Ample Reserves for Whom?
The Role of Foreign Banks in U.S. Monetary Policy Implementation *

Junko Oguri† Cristoforo Pizzimenti‡

This draft: January 2026

Preliminary & Comments Welcome

Abstract

This paper studies how foreign banks affect reserve demand and balance sheet normalization under the Federal Reserve's ample-reserves framework. Although foreign banks account for a small share of U.S. banking activity, they hold a disproportionate share of reserve balances. Using regulatory data and high-frequency reserve supply shocks, we document pronounced heterogeneity in reserve management across banks and a strong QE–QT asymmetry: foreign banks absorb most reserve inflows during QE but do not symmetrically release reserves during QT, shifting adjustment toward large domestic banks and the ONRRP facility. We decompose foreign bank reserve demand into a spread-sensitive arbitrage buffer and a structural component anchored by internal capital markets and global interest rate differentials. Embedding these mechanisms in a heterogeneous-demand model, we show that uncertainty about foreign banks' reserve behavior raises the reserve buffer required for effective interest rate control during balance sheet normalization.

*We would like to thank our advisors Efraim Benmelech, Nicolas Crouzet, Martin Eichenbaum, Gregor Matvos, Andrew Metrick, and Matthew Rognlie for their guidance and support. We are grateful to Stefania D'Amico, Darrell Duffie, Daniel Feinberg, Craig Furfine, Enrique Martínez García, Zhengyang Jiang, Thomas King, Rosie Levy, Benjamin Munyan, Ralf Meisenzahl, Ali Ozdagli, Jonathan Rose, Chase Ross, Sharon Ross, Jung Sakong, Alessio Saretto, Sam Schulhofer-Wohl, Seth Searls, Nitzan Tzur-Ilan, Mark A. Wynne, Xiaoqing Zhou, and seminar participants at Kellogg Finance Brown Bag, Northwestern Macro Lunch, Dallas Fed Brown Bag, and Chicago Fed Brown Bag for helpful discussions and comments. Junko acknowledges the generous support from the Federal Reserve Bank of Dallas and the Federal Reserve Bank of Chicago. We gratefully acknowledge financial support from the 2025 ECB Lamfalussy Research Fellowship.

†Kellogg School of Management, Northwestern University. junko.oguri@kellogg.northwestern.edu

‡Northwestern University cristoforopizzimenti2025@u.northwestern.edu

"...[the normalization plan] is consciously intended to avoid creating market strains...and will just run quietly in the background... We think this is a workable plan, and it will...be like watching paint dry..." (Yellen, 2017)

"...the ultimate size of our balance sheet will be driven principally by financial institutions' demand for reserves.... Estimates of the level of reserve demand are quite uncertain, but we know that this demand in the post-crisis environment is far larger than before..."
(Powell, 2019)

1 Introduction

For the past two decades, a series of extraordinary events have fundamentally reshaped the conduct of monetary policy. Confronted with the effective lower bound on nominal interest rates, central banks turned to unconventional tools to stabilize markets and support economic activity. Most prominently, many deployed balance sheet policies, including large-scale asset purchases and targeted lending operations, financed by a corresponding expansion of bank reserves on the liability side of central banks' balance sheets. These measures expanded central bank balance sheets to unprecedented levels. For instance, in the U.S., the Federal Reserve's balance sheet expanded from about 6 percent of GDP in 2007 to over 33 percent by 2023 (Figure 1). Over the same period, reserve balances rose from \$5.6 billion in December 2007 to a peak of \$4.27 trillion in December 2021 (Figure 2).

As central banks unwind their balance sheets and proceed with quantitative tightening (QT), an active debate has emerged in both academic and policy circles: What is the optimal long-run size of a central bank balance sheet?[1] Outside of crisis periods, this debate typically reflects a shared preference for balance sheets that are large enough to ensure efficient policy implementation, yet not so large as to constrain future crisis interventions or generate distortions in financial markets, including Treasury markets. Indeed, the Federal Reserve has

[1]Arguments for a larger balance sheet often highlight benefits such as reduced reliance on private-sector maturity transformation (Greenwood et al., 2016) and enhanced liquidity management (Gagnon and Sack, 2014); see Logan (2019) for a policy perspective. In contrast, arguments for a smaller balance sheet emphasize the need to mitigate risks to central bank net worth and monetary policy independence; for example by Sims (2013) and by Hall and Reis (2013); see Bowman (2025) for a recent policy perspective. A complementary line of work examines composition rather than size (Eren et al., 2024), while Wallace (1981) argued for irrelevance of the composition.

repeatedly emphasized its intention to "hold no more securities than necessary to implement monetary policy efficiently and effectively."[2]

A less explored question is whether institutional constraints prevent central banks from reaching this optimal balance sheet size. While balance sheet normalization is often discussed in terms of asset holdings, we argue that the binding constraint lies on the liability side of the balance sheet—specifically, in banks' demand for reserve balances. In the U.S., because reserves are the instrument through which Federal Reserve implement monetary policy in the ample-reserves framework, the level of reserves that banks are willing to hold ultimately limits how far balance sheets can be reduced without losing control of short-term interest rates. As banks are heterogeneous in their reserve management, this limit depends not only on the aggregate quantity of reserves, but also on how reserves are distributed across institutions. In the U.S., a small set of institutions holds a disproportionate share of reserves and plays an outsized role in determining marginal reserve demand.

This paper focuses on the distribution of reserves across banks rather than their aggregate level. In particular, we study how foreign banks' reserve holdings constrain the Federal Reserve's ability to shrink its balance sheet without destabilizing policy rate control. Surprisingly, foreign banks hold a disproportionately large share of reserve balances at the Federal Reserve—a pattern that cannot be explained by their major banking activities in the U.S. banking system. Figure 3 shows the shares of total assets, deposits, loans, and reserves held by foreign and domestic banks from 1980 to 2024.[3] While foreign banks consistently account for only about 10 to 25 percent of assets, loans, and deposits, their share of reserves diverges sharply after the Global Financial Crisis in 2008. Since then, foreign banks' reserve share has at times approached or exceeded 50 percent, far above their footprint in other major balance sheet categories.

We argue that foreign banks' reserve demand plays a central role in shaping the aggregate reserve demand at the margin. Because foreign banks hold a large and significant share of

[2]See Federal Open Market Committee (2014), which states that in the longer run the Federal Reserve intention for security holdings. See also Board of Governors of the Federal Reserve System (2025) for a historical overview of the FOMC's policy normalization discussions and communications.

[3]As documented by Afonso et al. (2025), reserve balances at foreign banking organization (FBO) branches are, on average, approximately 20 percent lower at quarter-end than on other business days, reflecting balance sheet adjustments around regulatory reporting dates. For the corresponding reserve shares measured at quarter-end (without window dressing adjustment) see Figure A1. We interpret these estimates as providing a lower bound ("structural component") on foreign banks' reserve holdings.

reserve balances, their behavior affects how quickly aggregate reserve demand becomes steep as reserves are withdrawn. Uncertainty surrounding their reserve demand, stemming from regulatory asymmetry and global funding, therefore constrains the Federal Reserve's ability to shrink its balance sheet without risking a loss of policy rate control.

We proceed in three parts. First, we document systematic differences in reserve management between foreign and domestic banks and provide causal evidence of heterogeneous reserve demand. We show that the relationship between overnight funding spreads and aggregate reserve supply varies systematically across bank types and across QE and QT regimes. Exploiting high-frequency, plausibly exogenous variation in reserve supply arising from fluctuations in the Treasury General Account (TGA), we estimate the causal effects of reserve supply shocks on funding spreads and bank balance-sheet adjustment. These estimates identify foreign bank branches and agencies as the marginal arbitrageurs in overnight funding markets: they respond strongly to instrumented changes in the federal funds rate–interest on reserves (FF–IOR) spread, whereas domestic banks are comparatively insensitive. We then document a pronounced QE–QT asymmetry in the cross-sectional absorption of reserve supply shocks. During balance sheet expansions, foreign banks absorb the majority of reserve inflows (approximately 76 percent during the 2009–2014 QE period), whereas during QT they do not symmetrically shed reserves, shifting the burden of adjustment toward large domestic banks and the ONRRP facility.

Second, we examine the institutional drivers of these differences across bank types and policy regimes. We decompose foreign branch reserve holdings into (i) an arbitrage buffer component that is elastic to FF–IOR spreads and therefore expands during QE and compresses during QT, and (ii) a structural component anchored by global liquidity governance and net funding from foreign headquarters through internal capital markets. We use quarter-end balance-sheet compression ("window dressing") by foreign banks as an empirical proxy for estimating the arbitrage buffer component. This behavior is motivated by differences in regulatory reporting conventions between the U.S. and other major foreign jurisdictions. Using this proxy, we show that the arbitrage buffer responds to FF–IOR spreads whereas the structural component is largely insensitive to domestic arbitrage incentives. To characterize the determinants of the structural component, we estimate bank holding company (BHC)-level panel regressions relating reserve holdings to alternative funding sources, drawing on the liquidity-dependence framework of domestic banks by Acharya and Rajan (2022); Acharya et al. (2023). Net liabilities to foreign headquarters (dollar funding from FBO headquarter) robustly predict reserve holdings across specifications, with an elasticity of approximately 0.44 that is stable with various specifications. Other funding measures, including wholesale funding and U.S. based funding, have limited explanatory power. The relationship holds across QE and QT episodes. Furthermore, Net liabilities to foreign headquarters as well as reserve holdings responds to global interest rate differentials: when U.S. IOR exceeds policy rates in parent banks' home jurisdictions, foreign branches receive more internal funding and hold higher reserve balances. Taken together, these results imply that during QT foreign banks primarily compress the arbitrage buffer, while the structural component persists because it reflects global internal capital allocation rather than U.S. deposit dynamics. This mechanism accounts for the incomplete unwinding of foreign-branch reserve holdings during QT.

Third, we develop a two bank type model of reserve demand with heterogeneous banks to interpret our empirical findings and to quantify their implications for balance sheet policy. In the model, domestic and foreign banks differ in balance sheet costs, liquidity risk, and access to funding markets, generating distinct reserve demand schedules. Aggregating across banks yields a nonlinear aggregate reserve demand curve, with a kink that emerges endogenously when the identity of the marginal holder of reserves shifts. On the supply side, we model a central bank that reduces reserves under quantitative tightening while seeking to maintain control of short-term interest rates. In a deterministic setting, the central bank can reduce reserves to the point where the aggregate demand curve becomes steep. When we introduce uncertainty in foreign banks' reserve demand, the optimal policy features a precautionary buffer of reserves above this deterministic threshold. We calibrate the model to the 2018–2019 QT episode. The calibration matches the cross-sectional distribution of reserves and implies that foreign banks increasingly determine marginal reserve demand as aggregate reserves decline. The resulting ample-reserves threshold closely aligns with the onset of money market stress in September 2019. Accounting for uncertainty in foreign banks' behavior raises the optimal reserve supply by a meaningful margin, quantifying the precautionary buffer required for stable policy implementation under QT.

Taken together, our findings show that institutional heterogeneity, and in particular the behavior of foreign banks, plays a central role in shaping aggregate reserve demand and the Federal Reserve's ability to maintain interest rate control during balance sheet normalization under the ample-reserves framework. Whereas Acharya and Rajan (2022); Acharya et al.

(2023) document that domestic banks' liquidity claims persist as reserves decline, generating liquidity dependence on the central bank, we show that foreign banks also contribute to an increase in the minimum level of reserves required for smooth policy implementation. At times, foreign banks hold a majority of reserve balances at the Federal Reserve, giving their reserve demand a first-order influence on the effective lower bound of reserve supply.

Foreign banks face lower balance sheet costs, which allow them to arbitrage the FF–IOR spread during QE and absorb a substantial share of aggregate reserve expansions. Beyond this spread-based activity, however, foreign banks exhibit a persistent structural demand for reserves that is anchored in internal capital markets within global banking organizations. As a result, during QT foreign banks are not necessarily the institutions that release accumulated reserves. Instead, the burden of balance sheet adjustment falls disproportionately on large domestic banks that are more closely integrated into domestic financial intermediation.

These features help explain why reserve holdings unwind asymmetrically during QT and why the minimum level of reserves consistent with stable policy implementation may increase. Moreover, because foreign banks' persistent reserve demand is not primarily tied to U.S. domestic deposit fragility, but instead reflects global factors such as foreign interest rates, regulatory environments, and internal liquidity management, it introduces additional uncertainty into the Federal Reserve's assessment of aggregate reserve demand. As a result, the level of reserves required to maintain effective interest rate control may remain elevated and uncertain even as domestic arbitrage incentives compress.

The rest of the paper is structured as follows. Section 2 reviews the related literature. Section 3 provides an overview of the Federal Reserve's balance sheet and the evolution of reserve demand. Section 4 presents the data sources and sample construction. Section 5 documents reserve demand heterogeneity using TGA shock identification and establishes the QE–QT asymmetry. Section 6 examines the institutional drivers of heterogeneous demand, introducing the decomposition into structural component and arbitrage buffer, and presenting evidence on headquarters funding through internal capital markets. Section 7 develops the model with heterogeneous demand and presents calibration results. Section 8 concludes.

2 Related Literature

This paper contributes to several strands of the macro-finance literature on reserve demand, global banking, and unconventional monetary policy.

First, the paper is most closely related to the rapidly growing literature on the liability side of central bank balance sheets and the challenges associated with estimating the demand for reserves. Lopez-Salido and Vissing-Jorgensen (2023) use deposit aggregates to identify the point at which reserve demand begins to slope downward. Acharya and Rajan (2022); Acharya et al. (2023) argue that reserve abundance paradoxically increases systemic liquidity risk and explore the "ratchet-up" effect. A broader set of studies examines monetary policy implementation under an ample-reserves regime and the functioning of the federal funds market, including Afonso et al. (2022a,b,c); Bech and Klee (2011); Bianchi and Bigio (2022); Bigio and Sannikov (2021); Darst et al. (2025); Dubois and Rintamäki (2025); Kashyap and Stein (2012); Keating and Macchiavelli (2017); Lagos and Navarro (2023); Smith and Valcarcel (2023). In particular, closely related to our work, Afonso et al. (2023) develop a theoretical framework to analyze the optimal supply of central bank reserves when the central bank faces uncertainty about banks' demand for reserves. Anbil et al. (2025) emphasizes heterogeneity in reserve demand across banks and studies its implications for federal funds market functioning. Our contribution differs from and complements this literature in several dimensions. While existing work primarily focuses on aggregate reserve demand, we emphasize systematic heterogeneity across banks. By decomposing reserve demand by bank type, we show that demand curves and elasticities differ markedly across bank groups and evolve over time. In particular, we document that the reserve demand of foreign banks is both more elastic and more volatile, introducing a source of uncertainty in aggregate reserve demand.

Second, the paper contributes to the literature on global banking and cross-border dollar intermediation. Bräuning and Ivashina (2020) show that global banks respond to cross-country policy rate differentials by reallocating funds across jurisdictions, often at the expense of lending in the home economy. Correa et al. (2020) document how U.S. G-SIBs and foreign banking organizations engage in reserve-draining intermediation in response to dollar funding pressures, with these behaviors shaped by differences in regulatory leverage ratio reporting frameworks. Kim (2025) sets up two-country model and show foreign banks active reserve holdings and interbank market activity leads to reducing welfare for domestic and foreign consumers. A broader literature studies global banking, cross-border capital flows, and arbitrage in dollar funding markets, including Aldasoro et al. (2022); Anderson et al. (2021); Berrospide et al. (2016); Cetorelli and Goldberg (2011); Ivashina et al. (2015); Kalemli-Özcan (2019); Morelli et al. (2022); Niepmann (2023); Niepmann and Shen (2024); Rime et al. (2022). While much of this literature focuses on the lending and funding activities of global banks, we contribute by shifting attention to the reserve holdings of foreign banks operating in the U.S. and by analyzing how these holdings interact with the Federal Reserve's balance sheet policies. In addition, we examine how differences in organizational form—specifically, branches and agencies versus subsidiaries—shape reserve management behavior and participation in U.S. dollar funding markets.

Lastly, the paper contributes to the literature on unconventional monetary policy, with a particular focus on the design and transmission of quantitative tightening (QT). Diamond et al. (2023) use structural estimation to show that an ample-reserves environment can crowd out private lending. Copeland et al. (2021); Yang (2020) emphasizes the role of reserves in supporting payment systems during QT, showing that low or unevenly distributed reserves can lead to payment delays, liquidity hoarding, and stress in short-term funding markets. Borio (2023) argue that the costs associated with floor systems and abundant reserves are often underestimated. Other related contributions include D'Amico and Feldman (2024); d'Avernas et al. (2025); Eisenschmidt et al. (2024); Goldstein et al. (2023); Ihrig and Wolla (2020); D'Amico and Seida (2024); Carlson et al. (2020); Vandeweyer et al. (2024). Our study adds to this literature by emphasizing the heterogeneous costs and incentives banks face when holding reserves, particularly for foreign institutions operating in the U.S. We show that such heterogeneity can lead to uneven responses to balance sheet normalization and may generate uncertainty during the QT process. More broadly, because the Federal Reserve does not coordinate policy rates or balance sheet policies with other central banks, divergence in global monetary conditions may further complicate the implementation and transmission of U.S. monetary policy.

3 Brief overview of Fed's monetary policy implementation scheme and reserves market

The surge in reserve supply since the Global Financial Crisis led to a fundamental shift in the Fed's monetary policy framework.[4] The Fed introduced interest on reserves (IOR) in October 2008, marking the transition to a floor system.[5] Under this system, IOR was intended to serve as the effective lower bound for the FFR, as banks would have no incentive to lend at rates below the risk-free IOR. The simpler mechanics of this framework, compared to the scarce reserves system, were expected to provide simpler control over short-term rates (Mester, 2024). By 2019, the Fed formally committed to maintaining an ample reserves regime. In March 2020, the Fed eliminated reserve requirements, setting them to zero.

However, despite IOR was intended to act as a floor, the FFR has persistently traded below IOR since its introduction (Figure A4). This deviation stems largely from Government-Sponsored Enterprises (GSEs), which may hold reserves at the Fed but are ineligible to earn interest on reserves. The Federal Home Loan Banks (FHLBs), in particular, face liquidity requirements that encourage them to lend excess balances below IOR, effectively pushing down the FFR. This structural feature of the market has persistently weakened the transmission of IOR to short-term rates, placing FFR below IOR, creating challenges for the Fed's monetary policy implementation.[6]

Responding to the leaky floor, the Fed has implemented several measures to complement the floor system. First, in 2013, the Fed launched the Overnight Reverse Repurchase Agreement (ONRRP) facility, allowing non-bank entities—such as GSEs and money market

[4]Before 2008, the Fed operated under a scarce reserves system, adjusting reserve supply through open market operations to steer the federal funds rate (FFR). Banks, constrained by reserve requirements and liquidity needs, actively managed their balances by borrowing and lending in the federal funds market and using intraday credit at the Fed. Small changes in reserve supply had large effects on interbank rates, placing the supply curve at the slope part of the reserve demand curve (Figure A2). This system collapsed in late 2008 as reserves flooded the banking system, rendering reserve requirements non-binding—fluctuations in reserve supply no longer moved the FFR, as liquidity had become abundant and supply now intersects with the flat part of the demand curve. For further details, see Ihrig and Wolla (2020).

[5]In 2006, Congress passed a series of regulatory relief measures for banks and other financial institutions. This package, known as the Financial Services Regulatory Relief Act of 2006, included the authorization for the Fed to pay interest on bank reserves. Although this was intended to take effect in 2011, the Fed's response to the GFC expedited the implementation of the IOR. The idea of paying interest on reserves holding had been proposed by economists including Friedman (1959) and the implementation concept was further developed by Goodfriend (2002). For further historical background, see Wall (2017); Borio (2023).

[6]For further details on the intraday federal funds dynamics, see Appendix A2. and Bech and Klee (2011).

mutual funds (MMFs)—to lend overnight to the Fed in exchange for Treasury securities. By providing an alternative investment option with wider counterparties, the ONRRP helped establish a firmer lower bound on short-term rates, mitigating the persistent gap between the FFR and IOR (Figure A3). However, this facility alone was not sufficient to prevent rate spikes in times of stress.

In September 2019, a series of factors triggered severe dislocations in the money market, exposing fragilities in the ample reserves framework (Anbil et al., 2020; Kahn et al., 2023). A corporate tax payment deadline drained liquidity from the banking system, while an unusually large settlement of Treasury securities from prior auctions forced primary dealers and investors to secure short-term financing. At the same time, reserves had been steadily declining due to the Fed's balance sheet normalization, leaving banks less willing to lend in overnight markets. These pressures culminated in a sharp and unexpected spike in repo rates, with overnight borrowing costs surging well above the Fed's target range. The episode underscored the limitations of the floor system in an environment where reserves, though still ample in aggregate, were unevenly distributed across institutions.[7]

Uncertainty about reserve demand, together with episodes of money market stress, has become a central consideration in determining the extent of the Federal Reserve's balance sheet normalization. While various tools have helped limit volatility in the FFR, fragilities remain, particularly in the repo market around financial reporting dates like quarter-ends. In September 2024, for instance, the Secured Overnight Financing Rate (SOFR) spiked 21 basis points above its weekly average, reflecting financing pressures. The ongoing QT has further intensified these issues, as private investors absorb a growing supply of Treasuries, often requiring repo financing, while the reduction in Fed-supplied liquidity tightens overall market conditions. Structural constraints—including regulatory requirements, counterparty credit limits, and balance sheet costs—continue to impede the smooth redistribution of liquidity, exacerbating rate volatility.[8] These dynamics may suggest the limits of relying solely on

[7]In response, the Fed introduced the Standing Repo Facility (SRF) for primary dealers in 2021, designed to stabilize short-term funding markets by offering eligible banks and primary dealers access to secured borrowing against high-quality collateral like Treasury securities. The SRF is intended to act as an automatic stabilizer, ensuring that liquidity shortages do not lead to disorderly increases in borrowing costs. By creating a mechanism through which the Fed can directly inject liquidity into the repo market, the SRF was intended to improve rate control and reduce the likelihood of extreme funding pressures disrupting monetary policy implementation.

[8]For further details, see Fed official speeches such as Perli (2024a,b).

administered rates to control short-term markets, highlighting the interrelationships between reserves, policy tools, and the evolving structure of funding markets.

The appropriate level of reserves for smooth policy implementation remains uncertain, as the Fed struggles to accurately gauge the minimum threshold necessary to prevent market disruptions. In October 2024, the New York Fed introduced the Reserve Demand Elasticity (RDE) indicator to monitor shifts in reserve demand and assess the sensitivity of short-term interest rates to fluctuations in reserves. Constructed using forecast errors from a daily time-varying VAR model of aggregate reserves and federal funds rates, this tool provides a system-wide measure of liquidity conditions (Afonso et al., 2022c). However, it does not account for differences in reserve demand across institutions. The model assumes a single reserve demand curve, implicitly treating all banks as homogeneous. In reality, reserve demand varies significantly with factors such as bank size, regulatory constraints, access to alternative funding markets, and institutional structure. In particular, foreign banks would respond to a different set of incentives, including global funding conditions and home-country regulations. By aggregating reserve demand into a single elasticity measure, the RDE risks overlooking key frictions in liquidity distribution, limiting its effectiveness in assessing the true availability of reserves in the banking system.

This measurement challenge is further complicated by the evolving nature of reserve demand, which has forced the Fed to repeatedly revise its estimates of the lowest comfortable level of reserves (LCLOR). For instance, over the past five years, these estimates have steadily increased—from $1.5 trillion in October 2019 to approximately $3 trillion in April 2024 (Nelson, 2024a). Rather than being a fixed threshold, the LCLOR is an evolving concept, reflecting the dynamic nature of reserve demand across different types of institutions. This uncertainty complicates the Fed's ability to forecast the effects of balance sheet normalization. Unlike QE, which injected reserves into a system already operating under clear constraints, QT unfolds in an environment where reserve demand is shifting in unpredictable ways. As the Fed continues to reduce its balance sheet, the interaction between reserve supply, market structure, and institutional heterogeneity presents challenges for both policy transmission and financial stability. To better understand reserve demand, it is important to consider cross-bank heterogeneity rather than relying solely on aggregate measures.

Aggregate reserve levels alone do not fully determine liquidity conditions in the financial system. The distribution of reserves across banks plays a central role in shaping market functioning and interest rate dynamics. We argue that foreign banks play a significant role in the heterogeneity of reserve distribution, introducing an additional layer of uncertainty to monetary policy implementation. Foreign banks operating in the US are not directly regulated or supervised by US authorities in the same way as domestic banks. The majority of foreign banks' reserves are held through branches and agencies (hereafter branches), rather than subsidiaries, and these branches have greater flexibilities in their liquidity allocation. As a result, the Federal Reserve has limited control over how foreign banks adjust their reserve holdings.

Beyond the financial stability concerns, the presence of foreign banks in the reserve market could raise political and fiscal considerations. interest on reserves (IOR) is paid to banks holding reserve balances at the Fed, including foreign institutions.[9] Let us consider a case when foreign banks collectively hold around \$1.3 trillion in reserves, the level as of January 2025, and the IOR remains at 4.4%. Assuming bi-weekly compounding, foreign banks would earn approximately \$58.4 billion in interest payments by year-end for parking their deposits at the Fed.[10] Since the Fed's net income affects its remittances to the Treasury, these interest payments could raise concerns over the fiscal cost of monetary policy, particularly when reserve balances remain elevated. This dynamic further underscores why foreign banks' reserve behavior is not just a financial stability issue but also a broader policy concern.

In the following sections, we present a series of empirical facts that characterize the reserve behavior of foreign banks, documenting how regulatory structure and global internal capital markets shape their reserve holdings.

4 Data

This section describes the data sources used throughout the paper. We draw on two main types of data: (i) quarterly bank balance sheet data from regulatory filings, which provide detailed cross-sectional and time-series information on reserve holdings and bank characteristics; and (ii) high-frequency data on reserve supply shocks and money market conditions, which enable causal identification of reserve demand elasticities. Appendix A1. provides further details on

[9]For instance, in July 2025, Senator Ted Cruz introduced legislation to eliminate the Federal Reserve's authority to pay interest on bank reserves, citing that foreign banks receive roughly 40–50 percent of aggregate IOR payments (Cruz, 2025).

[10]Interest payments occur after each maintenance period, typically every 14 days.

data collection and cleaning procedures.

4.1 Quarterly Bank Balance Sheet Data

The primary dependent variable in this analysis is reserve holdings, which are exclusively held by depository institutions with a master account at the Federal Reserve.[11] Reserve data is available at the bank level in FFIEC 031, 041, and 002 reports. We include all banks that report positive reserve balances and total assets. To measure reserves at the holding company (BHC) level, we aggregate individual bank reserves using FR Y-9C data and attributes information on FFIEC website. We also link each institution to the country of its headquarters, identifying foreign banking organizations (FBOs) for non-U.S. banks.

To supplement balance sheet data, we merge regulatory filings, market data, and policy rate series. We collect data from Bloomberg, foreign central banks and supervisory authorities, and financial stability board websites, deriving quarterly averages to align with the Call Report data. In analyses involving deposit facility spreads across jurisdictions, we focus on foreign banks from the Euro area, Japan, the United Kingdom (UK), Canada, and Switzerland, as these institutions collectively account for over 60% of the reserve holdings of all foreign banks in our core analysis periods from 2009 to 2025 (Figure 4). We also collect deposit facility rates for each currency area from respective central banks. Figure 7 shows the spread between the U.S. IOR and foreign deposit facility rates, which exhibited substantial variation during our sample period.

Our sample spans 1980Q2 to 2025Q3, covering the longest available time horizon including foreign branches and agencies data. However, most analyses in the paper focus on selected sub-periods relevant to regulatory changes and shifts in reserve demand behavior.

Tables 1 and 2 provide summary statistics for key variables, while Figures A5 break down the balance sheet composition of domestic banks, foreign subsidiaries, and foreign branches. Domestic banks and foreign subsidiaries share similar balance sheet composition, both relying on deposit funding and maintaining sizable capital buffers. Meanwhile, foreign branches operate differently. They depend more on wholesale funding, hold a larger share of

[11] These institutions include nationally chartered banks, eligible state-chartered banks, and other entities classified under Section 19(b) of the Federal Reserve Act. As of March 2020, the Federal Reserve reduced reserve requirement ratios to zero percent, effectively eliminating reserve mandates for depository institutions. While banks with master accounts continue to hold reserves, these holdings are now discretionary rather than regulatory obligations.

reserves, and engage extensively in cross-border liquidity transfers with their headquarters (HQs). Figure 5 shows how foreign bank presence in the U.S. varies by jurisdiction, with heterogeneity in organizational structure. For instance, Canadian and UK banks primarily operate through subsidiaries. In contrast, Japanese banks lean heavily on branches. Figure 6 shows the reserve holdings allocation across major jurisdictions.

4.2 High-Frequency Data: Weekly Balance Sheets and Daily Money Market Conditions

4.2.1 Weekly Bank Balance Sheet Data (H.8)

In order to complement the low frequency of quarterly data, we also use weekly balance sheet data from the Federal Reserve's H.8 Report: Assets and Liabilities of Commercial Banks in the United States. This dataset provides a higher-frequency view of reserve dynamics than the quarterly reports, enabling us to trace behavior around key dates such as quarter-ends and in response to reserve supply shocks.

Because reserves are not separately reported in the H.8 data, we use cash assets as a proxy, a measure that closely tracks reserve balances.[12] The H.8 data distinguish between (i) large domestic banks (top 25 by asset size), (ii) smaller domestic banks, and (iii) U.S. branches and agencies of foreign banks, allowing us to trace heterogeneous responses across bank types. Figure 8 plots cash holdings from 2008 to 2025 for the three groups. Panel A shows aggregate cash levels, Panel B the share of total system cash, and Panel C the cash-to-assets ratio. Vertical lines indicate calendar quarter-ends. Foreign branches consistently hold a disproportionately large share of system-wide cash—typically 35–45 percent since 2015, rivaling the holdings of the largest domestic banks. Their share temporarily dipped during the COVID-19 period when domestic banks accumulated precautionary liquidity, but by 2024 it had returned to pre-pandemic levels. Branches and agencies of foreign banks exhibit significantly higher reserve-to-asset ratio, compared to Domestic large bank and small banks. As shown in panel C, foreign banks often far exceeed in terms of ratio, sometimes marking 50% reserve asset ratio or so. Foreign branches exhibit pronounced and recurrent drops in cash holdings at quarter-ends—a pattern absent among domestic banks. These declines,

[12] See Appendix A1.2 and Appendix Figure A7 for verification that cash assets are an accurate proxy for reserves and for additional data construction details.

known as “window dressing”, are particularly visible when the reporting Wednesday coincides with quarter-end dates.

