Money Market Funds and the Pricing of Near-Money Assets*

Sebastian Doerr† Egemen Eren‡ Semyon Malamud§

January 27, 2023

Abstract

US money market funds (MMFs) play an important role in short-term markets as large investors of Treasury bills (T-bills) and repurchase agreements (repos) with banks and the Federal Reserve, some of the world's safest and most liquid assets. We build a theoretical model in which MMFs' strategic interactions generate a trade-off between their market power in the repo market and their price impact in the T-bill market. Empirically, we show that MMFs' portfolio allocation decisions between repos and T-bills have an economically significant impact on T-bill rates and market liquidity, and the liquidity premium on T-bills. Guided by our model, we devise instrumental variables to establish a causal effect. Using a granular holding-level dataset we confirm the model's prediction that MMFs internalize their price impact in the T-bill market when they set repo rates. Moreover, when Treasury market liquidity is low, MMFs tilt their portfolios away from T-bills towards repos with the Federal Reserve. Our results have broad implications.

Keywords: T-bills, repo, market power, price impact, liquidity premium, money market funds
JEL Classification Numbers: E44, G11, G12, G23

*We are grateful to Stijn Claessens, Wenxin Du, Darrell Duffie and Alp Simsek for helpful comments and suggestions. Alexis Maurin and Albert Pierres Tejada provided excellent research assistance. Semyon Malamud acknowledges the financial support of the Swiss National Science Foundation (Grant 100018_192692, "International Macroeconomics with Financial Frictions") and the Swiss Finance Institute. Parts of this paper were written when Malamud visited the BIS. The views in this article are those of the authors and should not be attributed to the BIS. All errors are our own.

†Bank for International Settlements (BIS); Email: `sebastian.doerr@bis.org`

‡*Corresponding author.* Bank for International Settlements (BIS); Email: `egemen.eren@bis.org`

§Swiss Finance Institute, EPF Lausanne, and CEPR; E-mail: `semyon.malamud@epfl.ch`

1 Introduction

US Treasury bills (T-bills) and repurchase agreements (repos) are among the most important financial instruments of global finance. T-bills are considered the world's safest and the most liquid government debt security (e.g. Krishnamurthy and Vissing-Jorgensen, 2012; Greenwood, Hanson and Stein, 2015; Nagel, 2016). Repos are safe investment vehicles instrumental for banks and other financial institutions to raise short-term capital and manage their liquidity needs (CGFS, 2017). The importance of repos is set to increase further as repo-based reference rates are replacing LIBOR in derivatives markets. Meanwhile, repos issued by the Federal Reserve through its reverse repo (RRP) facility constitute a critical monetary policy instrument (Afonso et al., 2022a).[1]

Aside from their safety and liquidity, T-bills, repos with banks, and repos with the Federal Reserve share another critical feature: US money market funds (MMFs) are key investors. US MMFs' total assets under management (AUM) are around $5 trillion, equal to roughly 20% of US GDP or total US commercial bank assets. Around $3 trillion of MMFs' AUM are held in safe near-money assets, such as T-bills and repos with banks and the Federal Reserve. At its peak, MMFs held around 45% of total marketable T-bills, and their holdings of repos with banks amount to around $500 billion. At its peak utilization at the end of September 2022, MMFs were by far the largest investor group in the RRP facility, with an investment of $2.2 trillion out of the total of $2.4 trillion. Their use of the RRP facility has skyrocketed from zero within a few months in 2022.

Our main contribution is to theoretically and empirically show that strategic interactions of MMFs with banks and each other significantly impact the pricing of the world's most liquid assets. We show theoretically that, when MMFs exercise market power in the repo market and banks demand fewer repos, MMFs have more "residual cash" left over from repo lending to banks.[2] As a result, their demand for T-bills increases and T-bill rates rise. Moreover, market liquidity worsens and the liquidity premium attached to T-bills increases. Empirically, we show that, on aggregate,

[1] From the point of view of the collateral providers, these transactions are called repos. From the point of view of cash lenders, these operations are called reverse repos. In both repos with banks and the Federal Reserve, MMFs are cash lenders, with banks and the Federal Reserve providing collateral. For brevity, we refer to these transactions as repos throughout the paper.

[2] Guided by our theory, we use the "residual cash share", which is the share of funds left over from repo lending to banks from the total invested between repos, T-bills and the RRP, as the right-hand side variable in our aggregate time-series regressions (see Section 5 for a detailed discussion).

MMFs' portfolio allocations have the predicted and economically significant impact on T-bill rates, market liquidity and the liquidity premium on T-bills.[3] We devise instruments guided by our model to show that the impact is causal. Next, using a granular, holding-level dataset, we show that MMFs indeed internalize their price impact in the T-bill market while setting repo rates. Moreover, when liquidity in the Treasury market is low, MMFs tilt their portfolios away from T-bills towards the RRP facility. This finding helps to rationalize MMFs' meteoric rise of the utilization of the RRP facility, at least in part, with the deterioration of liquidity conditions in the Treasury market.

Quantitatively, the partial impact of a standard deviation increase in the residual cash share on T-bill rates is equivalent to the partial effect of a one percentage point increase in the federal funds rate, or a third of a percent increase in the T-bill supply normalized by GDP. The partial impact of a standard deviation increase in the residual cash share reduces market liquidity by around half a standard deviation across our liquidity measures. Its partial impact on the liquidity premium is equivalent to the partial effect of slightly more than 1 percentage point rise in the federal funds rate, or around one-fifth of a percent rise in the T-bill supply normalized by GDP.

We begin our analysis by presenting a model of strategic interactions of MMFs with banks in the repo market and each other in the T-bill market. Funds optimally set repo rates with banks and allocate their funds between T-bills, repos with banks, and repos with the Federal Reserve (RRP). Consistent with the data, there is market concentration, and large funds simultaneously account for a significant share in the repo and T-bill markets, which creates the critical trade-off between *market power* in the repo market and *price impact* in the T-bill market. All else equal, funds with greater *market power* charge higher rates, so banks demand fewer repos. As a result, MMFs allocate more of their cash to T-bills, thereby lowering T-bill rates due to their *price impact*. In the model, MMFs on aggregate also have an impact on Treasury market liquidity.

