How abundant are reserves?
Evidence from the wholesale payment system

Gara Afonso
Federal Reserve Bank of New York

Darrell Duffie
Stanford GSB

Lorenzo Rigon
Stanford Economics Department

Hyun Song Shin
Bank for International Settlements

November 4, 2022

Abstract

Before the era of large central bank balance sheets, banks relied on incoming payments to fund outgoing payments in order to conserve scarce liquidity. Even in the era of large central bank balance sheets, rather than funding payments with abundant reserve balances, we show that outgoing payments remain highly sensitive to incoming payments. By providing a window on liquidity constraints revealed by payment behavior, our results shed light on thresholds for the adequacy of reserve balances. Our findings are timely, given the ongoing shrinking of central bank balance sheets around the world in response to inflation.

Keywords: Real-time gross settlement (RTGS) systems; Quantitative tightening; Balance sheet management; Reserve balances

JEL: E42, E44, E52, E58, G21

The views presented here are solely those of the authors and do not necessarily represent those of the Bank for International Settlements, the Federal Reserve Bank of New York, or the Federal Reserve System. Duffie is also a CDI Research Fellow and a Research Associate of the National Bureau of Economic Research. We are grateful to Morten Bech, Vítor Constâncio, Rodney Garrett, Gary Gorton, Scott Hendry, Gabriele La Spada, Jamie McAndrews, and Christine Parlour, and participants at the 2021 ECB Conference on Money Markets and the 2022 DNB-Riksbank-Bundesbank Macroprudential Conference for valuable feedback. We thank Will Arnesen, Steph Clampitt, Doug Leonard, and Peter Prastakos for excellent research assistance.

1 Introduction

As major central banks respond to heightened inflation by removing policy accommodation, including through large reductions of their balance sheets, a critical issue is the adequacy of reserves for the smooth functioning of money markets and the financial system as a whole.

We approach this issue from the perspective of the payment system. Before the era of large central bank balance sheets, banks relied to a large extent on incoming payments to fund outgoing payments. This was forced by the small aggregate quantity of reserve balances supplied by central banks. Because outgoing payments deplete a bank's reserve balances, banks relied heavily on incoming payments to fund outgoing payments, conserving scarce liquidity. The sensitivity of outgoing payments to incoming payments has served as a window on the shadow price of liquidity.

In the era of large central bank balance sheets, it has been conventional wisdom that the reliance of banks on their incoming payments to make outgoing payments would no longer apply because banks hold abundant reserves at their central banks. However, we find that, even in the era of "ample" reserves, the amount of payments that a bank makes in a given minute depends significantly on the amount of payments that it has received over preceding minutes, indicating a high degree of dependence between incoming payments and outgoing payments. In the terminology of game theory, this is evidence of "strategic complementarity," in that payments are increasing in the payments made by others.

The close coupling of interbank payments that we find suggests that banks act as though their intraday reserve balances are a scarce resource, even when total reserve balances in the US banking system are well over \$1 trillion, far in excess of aggregate reserve balances before the Global Financial Crisis (GFC) of 2008-2009. This in turn suggests that there is still a potential for strategic cash hoarding when reserve balances get sufficiently low. Indeed, this happened in mid-September 2019 and mid-March 2020, when large banks suddenly negotiated very high interest rates for providing overnight liquidity in the Treasury repo market.

In this paper, we study the coupling of interbank payments in the Fedwire Funds Service (Fedwire) system –the primary large-value payment system operated by the Federal Reserve– as a lens into the key question of the adequacy of reserve balances in the era of large central bank balance sheets. In 2020, the Fedwire Service handled a daily payment value of over \$3.3 trillion dollars, meaning that payments on the scale of annual US GDP were made roughly every 7 days. In spite of the scale and speed of payments that go through large-value payment systems, the principles underpinning account-based payment systems have remained largely unchanged for centuries. A payment is made by debiting the account of the payer and crediting the account of the receiver. In the case of a wholesale payment system such as Fedwire, the system's participants hold balances at the Federal Reserve, but the ledger operations are common among most bank-railed payment systems. The sender of a payment must have sufficient deposit balances to initiate a payment. Failing that, the payer must first borrow the needed amount (and then receive the loaned funds) or must wait until other incoming payments replenish its deposit balance sufficiently to send the payment.

Before the GFC, banks in most jurisdictions maintained small reserve balances at the central bank compared to their outgoing payment volumes. They relied heavily on incoming payments to accumulate sufficient balances to make outgoing payments. The reliance on incoming payments implied a high velocity of payments and a high degree of strategic complementarity among system participants, whereby the willingness to make payments promptly was greater when other system participants did so. Strategic complementarity gives rise to potential gridlock scenarios whenever a cautious stance by banks that wait for sufficient incoming payments before making outgoing payments slows down the system as a whole. McAndrews and Potter (2002) studied the gridlock that followed the 9/11 attacks in 2001. Afonso and Shin (2011) and Yang (2022) provide models calibrated for the Fedwire system to study the impact of liquidity hoarding on the functioning of the payment system. In order to avoid potential gridlocks, the Federal Reserve provides daylight overdrafts, intraday loans to banks that they can use to make outgoing payments in the expectation that positions would be positive before the end of the day.[1]

Discussion of strategic payment timing complementaries faded from view after the GFC, with the large expansion of reserve balances at the central bank in many jurisdictions caused by central bank asset purchase programs. The conventional wisdom has been that large reserve balances maintained by banks reduced or eliminated the reliance on incoming payments to make outgoing payments.

This conventional wisdom turns out to be false. In an empirical investigation of the Fedwire system in the United States, we find that strategic complementarity of payments is alive and well. A 1-percent increment received in incoming payments by a bank in a 15-minute window predicts an additional 0.4 percent of outgoing payments in the subsequent

[1]Further modeling of strategic payment timing games is provided by Bech (2008), Bech and Garratt (2003), and Nellen (2019).

one-minute window, even after we control for banks' balances at the beginning of the day. This effect holds during normal times that exclude stresses such as those of mid-September 2019. Our results shed light on the shifting nature of bank balance sheet constraints in the era of ample reserves and point to the relative scarceness of these reserve balances.

Our findings are especially notable given the contrary indications that came from the diminished role of intraday liquidity provided by central banks in real-time gross settlement (RTGS) systems. In the United States, daylight overdrafts refer to the intraday loans granted to banks by the Federal Reserve that allow banks to fulfill their payment obligations without waiting for incoming payments to replenish their reserve balances. Before the GFC, daylight overdrafts were used actively to oil the wheels of the wholesale payment systems and prevent gridlocks.

