_{2019}^M$, i.e., the average quantity of reserves in excess of LCR and Regulation D *in our subsample of active banks* for the two base years. For some exercises (e.g., the top-right panel of Figure 18) we want to show—in the same graph—the model output along with actual daily data observations of total reserves and interest rates, but the observations that we have

[118]We used grid sizes of: 50 points for $\underline{s}$, 50 points for $\tilde{s}$, 123 points for Q_0, and 63 points for ξ. This gave a combination of 19,372,500 values for ν. The bounds for each grid were: -0.50 and 0.01 for $\underline{s}$, 0.00 and 1.00 for $\tilde{s}$, $-3 \times Q_{2019}$ and $3 \times Q_{2019}$ for Q_0, and $1e-6$ and 0.10 for ξ. We always found the the optimal value for ν well within our bounds.

[119]That is, $\{Q_t^M\}$ is the time series for the aggregate quantity of reserves for the subsample of banks that were active in fed funds trading during the years under study, net of Regulation D and LCR requirements, as explained in Section 3.3.

available at a daily frequency are $\{Q_t^D\}_{t\in\mathbb{T}}$, not $\{Q_t^M\}_{t\in\mathbb{T}}$. So we need a way to "transform" each daily observation, Q_t^D, into an estimate of Q_t^M.

We adopt a transformation $\mathcal{G}$, such that $Q_t^M = \mathcal{G}\left(Q_t^D;\mathbb{T}\right)$ for all $t\in\mathbb{T}$, which satisfies two properties for any sample $\mathbb{T}$: (1) daily variation in reserves in the full sample of banks is the same as daily variation in reserves in the subsample of banks, i.e., $Q_{t+1}^M - Q_t^M = Q_{t+1}^D - Q_t^D$ for all $t\in\mathbb{T}$ (this is consistent with our strategy of calibrating the slope of our model-generated reserve demand to match the liquidity effect associated with variation in the quantity of reserves of the full sample of banks); and (2) $\bar{Q}_{\mathbb{T}}^M = \mathcal{F}\left(\bar{Q}_{\mathbb{T}}^D\right)$, where $\mathcal{F}$ is a linear function that satisfies $\mathcal{F}\left(\bar{Q}_{2017}^D\right) = \bar{Q}_{2017}^M$ and $\mathcal{F}\left(\bar{Q}_{2019}^D\right) = \bar{Q}_{2017}^M$ (the subscript "2017" denotes the sample of all trading days in the year 2017, and the subscript "2019" denotes the sample of all trading days in the year 2019). For any sample $\mathbb{T}$ of trading days, we posit

$$\begin{aligned} Q_t^M &= \mathcal{G}\left(Q_t^D;\mathbb{T}\right) \\ &\equiv Q_t^D - \bar{Q}_{\mathbb{T}}^D + \bar{Q}_{\mathbb{T}}^M, \end{aligned} \tag{41}$$

where

$$\bar{Q}_{\mathbb{T}}^M \equiv \omega_{\mathbb{T}}^D \bar{Q}_{2019}^M + \left(1-\omega_{\mathbb{T}}^D\right)\bar{Q}_{2017}^M, \tag{42}$$

with

$$\omega_{\mathbb{T}}^D \equiv \frac{\bar{Q}_{\mathbb{T}}^D - \bar{Q}_{2017}^D}{\bar{Q}_{2019}^D - \bar{Q}_{2017}^D}. \tag{43}$$

Intuitively, for any sample $\mathbb{T}$ of trading days, the transformation (41) adds the daily deviations relative the sample mean of total reserves *of all banks*, to an *imputed mean* corresponding to the *subset of active banks* (the imputed mean is the linear transformation, $\mathcal{F}$, that is consistent with the observations $\bar{Q}_{2017}^M$ and $\bar{Q}_{2017}^M$ in the sense that $\mathcal{F}\left(\bar{Q}_{2017}^D\right) = \bar{Q}_{2017}^M$ and $\mathcal{F}\left(\bar{Q}_{2019}^D\right) = \bar{Q}_{2017}^M$).[120]

D.3 Events of September 13–20, 2019

In Section 8 we use our quantitative theory to analyze the fed-funds rate spikes of September 2019. In this section we give more background on the associated reserve-draining shocks, and

[120]To check that $\mathcal{G}$ and $\mathcal{F}$ satisfy the desired properties, notice that for any sample $\mathbb{T}$, (41) implies $Q_{t+1}^M - Q_t^M = Q_{t+1}^D - Q_t^D$ for all $t\in\mathbb{T}$, and (42)-(43) imply $\bar{Q}_{\mathbb{T}}^M = \mathcal{F}(\bar{Q}_{\mathbb{T}}^D)$, with

$$\mathcal{F}(\bar{Q}_{\mathbb{T}}^D) \equiv \frac{\bar{Q}_{2017}^D\bar{Q}_{2019}^M - \bar{Q}_{2019}^D\bar{Q}_{2017}^M}{\bar{Q}_{2017}^D - \bar{Q}_{2019}^D} + \frac{\bar{Q}_{2017}^M - \bar{Q}_{2019}^M}{\bar{Q}_{2017}^D - \bar{Q}_{2019}^D}\bar{Q}_{\mathbb{T}}^D,$$

which satisfies $\mathcal{F}\left(\bar{Q}_{2017}^D\right) = \bar{Q}_{2017}^M$ and $\mathcal{F}\left(\bar{Q}_{2019}^D\right) = \bar{Q}_{2019}^M$.