Our model also provides insights into why MMFs invest both in T-bills and the RRP facility, two

[3] In our time series regressions, we use the (1-month) T-bill-RRP spread, which measures the rates on T-bills after taking out the aggregate macroeconomic conditions reflected by the RRP, to show MMFs' impact on T-bill rates. We use the Bloomberg Liquidity Index and an Amihud liquidity measure for 1-month T-bills (Amihud, 2002) to show their impact on market liquidity. We use the 1-month General Collateral (GC) repo-T-bill spread to measure the liquidity premium. This is the measure typically used in the literature as term GC repos are as safe as T-bills since they have US Treasury collateral but are much less liquid than T-bills (e.g. Duffee, 1996; Longstaff, 2000; Nagel, 2016; d'Avernas and Vandeweyer, 2021).

close substitutes for cash management purposes. The RRP facility provides financial institutions ineligible to receive interest on reserve balances held at the Federal Reserve, particularly MMFs, with a risk-free overnight investment option. The rate is set administratively by the Federal Reserve and is fixed. Even though there are counterparty limits, all MMFs are comfortably below this limit as of the end of 2022. Therefore, contrary to the T-bill market, MMFs do not need to worry about price impact when they invest in the RRP. This leads to an interesting problem for the allocation between T-bills and RRP. Without any price impact considerations in the T-bill market, one might expect that MMFs would allocate all of their "residual cash" (i.e., funds left over from repo lending to banks) either entirely in T-bills or the RRP, depending on which instrument offers a higher rate. However, a significant fraction of MMFs split their "residual cash" between T-bills and the RRP facility. In our model, funds choose an interior solution where they invest both in T-bills and the RRP facility. Therefore, the existence of the RRP does not entirely eliminate the price impact considerations for MMFs in T-bill markets, even though it alleviates them. The model predicts that an individual MMF's price impact considerations intensify when liquidity in Treasury markets is low as they tilt their demand towards the RRP compared to T-bills when the Treasury market is illiquid.

The model generates predictions for how MMFs impact the pricing of T-bills. For one, if MMFs have more "residual cash" left over from repo lending, this should reduce T-bill rates due to funds' price impact in this market. We start by demonstrating a significant negative effect of MMFs' "residual cash share" on the T-bill-RRP spread in regressions at the monthly level. By computing the spread of T-bill rates over the RRP rate, we account for out the impact of aggregate macroeconomic conditions, as they are reflected in the interest rate on RRP set by the Federal Reserve. Similarly, we show that a higher residual cash share negatively effects market liquidity. Moreover, our measure of the liquidity premium, the GC-T-bill spread, is higher when MMFs allocate less of their funds to repos with banks, suggesting a higher liquidity premium attached to T-bills. Importantly, adding the residual cash share to standard specifications in the literature that explain the spread improves the explanatory power markedly.

In our model the residual cash share variable is an equilibrium outcome, so regressing rates on quantities suffers from endogeneity. Isolating variation in the residual cash that is not driven by the T-bill rate itself hence requires an instrumental variable.

To estimate the causal effect of MMFs' portfolio allocation on T-bill rates we construct instrumental variables guided by our theory. In the model, higher market concentration in the repo market leads to higher repo rates, which lowers the aggregate bank demand for repos, which we observe in the micro data. As a consequence, MMFs have more "residual cash" to invest in the T-bill market. Because of funds' price impact, T-bill rates decline relative to the RRP rate, and increase relative to the GC repo rates, reflecting the rising liquidity premium of T-bills.

Our main instrumental variable is the HHI index of market concentration in the bank repo market, constructed from funds' market shares in the repo market. We argue that the instrument has several desirable features to isolate variation in residual cash that is exogenous to conditions in the T-bill market. In our model, the HHI only impacts T-bill rates through its effect on residual cash: greater concentration in the repo market means funds charge higher rates and face lower demand by banks, so they must allocate more of their cash to the T-bill market. Note that the presence of the RRP does not fully eliminate this given that due to frictions of utilizing the RRP facility, MMFs optimally choose an interior solution between T-bills and the RRP facility, which is what we observe in the data.

We use two alternative instruments to address any remaining concerns with our identification strategy. A concern could be that banks' demand for repos might jointly affect the HHI and T-bill markets, violating the exclusion restriction. To isolate variation in the HHI that is not driven by banks' demand, we take out each bank's demand at a given time by first regressing the size of each individual contract on bank-time fixed effects. We then reconstruct the HHI from the residuals. Second, we exploit the 2016 US MMF reform that drastically increased concentration in the repo market, but that was plausibly exogenous to developments in the Treasury market and bank market (see Aldasoro, Ehlers and Eren (2022)). Our estimates using these alternative instruments are similar to our those with our baseline instrument.

After establishing the aggregate impact of MMFs in the T-bill market, we turn to the trade-offs they face when setting repo rates and the impact of overall liquidity conditions in the Treasury markets on their individual portfolio allocations. We test the model predictions using a detailed dataset of US MMFs' portfolio holdings at individual holding levels at month-ends. The data, obtained from MMFs' regulatory filings, cover the universe of US MMF funds between February 2011 to October 2022 and provide detailed information on contract characteristics for each holding. We construct counterparts of key model variables in the data for funds' market shares in the repo and T-bill markets and their residual cash, overall market concentration in the repo market, and Treasury market liquidity.

We find evidence for the predictions of the model on the market power-price impact trade-off and the impact of liquidity conditions in the Treasury market on MMF portfolio allocations. First, funds' market share (i.e. market power) in the repo market affects repo rates between MMFs and banks positively, whereas funds' market share in the T-bill market affects repo rates negatively. This result is consistent with the interpretation that MMFs internalize their price impact in the T-bill market when they set repo rates. Due to the granular nature of our data, we can control for a battery of time-varying fixed effects to rule out alternative explanations arising from potential differences in time-varying unobservable bank, instrument, or fund-type characteristics. Second, we show that funds allocate more of their "residual cash" to the RRP if the Treasury market liquidity is low, even when controlling for fund-fixed effects and time-fixed effects (which absorb the impact of movements in the interest rates of T-bills and the RRP facility).