(a) (b)

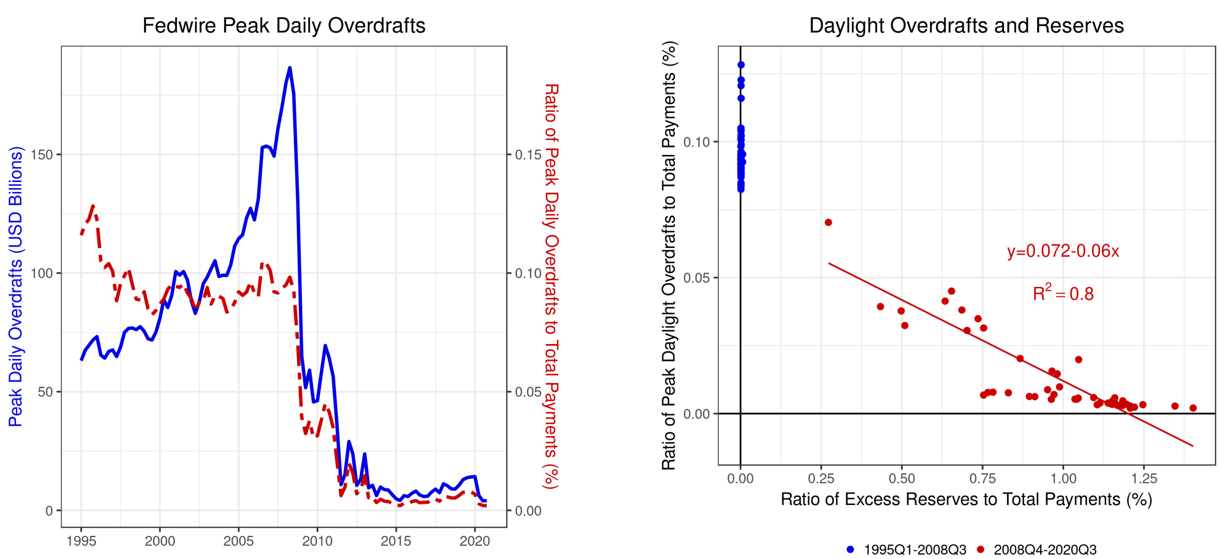

Figure 1: **Daylight overdrafts in Fedwire** - Peak daylight overdrafts (blue) and ratio of peak daylight overdrafts to total payments (dashed red) (a), and relationship between the ratio of peak daylight overdraft to total payments and excess reserves (b), 1995-2020Q3. Source: Federal Reserve, FRED and Fedwire Funds Services.

Figure 1 shows the peak daylight overdrafts in Fedwire over a long time span, going back to the 1990s. The left panel presents the value of peak daylight overdrafts both in dollar terms and as a proportion of total payments. We see the dramatic decline in peak daylight overdrafts in the regime of large reserve balances following the GFC of 2008. The right panel is a scatterplot of peak daylight overdrafts at different levels of excess reserves, both normalized by total daily payments. The blue dots are pre-GFC. The red dots are post-GFC, and show a strongly negative relationship between daylight overdrafts and excess reserves.

At face value, the evidence in Figure 1 might give rise to the hypothesis that strategic complementarity in payments has been eliminated or greatly diminished. We nevertheless find that strategic complementarity, in the form of a positive reaction function mapping outgoing payments into incoming payments, remains a robust and consistent feature of the data. The system as a whole is less reliant on daylight overdrafts provided by the Federal Reserve, but the decisions of individual banks continue to reveal significant balance-sheet liquidity constraints.

Although the exact nature of these balance sheet constraints is beyond the scope of our paper, there are important clues from the recent literature. Afonso, Cipriani, Copeland, Kovner, La Spada, and Martin (2021) and Correa, Du, and Liao (2020) discuss how reserves are deployed to repo lending or to supplying dollars in the FX swap market when spreads in these markets are unusually high. Banks do so by channeling the funds to their broker-dealer subsidiaries while keeping the overall size of the bank holding company's balance sheet unchanged. The fact that the overall balance sheet remains unchanged, but with one class of liquid assets giving way to another, suggests that banks are conscious of their overall Basel III leverage ratio. To the extent that deviations from covered interest parity in the FX swap market are attributable to the same underlying forces constraining commercial bank balance sheets, bank balance sheet constraints are the common thread tying the two, as discussed in Avdjiev, Du, Koch, and Shin (2019).

Ihrig (2019), Copeland, Duffie, and Yang (2022), and d'Avernas and Vandeweyer (2021) additionally point to the impact on intraday reserves management of post-GFC regulatory and supervisory liquidity requirements. For example, under post-GFC rules for resolution planning, commercial banks are required to have sufficient liquidity in a failure-resolution scenario to distribute reserve holdings across legal entities and across international units of the bank holding company. These requirements tie up reserves, exacerbating the strong hoarding incentives during times of stress in the short-term funding markets, as was the case in September 2019 and in the early weeks of the Covid-induced stresses in March 2020. In particular, Copeland, Duffie, and Yang (2022) show that the likelihood of a significant jump in Treasury repo rates is closely associated with intraday delays in incoming payments to the largest banks that supply funding in the repo market.

We find a clear relationship between the aggregate supply of reserve balances and the payment reaction function that we estimate. Consistent with the constraints highlighted in the literature described above, we find that larger reserve balances are associated with a flatter payment reaction function, as one would expect given the reduced precautionary motive. We show that when reserve balances are relatively low, strategic complementarity is stronger: banks receive a higher fraction of their payments late in the afternoon as other banks delay payments to economize on intraday liquidity.

The payment behavior that we uncover in our paper is consistent with the work of Afonso, Giannone, La Spada, and Williams (2022), who estimate a non-linear demand for reserves in the era of large central bank balance sheets.[2] Afonso, Giannone, La Spada, and Williams (2022) show that the price of reserves becomes sensitive to changes in their quantity as reserves in the banking system decline, reflecting the relative scarcity of reserves. Over a similar time period, we show that the coupling of payments is strengthened as reserves decline and banks economize on liquidity.

Beyond the quantitative regularities that we uncover, our key qualitative message is that the strategic complementarity of the payment decisions is alive and kicking. This behavior remains an important component of the overall approach to central bank balance sheet operations even in a setting of large amounts of reserves.

We begin with a brief introduction of the institutional details behind the US Fedwire system and some descriptive evidence relevant to our study. The core of our analysis is in the subsequent three sections of the paper, where we derive empirical estimates of the reaction functions of payment system participants, the relationship between these reaction functions and reserve balances, and how these reaction functions have changed over time.

2 Data description

Our main dataset includes all payment transactions in Fedwire Funds Service (Fedwire Funds). Fedwire Funds is a real-time gross settlement system that settles transactions individually on an order-by-order basis, without netting. The system was developed in 1918 and is owned and operated by the Federal Reserve Banks. Participants, including

[2] Hamilton (1997) estimates the reserve demand curve prior to the GFC. Smith (2019) and Lopez-Salido and Vissing-Jorgensen (2022) provide further post-GFC estimates.

banks, businesses and government agencies use Fedwire Funds to make same-day transfers in central bank money, that is, transfers between accounts held at the Fed. Fedwire Funds operates 21.5 hours each business day (Monday through Friday), from 9:00 pm Eastern Time (ET) on the preceding calendar day to 6:30 pm ET.[3] Settlement of funds is immediate, final, and irrevocable. As shown in Figure 2, daily instances and values of transfers have increased during the last decade, from around 500,000 daily transfers and $2.4 trillion on an average day in 2010 to more than 700,000 transfers with $3.3 trillion in daily value in 2020.