4.2.2 Other Money Market and Treasury Data

We supplement bank balance sheet data with standard daily money market interest rates and Treasury cash management data. Policy and market rates—including the interest on reserves (IOR), the effective federal funds rate (EFFR), the federal funds target range, and the Secured Overnight Financing Rate (SOFR)—are obtained from the Federal Reserve Bank of St. Louis (FRED) and Bloomberg. These series are used to construct measures of domestic arbitrage incentives faced by depository institutions.

To measure exogenous variation in reserve supply, we use data on the U.S. Treasury General Account (TGA) at the Federal Reserve. Daily TGA balances are obtained from the U.S. Treasury's Daily Treasury Statement. We also collect data on Treasury securities issuance, including auction volumes and maturities, from the Treasury website, which are used to refine the measurement of TGA-driven reserve supply shocks.

The TGA serves as the primary cash management account of the U.S. Treasury. Changes in the TGA are an autonomous factor on the Federal Reserve's balance sheet: dollar-for-dollar increases in the TGA drain reserves from the banking system, while decreases inject reserves, independent of banks' balance sheet decisions. Although Treasury cash balances have always affected reserve supply, their quantitative importance increased substantially after the Global Financial Crisis. Prior to 2008, the Treasury managed most of its cash through the Treasury Tax and Loan (TT&L) program, which placed funds with private depository institutions. Following the introduction of interest on reserves in late 2008, this arrangement became less efficient, and the Treasury gradually shifted cash management to the TGA, effectively ending the TT&L program by 2012 (Santoro, 2012). As a result, post-crisis fluctuations in the TGA are large and frequent, making the TGA the dominant autonomous driver of reserve supply over our sample period.

5 Reserve Demand Heterogeneity

This section documents heterogeneity in banks' reserve demand and examines its implications for overnight funding markets. We first document a robust relationship between overnight funding spreads and aggregate reserve supply, and show that the strength of this relationship varies across bank types and across QE and QT phases. We then use high-frequency, plausibly exogenous variation in reserve supply arising from Treasury General Account (TGA) shocks to identify causal effects on funding spreads and bank balance-sheet adjustment. The results characterize which institutions operate at the margin of reserve demand and document a pronounced QE–QT asymmetry in the cross-sectional absorption of reserve fluctuations.[13]

5.1 Funding Spreads and Aggregate Reserve Supply

Figure 9 plots the spread between the federal funds rate and the interest on reserves (FF–IOR) against aggregate reserve balances held at the Federal Reserve at a weekly frequency. Observations are grouped into four monetary policy regimes: QE (January 2009–October 2014), QT (November 2014–October 2019), QE (November 2019–November 2021), and QT (December 2021–September 2025). Across regimes, the figure reveals a broadly downward-sloping relationship between reserves and funding spreads: as aggregate reserve balances decline, overnight funding conditions tighten.

This pattern is consistent with banks responding to arbitrage incentives created by the persistent gap between the federal funds rate (FFR) and the IOR. Since the introduction of IOR, the FFR has typically traded below IOR (Figure A4), creating an opportunity for banks to borrow in the federal funds market or other money markets and deposit the proceeds at the Federal Reserve. Interpreted through this lens, Figure 9 suggests that when the FFR–IOR spread widens (i.e., becomes more negative), banks optimally expand reserve holdings, while tighter spreads are associated with lower reserve demand.

For observations after November 2019, we also plot the SOFR–IOR spread (dashed lines), reflecting the growing importance of repo markets in overnight intermediation as well as due to the disappearance of FF–IOR spread variations. This highlights the increasing role of secured funding markets (in part driven by the expansion of sponsored repo arrangements) and broader access to the Federal Reserve's standing facilities.

While the aggregate relationship between funding spreads and reserves is strong, it

[13]Throughout, we use "arbitrage" as shorthand for spread-based balance-sheet adjustments that exploit the FF–IOR (and, in the post-2019 period, SOFR–IOR) differential. In practice, these adjustments are subject to balance-sheet costs, internal risk limits, and regulatory constraints, so observed responses need not be symmetric across regimes or shocks.

masks substantial heterogeneity across bank types and monetary policy regimes. Figure 10 disaggregates the spread–reserve relationship by bank type—foreign banks branches and agencies, large domestic banks (top 25 by assets), and small domestic banks—and by QE versus QT phases. Reserve holdings are proxied by cash balances and measured at a weekly frequency.

The responses differ markedly across institutions. Small domestic banks exhibit little sensitivity of reserve holdings to either funding spreads or changes in aggregate reserve supply. By contrast, large domestic banks and foreign branches display more responsiveness, with slopes that vary between QE and QT episodes. These institutions actively adjust reserve holdings as spread incentives evolve, but the strength of the response is regime-dependent.

5.2 Identifying Exogenous Reserve Supply Shocks Using the TGA shock

The correlations documented above are suggestive but do not establish causality. Funding spreads and reserve holdings might be jointly determined and may both respond to unobserved shocks to money markets or bank balance sheets. To isolate exogenous variation in reserve supply, we exploit fluctuations in the Treasury General Account (TGA), an autonomous factor on the Federal Reserve's balance sheet that generates high-frequency, plausibly exogenous movements in aggregate reserves.

Our identification strategy builds on the mechanical relationship between TGA movements and reserve supply in the post-2008 ample-reserves implementation framework. Increases in the TGA drain reserves from the banking system, while decreases in the TGA inject reserves, unless offset by changes in overnight reverse repo (ONRRP) usage, which we explicitly control for in the empirical specifications. Unlike in earlier reserve-targeting regimes, the Federal Reserve does not systematically sterilize these day-to-day TGA fluctuations, allowing high-frequency Treasury cash-management flows to pass through directly to aggregate reserve balances.

To isolate the unanticipated component of reserve supply variation, we construct TGA shocks following Bräuning (2017). Specifically, we regress daily changes in the TGA balance on beginning-of-month, end-of-month, beginning-of-quarter, end-of-quarter, and day-of-week fixed effects to remove predictable calendar-driven variation in Treasury cash management. We further control for anticipated financing flows using information on expected Treasury security issuance and maturities. The residual from this regression, shown in Figure 11, is interpreted as an unanticipated TGA shock. Our sample spans January 1, 2009, to September 30, 2025.

TGA movements have been used extensively in the prior literature (e.g., Afonso et al., 2022c; Bräuning, 2017; Correa et al., 2020; Hamilton, 1997) and the shock satisfies the standard instrumental-variable conditions of relevance and exogeneity in this setting. Relevance follows from the fact that TGA movements mechanically inject or drain reserves and therefore affect overnight funding conditions: a reserve drain tightens money markets, raising unsecured and secured borrowing rates relative to the interest on reserves. We verify this relationship empirically in the first-stage regressions. Plausible exogeneity follows from the institutional determinants of Treasury cash management. High-frequency TGA movements are driven by fiscal operations, such as tax receipts, benefit payments, and the settlement of Treasury securities, rather than by contemporaneous conditions in interbank or repo markets. Banks do not transact directly with the TGA; instead, they adjust borrowing and reserve positions in response to the funding market conditions induced by TGA-driven reserve fluctuations. The primary threat to the exclusion restriction is that issuance-driven TGA changes may directly affect repo rates through collateral supply.[14] We address this concern by removing anticipated issuance-related variation from TGA changes.

Using daily data, we estimate the first-stage relationship:

$$\Delta\left(i^F - i^R\right)_t = \beta_0 + \beta_1\,\Delta\text{TGA}_t + \beta_2\,\Delta\text{ONRRP}_t + \beta_3\,X_{f(t)} + \varepsilon_t, \tag{1}$$

where $\Delta(i^F - i^R)_t$ is the daily change in the funding spread (FF–IOR or SOFR–IOR), ΔTGA_t is the TGA shock, ΔONRRP_t controls for changes in ONRRP take-up, and $X_{f(t)}$ denotes FOMC-period fixed effects.

Panel A in Table 3 reports the first-stage estimates. A \$100 billion increase in the TGA balance (a reserve drain) raises the FF–IOR spread by roughly 35 basis points. The instrument remains relevant across policy regimes, with stronger effects during QE and weaker but still detectable effects during QT, reflecting reduced variation in spreads in the

[14]Recent policy discussions have emphasized that, in principle, the Federal Reserve could actively offset fluctuations in the Treasury General Account, for example, by adjusting its asset holdings or backing the TGA with short-maturity securities, to insulate reserve balances and money market conditions from Treasury cash-management operations. See Nelson (2024b), Vissing-Jorgensen (2025), and Miran (2025). However, under the current ample-reserves operating framework, such sterilization is not implemented on a routine basis, and predictable Treasury settlements continue to generate high-frequency variation in reserve supply.

most recent tightening episode. Using SOFR–IOR spreads, as shown in Panel B, yields similar conclusions, confirming that TGA shocks affect secured as well as unsecured funding markets.[15]

Annex Tables A1 and A2 decompose TGA shocks into reserve-draining and reserve-injecting episodes. The effects are concentrated on reserve drains: positive TGA shocks significantly tighten funding spreads, while negative shocks have small and statistically insignificant effects, consistent with a flatter reserve demand curve on the abundant-reserves side and with nonlinearities in the spread–reserves relationship.

5.3 Spread Sensitivity in Overnight Borrowing: Foreign Banks at the Margin

Having established that TGA shocks causally move funding spreads, we next examine which institutions adjust their overnight borrowing in response. Using the TGA shock as an instrument, we estimate the effect of spread changes on federal funds borrowing volumes by bank type using daily transaction data from the FR 2420 report. The FR 2420 data provide borrowing volumes separately for foreign banking organizations (branches and agencies) and domestic banks from March 2016 onward; thus the analysis primarily captures the post-2016 portion of the sample and provides limited coverage of the earlier QE period.

The second-stage specification is:

$$\Delta \ln\Big(\text{Borrowing}\Big)_{j,t} = \alpha_0 + \alpha_1\, \Delta\widehat{(i^F - i^R)}_t + \gamma_{f(t)} + u_{j,t}, \tag{2}$$

where $j \in \{\text{Foreign}, \text{Domestic}\}$ indexes bank type, and $\Delta\widehat{(i^F - i^R)}_t$ is the predicted spread change from the first stage.

Table 4 reports the second-stage results. Foreign banks respond strongly to instrumented changes in spreads, while domestic bank borrowing is comparatively insensitive. Column (3) suggests that 1 basis point decrease in the FFR-IOR spread (i.e., FF falling relative to IOR, widening arbitrage) is associated with a 4% increase in Fed Funds borrowing by foreign banks. Splitting the sample by policy regime (Table 5) shows that during QE, both bank types adjust borrowing in response to spreads, whereas during QT the response is weaker and is concentrated amongst foreign banks, with very limited adjustment by domestic banks. This

[15] SOFR-based spreads provide a useful complement in the post-2019 period, when secured funding markets play a more prominent role in overnight intermediation.

pattern indicates that the margin of overnight borrowing adjustment shifts across regimes and bank types, consistent with tighter balance-sheet constraints for domestic institutions when reserves become scarcer. These findings are robust to alternative measures. Using SOFR–IOR spreads as an instrument yields similar conclusions (Table A3). Across specifications, foreign banks appear to operate closer to the margin of spread-based balance-sheet adjustment in overnight funding markets.

5.4 Reserve Holdings and QE–QT Asymmetry

We next examine how exogenous changes in aggregate reserve supply transmit to banks' reserve holdings. Because reserve balances can only be held by institutions with accounts at the Federal Reserve, any changes in aggregate reserve supply must be absorbed by some banks or by the overnight reverse repo (ONRRP) facility. Identifying which institutions adjust their reserve positions in response to supply shocks sheds light on the distributional incidence of reserve fluctuations and helps locate the margin of reserve demand.

To quantify this adjustment, we regress changes in reserve holdings on TGA shocks. For each bank category i as well as ONRRP facility, we estimate

$$\begin{cases} \Delta \text{Reserves}_{i,t} = \beta_i \, \Delta \text{TGA}_t + \delta_t + \varepsilon_{i,t}, \\ \Delta \text{ONRRP}_t = \beta_{\text{ONRRP}} \, \Delta \text{TGA}_t + \delta_t + \varepsilon_t. \end{cases} \tag{3}$$

The index i denotes foreign bank branches, the top 25 domestic banks by assets, and smaller domestic banks. The regressions are estimated at a weekly frequency from January 1st, 2009 to September 30th, 2025, excluding quarter-end weeks. All specifications include FOMC-period fixed effects.

Table 6 reports the pass-through from TGA-induced reserve supply shocks to changes in reserve holdings by bank type and policy regime. The results reveal a pronounced QE–QT asymmetry in the cross-sectional distribution of reserve adjustment. During the first balance sheet expansion (QE1), foreign bank branches absorb the bulk of reserve injections: a \$1 billion decline in the TGA is associated with a \$0.63 billion increase in foreign branch reserves. In contrast, large and small domestic banks exhibit little response, and the ONRRP facility plays a negligible role. This pattern indicates that foreign branches operate at the margin of reserve demand during QE, absorbing most systemwide reserve expansions.

The pattern reverses during tightening episodes. In both QT1 and QT2, foreign branches become largely unresponsive to reserve drains, while large domestic banks absorb a substantial share of aggregate reserve reductions, reducing reserves by roughly \$0.25–\$0.33 billion per \$1 billion increase in the TGA. Small domestic banks contribute modestly, and the ONRRP facility absorbs an increasing share of the adjustment in the post-pandemic period. Taken together, these results show that reserve demand is highly heterogeneous and state-dependent: foreign branches absorb reserves during expansions but do not symmetrically shed them during contractions, shifting the burden of adjustment during QT toward domestic institutions.

We next visualize time variation in reserve absorption using rolling regression estimates. While the baseline coefficients summarize average elasticities over the full sample, reserve demand need not be stable: shifts in market structure, regulation, and balance sheet constraints can alter banks' capacity and willingness to absorb reserve shocks (Afonso et al., 2022c). Accordingly, we estimate 52-week rolling coefficients, $\hat{\beta}_{j,t}^{(52)}$, from the baseline specification and use them to construct time-varying absorption shares,

$$\theta_{j,t} \;=\; 100 \times \frac{-\,\hat{\beta}_{j,t}^{(52)}}{\sum_{k\in\{\text{Foreign branches},\,\text{Large domestic},\,\text{Small domestic},\,\text{ONRRP}\}}\left(-\,\hat{\beta}_{k,t}^{(52)}\right)}. \tag{4}$$

The negative sign ensures that institutions absorbing reserve drains receive positive shares. Because different components may move in opposite directions, individual shares can be negative and cumulative shares can exceed 100%.

Figure 12 plots the time-varying composition of reserve absorption shares, with vertical dashed lines indicating major policy regime changes. Three patterns emerge. First, during the early QE period (2009–2014), foreign branches absorb the majority of TGA-induced reserve fluctuations, consistent with their high reserve-demand elasticity and active cross-border liquidity management. Second, following the introduction and subsequent expansion of the ON RRP facility after 2014, reserve absorption becomes more distributed, with ON RRP increasingly accounting for the marginal adjustment and the contribution of foreign branches declining. Third, during QT episodes, absorption shifts toward large domestic institutions together with ON RRP, while foreign branches' absorption becomes muted. Small domestic banks play a limited role throughout the sample.

These dynamics align with the heterogeneity in FF–IOR arbitrage documented earlier.

During QE1, foreign branches absorb approximately 76% of aggregate reserve injections, despite representing a substantially smaller share of banking system assets, while large and small domestic banks absorb 15% and 9%, respectively. As shown in Figure 8, Panel C, foreign branches accumulate reserves aggressively during QE but release them only partially during subsequent QT episodes, resulting in persistently elevated reserve-to-asset ratios. Domestic banks, by contrast, exhibit more symmetric reserve adjustments across balance sheet expansions and contractions.

Our results indicate that the distribution of a given stock of reserves has also become increasingly skewed, with foreign branches holding a disproportionate share that does not readily recirculate during balance sheet normalization. Together, these mechanisms suggest that the effective "ample reserves" threshold is higher than pre-QE levels, and that an additional buffer above this threshold may be required to account for reserves that are unevenly distributed across institutions. For the implementation of monetary policy, this cross-sectional ratchet implies that balance sheet normalization need not mirror the expansionary phase of QE. The institutions that absorb reserves during balance sheet expansions are not the same institutions that release reserves during contractions. This asymmetry helps explain why money market stress indicators, such as elevated repo rate volatility and widening SOFR–IOR spreads, emerged during the 2019 QT episode and have reappeared during the current normalization, even at aggregate reserve levels that might appear ample by historical standards.

6 Institutional Drivers of Heterogeneous and Asymmetric Reserve Demand

The previous section documented substantial heterogeneity in reserve demand and a pronounced QE–QT asymmetry: U.S. branches and agencies of foreign banking organizations expand reserve holdings elastically during balance sheet expansions but do not symmetrically unwind these positions during contractions. This section examines the institutional features underlying this pattern. We propose a decomposition of reserve holdings into two components: a spread-sensitive "arbitrage buffer component" that expands during QE and contracts during QT, and a "structural component" determined by internal liquidity governance and headquarters funding. The evidence is primarily descriptive, but it clarifies the mechanisms that generate regime-dependent reserve absorption and motivates the modeling framework in subsequent sections.

6.1 Institutional Features of Foreign Branches and Agencies

Foreign banks have systematically held larger reserve balances than domestic banks since the introduction of interest on reserves (IOR) in late 2008.[16] This gap is not fully explained by size, systemic importance, or observable balance sheet composition.[17] A key institutional distinction is organizational form. Foreign banks in the United States may operate through branches and agencies, subsidiaries, or a combination of both. Branches and agencies are legal extensions of the parent bank rather than separately capitalized entities, whereas subsidiaries are stand-alone U.S. institutions. This distinction maps directly into the regulatory perimeter and the marginal balance-sheet cost of holding reserves.

Regulatory asymmetries. Branches and agencies face a lighter U.S. prudential burden than subsidiaries and domestic banks. Because they are not separately capitalized U.S. entities, they are not subject to the same capital and leverage requirements that constrain domestic institutions. Unlike subsidiaries, they cannot accept FDIC-insured retail deposits, but they also avoid deposit insurance fees and certain capital surcharges.[18] Branches can accept wholesale deposits broadly, while agencies face restrictions on deposits from U.S. citizens. Since July 2016, foreign banking organizations with more than $50 billion in U.S. assets (excluding branch assets) have been required to form an Intermediate Holding Company (IHC) under Regulation YY, but this requirement applies to subsidiary operations rather than branches. Taken together, these regulatory asymmetries substantially lower the balance-sheet cost of holding reserves for branches and agencies compared to subsidiaries and domestic banks.

Access to headquarters funding. Critically, branches and agencies can transfer liquidity between their U.S. operations and foreign headquarters (HQ) through internal capital markets,

[16]See Appendix Figure A6 for the evolution of foreign banks' reserve holdings relative to domestic banks.

[17]Appendix Table A5 reports regressions of reserve holdings on foreign status, G-SIB designation, and their interaction. The foreign indicator remains large and statistically significant after controls and fixed effects.

[18]Insured branches that were accepting domestic retail deposits as of December 19, 1991, were grandfathered. As of June 2025, only four such branches remain in our sample. According to Banegas and Tase (2020), only one of these participates in the federal funds market.

providing flexibility to reallocate dollar liquidity in response to global interest rate differentials and internal funding needs.[19] In contrast, subsidiaries must maintain capital and liquidity buffers consistent with domestic prudential standards. Figure 13 shows the quarterly data of HQ lending assets and borrowing across jurisdiction. Before the crisis, branches and agencies accumulated net claims, however, that trend has shifted as net borrower from the HQ.

We quantify the importance of organizational form using pre-crisis cross-sectional variation in foreign banks' U.S. presence. Following Mian and Sufi (2022), we exploit the share of U.S. assets operated through branches in 2007Q1 as a predetermined measure of branch intensity, leveraging the high persistence of branch-versus-subsidiary structures documented by DiSalvo (2019). Specifically, we estimate:

$$\log(\text{Reserves})_{it} = \beta_1 \, \text{Foreign}_i + \beta_2 \, \text{BranchRatio}_i^{2007} + \Gamma X_{it} + \alpha_t + \epsilon_{it}, \tag{5}$$

where Foreign is an indicator for foreign ownership and $\text{BranchRatio}^{2007}$ is the share of U.S. assets in branch form as of 2007Q1.

Table 7 shows that organizational form is a strong predictor of reserve holdings. A higher pre-crisis branch share is associated with a large and statistically significant increase in reserves, even after controlling for size, funding composition, and fixed effects. Using the cross-sectional dispersion in branch share, the estimates imply that a one–standard-deviation increase in branch intensity is associated with roughly a 40 percent higher level of reserves. Within the foreign-bank sample, organizational form alone explains a substantial share of the cross-sectional variation in reserve holdings.

6.2 Decomposing Reserve Holdings: Structural Component and Arbitrage Buffer

These institutional characteristics of branches and agencies imply that their reserve holdings reflect both short-horizon arbitrage incentives enabled by lower balance sheet cost and internal liquidity management considerations given their close connection to HQs in their home countries. Accordingly, we decompose reserve holdings into two conceptually distinct

[19] Branches and agencies are subject to Sections 23A and 23B and Regulation W, but these rules primarily govern transactions among U.S. affiliates and impose few constraints on cross-border liquidity transfers.

components.

$$\text{Reserves} = \underbrace{\text{Arbitrage Buffer}}_{\text{elastic to IOR–market spreads}} + \underbrace{\text{Structural Component}}_{\text{HQ funding and internal liquidity governance}} \quad (6)$$

This decomposition is conceptual rather than mechanical. Empirically, we operationalize the arbitrage buffer using reporting-date balance sheet compression by foreign branches, while the residual component—quarter-end reserve holdings that do not respond to short-horizon spread incentives—is interpreted as structural.

The *arbitrage buffer component* captures incremental reserve holdings that arise when administered rates, such as interest on reserves (IOR), exceed market funding rates. These positions are attractive during balance sheet expansions, particularly under quantitative easing, but can be reduced relatively quickly when balance-sheet costs increase or when IOR–market-rate spreads compress. This decomposition provides an angle for the QE–QT asymmetry documented earlier: balance sheet expansions primarily operate through an expansion of the arbitrage buffer, whereas balance sheet contractions compress this buffer without necessarily inducing a commensurate decline in the structural component. We do not necessarily interpret the window-dressing gap as a precise measure of arbitrage volumes, but rather as a revealed-preference indicator of balance sheet capacity that can be rapidly adjusted in response to reporting and funding incentives.

The *structural component*, by contrast, captures reserve balances that foreign branches and agencies hold as part of their internal liquidity management. This component reflects precautionary dollar liquidity demand, internal risk limits, and the allocation of funds from foreign headquarters through internal capital markets. By construction, the structural component is not expected to respond one-for-one to short-horizon arbitrage incentives or to marginal changes in administered interest rates.

Empirically, we proxy for the arbitrage buffer using systematic quarter-end balance-sheet compression ("window dressing") by foreign branches. Window dressing refers to the adjustment of balance sheets around reporting dates to reduce reported leverage or other regulatory metrics. A key regulatory asymmetry under Basel III is that most non-U.S. jurisdictions, including the euro area and Japan, compute leverage ratios using quarter-end snapshots, whereas U.S. and U.K. banks report both quarter-end and quarterly-average balance sheet measures.[20] As a result, foreign branches have a strong incentive to temporarily shrink their balance sheets at reporting dates, while U.S. banks face little comparable benefit. We interpret the magnitude of the quarter-end contractions in reserves as a revealed-preference measure of reserve holdings that are discretionary and can be reduced without materially disrupting ongoing operations. Specifically, we combine quarterly reporting with weekly data on reserve holdings. For each quarter, we compute the difference between the weekly reserve balance (proxied by cash holdings) reported in the immediately preceding quarter-end and the quarter-end reserve balance reported in the FFIEC 002. This difference provides an estimate of the reserve positions that are temporarily shed at reporting dates.

As daily reserve data for foreign branches are not publicly available, this measure is necessarily an approximation.[21] Nevertheless, the measure exhibits two salient empirical patterns that support its interpretation as a proxy for the arbitrage buffer.

First, systematic quarter-end balance-sheet compression is largely absent prior to the introduction of interest on reserves and the transition to an ample-reserves regime, consistent with limited arbitrage incentives in the pre-crisis period. Second, the magnitude of the window dressing gap declines steadily as quantitative tightening progresses (most notably around 2019), suggesting a gradual compression of the discretionary, spread-elastic component of reserve holdings (Figure 15).

Taken together, these patterns suggest that quarter-end balance-sheet compression provides a useful empirical window into the time-varying balance sheet adjustment capacity of foreign branches.

6.3 QE Expansion: The Arbitrage Buffer Component in Operation

Quantitative easing increases aggregate reserve balances while administered rates often maintain a positive spread between interest on reserves (IOR) and short-term market funding rates. In this environment, eligible institutions can borrow at market rates and hold reserves to earn IOR. The institutional features documented above—low marginal balance-sheet costs and access to internal headquarters funding—make foreign branches and agencies particularly

[20] See Appendix A3. for a detailed discussion of regulatory asymmetries between Basel III standards and U.S. regulatory requirements.

[21] Afonso et al. (2025), using confidential daily data, document average quarter-end declines in reserve balances of approximately 18–20 percent. Banegas and Tase (2020) report similar magnitudes. These estimates are broadly consistent with the window dressing gap constructed here.

well positioned to scale this arbitrage activity.

We evaluate the decomposition introduced in Section 6.2 by examining whether variation in interest rate spreads predicts the arbitrage buffer but not the structural component. This exercise is intended as a diagnostic validation of the decomposition rather than a stand-alone causal test. Because the analysis relies on quarterly observations, the number of data points is necessarily limited; the emphasis is therefore on the contrast between components rather than on precise estimation.

Table 8 Panel A reports the window dressing gap (the proxy for the arbitrage buffer) against the FFR–IOR spread. Two patterns emerge. First, during QE and early QT, the arbitrage buffer comoves strongly with the spread, even though this result should be taken cautiously given the limited number of observations. Second, and more importantly for the decomposition, the same spread variation has no explanatory power for the structural component of reserve holdings. When quarter-end reserve balances are used as the dependent variable, the estimated coefficients on FFR–IOR are small, statistically insignificant, and visually flat in the corresponding scatter plots.

These results support the interpretation of the window dressing gap as a spread-elastic arbitrage buffer layered on top of a structural component that is largely insensitive to short-term funding spreads. During QE, reserve accumulation by foreign branches reflects the expansion of this arbitrage buffer rather than movements in the structural component.

Federal funds market structure and the marginal arbitrageur. The concentration of buffer adjustment in foreign branches is consistent with the post-crisis segmentation of the federal funds market. Prior to 2008, the federal funds market facilitated broad-based interbank redistribution of reserves. In the ample-reserves regime, activity is concentrated among a narrower set of lenders and borrowers.

On the lending side, the Federal Home Loan Banks (FHLBs) are the dominant suppliers of overnight funds. Because FHLBs are ineligible to earn IOR, federal funds lending provides a natural outlet for their liquid asset holdings. Other government-sponsored enterprises largely withdrew from unsecured lending following the 2008 conservatorships, leaving FHLBs as the principal institutional lenders.

On the borrowing side, foreign branches account for the bulk of federal funds borrowing. In addition to the regulatory perimeter advantages documented in Section 6.1, their credit quality makes them eligible counterparties under FHLB unsecured lending standards, which can exclude smaller domestic institutions. This structure implies a concentrated market in which a small set of lenders and borrowers transact under institutional and operational constraints. Figure 14 shows that FHLB-to-foreign-bank transactions account for approximately 90 percent of federal funds volume.[22] In this setting, the effective federal funds rate (EFFR) reflects the behavior of marginal borrowers—predominantly foreign branches—rather than the system-wide marginal value of reserves. Consistent with this interpretation, EFFR typically trades below IOR (Figure A4).

6.4 QT Asymmetry: The Structural Component and Its Determinants

During quantitative tightening (QT), reserve balances decline and domestic arbitrage spreads compress, reducing the return to financing reserves through short-term funding markets. Yet reserve holdings at foreign branches unwind only partially. Our evidence indicates that this asymmetry arises because a substantial portion of reserves is anchored by internal capital market allocation rather than by domestic funding conditions. In particular, reserve holdings are closely tied to net funding from foreign headquarters, which remains stable across policy regimes and does not respond to domestic arbitrage spreads. We document that the central role of headquarters funding in shaping the structural component of reserve demand.

What level of reserves are required to maintain stable money market conditions? Our findings identify a distinct but complementary mechanism to the "liquidity dependence" ratchet documented by Acharya and Rajan (2022); Acharya et al. (2023) for domestic banks. Their analysis highlights a liability-side asymmetry of domestics banks, whereby during QE banks expand demandable deposits and credit lines more rapidly than the growth of liquid assets used to back them. This dynamic generates operational liquidity dependence that persists during QT, as domestic banks cannot easily shed these deposit relationships without impairing valuable client franchises.

We, focusing on foreign bank branches and agencies, propose closely related but distinct mechanism. Foreign banks' baseline reserve holdings are shaped primarily by internal capital market allocation within multinational banking organizations rather than by U.S. deposit dynamics.

[22] A small number of domestic correspondent banks remain active, but their volumes are limited and have little influence on the volume-weighted effective federal funds rate; see Anbil et al. (2025).

Global arbitrage spreads and headquarters funding. To test this hypothesis, we estimate the relationship between reserve holdings and balance-sheet funding components at the consolidated bank holding company (BHC) level:

$$\log(\text{Reserves}_{it}) = \beta \cdot \text{FundingSource}_{it} + \gamma \cdot X_{i,t} + \alpha_i + \delta_t + \epsilon_{it} \tag{7}$$

where i indexes BHCs and t indexes quarters. All specifications include BHC fixed effects (α_i) to absorb time-invariant bank characteristics and control for lagged log assets to account for mechanical size correlations. We progressively add time fixed effects and time-by-country fixed effects to absorb aggregate and country-specific shocks. Standard errors are clustered at the BHC level.

We focus on three funding sources that plausibly map into the structural component of reserve demand for foreign branches and agencies.

First, we consider Net HQ liabilities, which is calculated as liabilities to foreign headquarters minus claims on headquarters, which capture net internal dollar funding flows. As shown in Figure A5 and Figure 13, foreign branches and agencies have increasingly operated as net borrowers from headquarters since 2008. Given low frictions in cross-border intragroup transfers, these positions are a natural candidate for shaping baseline reserve holdings that are insensitive to short-horizon domestic arbitrage incentives.

Second, foreign branches hold sizable uninsured deposits in the United States, almost entirely wholesale in nature. While these liabilities are not retail deposits, their uninsured and potentially runnable character motivates a liquidity-management channel analogous to the mechanism in Acharya and Rajan (2022); Acharya et al. (2023).