All in all, our results have implications for the transmission of monetary policy, the robustness of repo-based benchmark rates and also for government debt issuance. MMFs typically receive inflows when the federal funds rate is higher (e.g. Duffie and Krishnamurthy, 2016; Drechsler, Savov and Schnabl, 2017; Xiao, 2020) and during flight-to-quality or dash-for-cash episodes (in the case of government and Treasury MMFs, (e.g. Eren, Schrimpf and Sushko, 2020b). Therefore, the market power-price impact trade-off could lead to downward pressure on T-bill rates during these episodes. This effect might also transmit to the long end of the yield curve due to the high sensitivity of long-term rates to short-term rates (Hanson, Lucca and Wright, 2021) or might create negative externalities through incentivizing private money creation (Greenwood, Hanson and Stein, 2015). These pressures could be exacerbated if a smaller central bank balance sheet results in a smaller role for the RRP facility and therefore a larger price impact in T-bill markets. Moreover, as the transition from credit-sensitive benchmark rates (LIBOR) to risk-free repo-based benchmark rates (SOFR) is underway (Huang and Todorov, 2022), our results imply that the market structure and liquidity conditions in repo and Treasury markets might impact benchmark rates and hence have spillover effects to many other markets.[4]

Related literature. Our paper contributes to a large literature that studies the relationship between the supply of and liquidity conditions in government debt markets, short-term rates, and money market conditions (e.g. Amihud and Mendelson, 1991; Duffee, 1996; Longstaff, 2004; Goldreich, Hanke and Nath, 2005; Krishnamurthy and Vissing-Jorgensen, 2012; Greenwood and Vayanos, 2014; Krishnamurthy and Vissing-Jorgensen, 2015; Sunderam, 2015; Greenwood, Hanson and Stein, 2015; Nagel, 2016; Lenel, 2017; Lenel, Piazzesi and Schneider, 2019; Li, Ma and Zhao, 2019; Klingler and Sundaresan, 2020; d'Avernas and Vandeweyer, 2021). In particular, liquidity conditions and market functioning issues in the Treasury market came to the fore during the Covid-19 crisis (e.g. Duffie, 2020; Schrimpf, Shin and Sushko, 2020; Eren and Wooldridge, 2021; Du, Hébert and Li, 2022; He, Nagel and Song, 2022). Moreover, interest rates on T-bills and the liquidity premium, that is the premium paid by investors for the liquidity services provided by US Treasuries, and in particular, T-bills, has fostered an important body of research both in macroeconomics (e.g. Krishnamurthy and Vissing-Jorgensen, 2012; Nagel, 2016) and in international finance (e.g. Jiang, Krishnamurthy and Lustig, 2021; Engel and Wu, 2021). Our paper contributes to this literature by highlighting the key role MMFs play in Treasury markets and how market frictions in repo markets and liquidity conditions in T-bill markets can affect key interest rates and spreads in the macroeconomy through MMFs' portfolio allocations with important consequences for the global economy.

Our paper also contributes to the literature on understanding the role of MMFs in money markets and the transmission of monetary policy. Several studies have documented the key role of MMFs in the repo and other short-term money markets, including during crisis episodes (e.g.

[4]MMFs also contributed to ructions in repo markets in September 2019 which led to significant moves in the SOFR, which is a case in point (e.g Avalos, Ehlers and Eren, 2019; Afonso, Cipriani, Copeland, Kovner, La Spada and Martin, 2020; Correa, Du and Liao, 2020; d'Avernas and Vandeweyer, 2020; Copeland, Duffie and Yang, 2021).

Kacperczyk and Schnabl, 2013; Chernenko and Sunderam, 2014; Krishnamurthy, Nagel and Orlov, 2014; Copeland, Martin and Walker, 2014; Schmidt, Timmermann and Wermers, 2016; Han and Nikolaou, 2016; Eren, Schrimpf and Sushko, 2020a,b; Hu, Pan and Wang, 2021; Cipriani and La Spada, 2021; Li, 2021; Aldasoro, Ehlers and Eren, 2022; Anderson, Du and Schlusche, 2022; Huber, 2022; Afonso, Cipriani and La Spada, 2022b). Our contribution is to jointly account for optimal price setting and allocations between T-bills, RRP, and repos, assets that makeup two-thirds of the aggregate money market fund sector's portfolios. MMFs have an important role in the transmission of monetary policy (e.g. Duffie and Krishnamurthy, 2016; Drechsler, Savov and Schnabl, 2017; Xiao, 2020). We show a channel through which frictions specific to MMFs can weaken the transmission of monetary policy.

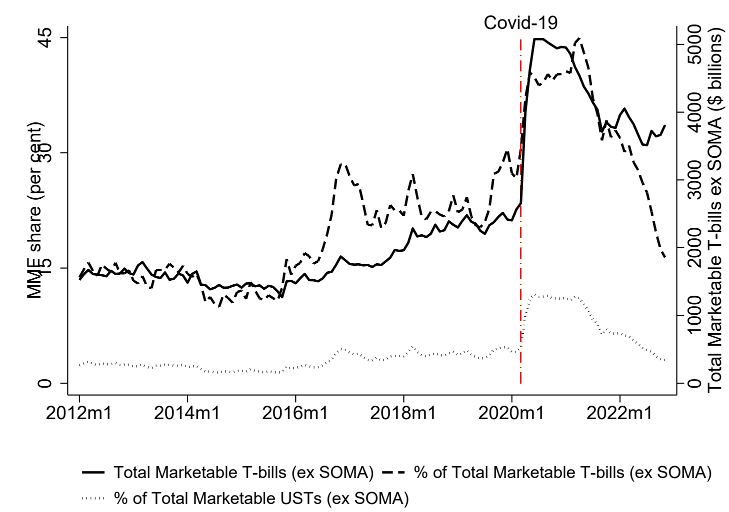
2 Facts on MMFs, T-bills, repos and the RRP facility

Description of MMFs. MMFs are short-term investment vehicles with a total AUM of around $5 trillion as of November 2022. They issue money market shares to investors and invest proceeds into short-term investments. For example, the weighted average maturity of the holdings of a median fund is around one month in our dataset.