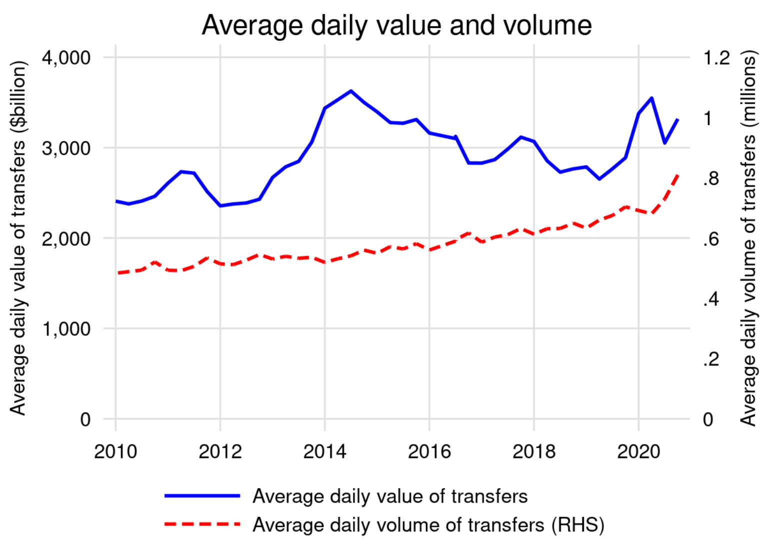

Figure 2: Fedwire Funds average daily value ($billion; blue) and average daily volume (millions of transactions; dashed red), 2010-2020. Source: Fedwire Funds Services.

More than 5,000 participants sent or received Fedwire Funds transfers in 2020. We exclude from our sample the master accounts of ancillary payment systems such as automated clearing houses (ACHs), other large value payment systems such as The Clearing House Interbank Payment System (CHIPS), settlement services such as Continuous Linked Settlement (CLS), and other "special" accounts including the Treasury General Account

[3]On March 8, 2021, the Fedwire Funds Service closing time was moved to 7:00 pm from 6:30 pm ET, expanding the number of operating hours to 22. Previously, operating hours were extended in December 1997 from ten hours to eighteen hours (12:30 am - 6:30 pm ET) and again in May 2004 to twenty-one and a half operating hours (9:00 pm - 6:30 pm ET).

(TGA).[4] For computational convenience, we limit our sample to the top 100 master accounts, ranked by the average daily total dollar value of payments sent in the first 100 business days in 2020.

Payments in Fedwire Funds are highly concentrated. On an average day in our sample, the top 100 entities (master accounts) are responsible for 89% of the dollar value of all payments sent through Fedwire Funds, which is around $3 trillion per day.

We focus particularly on the 15 largest master accounts and study their payment activity to and from any entity in the top 100. These 15 largest master accounts correspond to national and state member banks, as well as branches and agencies of foreign banks, among other entity types. For simplicity, we refer to all of these as "banks." On an average day in our sample, these top 15 banks are responsible for 76% of the dollar value of all payments sent by the top 100 entities, which is about $2 trillion per day.

Our second dataset covers the reserve balances of depository institutions, including commercial banks, credit unions, and branches and agencies of foreign banks, based on internal Federal Reserve accounting records. We use daily reserve balances that capture end-of-day closing balances for each entity in our sample. In our analysis, we proxy opening balances by the preceding day closing balances. As shown in Figure 3, reserve balances are also concentrated with the largest 15 banks, as measured by the average daily total dollar value of payments sent. These 15 banks held around 40% of the reserves in the US banking system during the first 100 business days in 2020.

3 Strategic complementarity in payments

Prior to the financial crisis of 2007-2009, US banks relied heavily on incoming funds to make their outgoing payments (McAndrews and Potter, 2002; Afonso and Shin, 2011). With the transition from an environment of scarce reserves to one of ample reserves, we question whether this strategic complementarity in payment timing still exists in real-time gross settlement (RTGS) systems such as Fedwire Funds.

To explore this question, we analyze the relationship between a bank's payments and its receipts for each minute in a business day. We focus our analysis on the top 15 most active banks as measured by their average daily total dollar value of payments during the

[4] We exclude entities that the National Information Center (NIC) entity type classifies as "Edge/Agreement Corporations" and "Domestic Entity Other," with the exception of the Federal Home Loan Banks. See https://www.ffiec.gov/npw/Help/InstitutionTypes for a list of NIC institution types.

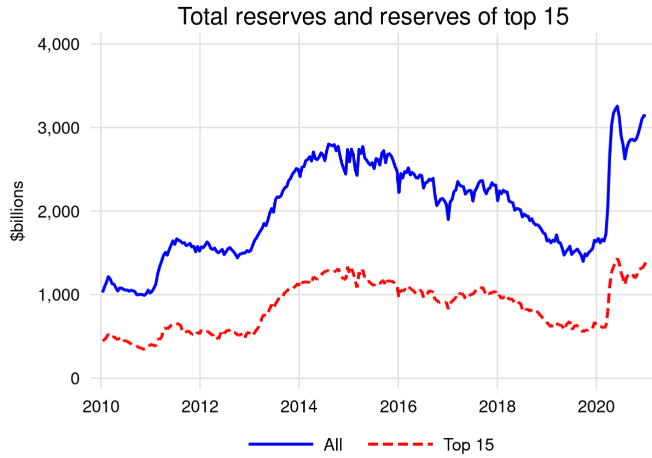

Figure 3: Aggregate reserve balances of US depository institutions (blue) and of top 15 largest entities as defined by average daily total dollar value of payments sent (dashed red), 2010-2020. Source: Internal Federal Reserve accounting records and Fedwire Funds Services.

first 100 business days in 2020, and study how the payments that a top bank makes in a one-minute interval is correlated with its receipts (incoming payments) over the previous 15 minutes.

The dollar value of payments in Fedwire Funds tends to follow a relatively predictable pattern over the course of a typical business day. As shown in Figure 4, there is almost no overnight payment activity. The pace of payments accelerates in the morning after 6:00 am and peaks in the late afternoon at around 5:00 pm, before the payment system closes at 6:30 pm ET.

In all of our model specifications, in addition to bank fixed effects and date fixed effects, we also include period-of-the-day fixed effects to control for these intraday payment dynamics. We include a dummy variable for payments that occur in the first half hour of the business day, between 9:00 and 9:30 pm ET ("opening-of-day payments") because payments may be queued prior to the start of the business day, for processing once the day begins. We also include a dummy variable for payments sent between 9:30 pm and 6:00 am ET ("early payments") to capture the fact that few payments are processed overnight. We also include dummies for each half hour between 2:00 pm and 6:00 pm ET ("afternoon payments") and a dummy variable for payments between 6:00 and 6:30 pm ET ("end-of-day payments") to account for payment dynamics during the last half hour of the business day, when banks can only make settlement payment orders but not payments on behalf of customers.