Third, we distinguish between total uninsured deposits and domestically sourced uninsured deposits to assess whether reserve holdings are more closely linked to funding obtained from U.S. counterparties, as opposed to offshore sources.

Table 9 shows the result. Net HQ funding is the robust predictor of reserve holdings, compared to other two variables. The estimated elasticity is approximately 0.44, implying that a 10 percent increase in net funding from headquarters is associated with a 4–5 percent increase in reserve balances. The estimate is stable across specifications and fixed effects. This stability is consistent with an interpretation in which the structural component reflects within-organization liquidity allocation rather than host-country regulation or common time-varying conditions. We also test robustness in subperiod analysis (Table A4).

The structural component of foreign banks' reserve holdings is not only shaped by internal capital market decisions but also responds to global interest rate differentials. When the U.S. interest on reserves exceeds the deposit facility rate in the parent bank's home country, multinational banking organizations face incentives to reallocate liquidity toward U.S. branches. We construct an international IOR spread as the difference between the Federal Reserve's IOR rate and the deposit facility rate of each foreign bank's home-country central bank (ECB, BOJ, BOE, BOC, and SNB).

Table 10 Panel A reports BHC-level panel regressions of log reserves on the international IOR spread, controlling for the domestic FF–IOR spread. The international spread coefficient is positive and statistically significant: a one percentage point increase in the US–home country IOR differential is associated with approximately 39 percent higher reserve holdings. Notably, the domestic FF–IOR spread coefficient is negative, suggesting that foreign branches do not accumulate reserves to arbitrage domestic spreads once global incentives are accounted for. Panel B confirms that net headquarters funding itself responds to the international spread, with an elasticity of approximately 0.26. These findings reinforce our interpretation: the structural component of foreign banks' reserve demand is driven by global factors—cross-border rate differentials and internal capital market allocation—rather than domestic funding conditions. This is consistent with Bräuning and Ivashina (2020), who document that international IOR differentials affect both reserve holdings and internal capital flows. Our analysis extends their findings to show that this channel has direct implications for balance sheet normalization: because headquarters funding does not respond symmetrically to domestic conditions during QT, the structural component of reserve demand introduces uncertainty that lies outside the Federal Reserve's direct policy control.

What ultimately drives headquarters funding flows to U.S. branches remains an open question. Our evidence that these flows respond to international IOR differentials suggests a connection to broader eurodollar market dynamics—foreign banks may use their U.S. branches as a node in global dollar intermediation, with headquarters funding reflecting intra-group positioning in response to relative policy rates across jurisdictions. We leave a deeper investigation of this channel, potentially using BIS cross-border banking data, to future research.

7 A Simple Model of Reserve with Demand Heterogeneity

This section presents a simple model of the reserves supply and demand that incorporates heterogeneity in reserve demand across domestic banks and foreign banks. The framework is motivated by the empirical findings documented in earlier sections and is designed to study the implications of heterogeneity of reserve demand for monetary policy implementation.

7.1 Demand

The reserve demand side consists of two types of banks: domestic banks, denoted by D, and U.S. branches of foreign banks, denoted by F.

Each bank of type $j \in \{D, F\}$ selects a non-negative desired reserve balance M_j, which comprises two components: a minimum reserves portion δ_j, and an optimally chosen excess reserves portion $X_j \geq 0$. That is,

$$M_j = \delta_j + X_j \geq 0 \tag{8}$$

The fixed component δ_j captures regulatory requirements, payment frictions, or settlement obligations. Conditional on δ_j, each bank chooses excess reserves X_j to minimize expected costs associated with end-of-day stochastic liquidity shocks. Let the actual end-of-day reserve balances be $X_j + \eta_j$. In the spirit of canonical reserves model by Poole (1968), after the interbank market closes and excess reserves are fixed, the bank is hit by an idiosyncratic shock, η_j which is mean zero random variable with continuous distribution function $F(\cdot)$. The bank's objective is to minimize expected costs:

$$\min_{X_j} \int_{-X_j}^{\infty} (X_j + \eta_j)(i - i_{\mathrm{IOR}} + \alpha_j)\, dF(\eta_j) + \int_{-\infty}^{-X_j} (-X_j - \eta_j)(i_{\mathrm{DW}} - i + \beta_j)\, dF(\eta_j), \tag{9}$$

where i is the federal funds rate, and the spreads $i - i_{\mathrm{IOR}} + \alpha_j$ and $i_{\mathrm{DW}} - i + \beta_j$ represent the effective return on reserves and the penalty for reserve shortfalls, respectively. We assume η_j follow the mean zero Gaussian $\eta_j \sim \mathcal{N}\left(0, \sigma_j^2\right)$, with type specific variance σ_j.

The parameter α_j captures a type-specific carry advantage or disadvantage. For example, positive α_j may arise due to favorable balance sheet treatment or internal capital pricing. Conversely, positive β_j captures the perceived or regulatory penalty (e.g., stigma) from borrowing via the discount window or from affiliated sources. Both parameters may be positive or negative, and are subject to corridor-consistency conditions to ensure that the aggregate market-clearing rate remains within the administered bounds. We impose the following assumptions reflecting institutional features based on the institutional setups we covered at the previous sections:

- $\sigma_F > \sigma_D$: Foreign banks' branches face greater idiosyncratic liquidity uncertainty due to their structural features. These include cross-border funding flows, intraday timing mismatches in USD funding markets, and limited access to stable retail deposit bases as they are not FDIC insured. Consequently, their reserve demand exhibits higher variance ex post.

- $\beta_F < \beta_D$: Foreign banks face a smaller effective penalty from reserve shortfalls. They are less affected by discount window stigma and branches can also access liquidity from their global headquarters.

- $\delta_F < \delta_D$: Foreign banks' branches maintain lower structural reserve buffers compared to domestic banks. This reflects the fact that foreign branches are exempt from certain regulatory requirements, including Liquidity Coverage Ratio (LCR) and intraday liquidity requirements.

and we impose state-contingent α_F assumptions as following:

- Regular Day $\alpha_F > \alpha_D$: Domestic banks face more binding balance sheet costs from holding reserves, due to regulatory constraints such as the leverage ratio, Supplementary Leverage Ratio (SLR) and FDIC assessment fees.

- Quarter-End Periods $\alpha_F \approx \alpha_D$: This reserves holding advantage disappears on key financial reporting dates as foreign branches must meet leverage ratio constraints at BHC level. Foreign banks face acute incentives to temporarily shrink their balance sheets, resulting into temporary increase in the balance sheet cost of holding reserves. This increase effectively eliminates their normal-period advantage, and we assume the foreign banks' advantage balances with that of domestic banks.

These assumptions imply that domestic banks face a steeper trade-off between reserve holdings and opportunity costs, while foreign banks serve as more elastic absorbers of reserves in equilibrium during regular days.

The demand curves implied by this structure are nonlinear in i. In the region where reserves are scarce, marginal valuation of liquidity is high, and demand is steeply downward-sloping. In contrast, once X_j is sufficiently large, additional reserves offer little marginal benefit, and the curve flattens—marking the transition to the ample/abundant reserves regime. The curvature and location of this transition depend on σ_j, α_j, and β_j, and play a central role in the aggregate elasticity of reserve demand.

Solving the first-order condition of (9) yields excess-reserve demand[23]

$$X_j^*(i) = -\sigma_j \Phi^{-1}\left(\frac{i - i_{\mathrm{IOR}} + \alpha_j}{c_j}\right). \tag{10}$$

where $c_j \equiv (i_{\mathrm{DW}} - i_{\mathrm{IOR}}) + \alpha_j + \beta_j > 0$ and Φ^{-1} is inverse of standard normal CDF. The total reserves demand for type j bank is

$$M_j^*(i) = X_j^*(i) + \delta_j. \tag{11}$$

The demand curve $M_j^*(i)$ is decreasing and nonlinear, reflecting the asymmetric cost structure induced by the corridor system and uncertainty in liquidity shocks.

The aggregate reserve demand curve is defined as the sum of type-specific total demands:

$$M_{\mathrm{agg}}(i) = M_D^*(i) + M_F^*(i) \tag{12}$$

which inherits a smooth, piecewise-convex shape from the individual curves. When the funds rate lies well above the IOR, both domestic and foreign banks face high opportunity costs of holding reserves, and demand is steep at scarce reserves. As the rate declines below IOR, reserves become more attractive, and the marginal value flattens out, indicating the transition into the ample regime. We now formalize the notion of ample reserves threshold.

Definition 1. Consider slope tolerance $\tau > 0$. For each bank type j, define ample threshold M_j^{flat} where the quantity of reserves at which the slope of its individual inverse demand curve equals τ:

$$M_j^{\mathrm{flat}} = \sup\{M \mid s_j(M) \geq \tau\} \tag{13}$$

[23] The full derivation of the first-order condition is provided in Online Appendix A5.1

where

$$s_j(M) = \left|\tfrac{di_j}{dM}(M)\right| = \frac{c_j}{\sigma_j}\,\varphi\!\left(\tfrac{M-\delta_j}{\sigma_j}\right), \qquad \varphi(z) := \tfrac{1}{\sqrt{2\pi}}\mathrm{e}^{-z^2/2} \tag{14}$$

For all $M \geq M_j^{\text{flat}}$, the individual demand curve is considered τ-flat. Then the aggregate demand curve threshold is defined as follows:[24]

Proposition 1. Let the aggregate inverse demand function be $i\,(M_{\text{agg}}\,)$ and its slope be $s_{\text{agg}}\;(M_{\text{agg}}\,)$. Then, for any slope tolerance $\tau > 0$, there exists a unique reserve level $M_{\text{agg}}^{\text{flat}}$ satisfying

$$s_{\text{agg}}\;\left(M_{\text{agg}}^{\text{flat}}\right) = \tau. \tag{15}$$

and the aggregate inverse-demand curve is τ-flat iff $M \geq M_{\text{agg}}^{\text{flat}}$. Moreover, this aggregate is less than or equal to the minimum of the individual threshold of bank j

$$M_{\text{agg}}^{\text{flat}} \;\leq\; \min\{\, M_D^{\text{flat}},\; M_F^{\text{flat}}\}, \tag{16}$$

where equality holds when $M_D^{\text{flat}} = M_F^{\text{flat}}$.

The intuition of Proposition 1 is that the banking sector as a whole stops caring about additional reserves as soon as the more elastic group has effectively satiated its liquidity needs, not when both groups are satiated. This feature tightens the ample-reserve threshold relative to taking the minimum of the two stand-alone thresholds. Next, we delve into which bank type pins down the aggregate kink:[25]

Proposition 2. Let $M_{\text{agg}}^{\text{flat}}$ be the unique reserve level defined in Proposition 1 and let us define the responsiveness of bank type j to the interest rate i changes

$$v_j(i) := \frac{dM_j^*}{di} = -\frac{\sigma_j}{c_j} \cdot \frac{1}{\varphi\left(\Phi^{-1}\left(\frac{i-i_{\text{IOR}}+\alpha_j}{c_j}\right)\right)} \tag{17}$$

[24]See A5.3 for the proof.
[25]See A5.4 for the proof.

Then type j pins the kink when $|v_j(i)|$ is the largest. Specifically, in the quarter-ends periods when $\alpha_F \approx \alpha_D$, foreign banks (type F) determines the aggregate τ-flat threshold (pins the kink).

Proposition 2 formalizes the intuition that the elasticity of the aggregate reserve demand curve is not static, but rather varies with institutionally-driven shifts in bank behavior. In normal periods, the identity of the marginal bank is not uniquely determined by qualitative assumptions, as the structural parameters governing each bank type's responsiveness have conflicting effects. This ambiguity is resolved at quarter-ends. During these periods, the leverage ratio, the major balance sheet constraints of holding reserves equalizes across domestic and foreign banks. This shift makes foreign banks become the dominant drivers of elasticity.

7.2 Supply

The supply-side framework follows the dual-mandate approach in Afonso et al. (2023). The New York Fed trading desk is charged with two operational goals exogenously given by the FOMC through "Implementation Note" released with each FOMC statement:

1. Rate objective: keep the effective federal funds rate close to the FOMC's point target i_{target}.[26]

2. Balance-sheet objective: maintain the size of the reserve balance sheet near a long-run benchmark, denoted M_{target}, which follows the Fed's continuation of balance sheet normalization, reducing the securities holdings.

Administered rates are set mechanically relative to the target:

$$\begin{aligned} i_{\text{DW}} &= i_{\text{target}} + s_1, \\ i_{\text{IOR}} &= i_{\text{target}} + s_2, \\ i_{\text{ONRRP}} &= i_{\text{target}} - s_3, \end{aligned} \tag{18}$$

with $s_1 \geq s_2 \geq 0$ and $s_3 \geq 0$. These spreads define a fixed corridor; throughout our model,

[26] We abstract from the federal funds target range for tractability.

we treat s_1, s_2, s_3 as exogeneously given.[27] With spreads fixed, the desk's single policy tool is the total quantity of reserves supplied, M. For any M, the market-clearing rate is given by the inverse aggregate demand

$$g(M) \;:=\; M_{\text{agg}}^{-1}(M), \qquad g'(M) < 0,$$

where $g'(M) < 0$ follows from $M'_{\text{agg}}(i) < 0$.[28]

Intuitively, increasing M shifts the operating point rightward along the aggregate curve, lowering the equilibrium rate; reducing M does the opposite. The Desk chooses M to balance rate precision against balance-sheet size using an upper-tail quadratic penalty that only punishes rate overshoots above the target:

$$\min_{M} \mathcal{L}(M) = \frac{1}{2}(1-\lambda)\mathbb{E}\left[(g(M) - i_{\text{target}}\,)_+^2\right] + \frac{1}{2}\lambda\left(M - M_{\text{target}}\,\right)^2, \tag{19}$$

where $x_+ := \max\{x, 0\}$. Consistent with the ample-reserves floor system, rate undershoots ($g(M) \leq i_{\text{target}}$) are not penalized while overshoots are. This is because rate overshoots carry disproportionately high social and operational costs: they reduce the credibility of the IOR/ONRRP floor, force ad-hoc open market operations, and entail political optics of losing control of the policy rate. By contrast, modest undershoots are readily absorbed by the ONRRP and the floor mechanism, with limited pass-through to money-market functioning.

When $\lambda \to 0$, the Desk prioritizes rate control and selects the smallest M such that $g(M) \leq i_{\text{target}}$; when $\lambda \to 1$, it prioritizes normalization and chooses $M \approx M_{\text{target}}$.

7.3 Deterministic optimum (without foreign banks' uncertainty)

We would like to see the optimal supply M^* under no foreign banks' uncertainty. Define

$$h(x) = \mathbb{1}\Big\{g(x) - i_{\text{target}}\Big\}(g(x) - i_{\text{target}})\, g'(x). \tag{20}$$

[27]Formerly, IOR is determined by the Board of Governors, while the ONRRP is set by the FOMC (Dawsey et al., 2023). In practice, the FOMC occasionally tightens or widens the corridor through "technical adjustments." These adjustments may occur independently of changes in the target range or alongside them, in which case the shift in administered rates may differ from the change in the target range. For example, on March 16, 2020, the spread between the discount window (DW) rate and the ONRRP rate was narrowed from 75 basis points to 25 basis points. Figure A8 documents the historical evolution of these spreads.

[28]See Appendix A5.2 for derivation of first and second derivatives of the aggregate inverse demand, $g(\cdot)$.

Then the first-order condition can be written as

$$(1-\lambda)h(M^*) \ + \ \lambda\Big(M^* - M_{\text{target}}\Big) \ = \ 0, \tag{21}$$

How does this optimal supply M^* change relative to λ?Define $F(M,\lambda) = (1-\lambda)h(M) + \lambda(M - M_{\text{target}})$. Away from the kink $g(M) = i_{\text{target}}$ (i.e., when the indicator is locally constant), the implicit-function theorem yields

$$\frac{\partial M^*}{\partial \lambda} = -\frac{\partial F/\partial \lambda}{\partial F/\partial M} = -\frac{-\,h(M^*) + \Big(M^* - M_{\text{target}}\Big)}{(1-\lambda)\,h'(M^*) + \lambda}. \tag{22}$$

Using the FOC (21) to substitute $M^* - M_{\text{target}} = -\frac{1-\lambda}{\lambda}\,h(M^*)$, the numerator becomes $h(M^*)/\lambda$, and we obtain the compact form

$$\frac{\partial M^*}{\partial \lambda} = \frac{\dfrac{h(M^*)}{\lambda}}{(1-\lambda)\,h'(M^*) + \lambda}. \tag{23}$$

This leads to two results:

- $g(M^*) \leq i_{\text{target}}$ (rate undershoots the target): $\partial M^*/\partial \lambda = 0$ as long as the penalty remains inactive.

- $g(M^*) > i_{\text{target}}$ (rate overshoots the target): $h(M^*) > 0$, while $h'(M^*) = [g'(M^*)]^2 + \Big(g(M^*) - i_{\text{target}}\Big)g''(M^*) \geq [g'(M^*)]^2 > 0$. Hence the denominator in Equation (23) is positive and $\partial M^*/\partial \lambda > 0$: the optimum moves toward M_{target} as the Desk places more weight on normalization.

7.4 Introducing Foreign Banks' Demand Uncertainty

We so far modeled reserve demand deterministically. However, in the real-world, the reserves market and repo market entails volatility and unexpected fluctuations, especially around month-ends and quarterly-ends.[29] In practice, structural buffers vary with reporting incentives,

[29] We focus on quarter-end "window-dressing" by foreign branches, but the framework is applicable to other kinds of foreign banks' demand shocks. Any transitory force that shifts foreign banks' demand (e.g., non-U.S. monetary policy moves, regulation changes, or foreign banks' unique funding market frictions) can be mapped into a reduced-form shock S. We concentrate on quarter-end window dressing because it is (i) institutionally

especially for branches and agencies of foreign banks. Quarter-end regulatory reporting is on a point-in-time basis for some jurisdictions outside the United States, most notably in Europe and Japan, giving foreign branches and agencies an incentive to compress their on-balance-sheet assets. Empirically, this "window-dressing" leads to a sharp, one-day contraction in the reserves they hold at the Federal Reserve, while domestic banks display little systematic change.

To capture this heterogeneity, we introduce stochastic variation in the foreign reserve floor while keeping the domestic floor fixed:

$$\delta_D \;=\; \bar{\delta}_D, \qquad \delta_F \;=\; \bar{\delta}_F \;-\; S, \tag{24}$$

where $S \geq 0$ is a one–sided (quarter-end) demand-reduction shock that lowers foreign banks' structural reserve buffers. We make no distributional assumption beyond $S \geq 0$ and finite moments.[30] Denote

$$\mu_S \;:=\; \mathbb{E}[S] \;\geq 0, \qquad \nu_S \;:=\; \mathrm{Var}(S) \;\geq 0. \tag{25}$$

A realization $S > 0$ reduces the foreign floor from $\bar{\delta}_F$ to $\bar{\delta}_F - S$, shifting the foreign demand curve $M_F^*(i)$ left by S at every i. Consequently, aggregate demand, $M_{\text{agg}}(i) = M_D^*(i) + M_F^*(i)$, also shifts left by the same amount because the domestic block is unchanged. Writing the market-clearing rate through the aggregate inverse demand $g(\cdot) := M_{\text{agg}}^{-1}(\cdot)$, the clearing rate on the shock date is $g(M - S)$.

Given the uncertainty introduction, we now turn to supply. We assume the FOMC does not adjust i_{target} or M_{target} in response to once in quarter regulatory window dressing by foreign banks, ant thus both targets remain unchanged relative to the deterministic benchmark. The Desk chooses the supply of reserves M to minimize the expected loss

$$\min_{M} \; \mathcal{L}(M) = \frac{1}{2}(1-\lambda)\,\mathbb{E}\Big[\Big(g(M-S) - i_{\text{target}}\Big)_+^2\Big] + \frac{1}{2}\lambda\,(M - M_{\text{target}})^2. \tag{26}$$

well-defined and (ii) empirically observable with clean timing, which facilitates calibration.

[30]The nonnegativity $S \geq 0$ reflects that window dressing reduces foreign reserve balances. The main results, however, do not rely on one-sided support: they continue to hold under mixed-sign shocks provided there is a nonzero probability of an upper-tail rate overshoot at the deterministic optimum, i.e., $\Pr\big(g(M^* - S) > i_{\text{target}}\big) > 0$.

When $S = 0$, the expectation in Equation (26) collapses and the loss function is equivalent to to the Equation (19). Under uncertainty, however, the expected penalty includes both the mean and the variance of the rate response to S. Differentiating (26) yields

$$(1-\lambda)\mathbb{E}[h(M-S)] + \lambda\,(M - M_{\text{target}}\,) = 0 \tag{27}$$

7.4.1 Comparative Statistics under Demand Uncertainty

Given this set up, we show that when foreign banks' show stochastic demand shocks, the optimal policy response by the Fed is to supply a larger quantity of reserves than would be necessary in a deterministic world, impeding balance sheet normalization. Let M^* be the optimal supply of reserves in the deterministic case (where $S \equiv 0$), which solves the first-order condition in Equation (21). Let M^{**} be the optimal supply in the stochastic case, which solves the first-order condition in Equation (27). We are interested in knowing the sign of $M^{**} - M^*$; whether the optimal reserves supply under stochastic foreign bank shock is larger or smaller than the deterministic supply.[31]

Proposition 3. Under demand uncertainty driven by foreign banks' window dressing where $S \geq 0$, the optimal supply of reserves is greater than or equal to the level in the deterministic case:$M^{**} \geq M^*$. Moreover, the inequality is strict ($M^{**} \ > \ M^*$), when the following is true:

$$\Pr\Big(g(M^* - S) > i_{\text{target}}\Big) \ > \ 0. \tag{28}$$

7.5 Calibration

We calibrate the model to the December 2018 to August 2019 phase of quantitative tightening (QT), during which aggregate reserves fell from about \$1.7 trillion to \$1.5 trillion, reaching closer to the aggregate kink (revealed in September 2019), while federal funds target range was unchanged. The objective of this exercise is quantitative: to assess, in dollar terms, where the potential kink could have been, or the additional reserve supply that would be required from the Fed under scenarios in which foreign banks introduce demand uncertainty. Certain

[31]See A5.5 for the proof.

parameters are identified directly from observed data, others are disciplined by established estimates in prior studies, and the remainder are inferred indirectly by matching model implications to the data. To ensure comparability across weeks, we exclude quarter-end observations from the data sample. Our choices of parameter values and specifications are summarized in Table11.

Directly observable data For the directly observable data, we calibrate the model using weekly (Wednesday-closing) averages of the federal funds rate i_t, the interest rate on reserves $i_{\mathrm{IOR},t}$, the discount window rate $i_{\mathrm{DW},t}$, and the ONRRP award rate $i_{\mathrm{ONRRP},t}$. Reserve quantities are constructed separately for domestic banks and U.S. branches of foreign banks, denoted $M_{D,t}$ and $M_{F,t}$, with aggregate reserves defined as $M_{\mathrm{agg}\ ,t} = M_{D,t} + M_{F,t}$.

Next, we set the following parameters using evidence in the previous literature: the liquidity shock scale σ_D and σ_F, the carry advantage on non quarter-end dates α_D and α_F, and the spread between shortfall penalty β_D and β_F. The remaining parameters are estimated by indirect inference on the QT sample.

Carry wedges We proxy the carry advantage for foreign banks, α_F, by combining estimates of balance sheet costs that disproportionately affect domestic banks. Following Banegas and Tase (2020), we incorporate an approximate 7 bps cost from FDIC assessments, which Afonso et al. (2019) and Schulhofer-Wohl and Clouse (2018) incorporate as a balance-sheet cost in their model calibration. Additionally, based on the 7.8 bps demand shift observed by Afonso et al. (2022c) around the SLR relief period, we add a conservative estimate for the SLR burden.[32] This results in a baseline carry advantage of $\alpha_F \approx 15$bps for foreign banks, relative to domestic banks for whom we normalize $\alpha_D \equiv 0$. At quarter ends when point-in-time reporting tightens foreign affiliates' balance-sheet constraints, we assume that the advantage disappears and impose $\alpha_F \approx \alpha_D = 0$.[33]

Reserve borrowing cost We normalize the foreign effective penalty as zero $\beta_F \equiv 0$ and measure the domestic reserve borrowing relative to this this base line. Under this convention,

[32]Comparing pre-relief estimates (Jan. 2015–Mar. 2020) with the period when the supplementary leverage ratio (SLR) relief was in place (Mar. 2020–Dec. 2021), Afonso et al. (2022c) report a vertical shift in the aggregate demand curve of about 7.8 bps.

[33]Industry analysis argues that daily leverage reporting can raise domestic banks' effective balance-sheet costs by considerably more (on the order of 50 bps), particularly for large institutions (Pozsar, 2017).

a positive β_D reflects a higher effective penalty for domestic banks; for example, stronger stigma associated with discount window borrowing.

Evidence suggests that foreign banks face weaker stigma relative to domestic banks. Although direct quantitative estimates of foreign-bank stigma are limited, several pieces of evidence are suggestive. First, during past episodes of market stress, foreign branches were among the most active users of the discount window, accounting for more than 70 percent of total borrowing (Keoun and Torres, 2011). Second, Goldberg and Skeie (2011) document that during the global financial crisis, branches' lending to their head offices increased by 78 percent between 2007Q3 and 2008Q3, arguing that branches drew on discount window credit to meet dollar funding needs of their headquarters. Finally, during the COVID-19 market stress in March 2020, foreign banks still accounted for roughly 32 percent of discount window lending.

For this reason, we interpret the stigma documented in the existing literature as applying primarily to domestic banks. A large body of work shows persistent reluctance to access the discount window. Armantier et al. (2024) show that during the September 2019 repo market disruption, the average "realized stigma spread" was approximately 43 basis points. They also mention that nearly three-quarters of federal funds transactions executed above the discount window rate during this period were settled at rates at least 25 basis points higher than the window rate. For this, we set the $\beta_D = 50$bps.

Liquidity shock scale and minimum required reserves We estimate the core demand parameters, the liquidity shock volatilities σ_j and minimum reserves δ_j, with a two-step procedure that ties market outcomes to bank-level heterogeneity.

First, using weekly data, for each bank type $j \in \{D, F\}$ and week t, define the type-specific corridor width $c_{j,t}$

$$c_{j,t} \equiv (i_{\mathrm{DW},t} - i_{\mathrm{IOR},t}) + \alpha_j + \beta_j \; > \; 0, \tag{29}$$

We then estimate the following:

$$M_{j,t} = \delta_j - \sigma_j \Phi^{-1}\left(\frac{i_t - i_{\mathrm{IOR},t} + \alpha_j}{c_{j,t}}\right) + u_{j,t}. \tag{30}$$

The slope recovers σ_j and the intercept recovers δ_j. This delivers the type aggregated levels.[34]

Next, using quarterly bank-level data, we allocate the type-level volatility across individual institutions by adapting the micro approach of Afonso et al. (2019), to our knowledge the only study that explicitly calibrates Poole-style liquidity shocks in the ample-reserves system.[35] For each bank i, we construct an 'effective size' for each bank, using uninsured deposit ratio and total assets to proxy payment-flow intensity. This feeds a saturating exponential map for the mean absolute shock, which we convert from their Laplace convention to our Gaussian parameter via mean-absolute-deviation matching; the resulting bank-level preliminaries $\sigma_{i,0}$ are then rescaled within type so that the correlation-adjusted aggregation exactly meets the weekly targets $\hat{\sigma}_j$:

$$\sigma_{i,0} = \theta_i \sqrt{\frac{\pi}{2}} \tag{31}$$

$$\tilde{\sigma}_j = \sqrt{(1-\rho_j)\sum_{i\in j}\sigma_{i,0}^2 + \rho_j\left(\sum_{i\in j}\sigma_{i,0}\right)^2} \tag{32}$$

Under our baseline sample, the estimates are $\sigma_D = \$62.9$billion and $\sigma_F = \$159.1$billion, consistent with $\sigma_F > \sigma_D$.

Finally, The intercepts of (30) recover minimum reserve buffers δ_j, interpreted as structural balances needed for settlement and institutional constraints, independent of precautionary motives. The resulting estimates, $\delta_D = \$840.8$ billion and $\delta_F = \$479.7$ billion, consistent with our directional assumptions about reserve demand heterogeneity.

[34] In Poole (1968), the probability a bank ends the day with a shortfall is a monotone function of the operating point of the policy rate within the corridor. With Gaussian liquidity shocks, the bank's optimal excess reserve choice implies $X_j(i) = \sigma_j \Phi^{-1}(\cdot)$, which maps to the linear specification above.

[35] Afonso et al. (2019) emphasize that as high-frequency, bank-specific reserve balances are not publicly available, making direct identification of payment-shock distributions not possible. They therefore calibrate bank-specific parameters and show that the implied behavior matches observed market rates in a Poole framework. We use their micro discipline to apportion risk across banks, while our weekly regressions pin down the aggregate level by type; the final within-type rescaling reconciles the two.

7.6 Numerical results and robustness

Figure 16 summarizes the calibrated demand curves. Panel A displays the aggregate curve, while Panels B and C show the domestic and foreign components. Solid lines denote regular days; dotted lines overlay quarter-end; $\alpha_D = \alpha_F$ (i.e., the foreign advantage disappears; we do not introduce quarter-end floor uncertainty here). The horizontal reference lines at 2.25%, 2.40 %, and 3.00% (ONRRP, IOR, DW). Two features stand out. First, the aggregate curve is composed by two very distinct demand curves. For instance, at aggregate reserves of 1.61 trillion, the model implies an equilibrium funds rate of 2.40% and the corresponding composition is 1.0 trillion held by domestic banks (about 62%) and 0.61 trillion held by foreign banks (about 38%). As reserves rise, the foreign curve flattens more quickly, so foreign banks increasingly absorb the marginal dollar; when the foreign advantage vanishes at quarter-end, that absorption weakens and the aggregate curve steepens. Second, quarter-end primarily shifts the foreign demand curve up , the aggregate curve correspondingly tilts upward, narrowing the ample-reserves region.

Figure 17 compares model-implied reserve shares with the observed weekly distribution. Over the range of aggregate reserves between 1.60 and 1.85 trillion, the model line tracks the empirical points closely for both domestic and foreign banks. At scarce reserve levels, for instance below 1.2 trillion, domestic banks dominate with a larger share, while foreign banks account for a smaller share. As reserves increase, the foreign curve flattens more quickly, so the model predicts that foreign banks absorb a growing fraction of marginal reserves. The scatter points plot the observed weekly reserve shares. The close fit between model and data indicates that the calibration captures the data closely.