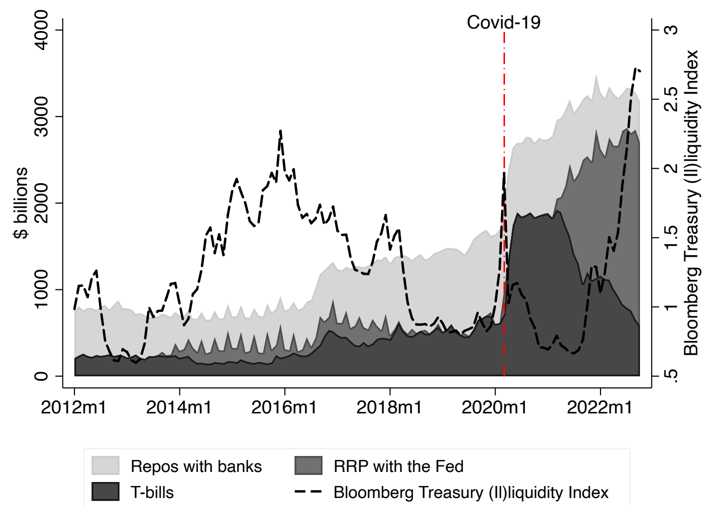
There are three major types of funds that we focus on in our study. *Treasury funds* are only allowed by regulation to invest in T-bills and repos backed by Treasury securities. *Government funds* are allowed to invest in a broader set of instruments, such as agency debt and repos backed by agency collateral in addition to what the Treasury funds can invest in; and *prime funds* that can also invest in unsecured instruments, such as commercial paper and certificates of deposits. Since we are interested in the pricing of near-money assets, we focus on MMF investments into T-bills and repos, both with banks and the Federal Reserve, as they are the safest, most liquid investments with a high degree of substitutability.

Fact 1: MMFs hold a substantial share of T-bills. T-bills are considered one of the safest and most liquid assets in the world. MMFs are significant players in T-bill markets throughout our sample period (since 2011). Their share of holdings of total marketable T-bills increased almost consistently between 2016 and 2022. During the Covid-19 crisis, the share of MMF holdings of total marketable T-bills (excluding those held by the Federal Reserve at the SOMA portfolio) rose to more than 45% as their holdings in the entire Treasury market (again excluding SOMA portfolio holdings) also rose to around 11% (see Figure 1(a)). In 2022, however, their share in the T-bill market declined as they substituted into the RRP facility of the Federal Reserve, one of the facts for which we set out to provide an explanation.

Figure 1: MMFs' role in the T-bill and repo markets (with banks and RRP)



(a) MMFs' comprise a substantial share of the T-bill market



(b) MMF portfolio allocation between T-bills, repo and RRP: RRP is preferred to T-bills when Treasury market liquidity is low

Notes: In Panel 1(a), the solid line is the time series of the total marketable T-bills excluding those held at the SOMA portfolio. The dashed line show the MMF share of holdings of the total represented in the solid line. The dotted line shows the MMF share of total holdings of all US Treasuries excluding those held at the SOMA portfolio. In Panel 1(b), the darkest area shows the total T-bill holdings of MMFs, the medium dark area shows the total investments in the RRP facility and the lightest area shows the total repos with banks. The dashed line shows the Bloomberg Liquidity Index, which is measured by the deviations from a fair value model across the yield curve. Higher values correspond to lower liquidity. Source: Crane Data, US Treasury, Bloomberg

Fact 2: There is significant time variation in the investments of MMFs between T-bills, RRP, and repos. Investments in the three money-like assets, T-bills, RRP, and repos, amount to around $3 trillion, or around two-thirds of the total AUM of MMFs. Other assets in MMF portfolios include commercial paper, certificates of deposits, and agency debt. Those assets are either not as liquid or include credit risks. Due to their close substitutability and being the safest and most liquid assets that impact benchmark rates, we restrict our attention to T-bills and repos with banks and the Federal Reserve in this paper. Repos with banks have been roughly stable since 2012. Before the Covid-19 crisis, at quarter-ends and regulatory reporting periods, some banks withdrew from repo markets and the MMFs would instead invest in the RRP facility (Aldasoro, Ehlers and Eren, 2022). During the first few months following the Covid-19 crisis, MMF holdings T-bills increased to more than $1 trillion (the median remaining maturity of T-bills held by MMFs is around two months). Since the end of 2021, their holdings of T-bills decline as their holdings in the RRP skyrocketed to more than $2 trillion as T-bill issuance has declined (see Figure 1(a) and the liquidity conditions in the Treasury market deteriorated (see the dashed line in Figure 1(b))).[5]

Fact 3: The repo market is concentrated and repo rates on average exceed T-bill and RRP rates. Measured by the Herfindahl-Hirschmann Index (HHI), holdings of MMFs of both T-bills and repos with banks are much more concentrated than if each fund would hold an equal share in these markets.[6] This concentration leads to imperfect competition, whereby MMFs exercise market power on banks, which is well-documented in the literature (e.g. Aldasoro, Ehlers and Eren, 2022; Huber, 2022). The concentration of funds in the repo market increased in response to the 2016 US MMF reform (see Aldasoro, Ehlers and Eren, 2022), which we use as part of our identification strategy. Partly a reflection of their market power in the repo market, MMFs typically obtain a higher interest rate by lending in the repo market compared to investing in T-bills or the RRP facility. Indeed, on average return over the sample period for funds is the highest from repo lending to banks followed by the average return on T-bills and the average return on the RRP facility. We document summary statistics about market concentration and average interest rates in Table 1.[7]

Fact 4: Large funds have a high market share both in the repo market with banks

[5]We use the Bloomberg index for Treasury market liquidity. Higher values of the index correspond to greater *illiquidity* (or lower liquidity). It measures deviations of yields from a fair-value model. It is constructed from the entire yield curve and is a standard gauge of government bond market illiquidity.

[6]With 462 unique funds in our dataset, an equal share of each fund would lead to an HHI of 21, which is more than 10 times lower than the average HHI over the sample period.