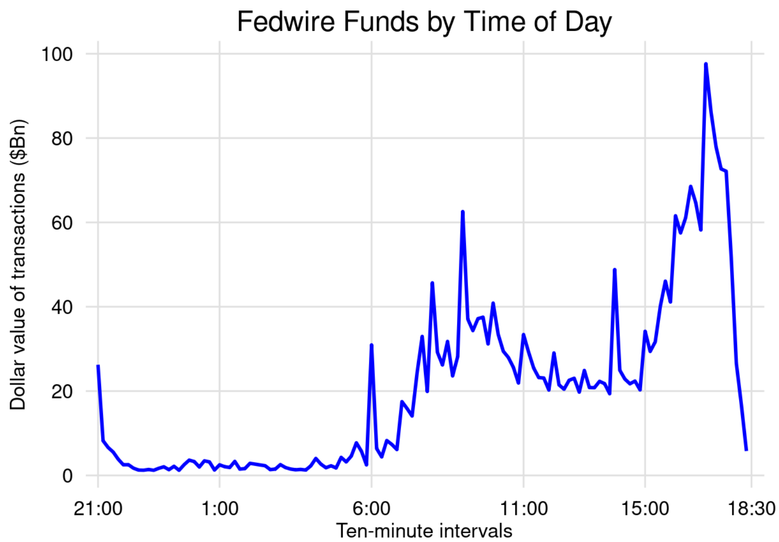

Figure 4: Daily dollar value of payments by time of the day, averaged over the first 100 days in 2020. Source: Fedwire Funds Services.

We estimate the following relationship between a bank's payments and its receipts:

$$\log\left(1+P_{imt}\right)=\beta_0+\beta_1\log\left(1+\sum_{s=m-15}^{m-1}R_{ist}\right)+\gamma_i+\gamma_t+\gamma_m^{open}+\gamma_m^{early}+\gamma_m^{afternoon}+\gamma_m^{eod}+u_{imt}, \tag{1}$$

where P_{imt} is the total dollar value of payments from bank i to its counterparties in minute m of day t; $\sum_{s=m-15}^{m-1}R_{ist}$ is the cumulative receipts of bank i during the previous 15 minutes; γ_i captures a (sender) bank fixed effect; γ_t is a date fixed effect; γ_m^{open} is an indicator variable equal to 1 between 9:00 pm and 9:30 pm ET on the preceding calendar day; γ_m^{early} is an indicator variable equal to 1 between 9:30 pm ET on the preceding calendar day and 6:00 am ET; $\gamma_m^{afternoon}$ is a series of indicator variables for successive 30-minute periods between 2:00 pm and 6:00 pm ET; γ_m^{eod} is an indicator variable equal to 1 between 6:00 pm and 6:30 pm ET; and u_{imt} is an "error" term. In all regressions, standard errors are corrected for heteroscedasticity and clustered at the bank (sender) level. Value amounts are in US dollars (USD).

Table 1 summarizes the results. Column (1) presents the least-squares estimates of the coefficients of our baseline specification and shows the relationship between a top-15 bank's payments over the next minute and its cumulative receipts in the previous 15 minutes, during the first 100 business days of 2020. The point estimate of the coefficient on cumulative receipts is positive (0.04), but not statistically different from zero. A key feature of the payments data is the presence of zeros, that is, minutes during which a bank does not make any payment. Overall, almost 45% of the observations in our dataset correspond to minutes when banks do not send payments. For comparison, column (2) shows our baseline relationship estimated only for those minutes in which banks make payments. The point estimate of the coefficient on cumulative receipts is now 0.119, with a standard error of 0.062, suggesting that a 1% increase in the cumulative payments that bank i receives in the previous 15 minutes translates into a 0.12% increase in the payments made by the bank over the next minute.

However, this linear-in-logs model might not be well-suited to explain the relationship between payments and receipts. First, to isolate the strategic complementarity effect, we focus on high-frequency data. In particular, we use minute-by-minute data for each of the top 15 most active banks. Second, payments in Fedwire Funds are concentrated late in the day, as shown in Figure 4. And while the payment system is open for 21.5 hours each business day, payments sent before 6:00 am ET (which covers 40% of operating hours) represent less than 7% of the total daily value sent. As a result, our data are characterized by minutes during which banks do not make payments (the zero values), and minutes in which banks make payments that take on a wide range of values. Because of this, the least squares estimators of the regression parameters are likely to be biased and inconsistent. So, we estimate a Tobit model, which in principle is able to capture the non-linearity in payments data.

Column (3) presents the results of the Tobit estimation. The point estimate of the coefficient for cumulative receipts is 0.575, with a standard error of 0.179. Since we are estimating a Tobit regression, the estimated coefficient cannot be directly interpreted as the marginal effect of a change in cumulative receipts on payments. To allow for that interpretation, the estimated coefficient can be adjusted by a scale factor that first captures the probability of observing whether a bank makes a (non-zero) payment in the next minute given its cumulative receipts (Wooldridge, 2001). The implied marginal effect, as reported in column (4), is that a 1% increase in the cumulative payments received by bank i in the previous 15 minutes translates into a 0.4% increase in the value of payments that the bank makes over the next minute. This relationship is robust to double-clustering the standard errors at the bank and date level. (See column (2) in Table 2.)

These results show a strong positive and significant relationship between a bank's receipts and the payments made by the bank to its counterparties. We find this relationship during a period when total reserves in the banking system exceeded \$2.5 trillion, suggesting that strategic complementarities in payments exist even in economies with abundant central bank reserves.

3.1 Robustness

Estimating complementarities directly is typically challenging because of omitted-variable concerns. All of our specifications include (i) bank fixed effects to control for unobserved factors that, while constant across days and minutes, vary across banks such as bank size; (ii) date fixed effects that may take into account factors that vary only over time, such as days with a high or low daily payment volume; and (iii) period-of-the-day fixed effects, to control for unobserved factors that vary across minutes, including the intraday payment dynamics shown in Figure 4.

Our analysis may also omit relevant variables that vary over time. For example, as the reserve balances that banks hold at the Federal Reserve increase, banks could rely more on these balances than on incoming payments to make outgoing payments. Duca-Radu and Testi (2021) prioivide some evidence of this in the Eurozone. Omitting balances may affect the relationship between payments and receipts. To assess this effect, specification (3) includes the opening balance B_{it} of bank i on date t. As shown in Table 2, column (3), controlling for opening balances does not alter the estimated relationship between a bank's recent receipts and the payments that it makes.

Our results could also be capturing a spurious relationship that might arise on days with high (or low) payment volumes, regardless of the existence of strategic complementarities in payments. All of our specifications include date fixed effects to capture unobserved factors such as days with high or low payment activity. In addition, to control for payment dynamics when a bank sends out payments, we include in specification (4) each bank's cumulative payments that day up to 15 minutes prior. As shown in Table 2, column (4), there persists a positive and significant relationship between a bank's recent receipts and the payments that it makes, although the estimated magnitude of the effect is reduced. A

1% increase in the cumulative payments that a bank receives in the previous 15 minutes translates, with this specification, into a 0.2% increase in the payments that it makes over the next minute. Model (5) presents an alternative specification in which we control for the payments sent by the bank in the prior minute. Consistent with our previous finding, we still find a positive and significant relationship between a bank's recent receipts and the payments that it makes.

Our results show the relationship between the payments that a bank makes in a given minute and the payments that it has received in the previous 15 minutes. We selected these time windows for consistency with the original work of McAndrews and Potter (2002). We have also considered alternative specifications with different time periods for cumulative receipts and for payments made. The results are qualitatively the same. For instance, column (6) in Table 2 presents the estimated relationship between payments made over the next minute and receipts in the previous 30 minutes. The marginal effect of a 1% increase in a bank's cumulative receipts now translates into a 0.6% increase in the payments it makes over the next minute. We also looked at the relation between receipts and payments made over time periods longer than 1 minute. Specification (7) shows that the relationship persists when based on the payments that a bank makes to its counterparties over 5-minute windows.