Figure 19 illustrates the model's quantitative implications for the Federal Reserve's optimal reserve supply under uncertainty. Panel A plots the aggregate reserve demand curve with the Monte Carlo uncertainty band reflecting foreign banks' quarter-end behavior. The vertical red lines mark the optimal reserve supply under full information (M^*, solid) and under uncertainty (M^{**}, dashed). When $\lambda = 0.5$, the calibration implies that the Fed must supply approximately \$108 billion more reserves ($M^{**} - M^* \approx \$108$ billion) to achieve comparable rate stability.

Panel B plots the resulting policy gap $\Delta(\lambda) = M^{**}(\lambda) - M^*(\lambda)$ across alternative policy weights λ. The gap declines monotonically as greater emphasis is placed on rate control, illustrating that uncertainty in foreign banks' reserve behavior is most costly when the central bank simultaneously values balance-sheet normalization. This represents a 6–12% reserve supply increase relative to the deterministic benchmark.

Overall, these results quantify how foreign banks' quarter-end uncertainty effectively shifts the Fed's optimal supply curve outward, reinforcing the need for a precautionary buffer of reserves under QT.

8 Conclusion

We argue that heterogeneity in bank balance-sheet structures, and in particular the role of foreign banks, is central to understanding reserve demand and the limits of balance sheet normalization under the ample-reserves framework. Although foreign banks account for a modest share of traditional U.S. banking activity, they hold a disproportionate share of reserve balances at the Federal Reserve and operate at the margin of reserve demand during key phases of the policy cycle.

Empirically, we document a pronounced asymmetry in how reserves are absorbed and released across policy regimes. During QE, foreign bank branches absorb the majority of reserve inflows, reflecting their low balance-sheet costs and their role as marginal arbitrageurs in short-term funding markets. During QT, however, these same institutions do not unwind reserve positions symmetrically. Instead, reserve reductions are borne disproportionately by large domestic banks and, increasingly, by the Federal Reserve's standing facilities. This asymmetry implies that the identity of the marginal reserve holder shifts endogenously over the policy cycle.

To explain this pattern, we decompose foreign banks' reserve holdings into two components. The first is a spread-sensitive arbitrage buffer component that expands when administered rates exceed market funding costs and contracts as those spreads compress. The second is a structural component anchored by internal capital market allocations within multinational banking organizations. Using bank holding company (BHC)–level panel data, we show that net funding from foreign headquarters is the dominant determinant of this structural component, with a stable elasticity across QE and QT periods, while domestic funding variables lose explanatory power once size is controlled for. As a result, even as arbitrage incentives fade during QT, a substantial portion of foreign banks' reserve demand persists.

These findings have direct implications for monetary policy implementation. Because the structural component of reserve demand is driven by global internal capital markets rather than domestic liquidity conditions, it lies largely outside the Federal Reserve's direct policy control. Monitoring aggregate reserve levels alone is therefore insufficient for assessing scarcity. What matters for rate control is not only how many reserves exist, but where they are held and which institutions are willing to adjust at the margin. The asymmetry we document implies that the aggregate reserve demand curve becomes steeper precisely during balance sheet normalization, when the burden of adjustment shifts toward institutions with higher balance-sheet costs.

We formalize these mechanisms in a model with heterogeneous banks, which generates a kinked aggregate reserve demand curve and captures the endogenous shift in the marginal holder of reserves. When uncertainty about foreign banks' reserve demand is introduced, optimal policy features a precautionary buffer of reserves above the deterministic ample-reserves threshold. Calibrating the model to the 2018–2019 normalization episode yields a buffer on the order of $100 billion, consistent with the onset of observed money market stress.

In summary, our results suggest that balance sheet normalization is constrained not only by aggregate liquidity needs, but by the cross-sectional distribution of reserves and the global nature of modern banking. The presence of large foreign banks with reserve demand tied to internal capital markets introduces an international dimension to liquidity dependence, increasing uncertainty about the minimum reserve supply consistent with effective rate control. Understanding the determinants of internal funding flows within multinational banks, and their interaction with global monetary conditions, remains an important direction for future research.

References

Viral V Acharya and Raghuram Rajan. Liquidity, liquidity everywhere, not a drop to use–why flooding banks with central bank reserves may not expand liquidity. Technical report, National Bureau of Economic Research, 2022.

Viral V Acharya, Rahul S Chauhan, Raghuram Rajan, and Sascha Steffen. Liquidity dependence and the waxing and waning of central bank balance sheets. Technical report, National Bureau of Economic Research, 2023.

Gara Afonso, Roc Armenter, and Benjamin Lester. A model of the federal funds market: yesterday, today, and tomorrow. *Review of Economic Dynamics*, 33:177–204, 2019.

Gara Afonso, Marco Cipriani, and Gabriele La Spada. Banks' Balance-Sheet Costs, Monetary Policy, and the ON RRP. *SSRN Electronic Journal*, 2022a. ISSN 1556-5068. doi: 10.2139/ssrn.4308896. URL `https://www.ssrn.com/abstract=4308896`.

Gara Afonso, Darrell Duffie, Lorenzo Rigon, and Hyun Song Shin. How abundant are reserves? evidence from the wholesale payment system. Technical report, National Bureau of Economic Research, 2022b.

Gara Afonso, Domenico Giannone, Gabriele La Spada, and John C Williams. Scarce, abundant, or ample? a time-varying model of the reserve demand curve. Technical report, Staff report, 2022c.

Gara Afonso, Gabriele La Spada, Thomas M Mertens, and John C Williams. The optimal supply of central bank reserves under uncertainty. *FRB of New York Staff Report*, (1077), 2023.

Gara Afonso, Marco Cipriani, JC Martinez, and Matthew Plosser. Reserves and where to find them. `https://libertystreeteconomics.newyorkfed.org/2025/06/reserves-and-where-to-find-them/`, June 2025. URL `https://libertystreeteconomics.newyorkfed.org/2025/06/reserves-and-where-to-find-them/`. Liberty Street Economics, Federal Reserve Bank of New York.

Iñaki Aldasoro, Torsten Ehlers, and Egemen Eren. Global banks, dollar funding, and regulation. *Journal of International Economics*, 137:103609, 2022.

Sriya Anbil, Alyssa Anderson, and Zeynep Senyuz. What happened in money markets in september 2019? Feds notes, Board of Governors of the Federal Reserve System, Washington, DC, February 2020. URL `https://doi.org/10.17016/2380-7172.2527`.

Sriya Anbil, Sebastian Infante, and Senyuz Zeynep. A tale of demand and supply for central bank reserves. *Working Paper*, 2025.

Alyssa G Anderson, Wenxin Du, and Bernd Schlusche. Arbitrage capital of global banks. Technical report, National Bureau of Economic Research, 2021.

Olivier Armantier, Marco Cipriani, and Asani Sarkar. Discount window stigma after the global financial crisis. Technical report, Staff Reports, 2024.

Ayelen Banegas and Manjola Tase. Reserve balances, the federal funds market and arbitrage in the new regulatory framework. *Journal of Banking & Finance*, 118:105893, 2020.

Basel Committee on Banking Supervision. Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools. Technical report, Bank for International Settlements, 2013. URL https://www.bis.org/publ/bcbs238.htm. Retrieved from Bank for International Settlements.

Basel Committee on Banking Supervision. Basel III leverage ratio framework and disclosure requirements. Technical report, Bank for International Settlements, 2014. URL https://www.bis.org/publ/bcbs270.htm. Retrieved from Bank for International Settlements.

Basel Committee on Banking Supervision. Revisions to the Basel III leverage ratio framework. Technical report, Bank for International Settlements, 2017a. URL https://www.bis.org/bcbs/publ/d365.htm. Retrieved from Bank for International Settlements.

Basel Committee on Banking Supervision. Finalising Basel III - In brief. Technical report, Bank for International Settlements, 2017b. URL https://www.bis.org/bcbs/publ/d424_inbrief.pdf. Retrieved from Bank for International Settlements.

Morten L Bech and Elizabeth Klee. The mechanics of a graceful exit: Interest on reserves and segmentation in the federal funds market. *Journal of Monetary Economics*, 58(5): 415–431, 2011.

Efraim Benmelech. An empirical analysis of the fed's term auction facility. Technical report, National Bureau of Economic Research, 2012.

Jose Berrospide, Ricardo Correa, Linda Goldberg, and Friederike Niepmann. International banking and cross-border effects of regulation: Lessons from the united states. Technical report, National Bureau of Economic Research, 2016.

Javier Bianchi and Saki Bigio. Banks, liquidity management, and monetary policy. *Econometrica*, 90(1):391–454, 2022.

Saki Bigio and Yuliy Sannikov. A model of credit, money, interest, and prices. Technical report, National Bureau of Economic Research, 2021.

Board of Governors of the Federal Reserve System. History of the fomc's policy normalization discussions and communications, 2025. URL https://www.federalreserve.gov/monetarypolicy/policy-normalization-discussions-communications-history.htm. Accessed January 2026; documents evolution of FOMC policy normalization discussions and communications.

Claudio Borio. Getting up from the floor. *SUERF Policy Notes*, 2023.

Michelle W. Bowman. Thoughts on monetary policy decisionmaking and challenges ahead. Speech at the Forecasters Club of New York Luncheon, New York, NY, 2025. URL https://www.federalreserve.gov/newsevents/speech/bowman20250926a.htm. Vice Chair for Supervision.

Falk Bräuning. The liquidity effect of the federal reserve's balance sheet reduction on short-term interest rates. *Federal Reserve Bank of Boston Research Paper Series Current Policy Perspectives Paper*, (18-1), 2017.

Falk Bräuning and Victoria Ivashina. Monetary policy and global banking. *The Journal of Finance*, 75(6):3055–3095, 2020.

Mark A Carlson, Stefania D'Amico, Cristina Fuentes-Albero, Bernd Schlusche, and Paul R Wood. Issues in the use of the balance sheet tool. 2020.

Nicola Cetorelli and Linda S Goldberg. Global banks and international shock transmission: Evidence from the crisis. *IMF Economic review*, 59(1):41–76, 2011.

Adam Copeland, Darrell Duffie, and Yilin Yang. Reserves were not so ample after all. Technical report, National Bureau of Economic Research, 2021.

Ricardo Correa, Wenxin Du, and Gordon Y Liao. Us banks and global liquidity. Technical report, National Bureau of Economic Research, 2020.

Sergio Correia, Stephan Luck, and Emil Verner. Failing banks. *Available at SSRN 4650834*, 2024.

Ted Cruz. Sens. cruz, scott lead bill to stop federal reserve from paying interest on bank reserves. Press Release, U.S. Senate, July 2025. URL https://www.cruz.senate.gov/newsroom/press-releases/sens-cruz-scott-lead-bill-to-stop-federal-reserve-from-paying-interest-on-bank-reserves. Fiscal Accountability for Interest on Reserves (FAIR) Act.

Stefania D'Amico and Corey Feldman. Balance sheet policy uncertainty and its aggregate implications. 2024.

Matthew Darst, Sotirios Kokas, Alexandros Kontonikas, José-Luis Peydró, and Alexandros Vardoulakis. Qe, bank liquidity risk management, and non-bank funding: Evidence from us administrative data. 2025.

Adrien d'Avernas, Damon Petersen, and Quentin Vandeweyer. The central bank's balance sheet and treasury market disruptions. 2025.

Kristopher Dawsey, William B English, and Brian Sack. The federal reserve balance sheet. In *Research Handbook of Financial Markets*, pages 6–32. Edward Elgar Publishing, 2023.

William F Diamond, Zhengyang Jiang, and Yiming Ma. The reserve supply channel of unconventional monetary policy. Technical report, National Bureau of Economic Research, 2023.

James DiSalvo. How foreign banks changed after dodd–frank. *Federal Reserve Bank of Philadelphia Economic Insights (USA)*, 3:1–6, 2019.

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

Itamar Drechsler, Alexi Savov, and Philipp Schnabl. Banking on deposits: Maturity transformation without interest rate risk. *The Journal of Finance*, 76(3):1091–1143, 2021.

Basile Dubois and Paul Rintamäki. Banking under large excess reserves. 2025.

Stefania D'Amico and Tim Seida. Unexpected supply effects of quantitative easing and tightening. *The Economic Journal*, 134(658):579–613, 2024.

Jens Eisenschmidt, Yiming Ma, and Anthony Lee Zhang. Monetary policy transmission in segmented markets. *Journal of Financial Economics*, 151:103738, 2024.

Egemen Eren, Timothy Jackson, and Giovanni Lombardo. Optimal central bank balance sheets. 2024.

Federal Open Market Committee. Policy normalization principles and plans, September 2014. URL `https://www.federalreserve.gov/monetarypolicy/files/FOMC_PolicyNormalization.pdf`. Adopted effective September 16, 2014.

José Fillat, Arthur Smith, and Stefania Garetto. To branch or not to branch? a quantitative evaluation of the consequences of global banks' organization. In *2018 Meeting Papers*, number 1079. Society for Economic Dynamics, 2018.

Milton Friedman. *A Program for Monetary Stability*. Fordham University Press, New York, 1959.

Joseph E Gagnon and Brian Sack. *Monetary policy with abundant liquidity: a new operating framework for the Federal Reserve*. Peterson Institute for International Economics, 2014.

Linda S. Goldberg and David R. Skeie. Why did u.s. branches of foreign banks borrow at the discount window during the crisis? Liberty Street Economics Blog, Federal Reserve Bank of New York, April 2011. URL `https://libertystreeteconomics.newyorkfed.org/2011/04/why-did-us-branches-of-foreign-banks-borrow-at-the-discount-window-during-the-crisis/`. Accessed: 2025-09-04.

Itay Goldstein, Ming Yang, and Yao Zeng. Payments, reserves, and financial fragility. *Available at SSRN 4547329*, 2023.

Marvin Goodfriend. Interest on reserves and monetary policy. *Economic Policy Review*, 8(1), 2002.

Robin Greenwood, Samuel G Hanson, Jeremy C Stein, et al. The federal reserve's balance sheet as a financial-stability tool. *Innovative Federal Reserve Policies During the Great Financial Crisis*, pages 63–124, 2016.

Robert E Hall and Ricardo Reis. Maintaining central-bank solvency under new-style central banking. *Hoover Institution, Stanford University, and Columbia University*, 2013.

James D Hamilton. Measuring the liquidity effect. *The American Economic Review*, pages 80–97, 1997.

Jane E Ihrig and Scott Wolla. Let's close the gap: Revising teaching materials to reflect how the federal reserve implements monetary policy. 2020.

Victoria Ivashina, David S Scharfstein, and Jeremy C Stein. Dollar funding and the lending behavior of global banks. *The Quarterly Journal of Economics*, 130(3):1241–1281, 2015.

R Jay Kahn, Matthew McCormick, Vy Nguyen, Mark Paddrik, and H Peyton Young. Anatomy of the repo rate spikes in september 2019. *Journal of Financial Crises*, 5(4):1–25, 2023.

Sebnem Kalemli-Özcan. Us monetary policy and international risk spillovers. Technical report, National Bureau of Economic Research, 2019.

Anil K Kashyap and Jeremy C Stein. What do a million observations on banks say about the transmission of monetary policy? *American Economic Review*, 90(3):407–428, 2000.

Anil K Kashyap and Jeremy C Stein. The optimal conduct of monetary policy with interest on reserves. *American Economic Journal: Macroeconomics*, 4(1):266–282, 2012.

Thomas Keating and Marco Macchiavelli. Interest on reserves and arbitrage in post-crisis money markets. 2017.

Bradley Keoun. Foreign banks used fed secret lifeline most at crisis peak. *Bloomberg Business*, April 2011. URL http://www.bloomberg.com/news/articles/2011-04-01/foreignbanks-tapped-fed-s-lifeline-most-as-bernanke-kept-borrowers-secret.

Bradley Keoun and Craig Torres. Foreign banks used fed secret lifeline most at crisis peak. *Bloomberg News*, April 2011. URL https://www.bloomberg.com/news/articles/2011-04-01/foreign-banks-tapped-fed-s-lifeline-most-as-bernanke-kept-borrowers-secret. Accessed: 2025-09-04.

Duhyeong Kim. Interbank markets, foreign banks, and monetary policy under a floor system. Available at SSRN, December 2025. URL https://ssrn.com/abstract=4922179.

Ricardo Lagos and Gastón Navarro. Monetary policy operations: Theory, evidence, and tools for quantitative analysis. Technical report, National Bureau of Economic Research, 2023.

Lorie K. Logan. Observations on implementing monetary policy in an ample-reserves regime. Technical report, Federal Reserve Bank of New York, April 2019. URL https://www.newyorkfed.org/newsevents/speeches/2019/log190417. Speech delivered on April 17, 2019.

David Lopez-Salido and Annette Vissing-Jorgensen. Reserve demand, interest rate control, and quantitative tightening. *Interest Rate Control, and Quantitative Tightening (February 27, 2023)*, 2023.

Loretta J Mester. The Fed's Ample Reserves Monetary Policy Operating Framework: It Isn't as Simple as It Looks. 2024.

Atif Mian and Amir Sufi. Credit supply and housing speculation. *The Review of Financial Studies*, 35(2):680–719, 2022.

Stephen I. Miran. Regulatory dominance of the federal reserve's balance sheet. Speech, November 2025. URL https://www.federalreserve.gov/newsevents/speech/miran20251119a.htm. Speech delivered November 19, 2025.

Juan M Morelli, Pablo Ottonello, and Diego J Perez. Global banks and systemic debt crises. *Econometrica*, 90(2):749–798, 2022.

Bill Nelson. From the floor back to the corridor: Why the choice of monetary policy implementation framework matters, September 2024a. URL https://bpi.com/from-the-floor-back-to-the-corridor-why-the-choice-of-monetary-policy-implementation-framework-matters/. Bank Policy Institute.

Bill Nelson. How the federal reserve got so huge, and why and how it can shrink. Technical Report Staff Working Paper 2024-1, Bank Policy Institute, February 2024b. URL https://bpi.com/wp-content/uploads/2024/02/How-the-Federal-Reserve-Got-So-Huge-and-Why-and-How-It-Can-Shrink.pdf. Staff Working Paper, February 7, 2024.

Friederike Niepmann. Banking across borders with heterogeneous banks, 2023.

Friederike Niepmann and Leslie Sheng Shen. Geopolitical risk and global banking. *Available at SSRN*, 2024.

Roberto Perli. Balance sheet normalization: Monitoring reserve conditions and understanding repo market pressures. Technical report, 2024a.

Roberto Perli. Facing quarter-end pressures: Understanding the repo market and federal reserve tools. Technical report, Federal Reserve Bank of New York, 2024b.

William Poole. Commercial bank reserve management in a stochastic model: implications for monetary policy. *The Journal of finance*, 23(5):769–791, 1968.

Zoltan Pozsar. Global money notes# 9, 2017.

Dagfinn Rime, Andreas Schrimpf, and Olav Syrstad. Covered interest parity arbitrage. *The Review of Financial Studies*, 35(11):5185–5227, 2022.

Paul Santoro. The evolution of treasury cash management during the financial crisis. *Current Issues in Economics and Finance*, 18(3), 2012.

Sam Schulhofer-Wohl and James A Clouse. A sequential bargaining model of the fed funds market with excess reserves. Technical report, working paper, 2018.

Christopher A Sims. Paper money. *American Economic Review*, 103(2):563–584, 2013.

A Lee Smith and Victor J Valcarcel. The financial market effects of unwinding the federal reserve's balance sheet. *Journal of Economic Dynamics and Control*, 146:104582, 2023.

Quentin Vandeweyer, Adrien d'Avernas, Antoine Hubert de Fraisse, and Liming Ning. The fiscal cost of quantitative easing. *Available at SSRN 5314990*, 2024.

Annette Vissing-Jorgensen. Fluctuations in the treasury general account and their effect on the fed's balance sheet. Feds notes, Board of Governors of the Federal Reserve System, August 2025. URL https://doi.org/10.17016/2380-7172.3873. FEDS Notes, August 6, 2025.

Larry Wall. Interest on reserves. *Notes from the Vault*, 2017.

Neil Wallace. A modigliani-miller theorem for open-market operations. *The American Economic Review*, 71(3):267–274, 1981.

Yilin David Yang. What quantity of reserves is sufficient? *Available at SSRN 3721785*, 2020.

Figures

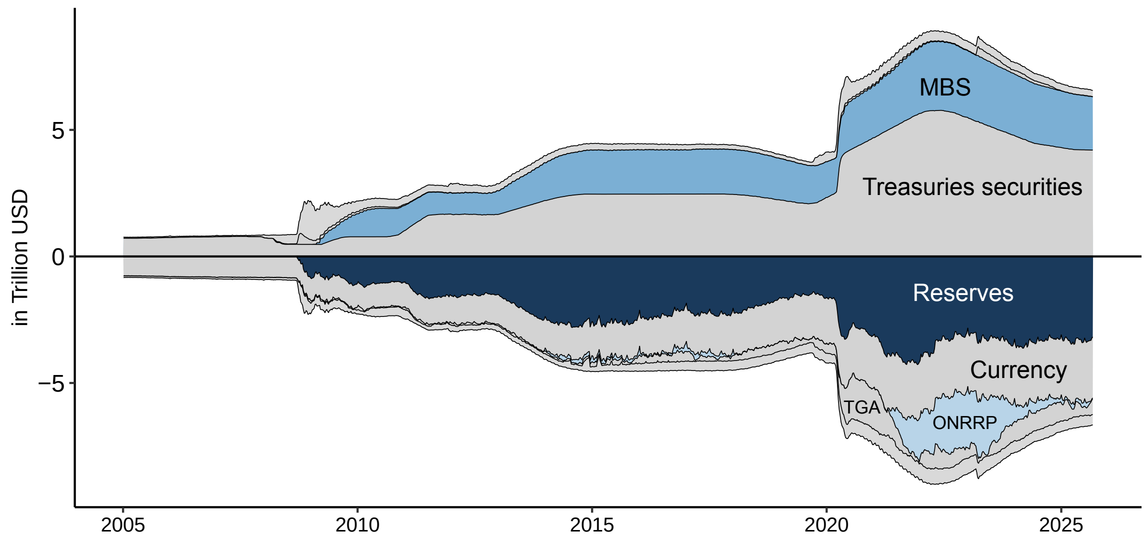


Figure 1: Federal Reserve Board Balance Sheet Breakdown This figure illustrates the composition of the Federal Reserve's balance sheet from 2005 to the present, highlighting the relevant components in this paper. Source: Federal Reserve Board H.4.1 Factors Affecting Reserve Balances

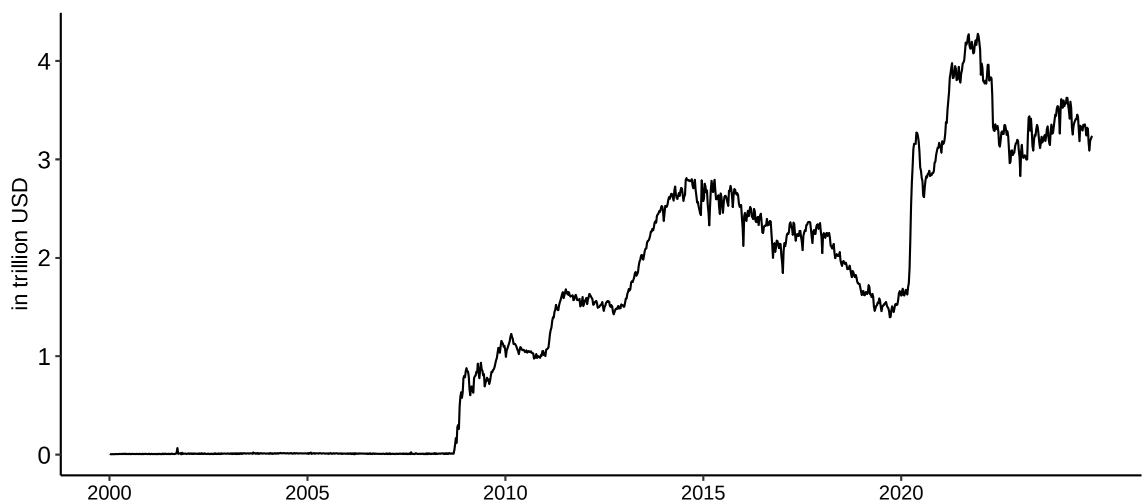


Figure 2: Reserves Outstanding This figure shows the outstanding reserves at Federal Reserve Board, at weekly level as of end Wednesday, from 2000 to 2024. Source: Federal Reserve Board H.4.1 Factors Affecting Reserve Balances

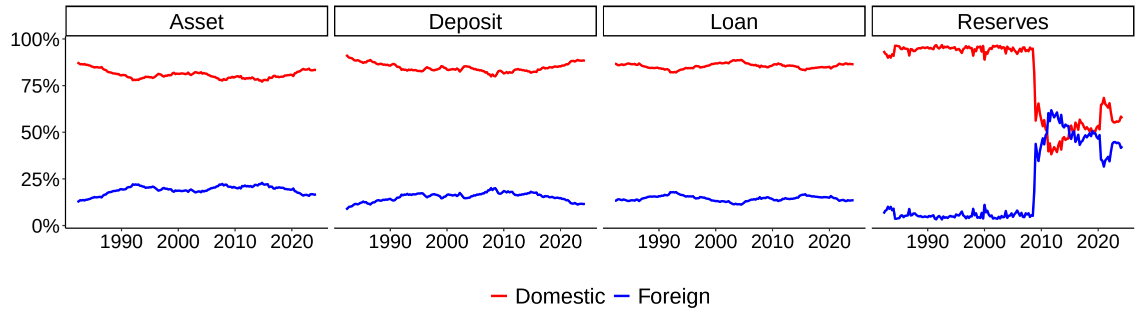


Figure 3: Share of Balance Sheet Components: Domestic Banks and Foreign Banks This figure shows the assets, loans, deposits, and reserves of the domestic banks and foreign banks (the aggregate sum of subsidiaries and branches and agencies) from 1980Q2 to 2024Q3 based on Call Report filings. The plots are adjusted for window-dressing. For non window dressing adjusted plot, see Figure A1. Source: FFIEC 002, 031, 041 reports

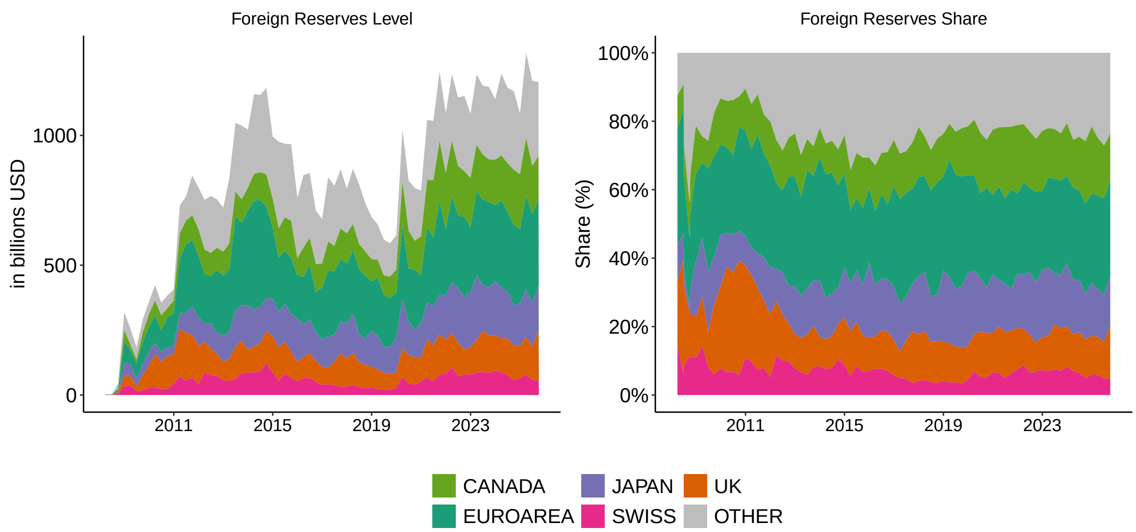


Figure 4: Reserves Holding by Foreign Banks Across Jurisdictions This figure presents the quarterly time series of reserve holdings by foreign banks (the aggregate of subsidiaries and branches) in the US, categorized by jurisdictions. Source: FFIEC 002, 031, 041 reports

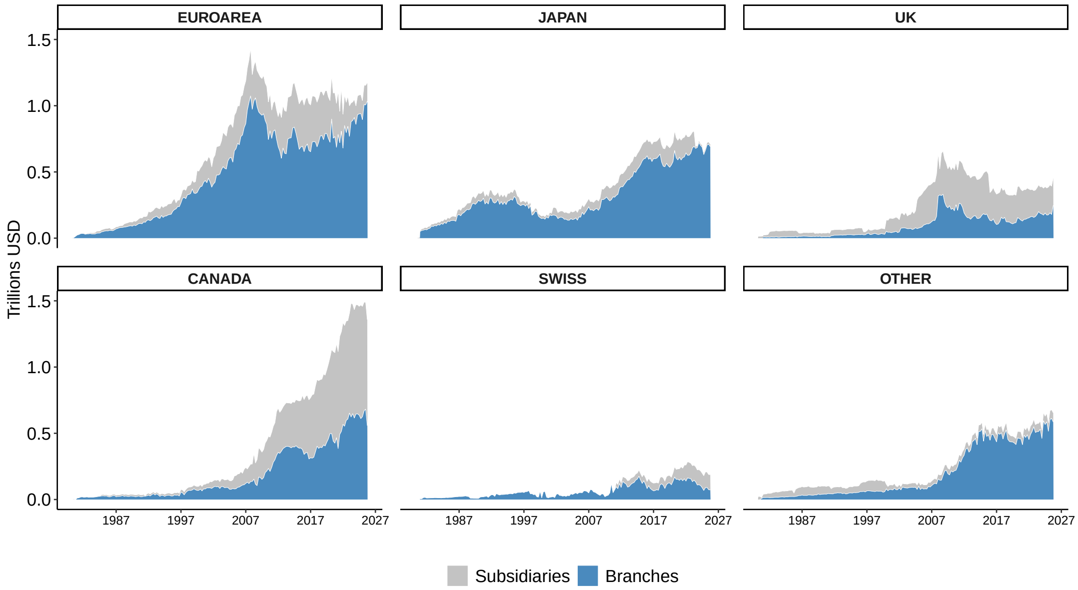


Figure 5: Asset Composition of Foreign Banks by Jurisdiction This figure displays the total assets of foreign banks in the U.S., disaggregated by jurisdiction and organizational structure. Each panel shows the time series of total assets for banks headquartered in a given jurisdiction, broken down into branches (gray) and subsidiaries (blue). Source: FFIEC 002, 031, 041 reports