the monetary-policy interventions that followed these rate spikes.[121]

The events unfolded as follows. On Friday, September 13 the beginning-of-day supply of reserves was about \$1.5 tn, and the EFFR printed at 214 bps. In the top panel of Figure 21, September 13 is the dark dot that sits on the demand for reserves generated by the theory—well within the FFR target range. On Monday, September 16 the beginning-of-day supply of reserves was \$51.5 bn lower than on the previous trading day (due to reserve-draining shocks that occured throughout Friday, September 13), and the EFFR printed at 225 bps (the upper limit of the target range). In the top panel of Figure 21, September 16 is the rightmost dark dot that sits on the upper limit of the target range for the EFFR. On Tuesday, September 17 the beginning-of-day supply of reserves was \$65.72 bn lower than on the previous trading day (due to reserve-draining shocks that occured throughout Monday, September 16), and the EFFR printed at 230 bps (5 bps above the upper limit of the target range). In the top panel of Figure 21, September 17 is the uppermost dark dot. Following an overnight repo operation that injected \$53 billion on Tuesday, September 17, the beginning-of-day supply of reserves on September 18 was \$46.3 bn higher than on the previous day, and the EFFR fell to 225 bps.[122]

The morning of Tuesday, September 17 was the first time since the GFC that the Desk conducted an open-market operation to manage the fed funds rate. That Tuesday afternoon the Desk announced it would conduct an overnight operation at 8:15 a.m. on Wednesday, September 18. This operation injected \$75 bn, which contributed to the beginning-of-day supply of reserves on Thursday, September 19, being \$3.67 bn higher than the previous day. Similar operations were used to inject \$75 bn every day until the end of the week. The EFFR printed at 190 bps on September 19 and September 20.[123]

[121]Table 2 summarizes the main facts. Most of the events we describe in this section are based on the detailed accounts provided Afonso et al. (2020a) and Anbil et al. (2020).

[122]On Monday afternoon (2019/09/16), in response to the observed upward pressure on the EFFR, the Desk announced an overnight repo operation to be conducted at 9:30 AM on Tuesday (2019/09/17), offering up to \$75 billion against Treasury, agency, and agency MBS collateral, of which only \$53 bn were subscribed.

[123]On Thursday, September 19, the Federal Reserve also made adjustments to administered rates and the FFR target range. The ONRRP was reduced from 200 bps to 170 bps, the IOR from 210 bps to 180 bps, and the DWR from 275 bps to 250 bps. The lower limit of the FFR target range was reduced from 200 bps to 175 bps, and the upper limit was reduced from 225 bps to 200 bps. On the morning of Friday, September 20, the Desk announced a series of operations over the quarter-end, which included three two-week operations covering the quarter-end and daily overnight operations of \$75 billion through October 10. The September 16–17 event seem to have had lasting an impact on the conduct of monetary policy. As Afonso et al. (2020a, p. 24) recount:

> *On October 11, 2019, the FOMC announced its intention to maintain an ample supply of reserve balances at or above the level that prevailed in early September. The FOMC instructed the Desk to purchase Treasury bills at least into the second quarter of 2020 (and to continue repo operations) in order to supply reserves and mitigate money market pressures that might impede policy imple-*

D.3.1 JPM earnings call for the period ending September 30, 2019

In this section we report the key excerpts of the earnings call of October 15, 2019, in which Jamie Dimon, Chairman and CEO of JPMorgan Chase (JPM), answered questions about why JPM was not more active lending in money markets during the week of September 16, 2019.

Question: Glenn Schorr (Analyst, Evercore ISI)

> *Curious your take on everything that went on in the repo markets during the quarter, and I would love it if you could put it in the context of maybe the fourth quarter of last year. If I remember correctly, you stepped in in the fourth quarter, saw higher rates, threw money at it, made some more money, and it calmed the markets down. I'm curious what's different this quarter that did not happen, and curious if you think we need changes in the structure of the market to function better on a go-forward basis.*

Answer: Jamie Dimon (Chairman and CEO, JPM)

> *So, if I remember correctly, you got to look at the concept of – we have a checking account at the Fed with a certain amount of cash in it. Last year we had more cash than we needed for regulatory requirements. So when repo rates went up, we went from the checking account, which [ph] was paying (00:14:10) IOER into repo. Obviously makes sense, you make more money. But now the cash in the account, which is still huge – it's $120 billion in the morning and goes down to $60 billion during the course of the day and back to $120 billion at the end of the day – that cash, we believe, is required under resolution and recovery and liquidity stress testing. And therefore, we could not redeploy it into repo market, which we would have been happy to do. And I think it's up to the regulators to decide they want to recalibrate the kind of liquidity they expect us to keep in that account. Again, I look at this as technical; a lot of reasons why those balances dropped to where they were. I think a lot of banks were in the same position, by the way. But I think the real issue, when you think about it, is what does that mean if we ever have bad markets? Because that's kind of hitting the red line in the Fed checking account, you're also going to*

> *mentation. The goal of the bill purchases was to ensure the smooth functioning of money markets at the current monetary policy stance, not to change the monetary policy stance.*

For details, see https://www.federalreserve.gov/newsevents/pressreleases/monetary20191011a.htm.

hit a red line in LCR, like HQLA, which cannot redeployed either. So, to me, that will be the issue when the time comes. And it's not about JPMorgan. JPMorgan will be fine in any event. It's about how the regulators want to manage the system and who they want to intermediate when the time comes.