McAndrews and Rajan (2000) find evidence during the pre-2000 scarce-reserves regime of a significant synchronous (same-minute) relationship between incoming and outgoing payments, suggesting that that this can arise from explicit coordination between banks. This is distinct from the phenomenon of strategic complementarity that we consider, which arises when a given bank reacts (without coordination) to incoming payments from other banks over a prior time window. McAndrews and Rajan (2000) also suggest that fostering coordinated timing of payments could be a policy approach to mitigating a lack of ample reserves.

Following the empirical approach in Chen, Goldstein, and Jiang (2010), we also explore differences in the level of strategic complementarities faced by banks. Banks face a higher degree of complementarity when aggregate balances in the banking system are low because they need to rely more on incoming funds to make their payments. Similarly, the degree of complementarity is lower for days on which aggregate balances are high, because banks may rely on their own balances to meet some payments and the benefit of coordinating their payments with other banks is reduced. We test for differences in the relationships between payments and past receipts for different levels of aggregate level of balances. Column (8) includes the interaction between a bank's cumulative receipts in the previous 15 minutes and aggregate opening balances for the same day. As shown in column (8), the relationship varies with the level of opening balances: On days with high aggregate opening balances, the link between payments and past receipts persists, but is dampened relative to days with low aggregate balances.

4 Complementarity and reserves

Section 3 finds complementarity in payments in 2020, consistent with reliance by the most active banks on incoming funds to make their payments, even in an era of "ample" reserves. The incentives of banks to adjust the timing of their payments to match that of their receipts creates conditions under which liquidity shocks can propagate through the US payment system. These dynamics are especially important if there are payment delays during days when system-wide reserve balances are relatively low, because banks might further adjust their liquidity management practices, reinforcing the strategic complementarity of payment timing (Yang, 2022).

In this section, we explore in more detail the relationship between strategic complementarity in payment timing and the aggregate level of reserves in the banking system. Column (9) of Table 2 shows an estimate of the relationship between a bank's cumulative receipts in the previous 15 minutes and its payments over the next minute, allowing for an interaction term between the bank's cumulative receipts and aggregate reserves on that day. We find that a 1% increase in the cumulative payments translates into a 0.4% estimated increase in the payments it makes over the next minute, and that this effect is stronger when aggregate reserves are lower.

Next, we look at payment activity on days with high and low aggregate reserve balances to identify potential dependence of the intraday timing of payments on the aggregate level of reserves. Figure 5 shows intraday payments received by the top 15 banks for days with high (blue) and low (dashed red) opening balances, as defined by the top and bottom deciles of the reserve balances distribution. These intraday receipts are aggregated over 20-minute periods, normalized by same-day total receipts, and averaged over the first 100 business days in 2020. As shown in Figure 5, the average share of daily payments that banks receive in the morning is higher on days with high opening balances than on days with low opening balances.

In order to examine this issue more closely, we estimate the relationship between these

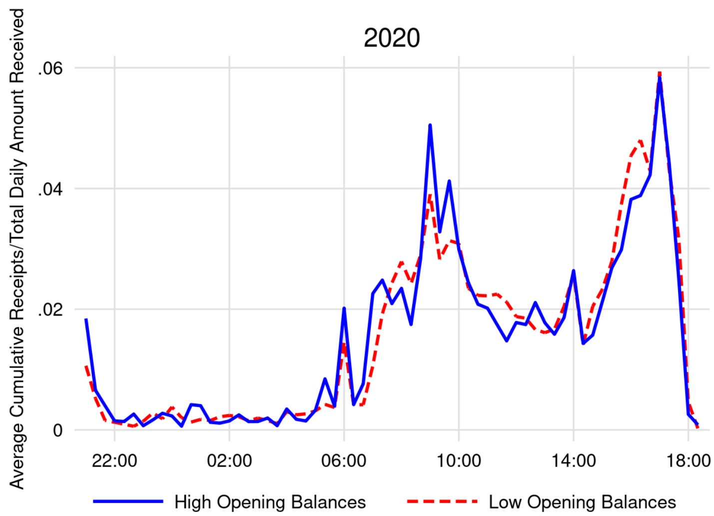

Figure 5: Share of daily payments received by the top 15 banks on top (blue) and bottom (dashed red) decile of days by opening balances, averaged over the first 100 days of 2020. Source: Internal Federal Reserve accounting records and Fedwire Funds Services.

shares of receipts during 20-minute periods throughout a business day and the level of reserve balances that day. For the first 100 business days in 2020, we aggregate the payments received by the top 15 banks over 20-minute periods, normalize these receipts by total receipts on the same day, and regress this ratio on time dummies for each of these 20-minute periods, and on the interaction of these period fixed effects with the total opening balances of the top 15 banks. We are interested in understanding if the relationship between the share of receipts and the level of reserves changes throughout the day. We estimate the specification

$$\frac{R_{st}}{R_t} = \beta_{1s} + \beta_{2s} B_t + u_{st}, \tag{2}$$

where R_{st} is the total dollar value of payments received by the top 15 banks in the 20-minute period s on day t; R_t is the total dollar value of payments received by the top 15 banks on day t; B_t is total opening balances of the top 15 banks on day t; and u_{st} is a residual. In all of our OLS regressions, standard errors are corrected for heteroscedasticity. Quantities are measured in US dollars.

Figure 6 summarizes the estimation of equation (2). Panel (a) presents estimates of β_{1s} for each 20-minute period during the business day. Consistent with the timing of payments shown in Figure 5, panel (a) of Figure 6 shows that the share of payments processed overnight is relatively low. This share increases in the morning at around 6:00 am ET and then again in the afternoon, with the majority of receipts between 3:30 pm and 5:30 pm ET. Panel (b) presents estimates of β_{2s}, showing a positive and statistically significant relationship between the share of receipts and reserve balances in the morning, and a significantly negative relationship in the afternoon. On days with high reserve balances, banks receive a higher fraction of their receipts in the morning, rather than later in the afternoon. McAndrews and Kroeger (2016), Copeland, Molloy, and Tarascina (2019), and Copeland, Duffie, and Yang (2022) find related evidence showing a negative relationship between the time of day by which half of the daily incoming payments to banks are received on the total opening balances of other banks.