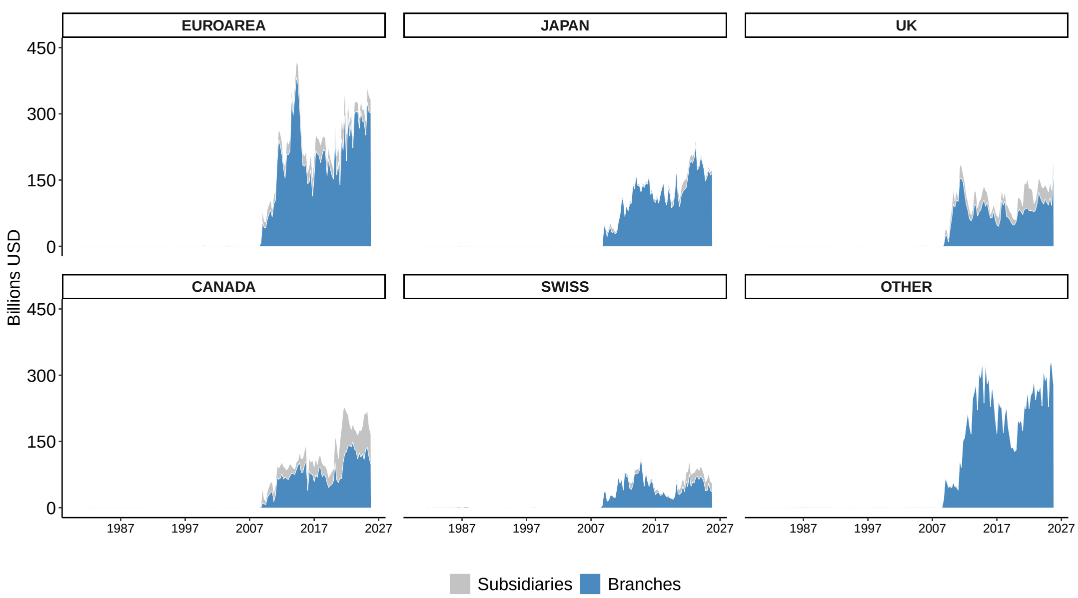


Figure 6: Reserves Composition of Foreign Banks by Jurisdiction This figure displays the total reserves of foreign banks in the U.S., disaggregated by jurisdiction and organizational structure. Each panel shows the time series of total assets for banks headquartered in a given jurisdiction, broken down into branches (gray) and subsidiaries (blue). Source: FFIEC 002, 031, 041 reports

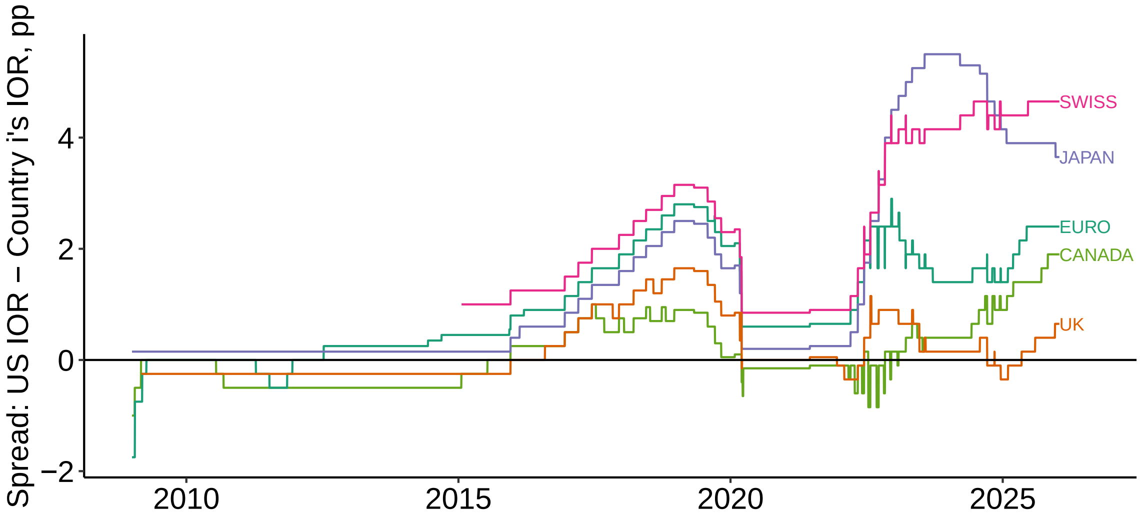
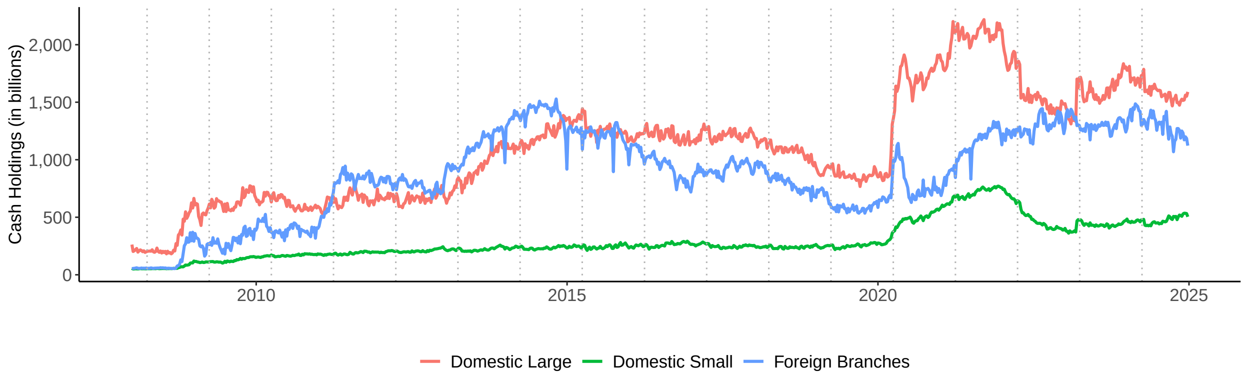
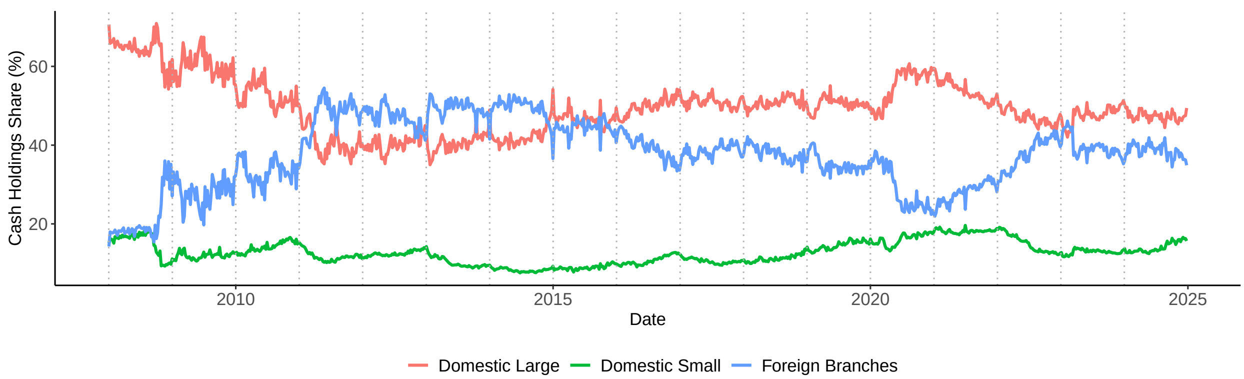


Figure 7: International Reserves Rate Spreads for Foreign Banks This figure displays the daily interest rate spread between the US interest on reserves (IOR) and the corresponding reserve interest rates in selected foreign jurisdictions. For the US, we use the interest rate on excess reserves (IOER) until July 28, 2021, after which we use the interest rate on reserve balances (IORB). For Japan, following the introduction of the three-tier reserve system in January 2016, we use the allocation to the policy rate balance as the relevant rate, representing the lowest interest rate applied to reserves held at the Bank of Japan. This rate was –0.1% through March 2024 and was raised to 0.1% thereafter. For the Euro area, we use the deposit facility rate, while for the UK, we use the Bank Rate. For Switzerland, we use the interest rate on sight deposits above the policy threshold, starting from January 22, 2015. Data sources: Bank of Canada, Bank of England, Bank of Japan, European Central Bank, Federal Reserve Board, and Swiss National Bank.

Panel A: Cash Holding (in billions USD)



Panel B: Share of Cash Holding



Panel C: Cash-to-Assets Ratio

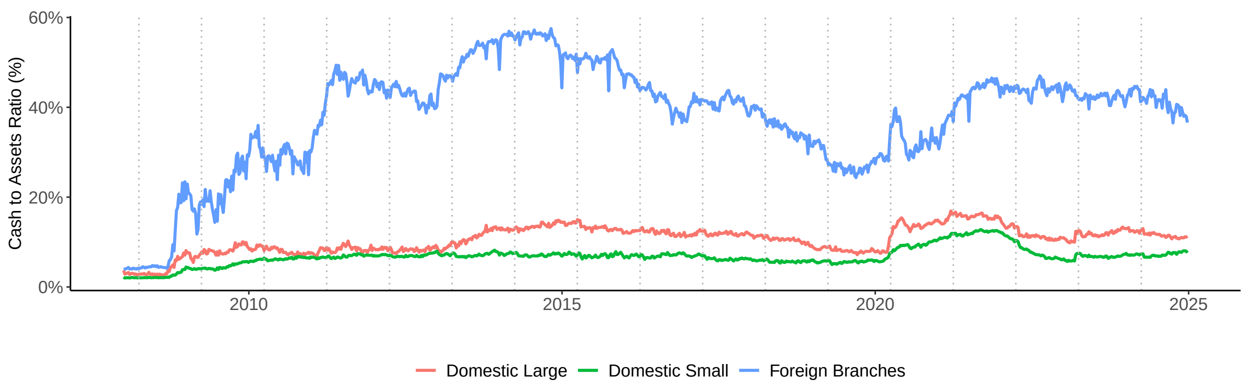


Figure 8: Cash Holdings by Bank Type This figure plots the evolution of cash balances among banks from 2008 to 2024, distinguishing between domestic large banks (Top 25 banks by asset size), domestic small banks, and U.S. branches of foreign banks. Panel A shows the total level of cash holdings (in billions of dollars). Panel B displays the share of total U.S. banking system cash held by each bank type. Panel C presents the ratio of cash to total assets. Vertical dotted lines indicate calendar year-ends. Source: Federal Reserve's H.8 release.

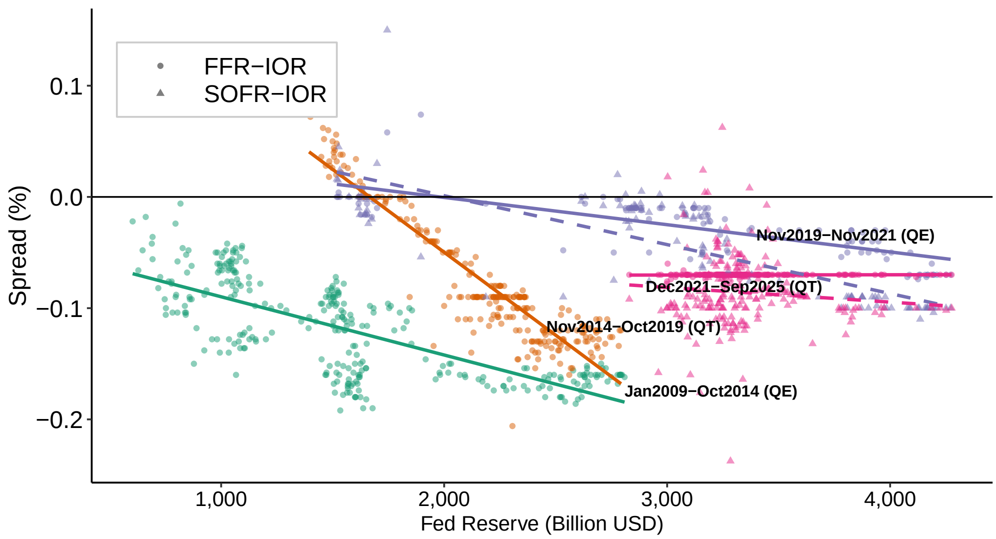


Figure 9: Funding Spreads and Aggregate Reserve Balances This figure plots the spread between the federal funds rate (FFR) and interest on reserves (IOR) against aggregate reserve balances at the Federal Reserve, measured at weekly frequency. Observations are grouped into four monetary policy regimes: QE (January 2009 to October 2014), QT (November 2014 to October 2019), QE (November 2019 to November 2021), and QT (December 2021 to September 2025). Solid lines show FFR-IOR spreads; dashed lines show SOFR-IOR spreads for post-November 2019 periods. Each point represents a weekly observation; fitted lines are estimated by OLS within each regime. Source: FRED (DFF, IORB, IOER, SOFR, WRESBAL)

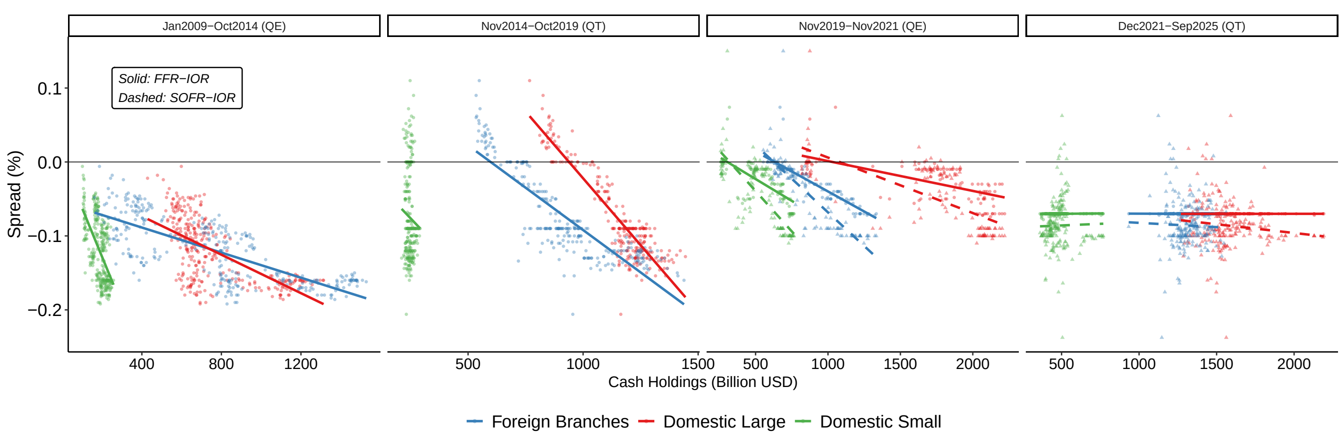


Figure 10: Cash Holdings and Funding Spreads by Bank Type This figure plots cash holdings against the FFR-IOR spread for three bank types: foreign branches (blue), large domestic banks (red), and small domestic banks (green). Each panel corresponds to a monetary policy regime. Solid lines represent FFR-IOR relationships; dashed lines represent SOFR-IOR relationships (available for post-2019 periods only). Cash holdings are measured weekly from the Federal Reserve H.8 release. Source: FRED (H.8 cash assets by bank type, DFF, SOFR, IORB, IOER).

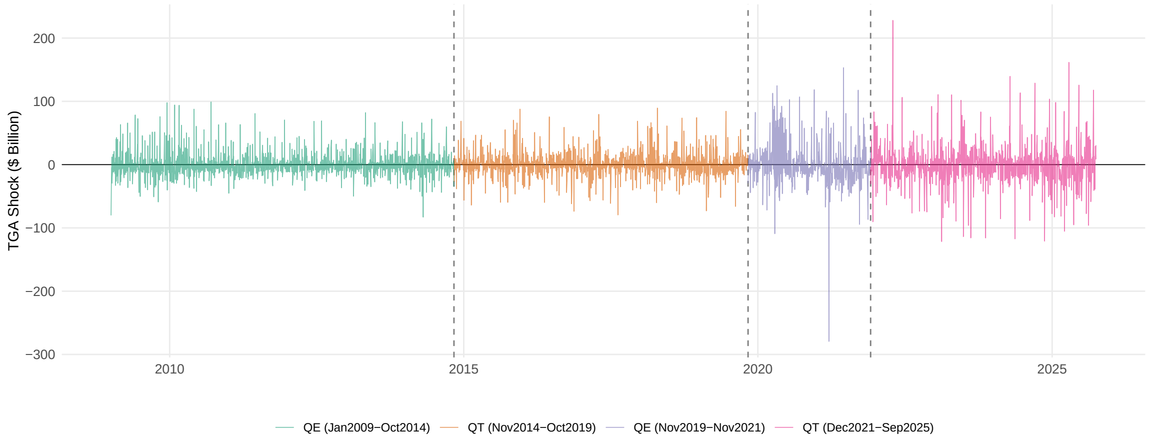


Figure 11: Daily Treasury General Account (TGA) Shocks This figure plots the daily unanticipated TGA shock, constructed as the residual from regressing daily changes in the TGA balance on calendar fixed effects (beginning/end of month, beginning/end of quarter, day of week) and anticipated Treasury issuance flows. Positive values represent reserve drains; negative values represent reserve injections. Vertical dashed lines indicate regime transitions. Colors denote monetary policy phases: QE (January 2009 to October 2014), QT (November 2014 to October 2019), QE (November 2019 to November 2021), and QT (December 2021 to September 2025). Source: U.S. Treasury Daily Treasury Statement.

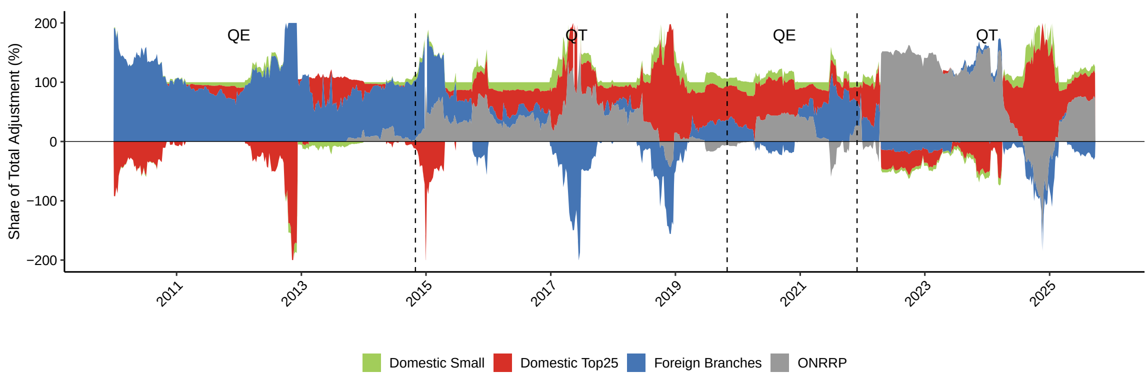


Figure 12: Composition of Reserve Adjustment to TGA Shocks over Time. This figure plots the time-varying composition of reserve adjustment across bank types, based on 52-week rolling regressions of changes in banks' reserve holdings (proxied by cash assets) on TGA shocks. Each stacked area represents the share of aggregate reserve adjustment accounted for by the corresponding bank group in a given week. Vertical dashed lines mark key policy regime transitions: the end of QE in 2014, the onset of balance sheet normalization in 2019, and the start of the 2021 QE episode. Values that exceed 200% or -200% are trimmed for the readability. The underlying regression specification follows Table 6.

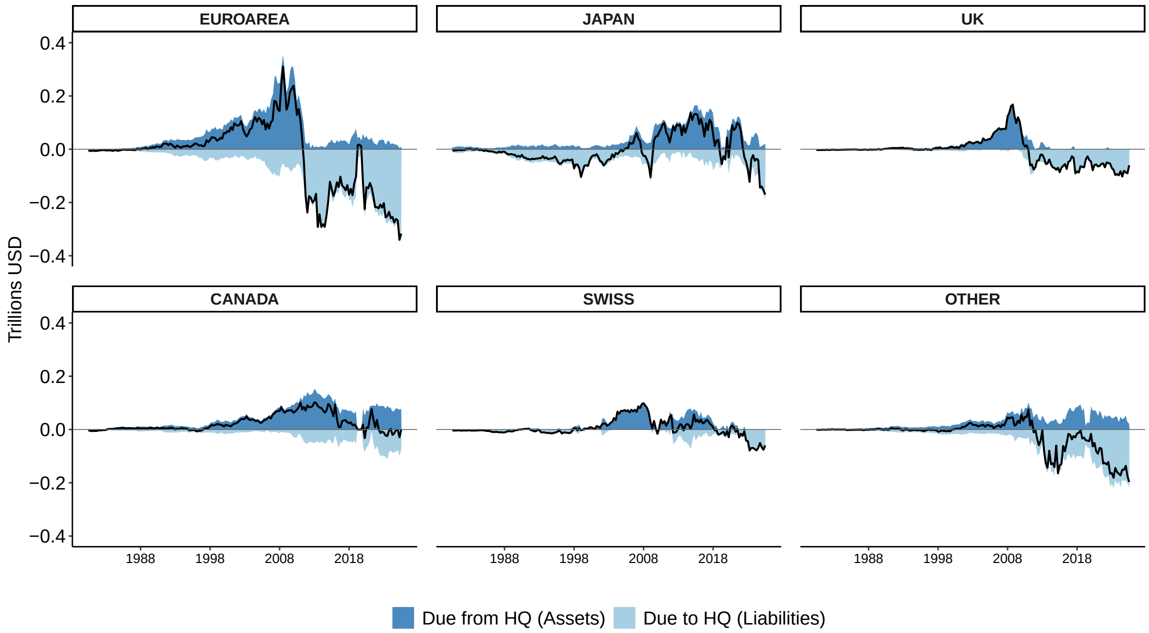


Figure 13: Headquarters (HQ) Funding Flows of Foreign Bank Branches This figure illustrates the intra-group funding flows between U.S. branches of foreign banks and their headquarters (HQ), disaggregated by jurisdiction. Positive values indicate lending from branches to HQ, while negative values represent borrowing from HQ to fund branch operations. The black solid line indicated the net balances. Source: FFIEC 002 reports

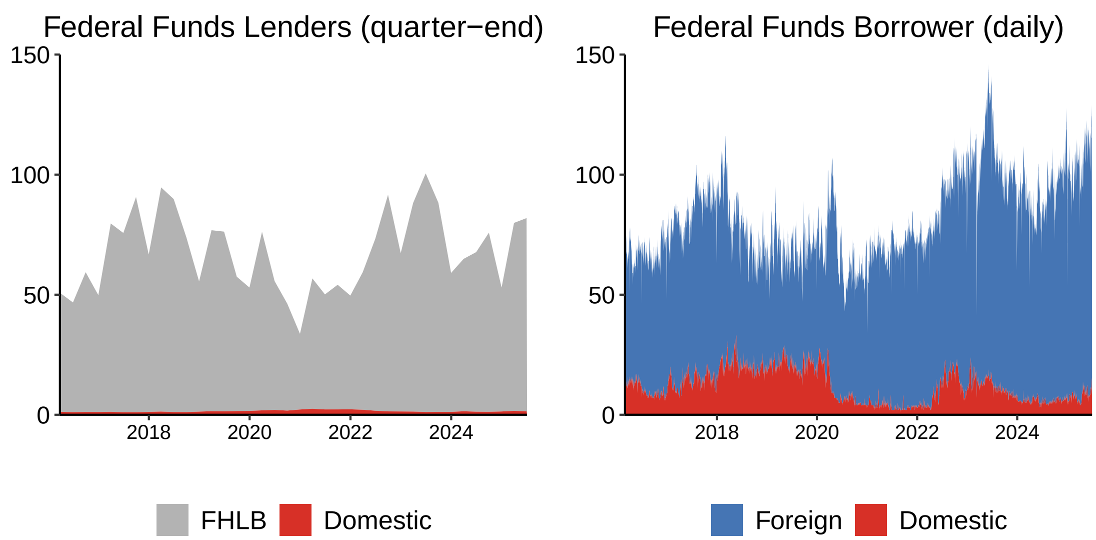


Figure 14: Federal Funds market dynamics Source: FFIEC 002, NY Fed, FHLB

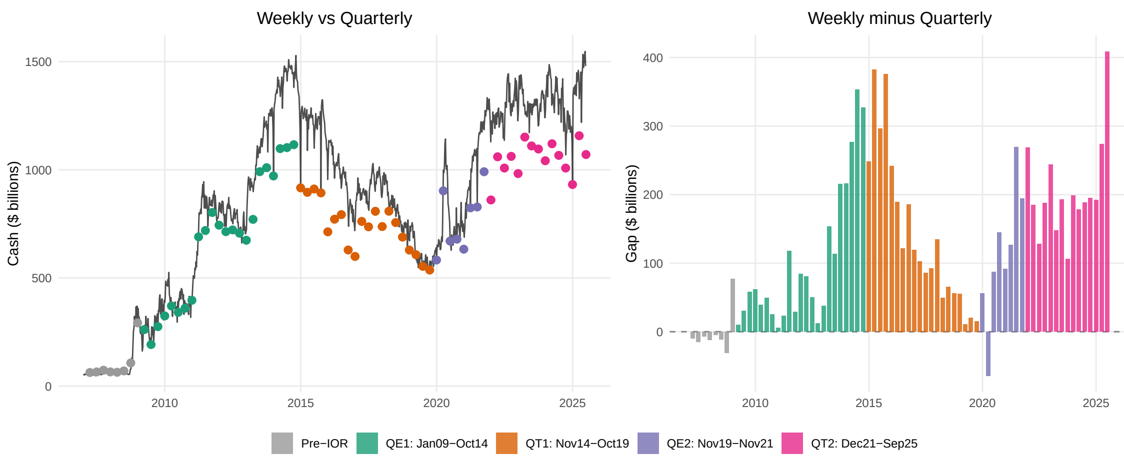


Figure 15: Window Dressing Gap Over Time: Foreign Bank Branches This figure plots the window dressing gap for U.S. branches of foreign banks, defined as the difference between weekly cash holdings (from the H.8 release, measured the week before quarter-end) and quarter-end cash balances (from FFIEC 002 filings). Panel A shows the window dressing amount in billions of dollars; Panel B shows the window dressing as a percentage of pre-quarter-end cash. Vertical dotted lines indicate regime transitions. Two patterns emerge: systematic quarter-end compression is largely absent before the introduction of interest on reserves in 2008, and the magnitude of the window dressing gap declines during quantitative tightening, particularly around 2019. Source: FFIEC 002 reports and Federal Reserve H.8 release.

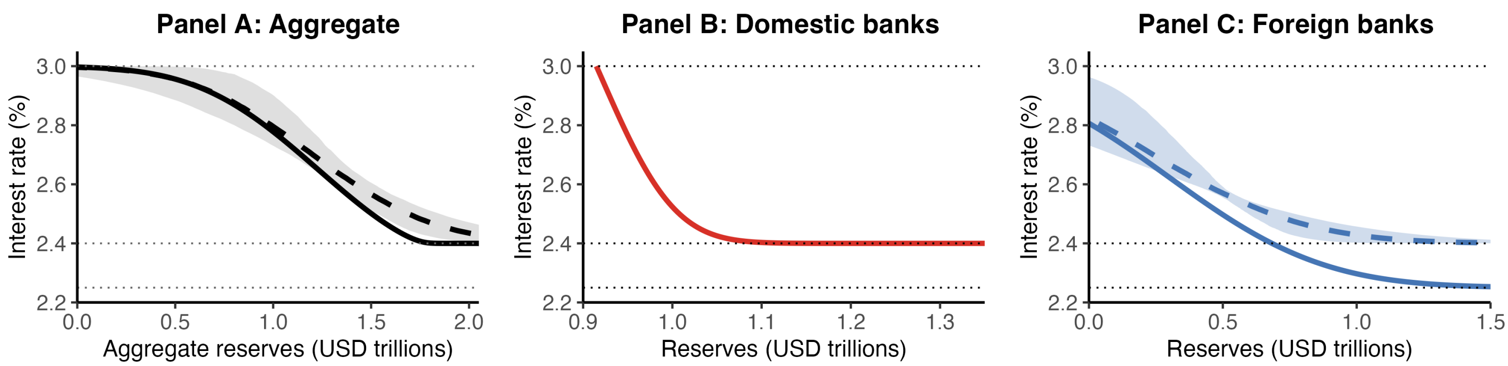


Figure 16: Model-Implied Inverse-Demand Curves This figure plots simulated reserve demand curves for (A) the aggregate, (B) domestic banks, and (C) foreign banks. The dashed lines show the mean simulated relationship between reserves and the short-term interest rate, while shaded areas represent Monte Carlo uncertainty bands arising from foreign banks' balance-sheet fluctuations due to quarter-end window dressing. The kink in Panel A reflects the transition between the ample- and scarce-reserves regimes, endogenously determined by the behavior of foreign banks. Horizontal dotted lines mark ONRRP (2.25%), IOR (2.40%), and DW (3.00%). Quarter-end mainly shifts the foreign schedule up/left and tilts the aggregate curve, while the domestic schedule is essentially unchanged.

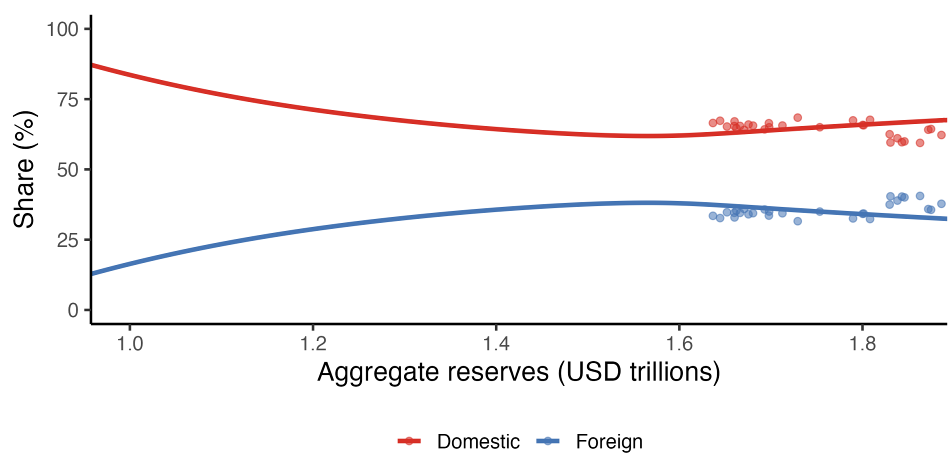


Figure 17: Reserve Shares: Model vs. Data Scatter points plot weekly reserve shares for domestic (red) and foreign (blue) banks over the QT sample (Dec 2018–Jul 2019), excluding quarter-ends and restricted to observations inside the policy corridor. Solid lines show the model-implied shares using the calibrated parameters in Table 11 with no additional fitting. The x-axis is aggregate reserves (USD trillions); the y-axis is the share in percentage. Over the observed range of 1.60–1.85 trillion, the model tracks the data closely: domestic shares decline modestly while foreign shares rise and then level off, consistent with foreign banks absorbing the marginal dollar in the ample-reserves regime.