Question: Erika Najarian (Analyst, Bank of America Merill Lynch)

Yes, good morning. My first question is a follow-up to Glenn's question. As we think about the crosscurrents of resolution planning, LCR, and liquidity stress testing, could you help us – what is the level of excess deployable cash at JPMorgan?

Answer: Jamie Dimon (Chairman and CEO, JPM)

As I said, we have $120 billion in our checking account at the Fed, and it goes down to $60 billion and then back to $120 billion during the average day. But we believe the requirement under CLAR and resolution and recovery is that we need enough in that account, so if there's extreme stress during the course of the day, it doesn't go below zero. If you go back to before the crisis, you'd go below zero all the time during the day. So the question is, how hard is that as a red line? Was the intent of regulators between CLAR and resolution to lock up that much of reserves in the account with Fed? And that'll be up to regulators to decide. But right now, we have to meet those rules and we don't want to violate anything we've told them we're going to do.

For a full transcript of the call, visit: `https://tinyurl.com/29scwszt`.

E Quantitative analysis for the pre-GFC-regulation regime

Our quantitative analysis in the body of the paper focuses on the current post-GFC monetary-policy framework. For completeness, and because the pre-GFC period is of historical interest, in this section we also study the pre-GFC framework. The pre-GFC and post-GFC frameworks differ in two ways. First, the quantity of excess reserves was close to zero in the former, but is very large in the latter. Second, as discussed in Section 3, regulations introduced after the GFC have affected banks' payoffs from fed funds trading. For this reason, in this section we recalibrate the model for a base year before the GFC, which we choose to be 2006.[124]

[124]Our main motive for recalibrating the model is that the trading network, which in our theory is represented by the parameters $\{\beta_i, \kappa_i\}_{i\in\mathbb{N}}$, may not be stable across policy regimes. For example, it is reasonable to imagine

E.1 Calibration

We set ι_w to match the prevailing DWR, and $\iota_o = 0$ (since there was no ONRRP facility in 2006). The remaining nine parameters, ι_r and $\{\beta_i, \kappa_i\}_{i\in\mathbb{N}}$, are calibrated so that the equilibrium of the model matches the following nine empirical moments: (i) effective fed funds rate[125]; (ii)-(v) reallocation indices $\{\mathcal{R}_i\}_{i\in\mathbb{N}}$ (as defined in Section 3.1); (vi)-(viii) participation rates $\{\mathcal{P}_i\}_{i\in\mathbb{N}\setminus\{F\}}$ (as defined in Section 3.1)[126]; (ix) empirical estimates of the "liquidity effect" (at the average level of aggregate reserves outstanding in the base year, as reported in Section 3.5).

Table 3 reports the parameter values, empirical targeted moments, and the corresponding theoretical moments for the 2006 calibration. Banks of type F, M, and S, accounted for about 0.5%, 3%, and 95%, of all the institutions that were active in the fed funds market in 2006, respectively.[127] To interpret the frequencies of payment shocks, $\{\lambda_i\}_{i\in\mathbb{N}}$, recall that λ_i represents the probability that a bank of type i receives a payment shock in a one-second time interval, so for example, $\lambda_F = 0.901$ implies a bank of type F receives a payment shock approximately every 1.1 seconds, on average. Similarly, $\lambda_M = 0.402$ implies a bank of type M receives a payment shock approximately every 2.5 seconds, and $\lambda_S = 0.007$ implies a bank of type S receives a payment shock approximately every 2.38 minutes, on average. The rate ι_w corresponds to a DWR equal to 6.25% per annum, which was in effect in the second half of 2006. The calibrated value of ι_r is 4.81% per annum.[128]

that the trading patterns represented by the type-specific meeting rates may change in response to regulatory constraints, in particular those post-GFC regulations that increased the cost of borrowing, and therefore the cost of intermediating fed funds. We use 2006 as the baseline year for the pre-GFC period for two reasons. First, policy rates and total reserves remained stable for most of that year, and it was the last "normal" year before the GFC that spurred the policy interventions that changed the landscape of the fed funds market. Second, the 2006 calibration will allow us to assess the model fit in a pre-GFC-regulation environment, and it will also allow us to test the quantitative predictions of the theory as we vary the level of aggregate excess reserves from near zero (the level they had during 2006) to \$2.689 tn (the level they reached for all the banks in our sample in 2014, which was the last pre-GFC-regulation year).

[125]Our calibration strategy uses the effective fed funds rate as a calibration target unless the Federal Reserve pays interest on reserves (IOR) in the base year, in which case we simply set ι_r to match the IOR. For example, the IOR was 2.35% per annum in May-July 2019, so we set $\iota_r = 0.0235/360$ in the 2019 calibration. The Federal Reserve did not pay IOR before October 9, 2008, so in the 2006 calibration we regard ι_r as a proxy for a bank's unmodelled opportunity cost of lending reserves in the fed funds market, and calibrate it internally so that the average (volume-weighted) interest rate in the model is equal to 5.25% per annum, which was the effective fed-funds rate prevailing during the second half of 2006.

[126]The participation rate of type F banks is not an explicit calibration target because it is implied by the participation rates of the other three types, since $\sum_{i\in\mathbb{N}} \mathcal{P}_i = 2$.