[7]Note that the summary statistics reported here are at the level of time (i.e. monthly averages). Therefore the summary statistics reported here differ from the summary statistics reported at the contract level in Table 6.

Table 1: **Summary statistics for various rates and market concentration**

Variable	Obs	Mean	Std. Dev.	Min	Max	P50
repo rate	110	88.94	92.15	7.16	380.3	39.73
Tbill rate (1M)	110	75.24	92.4	1	387	23
RRP rate	110	71.1	87	0	373	25
HHI bank repo	110	281.25	69.46	160.12	384.69	300.48
HHI (pre-reform)	36	192.93	33.15	160.12	332.34	185.33
HHI (post-reform)	74	324.22	30.97	273.1	384.69	320.8

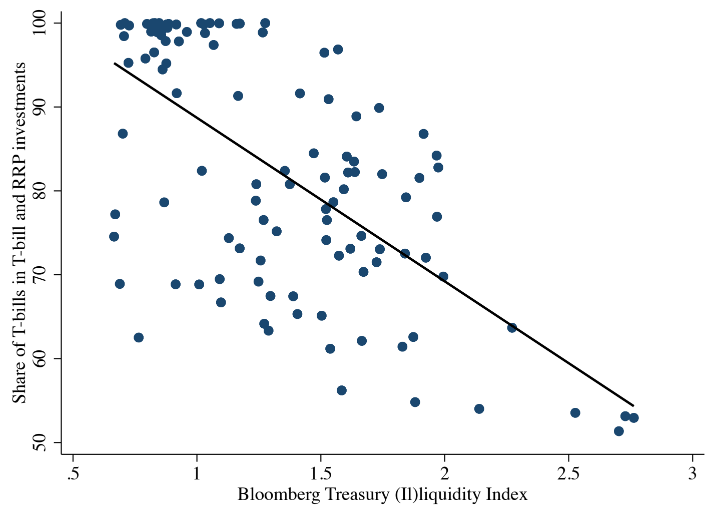
Note: The table reports the summary statistics of all variables across months. To make the summary statistics comparable, in this table, we only report the statistics since the introduction of the RRP facility in October 2013 until November 2022 (end of our sample period). *repo rate*, *Tbill rate* $(1M)$ and *RRP rate* are in basis points. and $GC - Tbill$ $(1M)$ are in basis points. We use the 1-month T-bill rate and the overnight RRP rate to calculate the $Tbill - RRP$ spread. HHI *bank repo* measures the HHI of funds in the repo market and is between 0 and 10,000 (constructed by summing the squared market share of each fund in the repo market). HHI $(pre - reform)$ refers to HHI *bank repo* before the implementation of the US MMF reform in October 2016. HHI $(post - reform)$ refers to HHI *bank repo* after the implementation of the US MMF reform in October 2016 (October 2016 is included in the HHI $(post - reform)$). *Data sources:* Crane Data, FRED

and in the T-bill market. We document this pattern in Figure 2(a). If these markets are not perfectly competitive, being a large player in the repo market potentially increases the market power of these funds. At the same time, being large in the T-bill market might also increase their price impact in this market. This trade-off is a key consideration in the rest of the paper.

Fact 5: Most MMFs invest both in T-bills and the RRP. Moreover, the portfolio allocation between T-bills and RRP is highly correlated with Treasury market liquidity. Despite the high degree of complementarity between T-bills and RRP, MMFs typically hold an interior mix between the two assets. Furthermore, we document in Figure 2(b) that the average share of assets allocated to T-bills versus the RRP correlates negatively with the liquidity of the Treasury market (measured by the Bloomberg liquidity index: higher values indicate lower liquidity).

Figure 2: Correlation of market shares in repo and T-bill markets, and MMF portfolio allocation

(a) Market shares of funds in repos with banks and T-bills are positively correlated: giving rise to the market power-price impact trade-off.



(b) Liquidity conditions correlate with the allocation between T-bills and RRP

Note: In Panel 2(a), we plot the average market share over time of each individual funds in the repo market on the y-axis against the average market share in the T-bill market. In Panel 2(b), we plot the share of investments in T-bills in total T-bill and RRP investments on the y-axis against the Bloomberg Liquidity Index on the x-axis. For the Bloomberg Liquidity Index, higher values correspond to lower liquidity in the Treasury market. Source: Crane Data, Bloomberg.

3 Theory

Motivated by these facts, we model how MMFs set repo rates to banks and allocate their funds between repos to banks, T-bills, and the RRP facility. The model accounts for strategic interactions between MMFs and banks as well as between MMFs. MMFs strategically interact with banks in the repo market and have market power. Whatever they don't lend to banks in equilibrium, they split between T-bills and the RRP. Given they are large players in the T-bill market, they can have a price impact on the T-bill market.

The key insight of the model is that market power in the repo market and liquidity conditions in the T-bill market affect equilibrium outcomes in both markets. This trade-off can be summarised by a simple example. Suppose a fund can only invest in repos with banks or T-bills. Market power in the repo market means funds can increase the rates they charge for repos reducing demand. As a result, funds have more cash left over from repo lending to be invested in the T-bill market. Due to their price impact, the more they invest in T-bills, the more they would put downward pressure on T-bill rates. If they are large in both markets, they would internalize their impact in the T-bill market and charge rates in repos that takes this into consideration.

The model generates multiple predictions on repo rates between MMFs and banks, portfolio allocations by MMFs, and key short-term rates in the macroeconomy, such as the rates on T-bills, which we test in the next two sections.

3.1 Model Setup: The Repo market

There are B banks indexed by $b = 1, \cdots, B$ and F funds indexed by $f = 1, \cdots, F$. Each fund f is characterized by its size, w_f.

In the repo market, fund f offers A rate $r_f(b)$ to the bank b. As is standard in the literature on the industrial organization, we assume that bank b chooses to borrow from fund f with probability[8]

$$\pi_f(r_f(b); b) \;=\; \frac{r_f(b)^{-\alpha_b} w_f}{\sum_\phi r_\phi(b)^{-\alpha_b} w_\phi}$$

Here, α_b is the bank-specific sensitivity of the bank's demand for loans concerning the offered rate $r_f(b)$. We use this sensitivity as a proxy for bank market power in negotiations with funds. The dependence on w is a reduced form model of market power: bigger funds have a higher chance of lending to a given bank.