(a) (b)

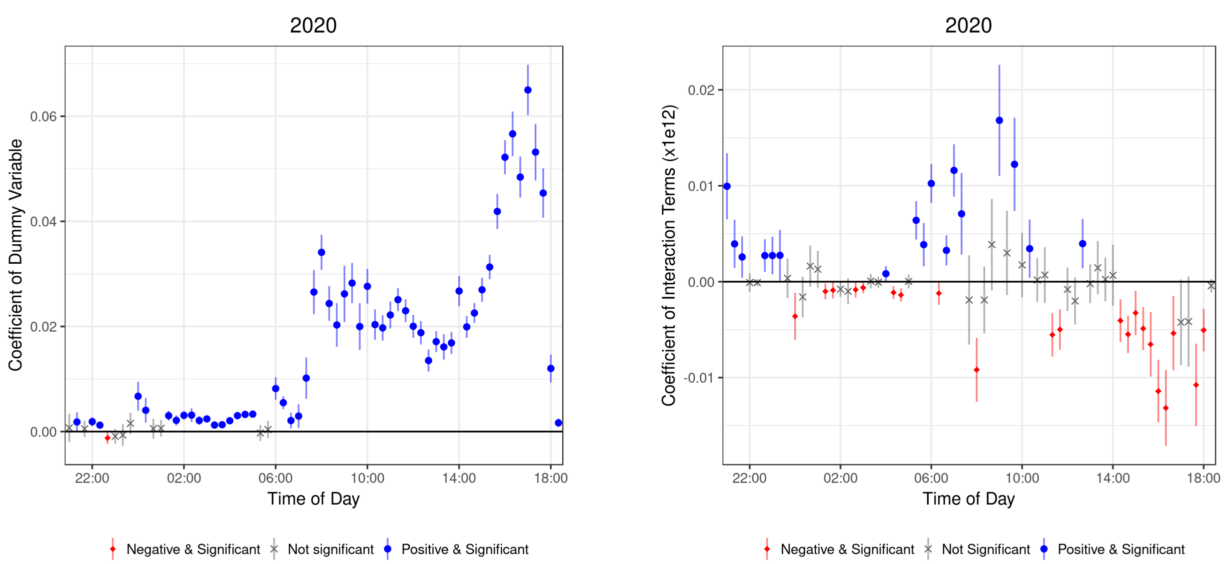

Figure 6: **Share of daily receipts and reserve balances** - Estimates of $\hat{\beta}_{1s}$ (a) and $\hat{\beta}_{2s}$ (b) in equation (2) estimated over 20-minute periods for the top 15 banks during the first 100 days in 2020. Solid blue circles (red diamonds) denote a positive (negative) coefficient that is statistically different from zero at the 90% confidence level. The lengths of the bars correspond to 90% confidence intervals around point estimates. Source: Internal Federal Reserve accounting records and Fedwire Funds Services.

5 Strategic complementarity over time

We showed in sections 3 and 4 that, in 2020, top banks react in the timing of their payments to the timing chosen by other banks, and that this relationship changes with the level of aggregate reserves. In this section, we look at the evolution of this relationship since 2010. We first estimate the relationship between the payments made by a bank over one-minute intervals and the payments that it received in the previous 15 minutes, using the baseline specification in equation (1). We estimate quarterly Tobit regressions for the top 15 most active banks as measured by their average total dollar value of payments in each quarter over 2010-2020. Figure 7 shows the Tobit coefficient for each quarter in our sample (blue), as well as the time evolution of the total reserves balances held by these top 15 most active banks (dashed red).

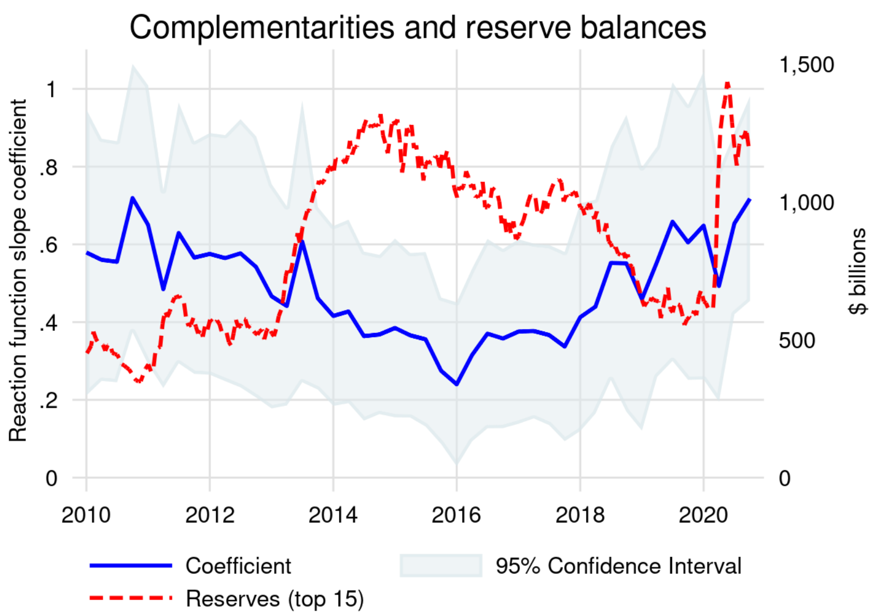

Figure 7: **Complementarities and reserve balances** - Quarterly Tobit coefficient estimates $\hat{\beta}_1$ (in equation (1)) and reserve balances of top 15 largest entities as defined by average daily total dollar value of payments sent (dashed red), 2010-2020. Source: Internal Federal Reserve accounting records and Fedwire Funds Services.

As shown in Figure 7, the positive relationship between a bank's outgoing payments and

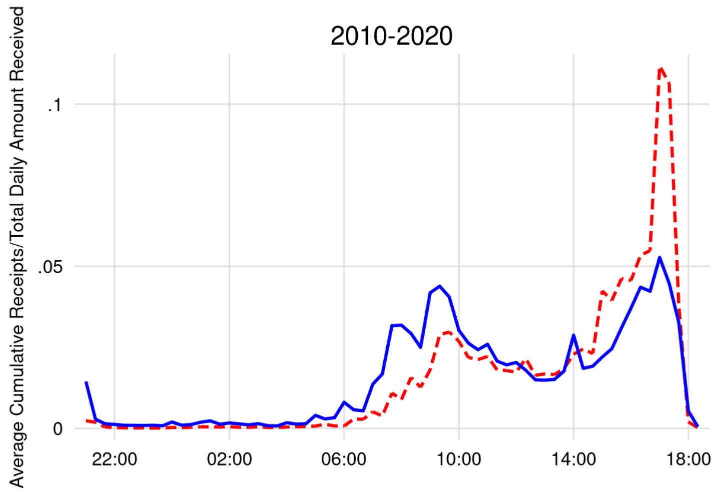

Figure 8: Share of daily payments received by the top 15 banks on top (blue) and bottom (dashed red) decile of days by opening balances, averaged over 2010-2020. Source: Internal Federal Reserve accounting records and Fedwire Funds Services.

its recent incoming payments is not just a feature of payments dynamics in 2020, but has been present since 2010. The evolution of this reactive approach to payment timing seems to mirror that of reserve balances, with higher reliance on incoming payments when reserve balances are lower. Table 3 replicates the analysis in Table 2 from 2010 to 2020. Results are qualitatively similar to those estimated in 2020, although of lower magnitude. These results capture the average effect over the ten-year period, where aggregate reserves ranged from \$1 trillion to over \$3 trillion. The point estimate of the coefficient for cumulative receipts in our baseline specification in equation (1) is 0.490, with a standard error of 0.115. In terms of marginal effects, a 1% increase in the cumulative payments received by bank i in the previous 15 minutes translates into a 0.24% increase in the value of payments that the bank makes over the next minute.

Figure 8 displays the average share of payments received by time of the day on days with high (blue) and low (dashed red) opening balances, as defined by the top and bottom deciles of the the distribution of aggregate reserve balances. On days with low reserve balances, banks are more reactive in their payment timing. Moreover, on low-balance days, banks receive a higher share of their daily payments later in the day.

Figure 9 summarizes the estimated relationship between the share of receipts and the level of reserves over 20-minute periods from 2010 to 2020 (equation (2)). The estimates of the coefficient $\hat{\beta}_{2s}$ in panel (b) show that on days with higher opening balances, a higher share of the daily payments is sent earlier in the day (in the morning).