[127]The main change in the bank population between 2006 and 2019 is the reduction in the total number of active banks in our sample, mostly due to the fact that almost half of the banks of type S that were fed funds market participants in 2006 did not trade fed funds during 2019.

[128]For comparison, 4.81% per annum is the 0.5 percentile of the volume-weighted distribution of rates observed

The frequency of trade, β_i, is the probability a bank of type i contacts a trading partner during a 42-second time interval. Thus, the calibrated values $\{\beta_i\}_{i\in\mathbb{N}}$ for 2006 imply that banks of type F, M, S, and GSE trade fed funds approximately every 1.75 minutes, 8 minutes, 20 minutes, and 3.5 minutes, respectively. The calibration also ensures that, when computed in the neighborhood of zero excess reserves, the magnitude of the "liquidity effect" in the theory is within the range of the empirical estimates reported in Section 3.5 (i.e., about a 1.7 bp increase in the fed funds rate per \$1bn reduction in the aggregate quantity of reserves).[129] The borrowing costs $\{\kappa_i\}_{i\in\mathbb{N}}$, which proxy for institutional and regulatory considerations that affect banks' incentives to buy fed funds, are nill for banks of type F, M, and S in the 2006 calibration.[130]

E.2 Validation

In this section we report the model fit of empirical observations that were not targeted in the calibration. We organize the material in two sections: the first focuses on the cross-sectional distribution of loan rates, and the second on the main features of the fed funds trading network.

E.2.1 Distribution of interest rates

Figure 23 shows the empirical and theoretical cumulative distribution functions of bilateral fed funds rates in the year 2006 (expressed in percent per annum).[131] The model delivers a reasonable fit for the distribution of bilateral fed funds rate, which was not a calibration target.

in the second half of 2006. That is, only half of one percent of the fed funds traded in the second semester of 2006 had a rate below 4.81%, so we regard 4.81% as a reasonable proxy for the unmodelled opportunity cost of an alternative use of reserves. We focus on July–December because in that period the administered rates (i.e., the Discount-Window rate and the fed funds rate target) were constant and equal to the rates targeted in the 2006 calibration (the administered rates had been gradually increasing in the first half of 2006).

[129] Figure 22 shows the magnitude of the liquidity effect in the model calibrated to 2006 (extracting reserves using the procedure described in Section 3.6), as well as the confidence bands for the estimates from Carpenter and Demiralp (2006) reported in Section 3.5. The model-generated liquidity effect is within the range of empirical estimates.

[130] In every calibration the value of κ_{GSE} is set large enough to match the observation that GSEs essentially do not borrow in the fed funds market.

[131] The empirical interest rates for 2006 are from the sample period July–December because throughout that period the Discount-Window rate and the fed funds rate target were constant and equal to the rates targeted in the 2006 calibration. To obtain the equilibrium rates for 2006, the model is calibrated as in Table 3.

E.2.2 Fed funds trading network

Figure 24 shows the empirical fed funds trading network for the year 2006 (top panel) and the corresponding trading network generated by the model for the 2006 calibration (bottom panel). As explained in Section 3.1, these network plots show the location of the four bank types in the coordinate axes defined by the reallocation index, $\mathcal{R}_i$, and the participation rate, $\mathcal{P}_i$, and convey information on the sizes of the flows of reserves associated with fed funds lending across and within bank types, as well as on the average interest rates on the underlying loans.[132]

The theoretical network matches several characteristics of the empirical one. For example, it replicates quite well the direction and volume of the loans between bank types (represented by the direction and width of the arrows between the nodes). In this regard, one shortcoming of the model is that it underpredicts the volume of loans within bank types S and M. The model is consistent with the empirical facts that banks of type S lend to each other at relatively high rates, while banks of type F can borrow at relatively low rates from banks of type M, S, and GSEs. In terms of shortcomings, the model predicts that banks of type S borrow at relatively high rates from GSEs, that loans between banks of type F carry relatively low rates, and that loans between banks of type M carry relatively high rates, as do loans from type F to type M, but these predictions do not match the empirical patterns.

E.3 Aggregate demand for reserves

Consider the model calibrated to the year 2006, as described in Table 3, but with $\iota_w = 0.0075/365$ and $\iota_r = 0.0025/365$, to match the DWR and IOR in the year 2014. Then, using the notation introduced in Section 3.6, let $\text{Y}_0 = 2006$ and $\text{Y}_1 = 2014$, i.e., Y_0 and Y_1 represent the years 2006, and 2014, respectively, with $\bar{n}^i_{2014}$ and $\bar{F}^i_{2014}$ given by the estimates reported in Section 3.3. Construct a grid, $\mathbb{G} \subset \mathbb{R}$ for ω, and for each $\omega \in \mathbb{G}$, use the interpolation procedure described by (7) and (8) to generate the sample $\left\{(\bar{n}^i_{\text{Y}_\omega}, \bar{F}^i_{\text{Y}_\omega})\right\}_{(i,\omega)\in\mathbb{N}\times\mathbb{G}}$. For each pair $(\bar{n}^i_{\text{Y}_\omega}, \bar{F}^i_{\text{Y}_\omega})$ of elements of this sample, use the model to compute the corresponding equilibrium