We start by describing and solving the rate choice problem of a bank. We assume that bank b has access to a decreasing-returns-to-scale technology that returns $\ell^{1-\xi} R_*$ for an investment amount of ℓ. Thus, the objective of a bank b borrowing ℓ at a rate $r_f(b)$ is given by

$$\max_\ell (\ell r_f(b) \;-\; \ell^{1-1/\xi} R_*)\,,$$

[8] It is straightforward to micro-found this demand with random preference shocks.

implying the following demand curve for repos:

$$\ell(r_f(b)) \; = \; r_f^{-\xi} R_*^{\xi} \, .$$

We assume that each f has an outside option to invest any available cash at a rate of ρ. For now, we treat this rate as exogenous and endogenize it below in the next section. The objective of fund f is thus to maximize the excess returns from lending to bank b

$$\Pi \; = \; \pi_f(r_f(b); b)\ell(r_f(b))(r_f(b) - \rho) \; = \; \frac{w_f r_f(b)^{-\alpha_b}}{\sum_\phi r_\phi(b)^{-\alpha_b} w_\phi} R_*^{\xi} (r_f(b)^{1-\xi} - \rho r_f(b)^{-\xi}) \tag{1}$$

over $r_f(b)$. Consider first the simpler case where the funds' market power is negligible (that is, w_f is sufficiently small). In this case, the fund ignores the competition with other funds in the repo market, and the first-order conditions for (1) take the form

$$(1 - \xi - \alpha_b) r_f(b)^{-\xi-\alpha_b} + (\xi + \alpha_b)\rho r_f(b)^{-\xi-\alpha_b-1} = 0 \, ,$$

implying a standard monopolist solution

$$r_*(b) \; = \; \rho \frac{\xi + \alpha_b}{\xi + \alpha_b - 1} \; = \; \rho \; + \; \underbrace{\rho \frac{1}{\xi + \alpha_b - 1}}_{markup} \, . \tag{2}$$

In the presence of market power, the fund's problem is significantly more complex because the fund internalizes its impact on the competition with other banks. Define

$$\Gamma_*(b) \; = \; \sum_\phi r_\phi(b)^{-\alpha_b} w_\phi \, .$$

This is a key quantity in our analysis, capturing the degree of competition in the repo market. The following result follows by direct calculation:

Lemma 3.1 (Optimal rates with imperfect competition) *The optimal rate satisfies*

$$r_f(b) \;=\; r_*(b) \;+\; \underbrace{\frac{1}{\alpha_b+\xi-1}\frac{\alpha_b r_f(b)^{-\alpha_b} w_f(r_f(b)-\rho)}{\Gamma_*(b)}}_{additional\ markup}\,. \tag{3}$$

In general, equation (3) is a complex, non-linear equation defining the equilibrium relationship between rates charged by different banks. However, when the number of funds is large and the size of each fund is small, it is possible to derive an approximate, closed-form solution for the repo market equilibrium. The following is true.

Proposition 3.2 (Equilibrium in the repo market) *Suppose that $w_f = w_f^*/F$, where w_f^* are uniformly bounded. For simplicity, we normalize $\sum_f w_f^* = F$.*[9] *Define*

$$H(W) \;=\; F^{-1}\sum_f (w_f^*)^2$$

to be the Herfindahl index of the fund size distribution. Then,

$$r_f(b) \;=\; r_*(b) \;+\; F^{-1} r_f^{(1)}(b) \;+\; F^{-2} r_f^{(2)}(b) \;+\; O(F^{-3})\,,$$

with

$$r_f^{(1)}(b) \;=\; \frac{w_f^*\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)$$

and

$$r_f^{(2)}(b) = \underbrace{\left(\frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\right)^2 (1-\alpha_b r_*(b)^{-1}(r_*(b)-\rho))(r_*(b)-\rho)}_{own\ market\ power\ convexity\ adjustment} + \underbrace{\frac{w_f^*\alpha_b^2 r_*(b)^{-1}}{(\alpha_b+\xi-1)^2}(r_*(b)-\rho)^2 H(W)}_{market\ concentration}$$

Proposition 3.2 shows how imperfect competition in the repo market makes it optimal for funds to charge additional markups, over and beyond the one in (2). First, there is a convexity adjustment

[9] E.g., the most competitive case corresponds to an equal distribution of size across funds, $w_f^* = 1/F$, with $H(W) = 1/F$, the lowest possible value.

because markups depend non-linearly on the fund's own market power, as captured by w_f. Second, it depends on the repo market concentration, as captured by $H(W)$.

3.2 T-bill Market

Now we introduce a simple model for rate determination in the T-bill market. The supply of treasuries is fixed, given by an exogenous number S. We assume that the T-bill market is populated by two types of agents: Liquidity providers and MMFs.

We model liquidity providers through an exogenous demand curve, $D(\rho) = a+\lambda\rho$. The behavior of MMFs is more subtle. We assume that a fund f has *residual cash,* Δ_f, that must be invested either in T-bills or in the RRP. The RRP rate is fixed at an exogenous ρ_*. Intuitively, we would expect that the demand D_f^T of fund f for Treasuries satisfies $D_f^T = \mathbf{1}_{\rho>\rho_*}\Delta_f$. That is, the fund would invest everything into Treasuries if the rate is above ρ_*, and invest all residual cash into RRP when $\rho_* > \rho$. However, this is not what we observe in the data. As Figure 1(b) shows, MMFs buy non-trivial amounts of treasuries even when ρ is significantly below ρ_*, suggesting that some frictions prevent MMFs from selecting this corner solution. One may then ask: Are MMFs responding *elastically* to changes in the T-bill rate? In the data, it is indeed the case: When ρ is above ρ_*, funds buy more T-bills. We model this price-elastic behavior by assuming demand curves

$$D_f^T(\rho) \;=\; (a_*(f) + \lambda_*(f)(\rho - \rho_*))\Delta_f \tag{4}$$

for some fund-specific coefficients $a_*(f),\ \lambda_*(f) > 0$. Funds with a higher $\lambda_*(f)$ are more elastic concerning T-bill rate changes and are, therefore, more aggressive in absorbing supply shocks and providing liquidity. Under the above assumptions, the T-bill rate is pinned down by the market-clearing condition

$$a + \lambda\rho \;+ \sum_f (a_*(f) + \lambda_*(f)(\rho - \rho_*))\Delta_f \;=\; S\,. \tag{5}$$