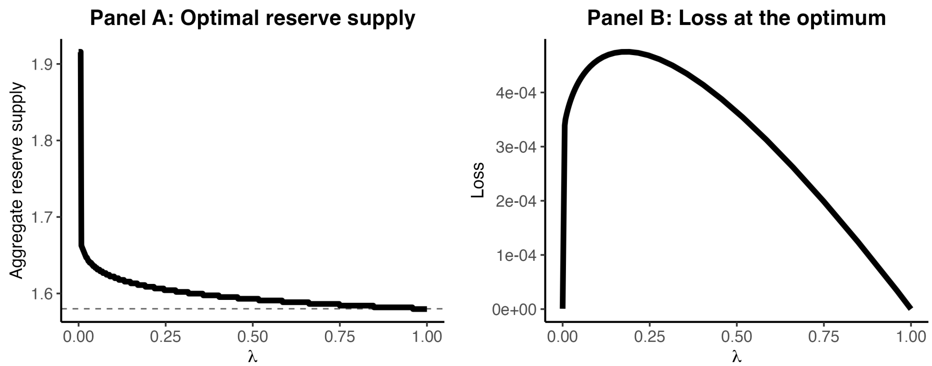


Figure 18: Optimal reserve supply and loss at the optimum under deterministic path. Panel A plots the optimal aggregate reserve supply $M^*(\lambda)$ as a function of the policy weight λ. The dashed line marks the ample-reserve kink threshold, $M_{\text{agg}}^{\text{flat}} = 1.58$ trillion. The policy target rate is set at $i_{\text{target}} = 2.375\%$, the midpoint of the FOMC's federal funds target range during the calibration window (lower limit: 2.25%, upper limit: 2.50%). Panel B shows the minimized loss $\mathcal{L}(M^*(\lambda))$, which is strictly convex and peaks at intermediate values of λ, illustrating the trade-off between rate precision and balance-sheet size.

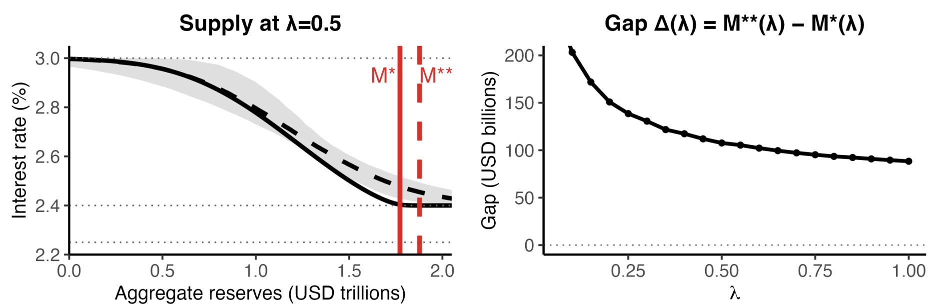


Figure 19: Optimal reserve supply and policy gap under uncertainty. This figure illustrates the effect of uncertainty in foreign banks' reserve demand on the Federal Reserve's optimal reserve supply. Panel A shows the calibrated aggregate reserve demand curve with Monte Carlo uncertainty bands reflecting quarter-end fluctuations in foreign banks' balance sheets. The solid red line (M^*) denotes the optimal supply under full information, while the dashed red line (M^{**}) denotes the optimal supply under uncertainty, implying a gap of approximately \$108 billion when $\lambda = 0.5$. Panel B plots the policy gap $\Delta(\lambda) = M^{**}(\lambda) - M^*(\lambda)$ across different weight parameters. The gap declines with higher weight on rate precision, indicating that uncertainty in foreign banks' behavior increases the precautionary reserve buffer required for rate stability during QT.

9 Tables

Table 1

Summary Statistics (All Banks)

This table presents summary statistics for the main variables used in this paper. The sample includes U.S. bank holding companies and standalone banks over the period 1980Q1–2025Q3. We report key balance sheet variables, including total assets, reserves, deposits, uninsured deposit ratios, and branch asset ratios. All monetary values are reported in millions of USD. The sample is split into two subperiods: 1980–2008 and 2009–2025, reflecting pre- and post-interest on reserves (IOR) introduction.

	Full Sample	1980–2008	2009–2025
Number of Samples	1,240,691	908,958	331,733
Foreign Banks (%)	3.89	4.32	2.72
G-SIBs (%)	-	-	0.54
Assets (million USD)			
Median	78.07	55.24	214.72
Average	1,831.75	769.31	4,742.89
SD	40,166.08	16,125.90	72,868.07
25 Percentile	33.10	26.12	98.52
75 Percentile	218.17	130.88	540.16
Reserves (million USD)			
Median	2.87	1.98	12.30
Average	123.14	7.12	441.03
SD	3,539.20	195.01	6,826.82
25 Percentile	1.13	0.84	5.12
75 Percentile	7.75	4.10	31.95
Deposit (million USD)			
Median	65.19	46.55	179.61
Average	1,167.25	472.24	3,071.60
SD	26,307.59	7,688.69	49,208.83
25 Percentile	27.76	22.00	82.18
75 Percentile	178.02	107.41	443.52
Uninsured Deposit Ratio (%)			
Median	0.00	0.00	0.00
Average	7.32	7.84	5.88
SD	21.81	23.04	17.92
25 Percentile	0.00	0.00	0.00
75 Percentile	2.69	5.35	0.00

Table 2

Summary Statistics (Foreign Banks)

This table presents summary statistics for the main variables used in this paper. The sample includes foreign bank holding companies (the aggregate of subsidiaries, branches, and agencies) over the period 1980Q1–2025Q3. We report key balance sheet variables, including total assets (within US), reserves, deposits, uninsured deposit ratios, and branch asset ratios. All monetary values are reported in millions of USD. The sample is split into two subperiods: 1980–2008 and 2009–2025, reflecting pre- and post-interest on reserves (IOR) introduction.

	Full Sample	1980–2008	2009–2025
Number of Samples	48,253	39,246	9,007
G-SIBs (%)	-	-	14.34
Assets (million USD)			
Median	486.16	332.61	3,689.07
Average	7,664.38	2,951.55	28,199.52
SD	29,247.56	9,998.24	60,241.13
25 Percentile	133.64	105.33	831.51
75 Percentile	2,547.60	1,512.74	17,446.45
Reserves (million USD)			
Median	0.61	0.31	542.90
Average	1,226.45	16.86	6,496.96
SD	6,404.91	378.93	13,601.75
25 Percentile	0.07	0.05	50.53
75 Percentile	9.54	1.93	5,326.66
Deposit (million USD)			
Median	168.45	109.44	1,358.86
Average	4,288.85	1,658.73	15,749.03
SD	18,803.75	6,008.27	39,693.83
25 Percentile	17.70	10.96	217.06
75 Percentile	1,182.92	663.76	8,639.88
Uninsured Deposit Ratio (%)			
Median	100.00	100.00	100.00
Average	85.65	86.00	84.09
SD	32.59	33.09	30.26
25 Percentile	100.00	100.00	80.31
75 Percentile	100.00	100.00	100.00
Branch Asset Ratio (%)			
Median	100.00	100.00	100.00
Average	84.52	84.37	85.19
SD	35.24	36.17	30.76
25 Percentile	100.00	100.00	96.77
75 Percentile	100.00	100.00	100.00

Table 3

First Stage: TGA Shock and Money Market Spreads

This table presents first-stage regressions of daily changes in money market spreads on TGA shocks. Panel A uses the federal funds rate minus IOR spread; Panel B uses the SOFR minus IOR spread (available from April 2018). The coefficient on ΔTGA represents the basis point change in the spread per $100 billion increase in TGA (reserve drain). Columns (1)–(3) use the full sample with progressively more controls. Columns (4) and (5) split by QE periods and QT periods. Column (1) reports IID standard errors; columns (2)–(5) report standard errors clustered by FOMC period. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Fed Funds Rate Spread

	Δ(FFR − IOR) (bps)				
	(1)	(2)	(3)	(4)	(5)
ΔTGA ($100bn)	0.3473***	0.3461***	0.3629***	0.5689***	0.1894
	(0.0915)	(0.1261)	(0.1290)	(0.1882)	(0.1846)
FOMC Period FE		✓	✓	✓	✓
ΔONRRP Control			✓	✓	✓
Sample	Full	Full	Full	QE	QT
Observations	4,208	4,208	4,208	1,968	2,240
R^2	0.003	0.024	0.034	0.080	0.017

Panel B: SOFR Spread

	Δ(SOFR − IOR) (bps)				
	(1)	(2)	(3)	(4)	(5)
ΔTGA ($100bn)	0.9753***	1.0068***	1.0579***	0.8689*	1.1175***
	(0.3143)	(0.2807)	(0.2832)	(0.4663)	(0.3455)
FOMC Period FE		✓	✓	✓	✓
ΔONRRP Control			✓	✓	✓
Sample	Full	Full	Full	QE	QT
Observations	1,814	1,814	1,814	480	1,334
R^2	0.005	0.013	0.014	0.023	0.022

Table 4

Second Stage: Money Market Spread and Fed Funds Borrowing

This table presents second-stage IV regressions of daily changes in interbank borrowing on money market spreads, instrumented by TGA shocks using Fed Funds borrowing volumes. The dependent variable is the log change in borrowing volume. Columns (1) and (4) report baseline IV estimates with IID standard errors. Columns (2)–(3) and (5)–(6) add FOMC period fixed effects with standard errors clustered by FOMC period. Columns (3) and (6) additionally control for ONRRP changes. The F-statistic tests the strength of the first-stage relationship between TGA shocks and the FFR-IOR spread. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Δ log(Fed funds borrowing)					
	Foreign			Domestic		
	(1)	(2)	(3)	(4)	(5)	(6)
ΔFFR-IOR	0.2679*	0.2714	-0.0425***	-0.2370	-0.2460	-0.0123
	(0.1524)	(0.1823)	(0.0131)	(0.1441)	(0.1766)	(0.0296)
FOMC Period FE		✓	✓		✓	✓
Δ ONRRP	✓	✓	✓	✓	✓	✓
Observations	2,484	2,484	2,484	2,484	2,484	2,484
R^2	-15.372	-15.184	-0.13521	-2.7672	-2.8672	0.00505
F-test (1st stage), d_DFFIOR	3.5246	3.5377	19.440	3.5246	3.5377	19.440

Table 5

Federal Funds Borrowing Response to Spread Changes by Regime

This table reports second-stage IV estimates of the effect of funding spread changes on federal funds borrowing volumes, separately by bank type and monetary policy regime. The dependent variable is the daily log change in federal funds borrowing. Columns (1)–(2) cover QE periods; columns (3)–(4) cover QT periods. ΔFFR-IOR is instrumented using TGA shocks. All specifications include FOMC period fixed effects. Standard errors are heteroskedasticity-robust. The F-test reports the first-stage F-statistic for the TGA shock instrument. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Δ log (Fed funds borrowing)			
	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)
ΔFFR-IOR	-0.1457***	-0.1547*	-0.0355***	0.0083
	(0.0289)	(0.0879)	(0.0132)	(0.0308)
FOMC Period FE	✓	✓	✓	✓
Δ ONRRP Control	✓	✓	✓	✓
Period	QE	QE	QT	QT
Observations	486	486	1,998	1,998
R^2	-1.0320	-0.05623	-0.07831	0.00204
F-test (1st stage)	21.495	21.495	14.283	14.283

Table 6

TGA Shock Absorption by Bank Type and Monetary Policy Regime

This table reports coefficient estimates from regressing weekly changes in cash holdings (H.8 data) on weekly TGA shocks, separately for each monetary policy regime. The dependent variable in columns (1) and (5) is the change in cash holdings of foreign bank branches; columns (2) and (6) use large domestic banks (top 25); columns (3) and (7) use small domestic banks; columns (4) and (8) use ON RRP take-up. Panel A covers the first QE/QT cycle (2009–2019); Panel B covers the second cycle (2019–2025). All specifications include FOMC period fixed effects. Standard errors are clustered by FOMC period. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	QE Periods				QT Periods			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: First QE/QT Cycle							
	QE1: Jan 2009 – Oct 2014				QT1: Nov 2014 – Oct 2019			
ΔTGA	-0.6315***	0.0719	-0.0225*	-0.0051	-0.0403	-0.2516***	-0.0787***	-0.1718
	(0.1202)	(0.0886)	(0.0126)	(0.0468)	(0.1109)	(0.0755)	(0.0204)	(0.1191)
Observations	304	304	304	304	261	261	261	261
R^2	0.21	0.12	0.10	0.13	0.02	0.14	0.19	0.02
	Panel B: Second QE/QT Cycle							
	QE2: Nov 2019 – Nov 2021				QT2: Dec 2021 – Sep 2025			
ΔTGA	-0.1064	-0.1364	-0.0418*	-0.1360	-0.0012	-0.3277***	-0.0723***	0.7255
	(0.0653)	(0.0848)	(0.0253)	(0.1014)	(0.0754)	(0.0977)	(0.0207)	(0.6913)
Observations	108	108	108	108	200	200	200	200
R^2	0.17	0.42	0.23	0.36	0.09	0.19	0.23	0.03
Dep. Variable	Foreign	Large	Small	ONRRP	Foreign	Large	Small	ONRRP

Table 7

2007 Branch Ratio in Foreign Banks and Reserve Holdings

This table reports regressions of log reserves on foreign-bank status and the 2007 branch-to-asset ratio: $Y_{it}) = \beta_1 \text{Foreign}_i + \beta_2 \text{BranchRatio}_i^{2007} + \Gamma X_{it} + \alpha_t + \epsilon_{it}$, where Foreign is an indicator equal to one if the bank is foreign-owned, and BranchRatio_07Q1 is the share of U.S. assets held in branch form as of 2007Q1. Specifications (1)–(3) use the full sample of banks, while specification (4) restricts to foreign banks only. All regressions control for log assets and log equity capital, and include quarter fixed effects as well as BHC fixed effects. Standard errors are clustered at the BHC level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	log(Reserves)			
	(1)	(2)	(3)	(4)
Foreign (F)	2.200***		2.126***	
	(0.2006)		(0.2010)	
$\text{BranchShare}_i^{2007}$		4.124***	2.001***	1.248***
		(0.2526)	(0.2687)	(0.1880)
BHC FE	✓	✓	✓	✓
Time FE	✓	✓	✓	✓
Balance Sheet Ctrl	✓	✓	✓	✓
Sample	All	All	All	Foreign
Observations	299,213	299,213	299,213	2,357
R^2	0.44	0.43	0.44	0.85

Table 8

Spread Sensitivity: Arbitrage Buffer vs Structural Component

Dependent variable is log(window dressing gap) in columns (1)–(3) and log(quarter-end reserves) in columns (4)–(6). Window dressing gap is defined as H.8 weekly cash minus Call Report quarter-end cash for foreign branches. QE includes QE1 (Jan 2009–Oct 2014) and QE2 (Nov 2019–Nov 2021); QT includes QT1 (Nov 2014–Oct 2019) and QT2 (Dec 2021–Sep 2025). Standard errors are clustered at the BHC level. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Arbitrage Buffer log(Window Dressing Gap)			Structural Component log(Quarter-End Reserves)		
	(1)	(2)	(3)	(4)	(5)	(6)
FFR–IOR	-4.461**	-4.375	-5.671***	-0.187	-1.234	-0.113
	(1.710)	(3.440)	(1.407)	(0.690)	(1.456)	(0.468)
R^2	0.10	0.05	0.33	0.00	0.02	0.00
Observations	66	31	35	66	31	35
Sample	All	QE	QT	All	QE	QT

Table 9

Reserve Holdings and Funding Sources: BHC-Level Analysis

This table compares the relationship between reserves and three funding sources at the BHC level for foreign bank branches. Panel A uses net HQ liabilities (HQ liabilities minus HQ assets), Panel B uses uninsured deposits, and Panel C uses domestic deposits. Net HQ liabilities use the inverse hyperbolic sine (IHS) transformation to accommodate potential negative values; all other variables use log. Size control uses lagged log assets. Standard errors clustered by BHC in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: log(Reserves)			
	(1)	(2)	(3)	(4)
Panel A: Net HQ Liabilities				
asinh(Net HQ Liab)	0.750***	0.441***	0.435***	0.443***
	(0.082)	(0.086)	(0.082)	(0.080)
R^2	0.806	0.826	0.841	0.857
Observations	6,051	5,853	5,853	5,853
Panel B: Uninsured Deposits				
log(Uninsured Dep)	0.373***	0.114	0.142*	0.173*
	(0.097)	(0.076)	(0.064)	(0.069)
R^2	0.710	0.748	0.790	0.801
Observations	7,946	7,753	7,753	7,753
Panel C: Domestic Deposits				
log(Domestic Dep)	0.315***	0.080	0.046	0.075
	(0.073)	(0.058)	(0.052)	(0.056)
R^2	0.710	0.747	0.788	0.800
Observations	7,849	7,660	7,660	7,660
BHC FE	✓	✓	✓	✓
Size Control		✓	✓	✓
Time FE			✓	✓
Time × Country FE				✓

Table 10

International Interest Rate Differentials and Foreign Bank Balance Sheets

This table examines how international interest rate differentials affect reserve holdings and headquarters funding at foreign bank branches. The sample includes foreign branches from five major currency areas (Euro Area, Japan, UK, Canada, and Switzerland) from 2009Q1 to 2024Q4. Panel A uses log reserves as the dependent variable; Panel B uses the inverse hyperbolic sine (IHS) of net headquarters liabilities (liabilities to foreign headquarters minus claims on headquarters). FF–IOR is the federal funds rate minus the interest on reserves. US IOR – Country$_i$ IOR is the international IOR differential, calculated as the US interest on reserves minus the deposit facility rate of each bank's home-country central bank. Column (1) includes only the domestic spread; column (2) includes only the international spread; column (3) includes both; column (4) adds time fixed effects (which absorb the domestic spread); column (5) adds country fixed effects. All specifications include BHC fixed effects and control for lagged log assets. Standard errors clustered by BHC in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Reserves

	log(Reserves)				
	(1)	(2)	(3)	(4)	(5)
FF–IOR	-1.334		-5.084***		-5.084***
	(1.497)		(1.565)		(1.566)
US IOR – Country$_i$ IOR		0.281**	0.390***	-0.099	0.390***
		(0.109)	(0.129)	(0.129)	(0.129)
Observations	4,159	4,159	4,159	4,159	4,159
R^2	0.573	0.577	0.580	0.620	0.580
Within R^2	0.298	0.305	0.310	0.307	0.310
BHC FE	✓	✓	✓	✓	✓
Time FE				✓	
Country FE					✓
Bank Control	✓	✓	✓	✓	✓

Panel B: Net HQ liabilities

	IHS(Net HQ Liabilities)				
	(1)	(2)	(3)	(4)	(5)
FF–IOR	0.587		-1.879		-1.879
	(1.550)		(1.516)		(1.517)
US IOR – Country$_i$ IOR		0.216***	0.256***	-0.187*	0.256***
		(0.072)	(0.071)	(0.107)	(0.071)
Observations	4,159	4,159	4,159	4,159	4,159
R^2	0.411	0.420	0.421	0.467	0.421
Within R^2	0.015	0.028	0.030	0.011	0.030
BHC FE	✓	✓	✓	✓	✓
Time FE				✓	
Country FE					✓
Bank Control	✓	✓	✓	✓	✓

Table 11

Calibration parameters (QT: December 2018 – August 2019)

Parameter	Value	Definition
Policy rates		
i_{DW}	3.00%	Discount window rate
i_{IOR}	2.40%	Interest rate on reserves
i_{ONRRP}	2.25%	ONRRP rate
Reserve holdings advantage and cost		
α_D	0 bps	Domestic carry advantage
α_F	15 bps	Foreign carry advantage
β_D	50 bps	Domestic shortfall penalty
β_F	0 bps	Foreign shortfall penalty
Shock scales and reserve floors (in billion USD)		
σ_D	62.9	Domestic liquidity shock scale
σ_F	159.1	Foreign liquidity shock scale
δ_D	840.8	Domestic required reserves
δ_F	479.7	Foreign required reserves

Appendix

A1. Data Details

A1.1 Quarterly Bank-BHC Panel Data

To construct quarterly panel dataset, we obtain Call Report data and Y-9C Reports from the National Information Center (NIC) and the Federal Reserve Bank of Chicago. FFIEC 031 and 041, the Call Report data for US commercial banks and foreign bank subsidiaries provides detailed balance sheet and income statement information.[36] FFIEC 002, 'Report of Assets and Liabilities of US Branches and Agencies of Foreign Banks' provide balance sheet data for foreign branches and agencies. The Y-9C reports contain similar data for bank holding companies (BHCs) level. The Call Report data is available from 1976Q1 for domestic banks and foreign bank subsidiaries, while data on foreign bank branches begins in 1980Q2. While balance sheet variables on foreign branches and agencies are similar to those for domestic banks, there are differences in RSSD variables and some contents. For instance, they do not contain income statement and capital accounts while providing insights on flows with their headquarters.[37] Y-9C reports are available starting in 1986Q2. To maintain consistency in our sample, we exclude reporters of FR 2886a (New York Investment Companies) and FR 2886b (Edge Act and Agreement Corporations) from our foreign samples and concentrate on branches and agencies and subsidiaries.

We link bank-level data to their respective BHCs using the following approach. When available, we rely on RSSD9347 and RSSD9348 to match RSSD9001 (individual bank IDs) to their parent holding companies. In cases where these identifiers are missing or inconsistent over the time, we supplement the linkage using the Relationship Data provided by the NIC. We also refer to 'Events & Changes' on FDIC website to account for some large mergers and liquidations. Foreign banking organizations (FBOs) samples post December 2008, we also rely on Federal Reserve's Structure and Share Data for mapping. For banks headquartered in Canada, the Eurozone, Japan, Switzerland, and the UK, we manually check and refine these assignments to ensure accurate representation at the top-tier BHC level.

To classify foreign banking organization's institution by country and type, we follow

[36] Typical economic analyses of Call Report data primarily use FFIEC 031, 041, and 051 reports. We exclude FFIEC 051 due to its focus on small banks. Designed for institutions with total assets under $5 billion, the FFIEC 051 report does not separately report reserve holdings, instead aggregating them with other cash assets, preventing us from observing reserves holdings. Additionally, smaller banks are less likely to engage in active reserve management, making them less relevant for studying reserve dynamics. To ensure accurate measurement of reserves and focus on institutions with significant reserve activity, we restrict our sample to FFIEC 031 and 041 filers.

[37] Previous literature that have utilized FFIEC002 include Aldasoro et al. (2022); Cetorelli and Goldberg (2011); Fillat et al. (2018).

Aldasoro et al. (2022), using RSSD9331 and RSSD9209.[38] When these variables are missing, we impute the country classification based on the bank's most recent filings or refer to the Federal Reserve's Structure and Share Data for verification. If a bank is headquartered in Austria, Belgium, Croatia, Cyprus, Estonia, Finland, France, Germany, Greece, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia, or Spain, we classify them as 'Euroarea'. To identify G-SIBs and domestic US banks subject to relevant regulations, we refer to data at Financial Stability Board (FSB) and Federal Reserve Board and other regulatory agencies. We then compile regulatory ratios from manually collected Basel III Pillar 3 Disclosure Reports and supplement this information with Bloomberg data before merging it into our dataset.

We construct our key variables—including reserves, assets, and balance sheet components—from Call Report data while ensuring consistency over time. For domestic banks and foreign subsidiaries, we adopt methodologies outlined in Kashyap and Stein (2000); Drechsler et al. (2017, 2021); Correia et al. (2024). Where applicable, we adhere to definitions used in these sources to maintain continuity. However, for foreign banks and more recent data, we develop new variable definitions to account for changes in FFIEC 002 reports. In our main analysis, we compute variables at both the BHC-merged level and the standalone bank level, ensuring robustness in our empirical results. We omit samples that report non positive reserves or assets. When calculating the shares of reserve holdings in Figure 3, we only compare the holdings of domestic and foreign (the aggregate of subsidiaries and branches and agencies holdings) banks, not considering the reserve holdings of other entities. In other words, we are not considering the shares of deposits holding referred as 'Term deposits held by depository institutions,' 'U.S. Treasury, General Account,' 'U.S. Treasury, Supplementary Financing Account,' 'Foreign official accounts,' and 'Other deposits' in H.4.1 report for our share analysis.

Our primary outcome is reserve holdings, which are held by depository institutions with a master account at the Federal Reserve. These institutions include nationally chartered banks, eligible state-chartered banks, and other entities classified under Section 19(b) of the Federal Reserve Act. In March 2020, reserve requirement ratios were reduced to zero, so reserve holdings in the post-2020 period reflect banks' discretionary liquidity management rather than statutory reserve mandates.

[38]We are deeply grateful to the authors for their invaluable guidance on classification methodologies and assistance.

A1.2 TGA Shock Related Data

For our analysis of supply shocks on short-term interest rates, we construct key variables from multiple sources. Daily Treasury General Account (TGA) balances are obtained from the Treasury website and Haver Analytics. Policy rates and market interest rates—including the Federal Funds Rate (FFR), the effective lower bound, and the Interest on Reserves (IOR)—are sourced from the Federal Reserve Bank of St. Louis (FRED). The general collateral (GC) repo rate and LIBOR rate come from Bloomberg.

To account for seasonal fluctuations in TGA shock, we clean Bräuning (2017); Correa et al. (2020). Specifically, we regress TGA changes on beginning-of-month, end-of-month, beginning-of-quarter, end-of-quarter, and day-of-week (Monday to Friday) fixed effects to isolate supply-driven fluctuations. To further refine this adjustment, we incorporate the expected maturity dates and volumes of Treasury bonds, extracting the residuals as a measure of supply shocks using net issuance data at Treasury website. To align with bank balance sheet reporting schedules, we aggregate daily TGA balances and Overnight Reverse Repurchase Agreement (ONRRP) balances into weekly averages. The aggregation period runs from Thursday to Wednesday, matching the weekly bank cash holding data from the H.8 report.

We rely on the FR 2420 for the daily transaction data on federal funds purchased, Eurodollar transactions, certificates of deposits. The breakdown of bank type on this data is only available from March 1st, 2016. Our primary data source for weekly bank balance sheet information is the H.8 Report, Assets and Liabilities of Commercial Banks in the United States, published by the Federal Reserve Board. The H.8 data, compiled from FR 2644 weekly reporting banks, provides an aggregate view of commercial bank assets and liabilities, allowing us to track trends in bank reserves. We use non-seasonally adjusted Wednesday values. Our sample period spans from January 1, 2009, to July 31, 2024, except for short term liabilities analysis which spans from March 1st, 2016 to July 31, 2024.

The classification of foreign banks in this section differs from our bank panel analysis. In the H.8 report, foreign-related institutions refer specifically to branches and agencies of foreign banks, as well as Edge Act and agreement corporations, but exclude foreign subsidiaries as well as International Banking Facilities (IBFs). As a result, the foreign banks examined in the supply shock section primarily capture the reserve holdings of foreign branches, not subsidiaries.[39]

Since reserves are not directly reported in the H.8 data, we use cash assets as a proxy variable. The H.8 cash holding category includes vault cash, cash items in process of collection, balances due from depository institutions, and reserves. While this measure closely tracks

[39] Edge Act and agreement corporations hold a negligible share of reserves.

reserves, it differs from the Call Report definition, which also includes balances due from foreign central banks and foreign financial institutions.[40]

We verify that the deviation between H.8 cash assets and actual reserve balances is minimal following the introduction of Interest on Reserves (IOR). Figure A7 left panel compares Call Report aggregate cash balances, Call Report aggregate reserves, H.4.1 reported Fed reserves, and H.8 bank cash holdings. Before the Global Financial Crisis in 2008, the discrepancy between cash assets reported in H.8 (purple line) and reserves at the Fed (light blue line) was substantial, but post-GFC, the two measures converge. Notably, weekly cash and reserve balances dip around quarter-ends, consistent with window-dressing, where banks adjusting their balance sheets for regulatory reporting purposes. This suggests that weekly data is better suited to capture banks' reserve management behavior and arbitrage sensitivity. Similarly, Figure A7 right panel shows that foreign-related banks' cash holdings in the H.8 data closely align with the reserve holdings of foreign branches reported in Call Report FFIEC 002 at quarter-end, while diverging at non-quarter-end periods, where H.8 cash holdings appear larger.

The H.8 dataset further classifies domestic banks into large domestically chartered banks, and small domestically chartered banks. Large domestically chartered banks refer to the top 25 banks ranked by domestic assets as of the most recent Call Report. These institutions include all U.S. Global Systemically Important Banks (G-SIBs) as well as most Category II to IV banks, which are subject to heightened prudential standards. For instance, in 2019Q1, the top 25 banks by total assets included all Category II to IV banks, making them a distinct regulatory group. Small domestically chartered banks consist of all other domestic banks reporting FR 2644 data. Since FR 2644 reporting is voluntary, the H.8 sample—approximately 850 institutions—is smaller than the full Call Report sample.

A2. Discussion on Intraday Liquidity

A potential determinant of reserve demand that we do not explicitly test in our analysis is intraday liquidity management. While intraday demands on reserves drive the distribution of transactions within the day across each settlement service (Copeland et al., 2021; Yang, 2020), we argue that their impact on foreign banks, particularly on branches, is limited due to regulatory and structural differences. For US banks, Dodd Frank as well as SR 14-1 have increased the need to hold reserves to meet both overnight and intraday liquidity obligations.

[40] The Call Report (FFIEC 031, 041, 002) defines cash to include balances due from foreign financial institutions (RCON0070). While a more precise measure of Call Report cash (excluding foreign balances) can be constructed, this breakdown is only reported by banks with total assets exceeding $300 million, limiting full comparability.

For foreign banks, the situation is different. Only IHCs (subsidiaries) are subject to the similar US supervision and regulation. Therefore, branches of foreign banks have been exempted thus not directly affected by the constraints. Large foreign banks are subject to LCR and other Basel III requirements, however, branches are not required to hold separate liquidity.

Meanwhile, foreign bank subsidiaries and branches operating in the United States have access to the Federal Reserve's daylight overdraft and discount window under the same eligibility criteria and terms as domestic depository institutions. Foreign banks with multiple subsidiaries and branches may access the discount window in more than one Federal Reserve District. In most emergency liquidity facilities, foreign banks generally receive treatment equivalent to that of domestic banks.[41]

For this reason, we believe that intraday liquidity requirements have limited impact on foreign banks' reserve demand. While domestic banks must actively manage reserves to meet intraday liquidity needs, foreign branches—unconstrained by the US liquidity requirements—rely more on parent-bank funding and the Federal Reserve's daylight overdraft facilities to meet short-term payment obligations. This structural distinction means that intraday liquidity considerations are less likely to be a primary driver of reserve holdings among foreign branches.