[132] In comparing the top and bottom panels of Figure 24, notice that while the positions of the four nodes (each of which represents the set of banks assigned to a particular type) in $\mathcal{R}_i$-$\mathcal{P}_i$ space have been used as calibration targets, the remaining collection of statistics that shape these network representations were not targeted. This includes the node sizes (each of which is proportional to the volume of trade between banks of a given type), the direction of each arrow (which indicates which bank type lends), the width of each arrow (which is proportional to the volume of trade between the bank types connected by the arrow), the colors of the arrows and nodes (which are light blue, dark blue, light red, or dark red, if the volume-weighted average interest rate on the loans between the two types of banks, expressed as a spread over the EFFR, falls in the first, second, third, or fourth quartile, respectively).

value-weighted fed funds rate, which we denote $\iota^*_{\Upsilon_\omega}$, and let $Q_{\Upsilon_\omega} \equiv \sum_{i \in \mathbb{N}} \bar{n}^i_{\Upsilon_\omega} \int a d\bar{F}^i_{\Upsilon_\omega}(a)$. This procedure delivers a sample of pairs, $\left\{\left(Q_{\Upsilon_\omega}, \iota^*_{\Upsilon_\omega}\right)\right\}_{\omega \in \mathbb{G}}$, which we represent with the mapping $\iota^*_{\Upsilon_\omega} = \mathcal{D}(Q_{\Upsilon_\omega}; \Pi)$. This mapping, which we interpret as the aggregate demand for reserves generated by the theory, is shown in Figure 25.[133]

[133]We use 2006 as one endpoint for our interpolation procedure since it was the last year of the scarce-reserve regime that prevailed until the GFC. We use 2014 as the other endpoint because it is the year when the quantity of reserves achieved its maximum historical level of the pre-2020 era. By varying ω on $[0,1]$ we can use (9) to span any aggregate level of excess reserves between 0 (roughly the pre-GFC level prevailing in 2006) and \$2.7 tn (roughly the level achieved in 2014).

Parameter	Target	Moment	
		Data	Model
$n_F = 0.005$	proportion of financial institutions of type F	4/754	0.005
$n_M = 0.029$	proportion of financial institutions of type M	22/754	0.029
$n_S = 0.950$	proportion of financial institutions of type S	716/754	0.950
$n_{GSE} = 0.016$	proportion of financial institutions of type GSE	12/754	0.016
$\lambda_F = 0.901$	bank-level share of unexpected payments per second for type F	0.901	0.901
$\lambda_M = 0.402$	bank-level share of unexpected payments per second for type M	0.402	0.402
$\lambda_S = 0.007$	bank-level share of unexpected payments per second for type S	0.007	0.007
$\lambda_{GSE} = 0$	bank-level share of unexpected payments per second for type GSE	0	0
$\iota_w = 0.0625/360$	Discount-Window rate (primary credit, 6.25% per annum)	0.0625/360	0.0625/360
$\iota_r = 0.0481/360$	effective fed funds rate (5.25% per annum)	0.0525/360	0.0525/360
$\beta_F = 0.401$	estimated liquidity effect around zero excess reserves (bps per \$bn)	$\in [1, 3]$	1.7
$\beta_M = 0.089$	participation rate of financial institutions of type M (i.e., $\mathcal{P}_M$)	0.53	0.54
$\beta_S = 0.033$	participation rate of financial institutions of type S (i.e., $\mathcal{P}_S$)	0.35	0.37
$\beta_{GSE} = 0.200$	participation rate of financial institutions of type GSE (i.e., $\mathcal{P}_{GSE}$)	0.14	0.13
$\kappa_F = 0$	reallocation index of financial institutions of type F (i.e., $\mathcal{R}_F$)	0.068	0.064
$\kappa_M = 0$	reallocation index of financial institutions of type M (i.e., $\mathcal{R}_M$)	−0.385	−0.268
$\kappa_S = 0$	reallocation index of financial institutions of type S (i.e., $\mathcal{R}_S$)	−0.268	−0.126
$\kappa_{GSE} = 1.25e\text{-}3$	reallocation index of financial institutions of type GSE (i.e., $\mathcal{P}_{GSE}$)	0.995	1

Table 3: Calibration for the year 2006.

Notes: Each non-shaded parameter is calibrated externally (i.e., to match a corresponding target moment, independently of the model and other parameters). Shaded parameters are calibrated internally (i.e., jointly, to match the set of shaded target moments, using the equilibrium conditions of the model, and given the values of the parameters calibrated externally). The calibration assumes a model period corresponding to approximately to 42 seconds in a trading day, $r = 0$, $\mathbb{N} = \{F, M, S, GSE\}$ (as discussed in Section 3.1), $\theta_i = 1/2$ for all $i \in \mathbb{N}$, $\{G_{ij}\}_{i,j\in\mathbb{N}}$ are estimated as described in Section 3.2, $\{F_0^i\}_{i\in\mathbb{N}}$ are estimated as described in Section 3.3, $u_i = 0$ for all $i \in \mathbb{N}$, $\{U_i\}_{i\in\mathbb{N}}$ are as in Section 4).