Solving (5), we arrive at the equation

$$\hat{\rho} \;=\; \rho_* \;+\; \underbrace{\frac{S - a - \sum_f a_*(f)\Delta_f}{\lambda + \sum_f \lambda_*(f)\Delta_f}}_{demand-supply\ imbalance} \tag{6}$$

showing explicitly how the demand pressure created by MMFs affects the equilibrium T-bill rate. The next big question is: Do MMFs internalize their price impact in the T-bills market when they set rates in the repo market? Here, one could consider two layers of internalization: The effect of the T-bill price impact of a given fund f on (1) the fund's rate setting in the repo market; (2) the rates set by other funds. We ignore the latter channel because it has lower-order effects and focuses on the channel (1).

3.3 Strategic Behavior Across the two Markets

We assume that each MMF starts with a deposit base, d_f, that is split between lending to banks and investments into T-bills and the RRP. Thus, the residual cash is given by

$$\Delta_f \;=\; d_f - \underbrace{\sum_b (R_*/r_f(b))^{\xi} \frac{r_f(b)^{-\alpha_b} w_f}{\Gamma_*(b)}}_{repo\ lending} \,. \tag{7}$$

Equation (7) defines the key mechanism of our model: In a fund chooses higher rates $r_f(b)$, repo lending drops, leaving the fund with more residual cash. The fund is then forced to invest this cash into T-bills; if the T-bill market is illiquid, buying a lot of T-bills creates a price impact, making this investment less attractive. When optimizing profits, funds should take this price impact into account. By (4), the total payoff that the fund receives from its T-bill/RRP investments is given by

$$\begin{aligned} D_f^T(\rho)\rho \;+\; (\Delta_f - D_f^T(\rho))\rho_* \;&=\; D_f^T(\rho)(\rho - \rho_*) \;+\; \Delta_f \rho_* \\ &=\; \Big((a_*(f) + \lambda_*(f)(\rho - \rho_*))(\rho - \rho_*) \;+\; \rho_*\Big)\Delta_f \\ &=\; \widetilde{\rho}\,\Delta_f \,, \end{aligned} \tag{8}$$

where we have defined

$$\widetilde{\rho} \;\equiv\; \Big((a_*(f) + \lambda_*(f)(\rho - \rho_*))(\rho - \rho_*) \;+\; \rho_*\Big)$$

to be the *effective rate* that the fund f earns on its investments across T-bills and RRP.

In equilibrium, the T-bill rate ρ in (8) is the market clearing rate $\hat{\rho}$ satisfying (6). As discussed above, we assume that the fund takes the repo rates charged by competitors as given and optimizes

$$\sum_b \frac{r_f(b)^{1-\alpha_b-\xi} - \widetilde{\rho} r_f(b)^{-\alpha_b-\xi}}{r_f(b)^{-\alpha_b} w_f \;+\; F_{-f}(b)} \;+\; \widetilde{\rho} d_f\,,$$

where $\hat{\rho}$ depends on $(r_f(b))_{b=1}^B$ directly through (6). Define

$$U_{-f} \;=\; S - a - \sum_{\phi \neq f} a_*(\phi)\Delta_\phi(r_\phi)$$

$$V_{-f} \;=\; \lambda + \sum_{\phi \neq f} \lambda_*(\phi)\Delta_\phi(r_\phi)$$

to be the two components of the *residual demand* of all other MMFs, defining the level and slope of their demand, as driven by their demand functions (4). The following is true.

Proposition 3.3 (Pass-through of repo rates into treasuries) *Suppose that F is large and $d_f = O(w_f)$, and that fund f takes $r_\phi, \phi \neq f$, as given. Let also*

$$\Xi_{-f} \;=\; \left(\frac{a_*(f)}{V_{-f} + \lambda_*(f) d_f} \;+\; \frac{U_{-f} - a_*(f) d_f}{(V_{-f} + \lambda_*(f) d_f)^2}\right)\,.$$

Then,

$$\frac{\partial \hat{\rho}}{\partial r_f(u)} \;=\; \hat{\rho}^{(1)}_{r_f(u)} F^{-1} \;+\; \hat{\rho}^{(2)}_{r_f(u)} F^{-2} \;+\; O(F^{-3}),$$

where

$$\hat{\rho}^{(1)}_{r_f(u)} \;=\; -\Xi_{-f} w_f^* R_*^\xi (\xi + \alpha_u) r_*(u)^{-\xi-1}$$

$$\hat{\rho}^{(2)}_{r_f(u)} \;= \Xi_{-f} w_f^* R_*^\xi r_*(u)^{-\xi-1}\Bigg((\xi + \alpha_u)\Big((\xi + \alpha_u + 1) r_*(u)^{-1} r_f^{(1)} + r_*(u)^{\alpha_u}\Gamma_*(u)^{(1)}\Big) \;+\; \alpha_u w_f^*\Bigg) \;-\; Q_f^*(u)$$

where we have defined

$$Q_f^*(u) \;=\; 2(w_f^*)^2 F^{-2}\frac{U_{-f} - a_*(f)d_f}{(V_{-f}+\lambda_*(f)d_f)^3}\Big(\sum_b (R_*/r_*(b))^\xi\Big)\left(R_*^\xi(\xi+\alpha_u)r_*(u)^{-\xi-1}\right)$$

and where

$$\Gamma_*(b)^{(1)} \;=\; -\alpha_b r_*(b)^{-1-\alpha_b} E[r_f^{(1)}(b)]\,, \tag{9}$$

and where $r_f^{(1)}(b)$ *is to be determined later in general equilibrium.*

Proposition 3.3 shows how a strategic change in the repo rate, through its impact on the residual cash and the demand function of the fund for T-bills, affects the equilibrium T-bill rate $\hat{\rho}$.