Additionally, the lack of granular daily data prevents a more precise examination of how intraday liquidity considerations might influence foreign banks' reserve management. Given that our analysis focuses on quarterly reserve adjustments in response to interest rate spreads, incorporating intraday dynamics would require a more detailed transaction-level dataset that is not publicly available. While intraday liquidity management is undoubtedly an important aspect of reserve behavior for some institutions, we find that its relevance for foreign banks—especially branches—is likely secondary to broader funding and arbitrage incentives.

A3. Relevant Regulations: Leverage Ratio and Liquidity Coverage Ratio

A3.1 Basel III standards

Basel III is a global regulatory framework developed by the Basel Committee on Banking Supervision (BCBS) to strengthen the resilience of internationally active banks. Introduced in response to the 2008 financial crisis, Basel III enhances capital adequacy, introduces stricter liquidity requirements, and limits excessive leverage to reduce systemic risk. The framework

[41] During the Global Financial Crisis, foreign banks made extensive use of the discount window and other emergency liquidity facilities. In October 2008, foreign banks accounted for at least 70% of outstanding discount window borrowing (Keoun, 2011). Benmelech (2012) finds that over 50% of Term Auction Facility (TAF) funds were borrowed by foreign banks.

builds on Basel I and Basel II by raising minimum capital requirements, implementing the Liquidity Coverage Ratio (LCR) and Net Stable Funding Ratio (NSFR), and incorporating a non-risk-based leverage ratio to serve as a backstop to risk-weighted capital measures. Basel III applies to banks operating across jurisdictions, with national regulators responsible for its phased implementation and enforcement.

Leverage Ratio: The Basel III leverage ratio framework has evolved over time to strengthen the resilience of the global banking system by constraining excessive leverage. Before its introduction, there was no globally standardized leverage ratio, as bank capital regulation primarily capital requirements. However, the Global Financial Crisis (GFC) exposed weaknesses in this approach, revealing that risk-weighted asset (RWA) calculations often understated actual risk exposures. The leverage ratio was introduced as a supplementary, non-risk-based measure to serve as a backstop against potentially miscalculated RWAs. Between 2009 and 2013, no formal requirements were in place as discussions around the framework were still ongoing. In January 2014, the Basel Committee on Banking Supervision (BCBS) introduced the Basel III leverage ratio framework and disclosure requirements, setting a minimum Tier 1 leverage ratio of 3% and mandating public disclosure to enhance market discipline. This marked the beginning of a parallel run period, allowing banks to report their leverage ratios without mandatory compliance (Basel Committee on Banking Supervision, 2014). In 2017, the BCBS refined the leverage ratio by revising the definition of exposure, incorporating insights from the parallel run to ensure the leverage ratio functioned as an effective backstop to risk-based capital requirements (Basel Committee on Banking Supervision, 2017a).

By January 2018, the leverage ratio transitioned from a reporting requirement to a binding Pillar 1 capital standard, requiring banks to maintain a minimum leverage ratio of 3% (Basel Committee on Banking Supervision, 2014). This transition was based on the original 2014 framework but incorporated the revised exposure definitions introduced in 2017. The 2017 revisions also introduced a leverage ratio buffer for Global Systemically Important Banks (G-SIBs), requiring them to hold an additional leverage ratio buffer equal to 50% of their risk-weighted higher loss absorbency requirement. Originally scheduled for implementation in January 2022, the redefinition of the exposure and the G-SIB leverage ratio buffer was postponed to January 2024 to provide banks with additional time to adapt to the new requirements (Basel Committee on Banking Supervision, 2017b).

Liquidity Coverage Ratio: Prior to the Global Financial Crisis (GFC), there was no globally standardized liquidity requirement. While the importance of liquidity risk was recognized, its measurement and regulation remained fragmented across jurisdictions, reflecting the structural differences in national financial markets. Basel II largely overlooked liquidity risk, focusing instead on capital adequacy. However, the crisis exposed significant weaknesses in banks' liquidity positions, highlighting the need for an internationally coordinated approach.

In December 2010, the Basel Committee on Banking Supervision (BCBS) introduced the Liquidity Coverage Ratio (LCR) as part of Basel III, requiring banks to hold sufficient high-quality liquid assets (HQLA) to withstand a 30-day period of severe liquidity stress. The LCR was initially scheduled for full implementation by January 1, 2015, but following concerns over its impact on financial markets and credit supply, the BCBS revised the framework in January 2013. The revisions adjusted the definitions of HQLA and net cash outflows while introducing a phase-in schedule, starting at 60% in 2015 and increasing annually to reach 100% by 2019 (Basel Committee on Banking Supervision, 2013).

With the full implementation of the LCR in 2019, banks are required to maintain a minimum 100% LCR under normal conditions, ensuring that short-term liquidity risk is adequately managed. During stress periods, banks are permitted to draw down their liquidity buffers, recognizing the role of HQLA in mitigating shocks rather than serving as a static regulatory minimum. This evolution marks a shift in Basel regulation, expanding its scope beyond capital adequacy to explicitly incorporate liquidity as a pillar of financial stability.

A3.2 US

The Basel III framework allows member countries a degree of national discretion in its implementation, enabling jurisdictions to tailor standards to their domestic banking systems. For this reason, the implementation schedule and the scope differ between the Basel III and the US standards.

Leverage Ratio: Leverage ratio requirements have long been integral to US banking regulation, predating international standards such as Basel III. While the definition of the numerator differs (Basel III being stricter), US banks have been historically subject to a minimum Tier 1 leverage ratio of 4%, with a 5% threshold for well-capitalized status since 1985.

In 2014, US regulators introduced the Supplementary Leverage Ratio (SLR) as part of the Basel III reforms, a year ahead of the Basel Committee's 2015 timeline. The SLR requires large, internationally active banks to maintain a minimum Tier 1 capital of 3% of their total leverage exposure, which includes both on-balance-sheet assets and certain off-balance-sheet exposures, which is consistent with Basel III definition. The US implementation did not include a phase-in period; banks were expected to comply with the SLR requirements upon the rule's effective date.

To further enhance the resilience of systemically important banks, US regulators implemented the Enhanced Supplementary Leverage Ratio (eSLR) in 2014. This rule mandates US G-

SIBs to maintain an additional buffer of 2% at the BHC level, raising the minimum SLR requirement to 5%. At the depository institution level, these banks are required to maintain a 6% SLR to be considered "well-capitalized."

In 2018, the Economic Growth, Regulatory Relief, and Consumer Protection Act (EGRRCPA) was enacted, mandating adjustments to the SLR for specific banking organizations.[42] This led banks under $250 billion total assets to be excluded from the SLR requirement. The final rule implementing this change became effective on April 1, 2020, aiming to more accurately reflect the leverage exposure of these specialized institutions.

In April 2020, responding to the outbreak of COVID-19, the Federal Reserve announced a temporary change to the Supplementary Leverage Ratio (SLR) requirements for large banks, intending to ease strains in the Treasury market and promote lending to households and businesses. This modification allowed banks to exclude U.S. Treasury securities and reserves at the Fed from the SLR calculation. This relief measure expired on March 31, 2021, at which point the previous SLR requirements were reinstated.[43]

Liquidity Coverage Ratio: In September 2014, US federal banking agencies finalized a rule implementing the Liquidity Coverage Ratio (LCR) for large and internationally active banking organizations. This rule established a standardized minimum liquidity requirement, mandating that these institutions hold sufficient high-quality liquid assets (HQLA) to cover projected net cash outflows over a 30-day stress period. The US LCR rule closely aligns with the Basel Committee on Banking Supervision's LCR standard but incorporates certain adjustments to address the specific characteristics of the US financial system.

A notable distinction between the US and Basel III LCR implementation lies in the phase-in schedules. The Basel III framework introduced the LCR starting at 60% on January 1, 2015, with annual increments of 10 percentage points, reaching 100% by January 1, 2019. In contrast, the US adopted a more accelerated timeline: institutions were required to maintain an LCR of 80% by January 1, 2015, increasing to 90% by January 1, 2016, and achieving full compliance at 100% by January 1, 2017.

[42] Furthermore, Section 402 of the EGRRCPA requires federal banking agencies to amend their capital regulations to allow custodial banks to exclude certain central bank deposits from the SLR calculation. This provision primarily benefits institutions predominantly engaged in custody, safekeeping, and asset servicing activities, such as The Bank of New York Mellon, Northern Trust Corporation, and State Street Corporation.

[43] Jurisdictions varied in their treatment of reserve exemptions in leverage ratio calculations. For instance, the ECB and Japan temporarily excluded reserves, while the UK had already permanently excluded them since 2016. Canada and Switzerland also implemented temporary exemptions, whereas the US reinstated its original SLR framework after March 2021.

A4. Institutional Details of the Federal Funds Market

In the post-crisis ample reserves regime, the federal funds market has become highly segmented and institutionally constrained. Although the federal funds rate (FFR) remains the Federal Reserve's primary operating target, it is no longer determined through broad-based trading among a wide range of depository institutions. Instead, as shown in Figure 14, the effective FFR is shaped by bilateral bargaining between a narrow set of participants: the Federal Home Loan Banks (FHLBs) on the lending side and branches of foreign banking organizations on the borrowing side. Understanding how the FF market evolved into this oligopoly-oligopsony structure requires highlighting two key factors: regulatory constraints and the intraday liquidity needs of the federal funds market participants.

On the lending side, the FHLBs have emerged as the dominant suppliers of federal funds. Like other GSEs, the FHLBs are ineligible to earn interest on reserve balances held at the Federal Reserve, which gives them a strong incentive to lend cash at rates below the interest on reserve balances (IOR). However, only the FHLBs remain active in the federal funds market. This is because Fannie Mae and Freddie Mac, which were placed into conservatorship in September 2008, curtailed and ultimately ceased participation in unsecured lending markets following supervisory guidance issued in the early 2010s.[44] As a result, the FHLBs are now the sole GSEs supplying liquidity to the federal funds market.

The FHLBs' participation in the federal funds market is shaped by both regulatory requirements and operational needs. Each of the 11 regional FHLB manages its own investment portfolio and liquidity position. To fund their activities, FHLBs issue debt through the Office of Finance and extend advances to member institutions. Critically, they must maintain sufficient intraday liquidity to meet debt service obligations, which are typically due around 12:00 p.m. ET. Among available short-term instruments, overnight federal funds lending is uniquely well-suited to this requirement. These transactions operate through the Fedwire Funds Service, which opens at 9:00 p.m. ET on the prior business day (T–1) and closes at 7:00 p.m. ET on the settlement day (T), offering a wide window for access.[45] More importantly, overnight fed funds loans are typically repaid by 9:00 a.m. on T+1, allowing the FHLBs to redeploy cash ahead of their mid-day funding obligations on the next day.

Federal funds also receive favorable regulatory treatment. Under contingent liquidity requirements, FHLBs must hold a minimum stock of liquid assets. Interest-bearing deposits at commercial banks are subject to unsecured credit exposure limits, whereas placements in

[44] Fannie Mae and Freddie Mac are permitted to place overnight cash in the Federal Reserve's ONRRP facility, which is a secured transaction.

[45] The Fedwire Funds Service closing time was extended from 6:30 p.m. to 7:00 p.m. ET on March 8, 2021.

the federal funds market are exempted from these constraints.[46] These features make federal funds lending not only operationally flexible but also regulatory-efficient, reinforcing its role as the FHLBs' preferred short-term investment vehicle.

Other short-term investment options available to the FHLBs are operationally or institutionally constrained in ways that limit their usefulness for intraday liquidity management. Interest bearing bank deposits, for example, are subject to strict internal single-name counterparty limits, which restrict unsecured exposure to individual institutions and make it difficult to deploy large volumes of cash. Although these placements may offer comparable returns to fed funds, they are typically less scalable under current regulatory risk frameworks. The ONRRP facility offers a risk-free overnight investment option, but its timing is misaligned with FHLBs' operational requirements. The ONRRP auction window only opens between 12:45 p.m. to 1:15 p.m. ET, and the transaction settles on a T+1 basis at 3:30 p.m. ET. This means that cash placed in the facility is unavailable on the same day, limiting its usefulness for mid-day liquidity obligations. Similarly, tri-party repo transactions unwind late around 3:30 p.m. ET. As a result, while these instruments may serve a role in broader liquidity management, they are not viable substitutes for federal funds.

On the borrowing side, foreign bank branches account for the majority of activity in the federal funds market. Several regulatory and structural features contribute to this oligopsonistic outcome. First, foreign bank branches are exempted from paying the FDIC deposit insurance assessment, which domestic banks must pay on their total liabilities, including fed funds borrowed. This exemption makes overnight borrowing less costly for foreign banks than for their domestic counterparts.

Second, many large domestic banks face binding constraints under the Supplementary Leverage Ratio (SLR) or enhanced SLR (eSLR), which raise the marginal cost of balance sheet expansion—especially when borrowing reserves. In contrast, branches of foreign banks are not subject to these leverage requirements at the branch level, allowing them to take on reserve without facing regulatory balance sheet cost.[47] This differential treatment creates a structural asymmetry in reserve demand, with foreign banks acting as more elastic borrowers in the overnight unsecured market.

In addition to these regulatory advantages, foreign bank branches are more likely to meet the credit standards required to access FHLB funding. The FHLBs impose internal counterparty rating requirements for fed funds lending, typically at or above investment

[46]This treatment was updated in January 2025, when the FHFA revised the liquidity framework for FHLBs: `https://www.fhfa.gov/news/news-release/fhfa-announces-final-rule-expanding-access-to-liquidity-for-the-federal-home-loan-bank-system`

[47]These exemptions apply to branches of FBOs. The consolidated parent may still be subject to leverage constraints, but these are often less strict compared to the Dodd Frank in the US.

grade, and in practice. Many smaller sized domestic banks lack formal credit ratings or prefer to borrow term advances rather than engage in the overnight market.[48] In contrast, the foreign banks with branches in the US are global institutions which tend to have more robust credit ratings, making them eligible counterparties for FHLB transactions.

This concentrated market structure has important implications for how the FFR is determined. In a frictionless environment with perfect competition among foreign banks, the FHLBs would be able to extract rates close to the borrowers' reservation price at IOR. In practice, however, the effective FFR in our data period are often below IOR (Figure A4). This spread reflects a bargaining process: while the FHLBs are constrained by operational needs and regulatory limits on the short-term investment, foreign bank branches retain enough pricing power to secure funding at a discount.[49] As a result, the FFR is not a market-clearing rate in the classical sense, but the negotiated outcome of bilateral interaction between a very small number of lenders and borrowers. It reflects the internal incentives and constraints of these institutions rather than system-wide marginal valuations of reserves. Moreover, this structure is fragile. It depends on the continued participation of both FHLBs as reliable lenders and foreign bank branches as eligible, low-cost borrowers. This structure is fragile: any disruption on either side could materially affect the level and behavior of the FFR. For example, FHLB short-term investment rule revision or divergence of global monetary policies could significantly alter participation and change bargaining dynamics.

A5. Model Appendix: Derivations and Proofs

A5.1 Derivation of the First-Order Condition for X_j^*

A bank of type $j \in \{D, F\}$ chooses its target end-of-day excess reserve balance $X_j \geq 0$. The realized balance is $X_j + \eta_j$, where the liquidity shock η_j is drawn from a Normal distribution with mean zero and variance σ_j^2, i.e., $\eta_j \sim \mathcal{N}(0, \sigma_j^2)$. Let $f_j(\cdot)$ and $F_j(\cdot)$ denote the probability density function (PDF) and cumulative distribution function (CDF) of this shock, respectively.

The bank's objective is to minimize the expected costs associated with its reserve balance. The cost function, $C_j(X_j)$, is given by the sum of the expected opportunity cost of holding positive reserves and the expected penalty for a reserve shortfall:

$$\min_{X_j \geq 0} C_j(X_j) = \mathbb{E}\left[(X_j + \eta_j)A_j\mathbf{1}\{X_j + \eta_j \geq 0\} + (-X_j - \eta_j)B_j\mathbf{1}\{X_j + \eta_j < 0\}\right], \quad (33)$$

where the indicator function $\mathbf{1}\{\cdot\}$ defines the relevant state, and the cost parameters are

[48] Foreign banks are not members of FHLBs thus do not have access to secured advances from FHLBs.

[49] For underlying bargaining process in the federal funds market, see Bech and Klee (2011).

defined as:

$$A_j := i - i_{\text{IOR}} + \alpha_j \quad \text{(Opportunity cost of funds)} \tag{34}$$

$$B_j := i_{\text{DW}} - i + \beta_j \quad \text{(Penalty cost of shortfall)} \tag{35}$$

We assume the policy rate i lies strictly within the corridor, which ensures $A_j > 0$ and $B_j > 0$ (that is, $i \in (i_{\text{IOR}} - \alpha_j, i_{\text{DW}} + \beta_j)$).

The expectation in (33) can be written in integral form:

$$C_j(X_j) = A_j \int_{-X_j}^{\infty} (X_j + z) f_j(z) dz + B_j \int_{-\infty}^{-X_j} (-X_j - z) f_j(z) dz. \tag{36}$$

To find the optimal level of reserves, X_j^*, we differentiate the cost function $C_j(X_j)$ with respect to X_j. Since X_j appears in the limits of integration, we apply the Leibniz rule:

$$\frac{d}{dx} \int_{a(x)}^{b(x)} g(x,t) dt = g(x, b(x)) \cdot b'(x) - g(x, a(x)) \cdot a'(x) + \int_{a(x)}^{b(x)} \frac{\partial g}{\partial x}(x,t) dt.$$

We apply this rule to each term in (36).

For the first term:

$$\frac{d}{dX_j} \left[A_j \int_{-X_j}^{\infty} (X_j + z) f_j(z) dz \right] = A_j \left[0 - (X_j - X_j) f_j(-X_j)(-1) + \int_{-X_j}^{\infty} f_j(z) dz \right] \tag{37}$$

$$= A_j \int_{-X_j}^{\infty} f_j(z) dz. \tag{38}$$

The boundary term evaluates to zero because $(X_j - X_j) = 0$.

For the second term:

$$\frac{d}{dX_j} \left[B_j \int_{-\infty}^{-X_j} (-X_j - z) f_j(z) dz \right] = B_j \left[(-X_j - (-X_j)) f_j(-X_j)(-1) - 0 + \int_{-\infty}^{-X_j} (-1) f_j(z) dz \right] \tag{39}$$

$$= -B_j \int_{-\infty}^{-X_j} f_j(z) dz. \tag{40}$$

This boundary term also evaluates to zero.

Combining the two parts, the full derivative is:

$$\frac{dC_j}{dX_j} = A_j \int_{-X_j}^{\infty} f_j(z) dz - B_j \int_{-\infty}^{-X_j} f_j(z) dz. \tag{41}$$

Using the definition of the CDF, $F_j(x) = \int_{-\infty}^{x} f_j(z) dz$, we can write (41) as:

$$\frac{dC_j}{dX_j} = A_j \left[1 - F_j(-X_j) \right] - B_j F_j(-X_j). \tag{42}$$

Assuming an interior solution ($X_j^* > 0$), we set the derivative to zero to find the FOC:

$$A_j \left[1 - F_j(-X_j^*)\right] = B_j F_j(-X_j^*). \tag{43}$$

We now solve the FOC (43) for X_j^*.

$$A_j = A_j F_j(-X_j^*) + B_j F_j(-X_j^*) \tag{44}$$

$$\Leftrightarrow F_j(-X_j^*) = \frac{A_j}{A_j + B_j}. \tag{45}$$

Given that $\eta_j \sim \mathcal{N}(0, \sigma_j^2)$, its CDF is related to the standard Normal CDF $\Phi(\cdot)$ by $F_j(x) = \Phi(x/\sigma_j)$. Substituting this to our equation yields:

$$\Phi\left(\frac{-X_j^*}{\sigma_j}\right) = \frac{A_j}{A_j + B_j}. \tag{46}$$

To isolate X_j^*, we apply the inverse standard Normal CDF, $\Phi^{-1}(\cdot)$, to both sides:

$$\frac{-X_j^*}{\sigma_j} = \Phi^{-1}\left(\frac{A_j}{A_j + B_j}\right). \tag{47}$$

Finally, solving for X_j^* gives the optimal excess reserve balance:

$$\begin{aligned} X_j^* &= -\sigma_j \Phi^{-1}\left(\frac{A_j}{A_j + B_j}\right) \\ &= -\sigma_j \Phi^{-1}\left(\frac{i - i_{\text{IOR}} + \alpha_j}{i_{\text{DW}} - i_{\text{IOR}} + \alpha_j + \beta_j}\right). \end{aligned} \tag{48}$$

A5.2 Derivation of $g'(M)$ and $g''(M)$

Recall that the aggregate (type–j) demand for reserves as a function of the rate i is

$$M_j^*(i) = \delta_j - \sigma_j \, \Phi^{-1}\left(\frac{i - i_{\text{IOR}} + \alpha_j}{c_j}\right), \qquad c_j := i_{\text{DW}} - i_{\text{IOR}} + \alpha_j + \beta_j > 0, \tag{49}$$

with Φ the standard normal CDF. Let us define $\varphi(z)$ and $z_j(i)$ as follow:

$$\begin{aligned} \varphi(z) &= \tfrac{1}{\sqrt{2\pi}} e^{-z^2/2} \\ z_j(i) &:= \Phi^{-1}\left(\frac{i - i_{\text{IOR}} + \alpha_j}{c_j}\right). \end{aligned} \tag{50}$$

Then differentiation uses the identity $(\Phi^{-1})'(p) = 1/\varphi(\Phi^{-1}(p))$.

Define the aggregate demand $M_{\text{agg}}(i) := M_D^*(i) + M_F^*(i)$. Its first and second derivatives are

$$M'_{\text{agg}}(i) = \sum_{j\in\{D,F\}} \frac{dM_j^*}{di}(i) = -\sum_j \frac{\sigma_j}{c_j}\frac{1}{\varphi\big(z_j(i)\big)} \;<\; 0, \tag{51}$$

$$M''_{\text{agg}}(i) = \sum_j \frac{d^2M_j^*}{di^2}(i) = -\sum_j \frac{\sigma_j}{c_j^2}\frac{z_j(i)}{\big[\varphi\big(z_j(i)\big)\big]^2}. \tag{52}$$

Hence $M_{\text{agg}}(i)$ is strictly decreasing in i, while its curvature $M''_{\text{agg}}(i)$ can be of either sign depending on the operating point via $z_j(i)$.

Finally, let the aggregate inverse demand be

$$g(M) := M_{\text{agg}}^{-1}(M), \qquad \text{so that} \quad M = M_{\text{agg}}\big(g(M)\big).$$

By the inverse function theorem,

$$g'(M) = \frac{1}{M'_{\text{agg}}\big(g(M)\big)} \;<\; 0. \tag{53}$$

Differentiating again gives

$$g''(M) = -\frac{M''_{\text{agg}}\big(g(M)\big)}{\big[M'_{\text{agg}}\big(g(M)\big)\big]^3}. \tag{54}$$

Since $M'_{\text{agg}}(g(M)) < 0$, its cube in the denominator is also negative. Therefore, the sign of $g''(M)$is dependent on numerator $M''_{\text{agg}}\,(g(M))$;

$$\begin{aligned}\operatorname{sgn}\left(g''(M)\right) &= \operatorname{sgn}\left(M''_{\text{agg}}(g(M))\right) = \operatorname{sgn}\left(-\sum_j \frac{\sigma_j}{c_j^2}\frac{z_j(g(M))}{\left[\varphi\left(z_j(g(M))\right)\right]^2}\right) \\ &= -\operatorname{sgn}\left(\sum_j \underbrace{\frac{\sigma_j}{c_j^2\left\{\varphi\left(z_j(g(M))\right)\right\}^2}}_{=w_j>0} z_j(g(M))\right)\end{aligned} \tag{55}$$

Therefore,

- If all $z_j(g(M)) < 0$ (low $i = g(M)$ relative to each type's midpoint), then $\sum_j w_j z_j < 0$ and the leading minus makes $g''(M) > 0$. This is the ample reserves regime: the inverse demand is convex.
- If all $z_j(g(M)) > 0$ (high i), then $\sum_j w_j z_j > 0$ and $g''(M) < 0$: the inverse demand is concave.

- If the z_j have mixed signs, the sign of $g''(M)$ is determined by the weighted sum $\sum_j w_j z_j$ and is a priori ambiguous.

Define "midpoint" rate for $j \in \{D, F\}$

$$i_j^{\text{mid}} := i_{\text{IOR}} - \alpha_j + \frac{c_j}{2} \quad \text{so that} \quad z_j\left(i_j^{\text{mid}}\right) = 0 \tag{56}$$

Then $z_j(i)$ is strictly increasing in i, hence:

- $i < \min(i_D^{\text{mid}}, i_F^{\text{mid}}) \Rightarrow z_j < 0, \quad \forall j \Rightarrow g''(M) > 0.$
- $i > \max(i_D^{\text{mid}}, i_F^{\text{mid}}) \Rightarrow z_j > 0 \quad \forall j \Rightarrow g''(M) < 0.$
- In between, $g''(M)$ changes sign exactly once at the unique i solving $\sum_j w_j z_j(i) = 0$ (weights $w_j > 0$ implicit in i).

A5.3 Proof 1.

The map $M \mapsto s_{\text{agg}}(M)$ is continuous and strictly decreasing for $M \geq \min\{\delta_D, \delta_F\}$, with $\lim_{M\to\infty} s_{\text{agg}}(M) = 0$ and $\lim_{M\searrow\min\delta_j} s_{\text{agg}}(M) = \infty$. By the intermediate-value theorem there is a unique $M_{\text{agg}}^{\text{flat}}$ solving $s_{\text{agg}} = \tau$. Next, the slope of the inverse aggregate demand curve is

$$s_{\text{agg}}(M_{\text{agg}}) = \left(\frac{1}{s_D(M_D)} + \frac{1}{s_F(M_F)}\right)^{-1} = \frac{s_D(M_D) \cdot s_F(M_F)}{s_D(M_D) + s_F(M_F)}, \tag{57}$$

For any positive a, b, the inequality $\frac{ab}{a+b} \leq \min\{a, b\}$ holds strictly unless $a = b$. Therefore, the aggregate slope satisfies

$$s_{\text{agg}} \leq \min\{s_D(M_D), s_F(M_F)\}. \tag{58}$$

thus the aggregate ample-reserve threshold is determined by the smaller of the two type-specific thresholds:

$$M_{\text{agg}}^{\text{flat}} \leq \min\left\{M_D^{\text{flat}},\ M_F^{\text{flat}}\right\}. \tag{59}$$

□

A5.4 Proof 2.

By definition, the marginal bank that pins the kink is the one whose behavior has the greatest influence on the slope of the aggregate demand curve at the threshold. The slope of the inverse aggregate demand curve, s_{agg}, is the reciprocal of the magnitude of the slope of the aggregate demand curve:

$$s_{\text{agg}}(i) = \left|\frac{di}{dM_{\text{agg}}}\right| = \left|\frac{dM_{\text{agg}}}{di}\right|^{-1} \tag{60}$$

and the slope of the direct demand curve is the sum of the individual responses:

$$\frac{dM_{\text{agg}}}{di} = \frac{dM_D^*}{di} + \frac{dM_F^*}{di} = v_D(i) + v_F(i) \tag{61}$$

The bank type that contributes more to the magnitude of this aggregate response is the one with the larger individual responsiveness magnitude. Therefore, the bank type with the largest $|v_j(i)|$ is the most elastic and pins the kink.

Notice that $|v_j(i)|$ is minimized when the term $\varphi(\cdot)$ in its denominator is maximized. The function $\varphi(z)$ is the standard normal PDF, with a maximum value of $\varphi(0) = 1/\sqrt{2\pi}$. The minimum value of $|v_j(i)|$ is therefore:

$$\min_i |v_j(i)| = \frac{\sigma_j\sqrt{2\pi}}{i_{\text{DW}} - i_{\text{IOR}} + \alpha_j + \beta_j} \tag{62}$$

Let $c_j := i_{\text{DW}} - i_{\text{IOR}} + \alpha_j + \beta_j$. Type F pins the kink if its minimum responsiveness is greater than Type D's, which holds if $\frac{\sigma_F}{c_F} > \frac{\sigma_D}{c_D}$.

For normal periods, the assumptions $\sigma_F > \sigma_D$ and $\beta_F < \beta_D$ push towards this inequality holding, but the assumption $\alpha_F > \alpha_D$ leaves the outcome ambiguous without specifying parameter magnitudes.

For quarter-end periods, we assume $\alpha_F \approx \alpha_D$. Then the comparison of the denominatior becomes foreign banks is strictly smaller: $c_F < c_D$ with $\alpha_F \approx \alpha_D$ and $\beta_F < \beta_D$.

Because the ratio for foreign banks (σ_F/c_F) has a larger numerator and a smaller denominator, it is guaranteed that $\min|v_F(i)| > \min|v_D(i)|$. Since the entire responsiveness curve for foreign banks is structurally more responsive than that of domestic banks, their responsiveness will be greater at the threshold, $|v_F(i^{\text{flat}})| > |v_D(i^{\text{flat}})|$. Thus, foreign banks pin the kink at quarter-ends. $\square$

A5.5 Proof 3.