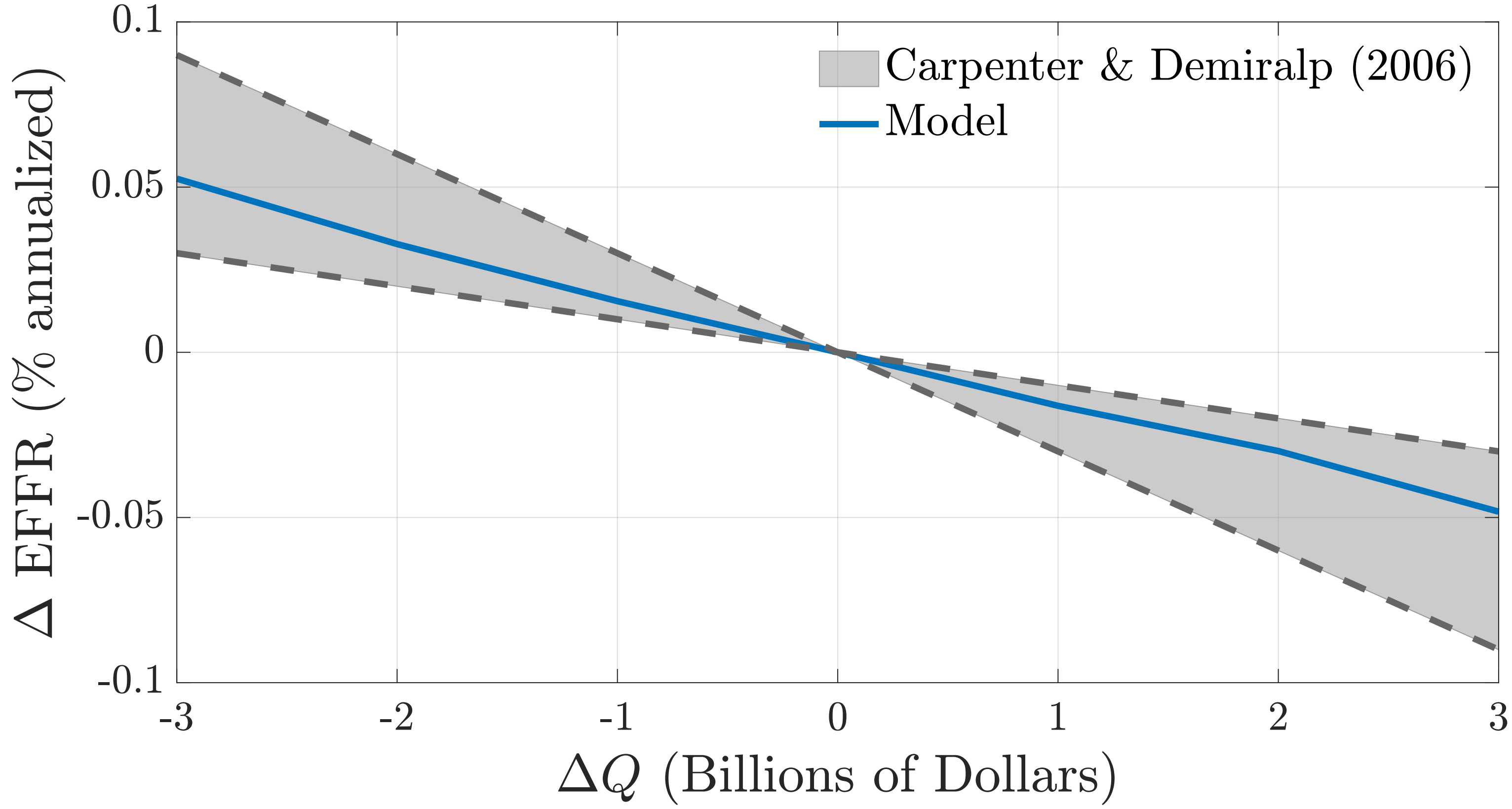

Figure 22: Liquidity effect: model and empirical estimates for the year 2006.

Notes: Rates in the vertical axis are in percent per annum. The shaded area represents the 95% confidence interval for the point estimate of the liquidity effect from Carpenter and Demiralp (2006). The solid line is the change in the equilibrium fed funds rate implied by the theory in response to changes in the total quantity of reserves (starting from the quantity of reserves corresponding to the 2006 calibration, and extracting reserves using the procedure described in Section 3.6.)