(a) (b)

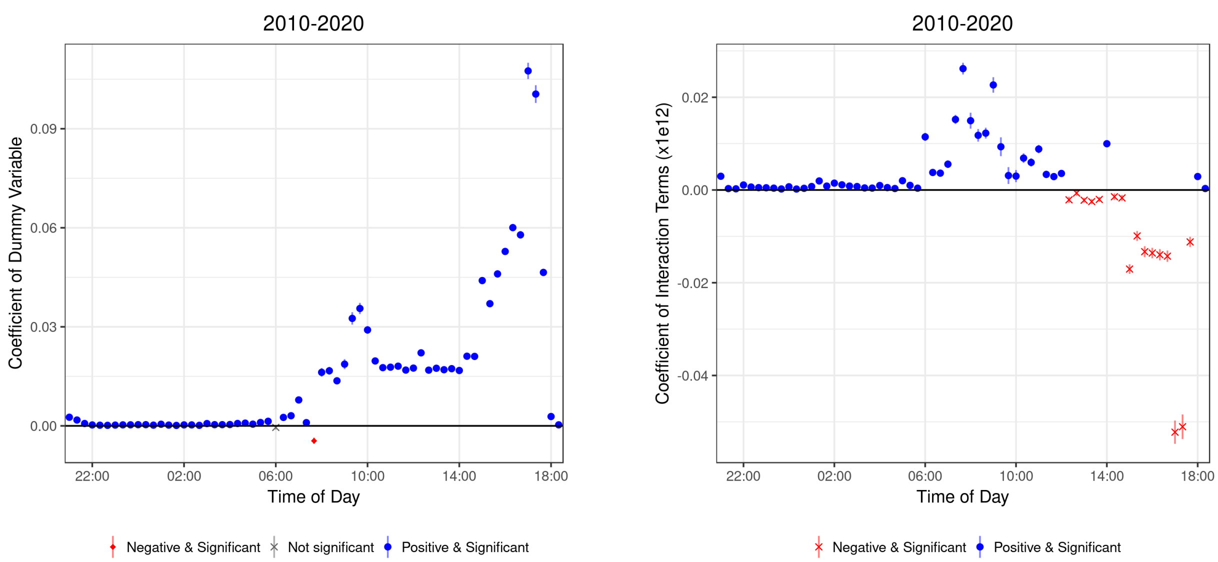

Figure 9: **Share of daily receipts and reserve balances** - Estimates of $\hat{\beta}_{1s}$ (a) and $\hat{\beta}_{2s}$ (b) over 20-minute periods for the top 15 banks during 2010-2020. Solid blue circles (red diamonds) denote a positive (negative) coefficient that is statistically different from zero at the 90% confidence level. The lengths of the bars correspond to 90% confidence intervals around point estimates. Data sources: Internal Federal Reserve accounting records and Fedwire Funds Services.

6 Concluding remarks

Even in the post-GFC era of ample reserves, the payment timing decisions of banks in the US wholesale RTGS system suggest that access to central-bank reserve balances is a constraint on funding liquidity. We find strong evidence that US banks still economize on intraday liquidity. They rely on incoming payments to make outgoing payments, showing a high degree of strategic complementarity in their payment decisions. These results persist despite large aggregate reserves balances in the banking system, well in excess of the aggregate balances prior to the Global Financial Crisis. Our results shed light into the ongoing discussion of the sizing of central bank balance sheets and the shifting nature of banks' demand for reserves. As central banks around the world respond to inflation by tightening their monetary stance and shrinking their balance sheets, the potential consequences for the wholesale payment system of the ongoing draining by central banks of reserves will likely be an important input into policy making.

Consistent with our interpretation of the data, our findings also show that the strength of the strategic complementarity of payment timing varies with the level of aggregate reserves, becoming stronger as aggregate reserve balances decline. This suggests a potential for strategic cash hoarding when reserve balances get sufficiently low, as was the case in mid-September 2019 and mid-March 2020. A shift in the business models of banks has led reserve balances to be used increasingly for short-term funding operations, including the intermediation of FX swap markets, Treasury repo markets, and other short-term funding markets. The shrinking supply of reserve balances may come to have potentially important implications for market functioning and financial stability.

Table 1: **Bank payments and receipts.** - The dependent variable is the logarithm of (1 plus) the total dollar value of bank i's payments to its counterparties in minute m of day t, P_{imt}. $\sum_{s=m-15}^{m-1} R_{ist}$ is bank i's cumulative receipts during the previous 15 minute. In columns (1) and (2), we estimate a linear model using OLS. In column (3), we estimate a Tobit regression. Standard errors are corrected for heteroscedasticity, clustered at the bank (sender) level. Variables are in USD. Specifications are run in logs. Sample includes payments of top 15 most active banks during the first 100 business days in 2020.

	$\log(1+P_{imt})$			
	Linear (OLS)		Tobit (MLE)	
	(1)	(2)	(3)	(4)
	$y \geq 0$	$y > 0$	Coefficient	Marginal
$\log(1+\sum_{s=m-15}^{m-1} R_{ist})$	0.040	0.119*	0.575***	0.395
	(0.083)	(0.062)	(0.179)	
Clustering	Bank	Bank	Bank	
Bank FEs	Y	Y	Y	
Date FEs	Y	Y	Y	
Early dummy	Y	Y	Y	
EOD dummy	Y	Y	Y	
Afternoon dummies	Y	Y	Y	
Open dummy	Y	Y	Y	
N	1,935,000	1,059,902	1,935,000	
Left-censored			875,098	
R^2	0.597	0.355		
Pseudo R^2			0.228	
Log-likelihood			-3,157,609.4	

Table 2: **Bank payments and receipts. Robustness -** The dependent variable is the logarithm of (1 plus) the total dollar value of bank i's payments to its counterparties in minute m of day t, P_{imt}. $\sum_{s=m-15}^{m-1} R_{ist}$ is bank i's cumulative receipts during the previous 15 minute, and B_{it} its opening balance on day t. We estimate a Tobit regression. Standard errors are corrected for heteroscedasticity, clustered at the bank (sender) level in specifications (1) and (3)-(9), and double-clustered at the bank and date level in specification (2). Variables are in USD. Specifications are run in logs. Sample includes payments of top 15 most active banks during the first 100 business days in 2020.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+\sum_{s=m}^{m+4} P_{ist})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$
$\log(1+\sum_{s=m-15}^{m-1} R_{ist})$	0.575***	0.575***	0.575***	0.302***	0.340***		0.519***	2.891**	3.131**
	(0.179)	(0.019)	(0.179)	(0.117)	(0.086)		(0.145)	(1.147)	(1.323)
$\log(1+\sum_{s=m-30}^{m-1} R_{ist})$						0.842***			
						(0.306)			
$\log B_{it}$			0.089*						
			(0.050)						
$\log(1+\sum_{s=1}^{m-16} P_{ist})$				0.899***					
				(0.121)					
$\log(1+P_{im-1t})$					0.681***				
					(0.065)				
$\log B_t$								1.474	
								(1.117)	
$\log B_t \times \log(1+\sum_{s=m-15}^{m-1} R_{ist})$								-0.194*	
								(0.105)	
$\log \text{Reserves}_t$									1.573
									(1.228)
$\log \text{Reserves}_t \times \log(1+\sum_{s=m-15}^{m-1} R_{ist})$									-0.207*
									(0.117)
Marginal effect of receipts	.395	.395	.395	.189	.257	.57	.469	.395	.395
Clustering	Bank	Bank × Day	Bank	Bank	Bank	Bank	Bank	Bank	Bank
Bank FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Date FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Early dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
EOD dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
Afternoon dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Open dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	1,935,000	1,935,000	1,935,000	1,935,000	1,935,000	1,935,000	1,935,000	1,935,000	1,935,000
Left-censored	875,098	875,098	875,098	875,098	875,098	875,098	656,455	875,098	875,098
Pseudo R^2	0.228	0.228	0.228	0.258	0.277	0.230	0.214	0.228	0.228

Values are log dollars. Sample is top 15 entities by average daily payment value in first 100 days of 2020.