Define

$$h(x) \;:=\; \mathbf{1}\{g(x) > i_{\text{target}}\}\,\Big(g(x) - i_{\text{target}}\Big)\,g'(x), \qquad \Psi(M) \;:=\; (1-\lambda)\Big[h(M-S)\Big] + \lambda\,(M - M_{\text{target}}).$$

The first-order conditions (FOCs) for the deterministic and stochastic problems can be written, respectively, as

$$\mathbf{1}\{g(M^*) > i_{\text{target}}\} \Big(g(M^*) - i_{\text{target}}\Big)\, g'(M^*) + \frac{\lambda}{1-\lambda} \Big(M^* - M_{\text{target}}\Big) = 0, \tag{63}$$

and

$$\Psi(M^{**}) = 0. \tag{64}$$

Step 1 (Sign of $\Psi(M^*)$*).* From (63) we have

$$(1-\lambda)\, h(M^*) + \lambda\, (M^* - M_{\text{target}}) = 0.$$

Since $S \geq 0$ and h will be shown to be increasing (Step 2), it follows that $h(M^* - S) \leq h(M^*)$ pointwise; taking expectations yields $[h(M^* - S)] \leq h(M^*)$. Therefore

$$\Psi(M^*) = (1-\lambda)\, [h(M^* - S)] + \lambda\, (M^* - M_{\text{target}}) \;\leq\; (1-\lambda)\, h(M^*) + \lambda\, (M^* - M_{\text{target}}) = 0.$$

If moreover $\Big(g(M^* - S) > i_{\text{target}}\Big) > 0$, then $h(M^* - S) < h(M^*)$ on a set of positive probability (because $g'(M^* - S) < 0$ and the indicator is 1 there), so the inequality is strict: $\Psi(M^*) < 0$.
Step 2 (Monotonicity of Ψ*).* On the active set $\{x : g(x) > i_{\text{target}}\}$, h is differentiable with

$$h'(x) = [g'(x)]^2 + \Big(g(x) - i_{\text{target}}\Big)\, g''(x).$$

By assumption, $g''(x) \geq 0$ on the active set and $[g'(x)]^2 > 0$, hence $h'(x) > 0$ whenever $g(x) > i_{\text{target}}$. On the inactive set $h \equiv 0$. Thus h is increasing, and so $M \mapsto [h(M - S)]$ is increasing. Consequently,

$$\Psi'(M) = (1-\lambda)\Big[h'(M - S)\Big] + \lambda \;\geq\; \lambda \;>\; 0,$$

showing that Ψ is strictly increasing on $\mathcal{M}$. *Step 3 (Order of the optimizers).* We have $\Psi(M^*) \leq 0$ by Step 1 and Ψ strictly increasing by Step 2. Hence the unique root of $\Psi(M) = 0$, namely M^{**} in (64), satisfies $M^{**} \geq M^*$. If $\Big(g(M^* - S) > i_{\text{target}}\Big) > 0$, then $\Psi(M^*) < 0$ (Step 1), so strict inequality $M^{**} > M^*$ follows. □

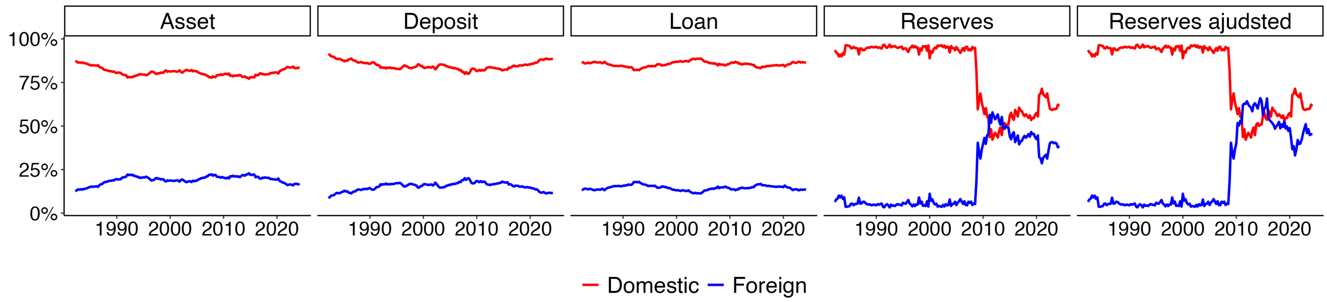


Figure A1: Share of Balance Sheet Components: Domestic Banks and Foreign Banks This figure shows the assets, loans, deposits, and reserves of the domestic banks and foreign banks (the aggregate sum of subsidiaries and branches and agencies) from 1980Q2 to 2024Q3. The "Reserves" show share based on Call Report quarterly-end reportings where "Reserves adjusted" show the reserves holdings based after adjusting for window dressing. Specifically, we replace quarter-end reserve reports of foreign branches and agencies with their daily cash balances from the H.8 release, carrying forward the most recent non-quarter-end value for March 31, June 30, September 30, and December 31. This adjustment removes artificial dips and yields a smoother, more accurate measure of foreign branches' true reserve holdings. No adjustment is applied to foreign bank subsidiaries or to domestic banks. Source: FFIEC 002, 031, 041 reports, Federal Reserve H.8 release.

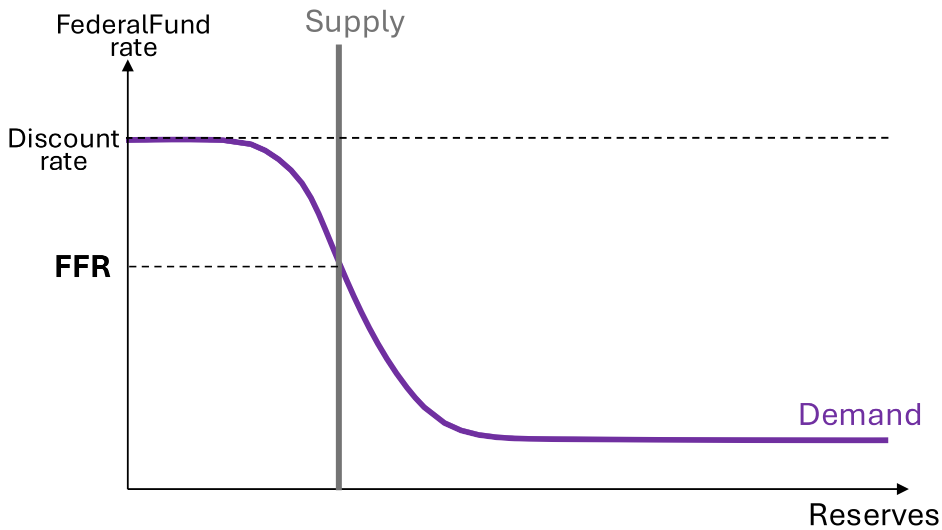


Figure A2: Illustration of Scarce Reserves System

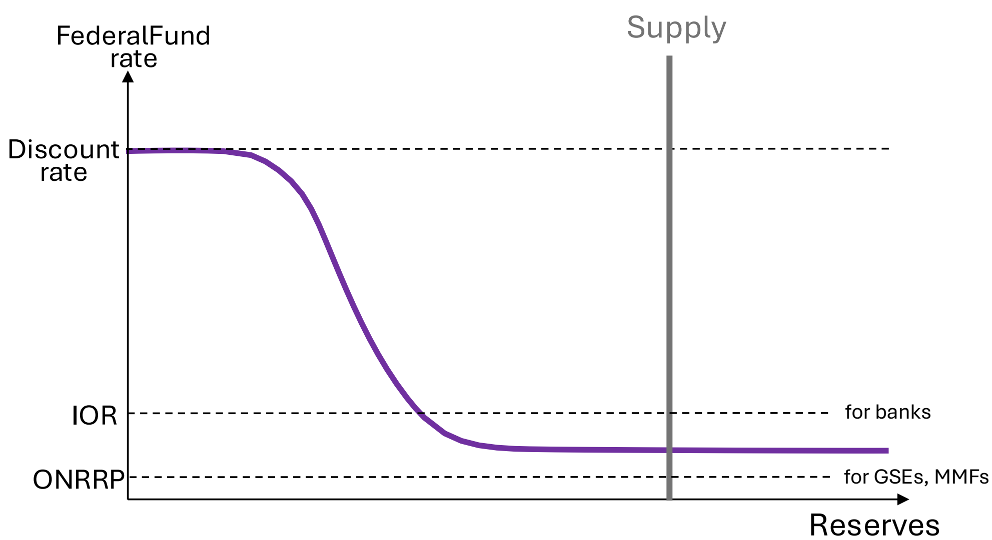


Figure A3: Illustration of Introduction of ONRRP

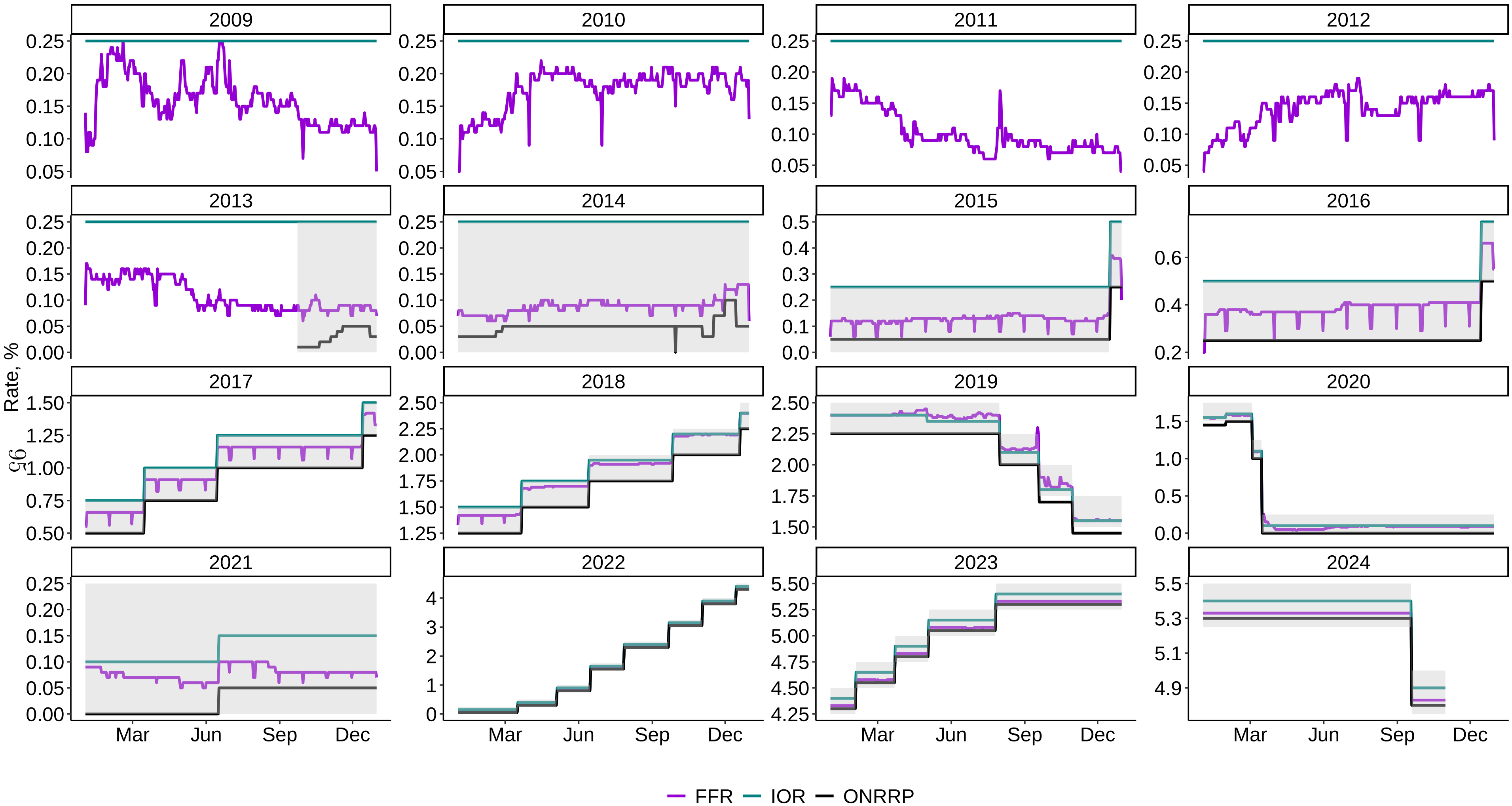


Figure A4: Federal Fund Rate (FFR), Interest on Reserves (IOR), Overnight Reverse Repurchase Agreements Award Rate (ONRRP), and the Federal Fund Target Range (2009-2024) The Federal Fund Target Range shaded in gray. Source: FRED

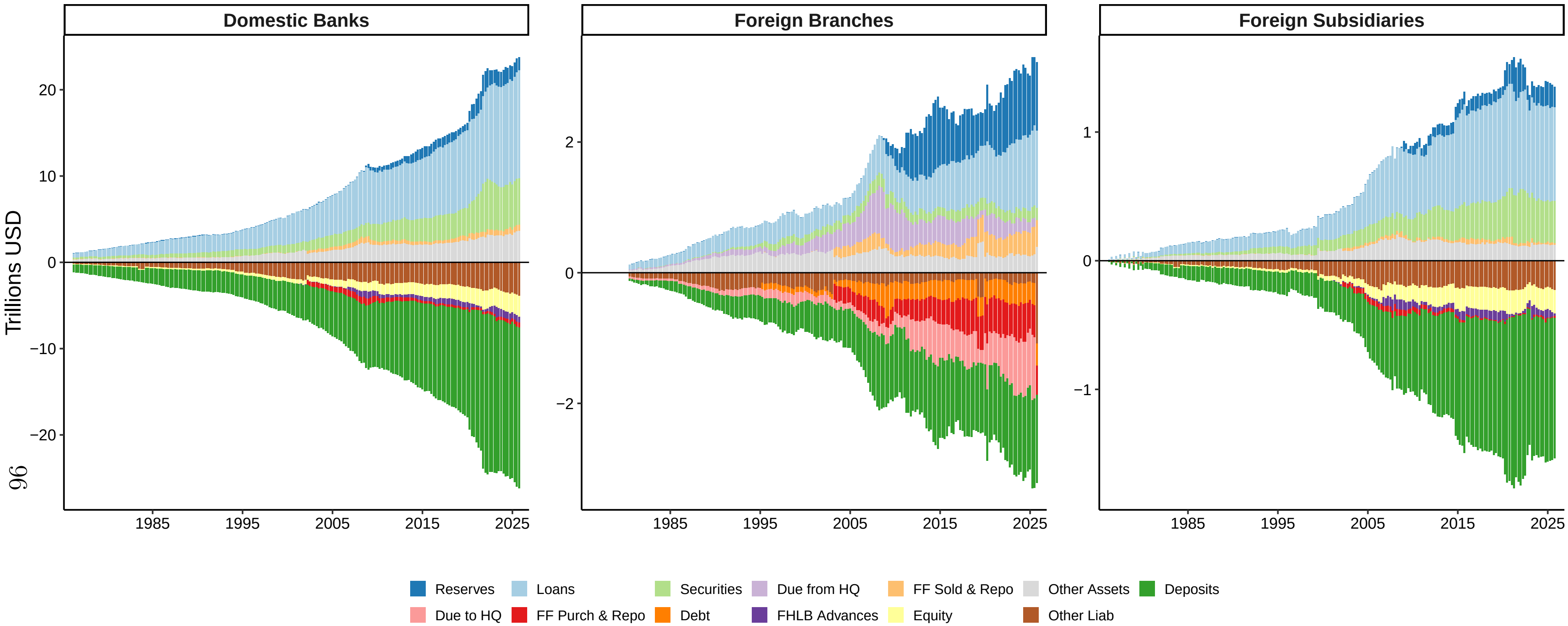


Figure A5: Balance Sheet Breakdown Source: FFIEC 002, 031, 041 reports, FR Y-9C reports

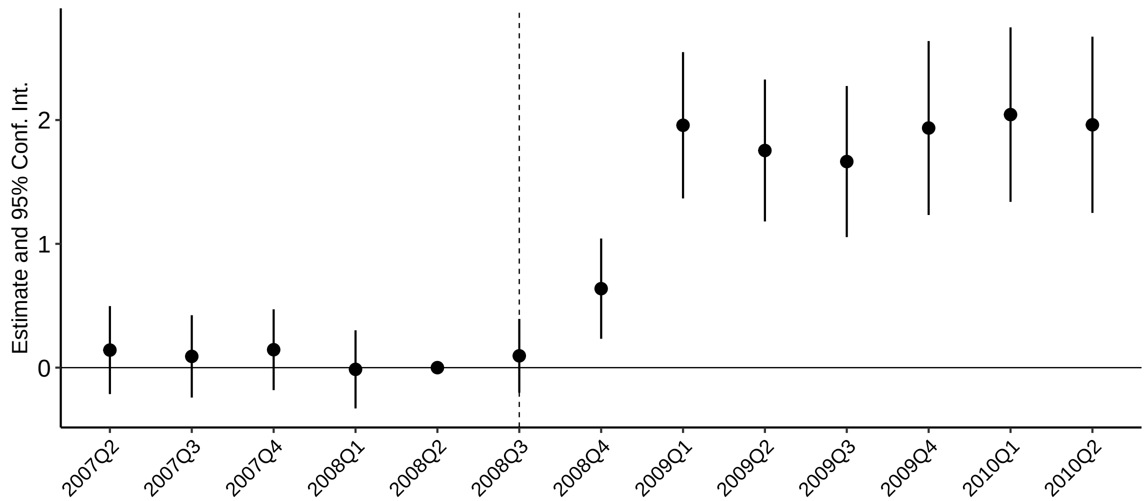


Figure A6: Event study around 2008Q3 This figure presents an event study analysis of the reserve holdings of bank holding companies (BHCs) around 2008Q3 of equation $Y_{it} = \sum_{k \neq 0} \gamma_k \cdot \mathbb{1}\{T_i = k\} + \beta X_{it} + \alpha_i + \delta_t + \epsilon_{it}$, where Y_{it} denotes the log of reserves held by bank holding company i at time t. The event, which corresponds to the third quarter of 2008, is normalized to $k = 0$, and $\mathbb{1}\{T_i = k\}$ is an indicator for periods k relative to this event. The coefficients γ_k capture the differential reserve accumulation of foreign banks relative to domestic banks before and after the event. The model includes BHC fixed effects (α_i), and time fixed effects (δ_t), which control for macroeconomic shocks common to all banks. Additionally, we control for bank-specific characteristics (X_{it}), such as the lagged value of total assets, which account for differences in bank size and balance sheet structure. The horizontal axis represents time in quarters relative to this event, while the vertical axis shows the estimated effect on log reserve holdings. Confidence intervals at the 95% level are shown for each estimate. Source: FFIEC 002, 031, 041 reports, FR Y-9C reports

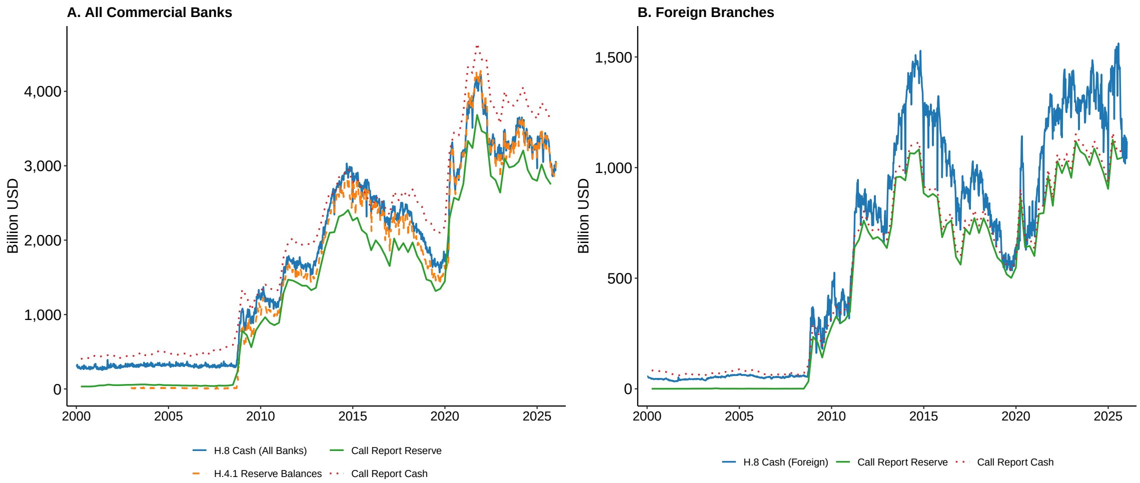


Figure A7: Cash and reserve data source and frequency comparison Source: FFIEC 002, 031, 041 reports, Assets and Liabilities of Commercial Banks in the United States (H.8), Federal Reserve Balance Sheet: Factors Affecting Reserve Balances (H.4.1)

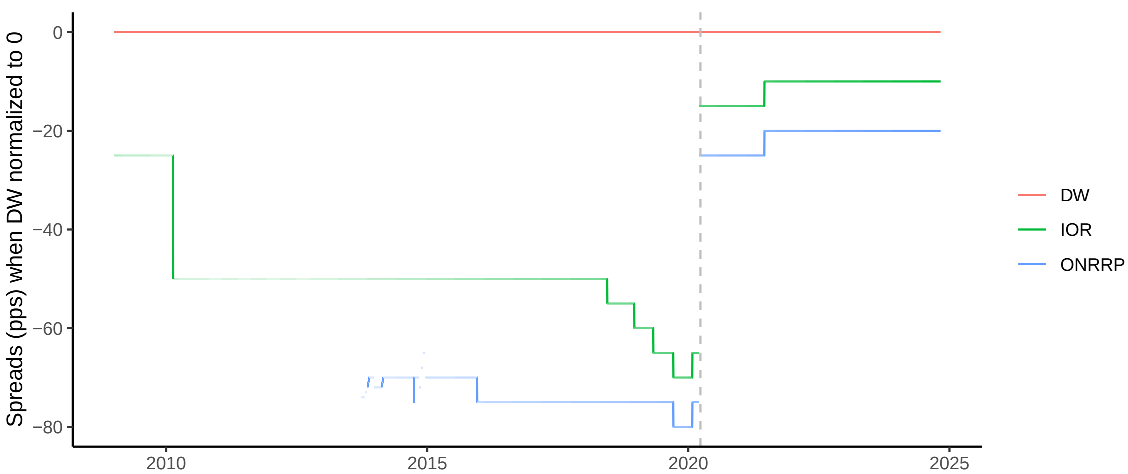


Figure A8: Historical Spreads between discount window rate (DW), interests on reserves (IOR), and overnight reverse repurchase agreements rates (ONRRP) This figure shows the spreads between the DW, IOR, and ONRRP facility administered rates. The rates are normalized at $i_{\mathrm{DW}} = 0$. The vertical dotted line indicates March 16th, 2020, where the DW-IOR spread has narrowed down from 75bps to 25bps.

Table A1

First Stage: TGA Shock and Money Market Spreads Positive TGA Shocks

This table presents first-stage regressions using only positive TGA shocks (reserve drains, when TGA balance increases). Panel A uses the federal funds rate minus IOR spread; Panel B uses the SOFR minus IOR spread (available from April 2018). The coefficient on ΔTGA represents the basis point change in the spread per $100 billion increase in TGA. Columns (1)–(3) use the full sample with progressively more controls. Columns (4) and (5) split by QE periods and QT periods. Column (1) reports IID standard errors; columns (2)–(5) report standard errors clustered by FOMC period. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Fed Funds Rate Spread

	Δ(FFR − IOR) (bps)				
	(1)	(2)	(3)	(4)	(5)
ΔTGA ($100bn)	0.4225***	0.4952***	0.5240***	0.9847***	0.1639
	(0.1420)	(0.1576)	(0.1570)	(0.2758)	(0.1801)
FOMC Period FE		✓	✓	✓	✓
ΔONRRP Control			✓	✓	✓
Sample	Full	Full	Full	QE	QT
Observations	1,922	1,922	1,922	821	1,101
R^2	0.005	0.114	0.119	0.166	0.088

Panel B: SOFR Spread

	Δ(SOFR − IOR) (bps)				
	(1)	(2)	(3)	(4)	(5)
ΔTGA ($100bn)	2.6055***	2.8487***	2.9506***	0.8991	3.8027***
	(0.4318)	(0.5669)	(0.6063)	(0.6592)	(0.7885)
FOMC Period FE		✓	✓	✓	✓
ΔONRRP Control			✓	✓	✓
Sample	Full	Full	Full	QE	QT
Observations	871	871	871	192	679
R^2	0.040	0.224	0.225	0.643	0.166

Table A2

First Stage: TGA Shock and Money Market Spreads
Negative TGA Shocks

This table presents first-stage regressions using only negative TGA shocks (reserve injections, when TGA balance decreases). Panel A uses the federal funds rate minus IOR spread; Panel B uses the SOFR minus IOR spread (available from April 2018). The coefficient on ΔTGA represents the basis point change in the spread per $100 billion decrease in TGA. Columns (1)–(3) use the full sample with progressively more controls. Columns (4) and (5) split by QE periods and QT periods. Column (1) reports IID standard errors; columns (2)–(5) report standard errors clustered by FOMC period. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Fed Funds Rate Spread

	Δ(FFR − IOR) (bps)				
	(1)	(2)	(3)	(4)	(5)
ΔTGA ($100bn)	0.0196	0.1183	0.0845	0.0893	0.0644
	(0.1914)	(0.2702)	(0.2604)	(0.2729)	(0.4024)
FOMC Period FE		✓	✓	✓	✓
ΔONRRP Control			✓	✓	✓
Sample	Full	Full	Full	QE	QT
Observations	2,286	2,286	2,286	1,147	1,139
R^2	0.000	0.036	0.065	0.110	0.055

Panel B: SOFR Spread

	Δ(SOFR − IOR) (bps)				
	(1)	(2)	(3)	(4)	(5)
ΔTGA ($100bn)	-0.6138	-0.6329	-0.6231	0.2932	-1.1879*
	(0.7074)	(0.4645)	(0.4501)	(0.4219)	(0.6426)
FOMC Period FE		✓	✓	✓	✓
ΔONRRP Control			✓	✓	✓
Sample	Full	Full	Full	QE	QT
Observations	943	943	943	288	655
R^2	0.001	0.115	0.118	0.777	0.016

Table A3

Second Stage: Fed Funds Borrowing Response to SOFR-IOR Spread

This table presents second-stage IV regressions of daily changes in Fed Funds borrowing on the SOFR-IOR spread, instrumented by TGA shocks. The dependent variable is the log change in Fed Funds borrowing volume. Panel A reports results for the full sample with progressively more controls. Panel B splits by monetary policy regime (QE vs QT periods). The coefficient on ΔSOFR-IOR represents the semi-elasticity: a 1 basis point increase in the spread is associated with a $\beta \times 100\%$ change in borrowing. Standard errors are clustered by FOMC period where fixed effects are included. The F-test reports the first-stage F-statistic for instrument strength. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

Panel A: Full Sample

	Δ log (Fed funds borrowing)					
	Foreign			Domestic		
	(1)	(2)	(3)	(4)	(5)	(6)
ΔSOFR-IOR	-0.0513***	-0.0507***	-0.0914***	0.0475**	0.0499**	0.0152
	(0.0163)	(0.0133)	(0.0245)	(0.0210)	(0.0224)	(0.0159)
FOMC Period FE		✓	✓		✓	✓
Δ ONRRP Control			✓			✓
Observations	1,969	1,969	1,969	1,969	1,969	1,969
R^2	-3.2875	-3.1833	-11.418	-0.77123	-0.83924	-0.04983
F-test (1st stage)	10.548	10.704	6.9243	10.548	10.704	6.9243

Panel B: By Monetary Policy Regime

	Δ log (Fed funds borrowing)			
	QE		QT	
	Foreign	Domestic	Foreign	Domestic
	(1)	(2)	(3)	(4)
ΔSOFR-IOR	-0.0640	0.0707	-0.0960***	0.0046
	(0.0383)	(0.0435)	(0.0294)	(0.0170)
FOMC Period FE	✓	✓	✓	✓
ΔONRRRP	✓	✓	✓	✓
Observations	480	480	1,489	1,489
R^2	-4.0177	-0.55845	-13.573	0.00666
F-test (1st stage)	2.1784	2.1784	5.0237	5.0237

Table A4

Reserve Holdings and Funding Sources: Subperiod Analysis

This table reports subperiod analysis of the relationship between reserves and funding sources at the BHC level for foreign bank branches. QE1 covers 2009–2014, QT1 covers 2017–2019, QE2 covers 2020–2022Q1, and QT2 covers 2022Q2–2025. QE Combined pools QE1 and QE2; QT Combined pools QT1 and QT2. All specifications include BHC fixed effects, lagged log assets as size control, and time fixed effects. †Time × Country FE used for combined periods only. Standard errors clustered by BHC in parentheses; observations in brackets. ***, **, and * denote significance at the 1%, 5%, and 10% levels, respectively.

	Dependent Variable: log(Reserves)		
	Net HQ Liab (IHS)	Uninsured Dep (log)	Domestic Dep (log)
Individual Subperiods			
QE1 (2009–2014)	0.392***	0.177*	0.048
	(0.068)	(0.103)	(0.073)
	[2,095]	[2,949]	[2,880]
QT1 (2017–2019)	0.550***	0.162**	0.085
	(0.108)	(0.066)	(0.067)
	[971]	[1,409]	[1,402]
QE2 (2020–2022Q1)	0.721***	0.222**	0.114
	(0.161)	(0.106)	(0.081)
	[819]	[1,003]	[1,000]
QT2 (2022Q2–2025)	0.459***	0.183***	0.115**
	(0.081)	(0.060)	(0.045)
	[1,201]	[1,406]	[1,407]
Combined Periods			
QE Combined	0.535***	0.219**	0.068
	(0.088)	(0.091)	(0.069)
	[2,914]	[3,952]	[3,880]
QT Combined	0.441***	0.090*	0.087**
	(0.070)	(0.049)	(0.039)
	[2,172]	[2,815]	[2,809]
BHC FE	✓	✓	✓
Size Control	✓	✓	✓
Time FE	✓	✓	✓
Time × Country FE	✓†	✓†	✓†

Table A5

Bank Characteristics and Reserve Holdings

Notes: This table presents the results of estimating: $Y_{it} = \beta_1 F_i + \beta_2 G_i + \beta_3 (F_i \times G_i) + \Gamma X_{it} + \epsilon_{it}$, using quarterly data for U.S. bank holding companies (BHCs), 2009:Q1–2024:Q4. Here Y_{it} is the log of reserves held by BHC i at time t. The vector X_{it} includes lagged bank size, the aggregate supply of reserves, and other balance-sheet controls. Standard errors are clustered at the BHC level. Columns (1)–(3) relate reserve holdings to (i) a foreign-bank indicator F_i, (ii) a G-SIB indicator G_i, and (iii) their interaction. Columns (4)–(6) add controls for log assets and log deposits; column (5) restricts the sample to foreign banks, and column (6) excludes G-SIBs. G-SIBs face stricter post-GFC liquidity and capital rules—most notably the Liquidity Coverage Ratio (LCR) and CCAR stress tests—which raise their demand for high-quality liquid assets such as reserves. Moreover, their central role in wholesale funding and payment systems necessitates larger settlement balances. ***, **, and denote significance at the 1 %, 5 %, and 10 % levels, respectively.

	log(Reserves)					
	(1)	(2)	(3)	(4)	(5)	(6)
F	3.361***		2.763***	1.661***		1.665***
	(0.2209)		(0.2127)	(0.1513)		(0.1516)
G		7.382***	8.555***	3.316***	0.5838**	
		(0.2182)	(0.2338)	(0.1344)	(0.2363)	
F × G			-4.299***	-1.508***		
			(0.3891)	(0.2224)		
Balance Sheet Control				✓	✓	✓
Reserve Supply Control	✓	✓	✓	✓	✓	✓
Sample	All	All	All	All	Foreign	Non-GSIBs
Observations	305,310	305,310	305,310	295,643	7,873	294,034
R^2	0.15	0.15	0.20	0.46	0.58	0.40