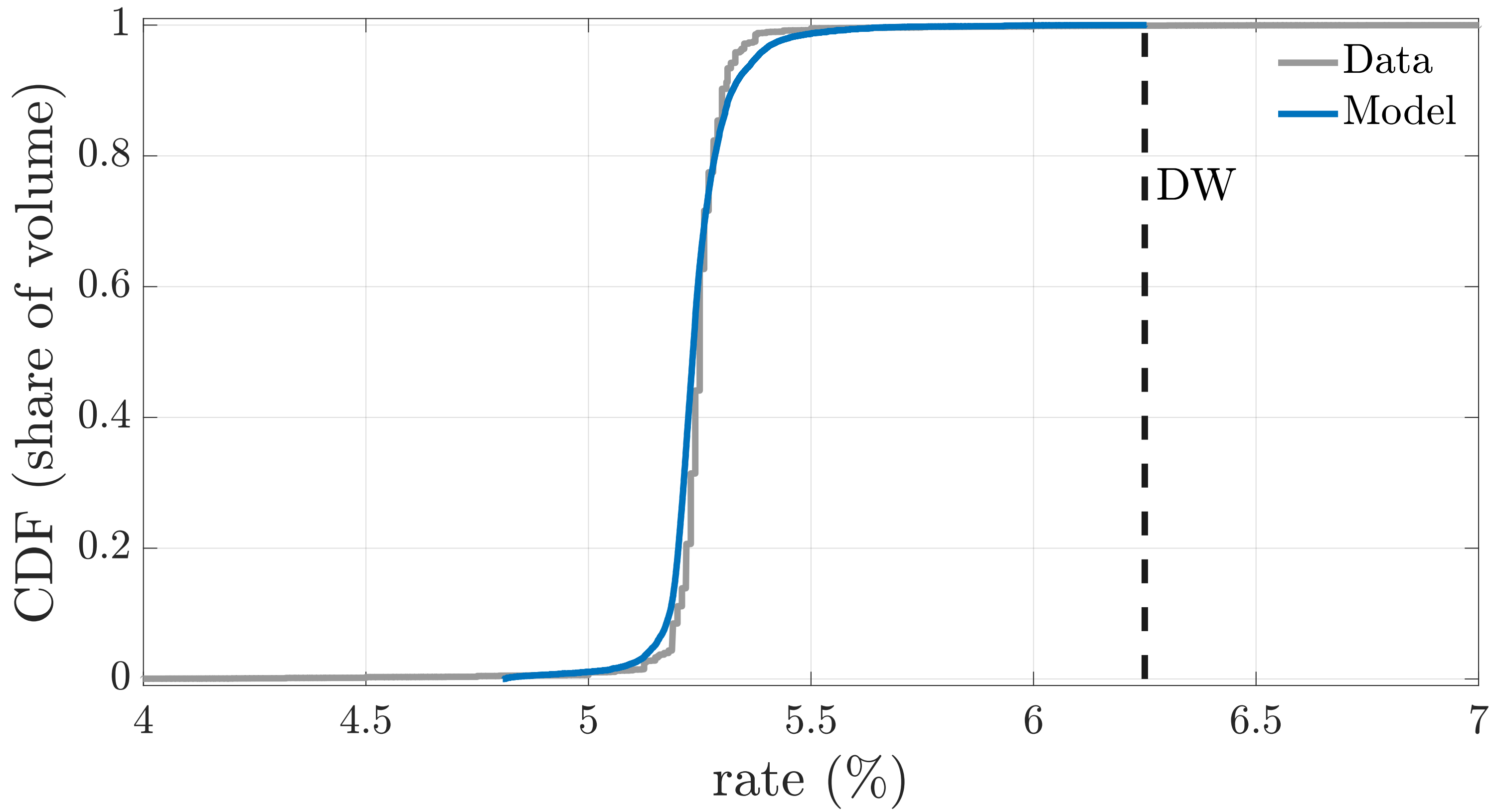

Figure 23: Empirical and theoretical cumulative distribution functions of bilateral fed funds rates for the year 2006.

Notes: Rates in the horizontal axis are in percent per annum.