Table 3: **Bank payments and receipts. Robustness -** The dependent variable is the logarithm of (1 plus) the total dollar value of bank i's payments to its counterparties in minute m of day t, P_{imt}. $\sum_{s=m-15}^{m-1} R_{ist}$ is bank i's cumulative receipts during the previous 15 minute, and B_{it} its opening balance on day t. We estimate a Tobit regression. Standard errors are corrected for heteroscedasticity, clustered at the bank (sender) level in specifications (1) and (3)-(9), and double-clustered at the bank and date level in specification (2). Variables are in USD. Specifications are run in logs. Sample includes payments of top 15 most active banks over 2010-2020.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$	$\log(1+\sum_{s=m}^{m+4} P_{ist})$	$\log(1+P_{imt})$	$\log(1+P_{imt})$
$\log(1+\sum_{s=m-15}^{m-1} R_{ist})$	0.490***	0.490***	0.490***	0.323***	0.295***		0.476***	3.865**	4.290**
	(0.115)	(0.003)	(0.115)	(0.067)	(0.046)		(0.074)	(1.635)	(2.051)
$\log(1+\sum_{s=m-30}^{m-1} R_{ist})$						0.640***			
						(0.196)			
$\log B_{it}$			-0.045						
			(0.112)						
$\log(1+\sum_{s=1}^{m-16} P_{ist})$				0.804***					
				(0.107)					
$\log(1+P_{im-1t})$					0.732***				
					(0.080)				
$\log B_t$								2.367**	
								(1.156)	
$\log B_t \times \log(1+\sum_{s=m-15}^{m-1} R_{ist})$								-0.284**	
								(0.139)	
$\log \text{Reserves}_t$									2.584*
									(1.396)
$\log \text{Reserves}_t \times \log(1+\sum_{s=m-15}^{m-1} R_{ist})$									-0.310*
									(0.169)
Marginal effect of receipts	.234	.234	.234	.128	.17	.29	.366	.232	.233
Clustering	Bank	Bank × Day	Bank	Bank	Bank	Bank	Bank	Bank	Bank
Bank FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Date FEs	N	N	N	N	N	N	N	N	N
Early dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
EOD dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
Afternoon dummies	Y	Y	Y	Y	Y	Y	Y	Y	Y
Open dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
Day-of-week FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-of-year FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
Month-end dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
Mid-month dummy	Y	Y	Y	Y	Y	Y	Y	Y	Y
Quarter-year FEs	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	53,270,550	53,270,550	53,270,550	53,270,550	53,270,550	53,270,550	53,270,550	53,270,550	53,270,550
Left-censored	28,541,241	28,541,241	28,541,241	28,541,241	28,541,241	28,541,241	22,790,543	28,541,241	28,541,241
Pseudo R^2	0.263	0.263	0.263	0.288	0.316	0.264	0.250	0.263	0.263

Values are log dollars. Sample is top 15 entities by average daily payment value in 2010-2020.

References

Gara Afonso and Hyun Song Shin. Precautionary demand and liquidity in payment systems. *Journal of Money, Credit and Banking*, 43(s2):589–619, 2011.

Gara Afonso, Marco Cipriani, Adam Copeland, Anna Kovner, Gabriele La Spada, and Antoine Martin. The market events of mid-September 2019. *Federal Reserve Bank of New York Economic Policy Review*, 27(2), 2021.

Gara Afonso, Domenico Giannone, Gabriele La Spada, and John C. Williams. Scarce, abundant, or ample? A time-varying model of the reserve demand curve. *Federal Reserve Bank of New York Staff Reports*, (1019), 2022.

Stefan Avdjiev, Wenxin Du, Catherine Koch, and Hyun Song Shin. The dollar, bank leverage, and deviations from covered interest parity. *American Economic Review: Insights*, 1:193–208, 2019.

Morten Bech. Intraday liquidity management: A tale of games banks play. *Economic Policy Review*, 14:7–23, 2008.

Morten Bech and Rod Garratt. The intraday liquidity management game. *Journal of Economic Theory*, 109:198–219, 2003.

Qi Chen, Itay Goldstein, and Wei Jiang. Payoff complementarities and financial fragility: Evidence from mutual fund outflows. *Journal of Financial Economics*, 97:239–262, 2010.

Adam Copeland, Linsey Molloy, and Anya Tarascina. What can we learn from the timing of interbank payments? *Liberty Street Blog, Federal Reserve Bank of New York*, 2019.

Adam Copeland, Darrell Duffie, and Yilin Yang. Reserves were not so ample after all: Evidence from the US Treasury repo market. *Working Paper, Stanford Graduate School of Business*, 2022.

Ricardo Correa, Wenxin Du, and Gordon Liao. US banks and global liquidity. *NBER Working Paper*, 2020.

Adrien d'Avernas and Quentin Vandeweyer. Intraday liquidity and money market dislocations. *Working paper, Stockholm School of Economics, August*, 2021.

Ioana Duca-Radu and Sara Testi. Liquidity usage in TARGET2. *ECB Economic Bulletin*, (3), 2021.

James D. Hamilton. Measuring the liquidity effect. *The American Economic Review*, 87 (1):80–97, 1997.

Jane Ihrig. Banks' demand for reserves in the face of liquidity regulations. *On the Economy Blog, Federal Reserve Bank of St. Louis, March*, 2019.

David Lopez-Salido and Annette Vissing-Jorgensen. Reserve demand and balance sheet run-off. *Federal Reserve Board Working Paper*, 2022.

James McAndrews and Alexander Kroeger. The payment system benefits of high reserve balances. *Journal of Payments Strategy and Systems*, 10(1):72–83, 2016.

James McAndrews and Simon Potter. Liquidity effects of the events of September 11, 2001. *Economic Policy Review*, 8, 2002.

James McAndrews and Samira Rajan. The timing and funding of Fedwire funds transfers. *Federal Reserve Bank of New York Economic Policy Review*, 6(2):17–32, 2000.

Thomas Nellen. Intraday liquidity facilities, late settlement fee and coordination. *Journal of Banking and Finance*, 106:106–131, 2019.

A. Lee Smith. Do changes in reserve balances still influence the federal funds rate? *Federal Reserve of Kansas City Economic Review*, 2019.

Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, first edition, 2001.

Yilin Yang. What quantity of reserves is sufficient? *Working Paper, Stanford Graduate School of Business*, 2022.