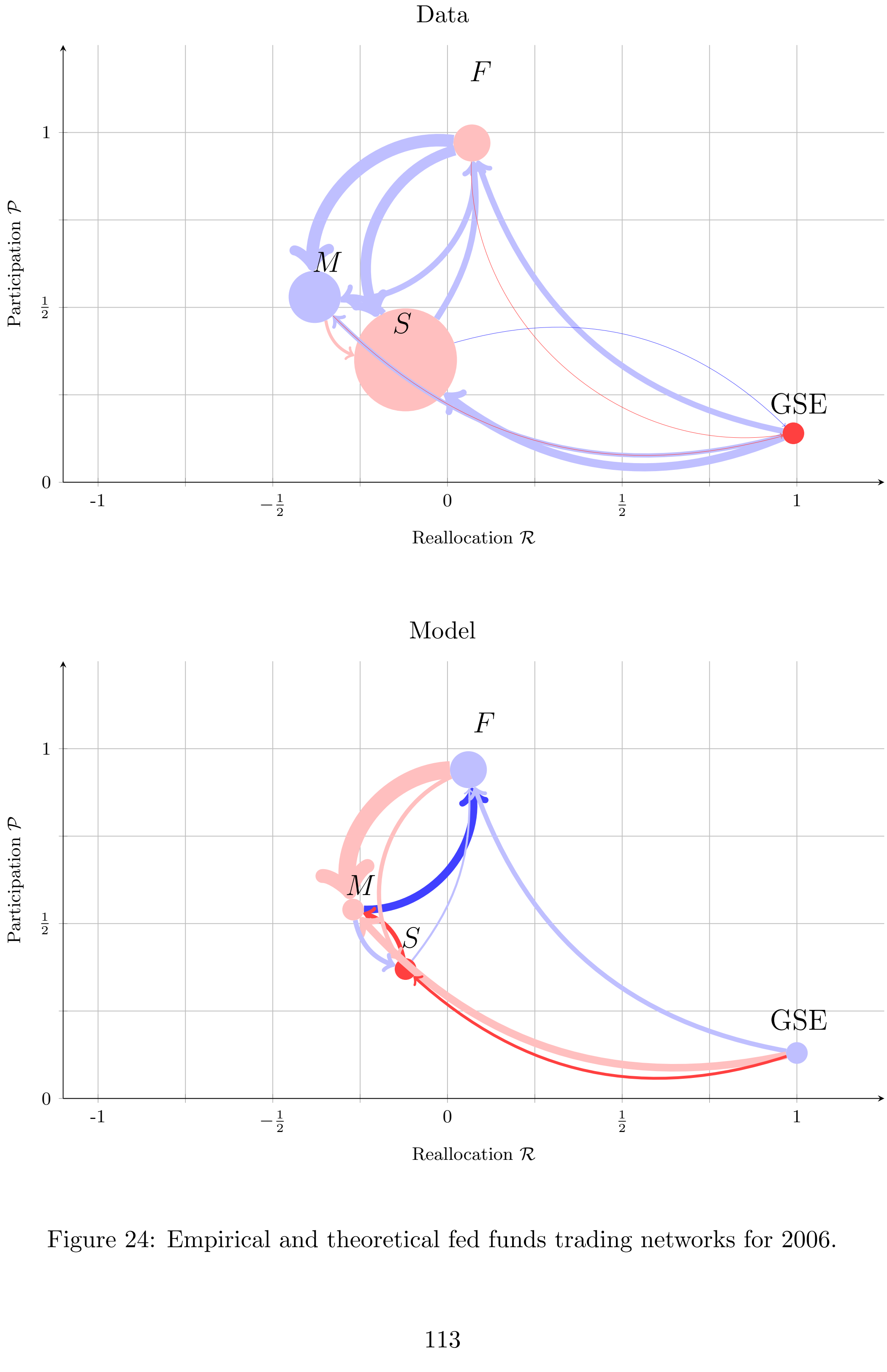

Figure 24: Empirical and theoretical fed funds trading networks for 2006.

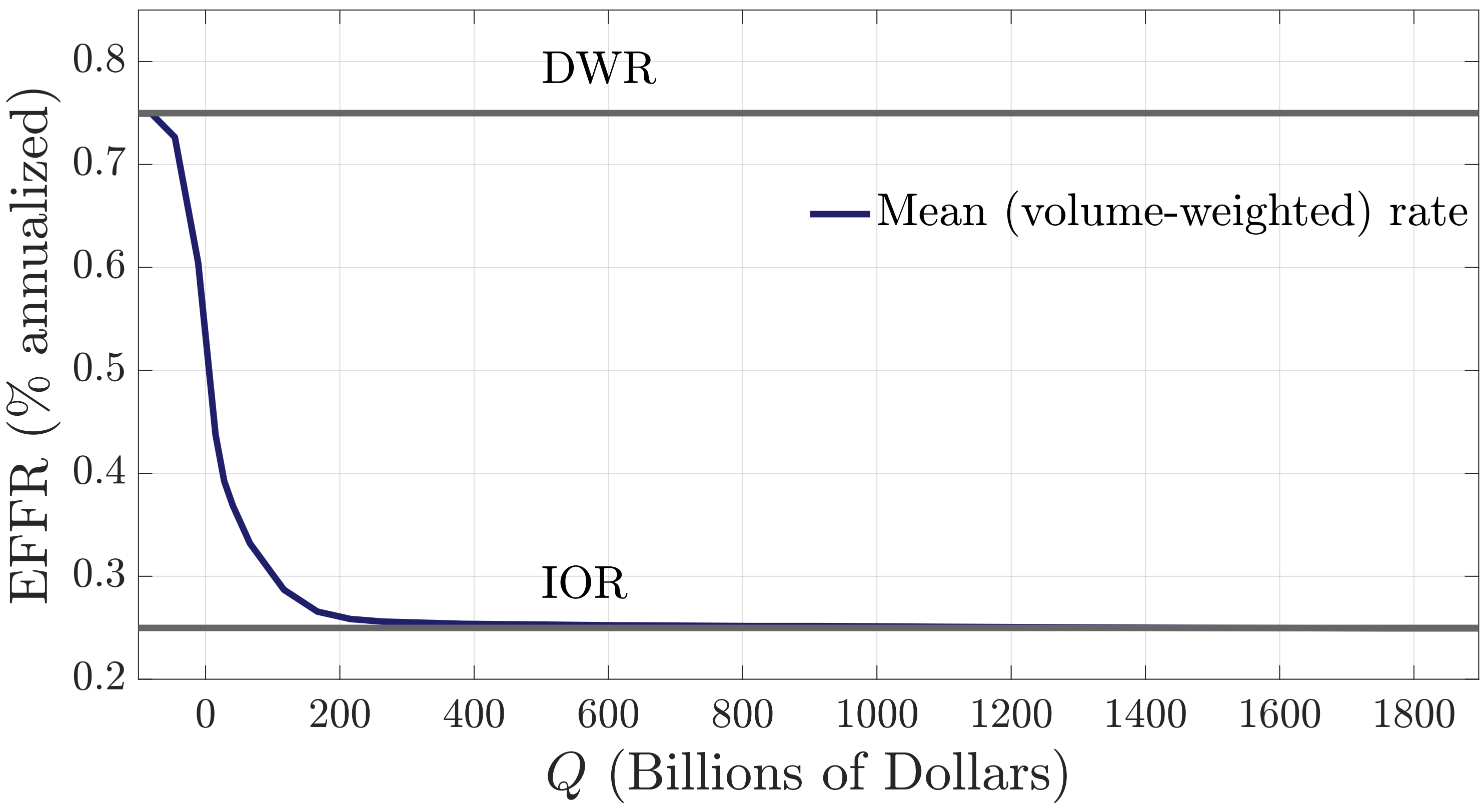

Figure 25: Theoretical aggregate demand for reserves for the year 2006 calibration.

Notes: Theoretical aggregate demand $\iota^*_{\Upsilon\omega} = \mathcal{D}(Q_{\Upsilon\omega};\Pi)$ for the model calibrated as in Table 3, and with $\iota^*_{\Upsilon\omega}$ and $Q_{\Upsilon\omega}$ computed with the interpolation procedure described in Section 3.6, for $\Upsilon_0 = 2006$ and $\Upsilon_1 = 2014$.