Having identified this pass-through coefficient, we can now write down the first-order condition of each fund concerning the rate $r_f(b)$ it charges to a particular bank. Solving these first-order conditions leads to the equilibrium link between the repo rates set by a fund and the T-bill rate, $\hat{\rho}$. Recall that

$$D_f^T(\rho) \;=\; (a_*(f)+\lambda_*(f)(\rho-\rho_*))\Delta_f$$

is the demand for T-bills by the fund f, and let

$$\widetilde{\rho}_f \;=\; \underbrace{\rho_*}_{RRP\ rate} + \frac{D_f^T(\rho)}{\Delta_f}\underbrace{(\hat{\rho}(r_f)-\rho_*)}_{excess\ return\ on\ the\ T-bills}$$

be the effective rate earned by the fund f on its residual cash Δ_f. Let also

$$\Delta_f^* \;=\; \frac{\Delta_f}{w_f}$$

be the ratio of the residual cash to fund size in the repo market. Let also

$$\Lambda_f \;=\; \frac{(\xi+\alpha_u)R_*^\xi}{\xi+\alpha_u-1}\Xi_{-f}\left(2\frac{D_f^T(\hat{\rho})}{\Delta_f} - a_*(f)\right)$$

be a measure of funds' own price impact in the T-bill market. We will also use $E[x_f]$ to denote cross-sectional averages, weighted with w_f^* :

$$E[x_f] \; = \; F^{-1} \sum_f w_f^* x_f \, .$$

Proposition 3.4 (Equilibrium Repo Markups) *The optimal repo rate set by fund f for bank b satisfies*

$$r_f(u) \;=\; \frac{\alpha_u + \xi}{\alpha_u + \xi - 1} \widetilde{\rho}_f \;+\; F^{-1} r_f(u)^{(1)} \;+\; F^{-2} r_f(u)^{(2)} \;+\; O(F^{-2})$$

with

$$r_f(u)^{(1)} \;=\; \underbrace{\frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} \left(1 + \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} F^{-1}\right) (r_*(u) - \widetilde{\rho}_f)}_{repo\ market\ power} \;-\; \underbrace{w_f^* \Lambda_f \Delta_f^*}_{T-bill\ price\ impact}$$

and

$$\begin{aligned} r_f(u)^{(2)} \;=&\; \underbrace{(w_f^*)^2 C_f(u)}_{convexity\ price\ impact\ adjustment} \\ &+\; \frac{\alpha_u^2 w_f^* r_*(u)^{-1}}{\alpha_u + \xi - 1} (r_*(u) - \widetilde{\rho}_f) \underbrace{(E[r_f(u)^{(1)}] - r_f(u)^{(1)})}_{heterogeneity} \end{aligned}$$

where

$$C_f(u) \;=\; -2((\alpha_u + \xi - 1))^{-1} r_*(u)^{\xi+1} \Delta_f^* \frac{U_{-f} - a_*(f) d_f}{(V_{-f} + \lambda_*(f) d_f)^3} \Big(\sum_b (R_*/r_*(b))^{\xi} \Big) \left(R_*^{\xi} (\xi + \alpha_u) r_*(u)^{-\xi-1} \right)$$

$$+\; (\alpha_u + \xi - 1))^{-1} \Delta_f^* \Xi_{-f} R_*^{\xi} \alpha_u$$

is a convexity adjustment for the price-impact effects.

3.4 Equilibrium T-bill Rate

In this section, for simplicity, we assume that all banks have the same elasticity coefficient: $\alpha_b = \alpha$ is independent of b.[10] In this case, absent market power, fund f charges the rate

$$r_*(f) \;=\; \frac{\alpha + \xi}{\alpha + \xi - 1} \widetilde{\rho}_f^*$$

[10] All our results hold when banks are heterogeneous, but the equilibrium expressions become more complex.

to all banks in the repo market, where

$$\widetilde{\rho}_f^* \;\equiv\; \Big((a_*(f)+\lambda_*(f)(\rho^*-\rho_*))(\rho^*-\rho_*) \;+\; \rho_*\Big)$$

is the effective rate earned by fund f on its residual cash, and where ρ^* is the frictionless equilibrium rate, given by

$$\hat{\rho}^* \;=\; \rho_* \;+\; (\lambda+\bar{\lambda})^{-1}\Big(\underbrace{S}_{supply} - \underbrace{\Big(a+\sum_f a_*(f)\Delta_f(0)\Big)}_{demand}\Big)\,, \tag{10}$$

where the frictionless level of residual cash of fund f is given by

$$\Delta_f(0) \;=\; \underbrace{d_f}_{deposits} - \underbrace{B(R_*/r_*(f))^{\xi} w_f}_{repo\ lending}\,,$$

and where we have defined

$$\bar{\lambda} \;=\; \sum_f \lambda_*(f)\Delta_f(0)\,.$$

In the presence of market power and imperfect competition in both repo and T-bill markets, equilibrium T-bill rate, $\hat{\rho}$, deviates from its frictionless level (10). To characterize this deviation, we introduce an important quantity,

$$\psi_f \;=\; (\alpha+\xi)r_*(f)^{-\xi-1}R_*^{\xi}(\lambda+\bar{\lambda})^{-1}\Big(\underbrace{a_f}_{inelastic} + (S-a-A)(\lambda+\bar{\lambda})^{-1}\underbrace{\lambda_f}_{elastic}\Big)\,,$$

capturing the total pass-through of residual cash flow shocks of fund f into the fund's demand for T-bills. We then define

$$\mathcal{E} \;=\; 1 \;+\; 2(\hat{\rho}^*-\rho_*)\frac{\alpha+\xi}{\alpha+\xi-1}\Big(\frac{\xi}{\alpha+\xi}E[\psi_f]E[\lambda_*(f)] \;+\; \mathrm{Cov}(\psi_f,\lambda_*(f))\Big)$$

This quantity is the equilibrium elasticity of the T-bill rate to shocks originating from imperfect competition. The following is true.

Proposition 3.5 (Equilibrium T-bill rate) *In equilibrium,*

$$\hat{\rho} \;=\; \hat{\rho}^* \;+\; \hat{\rho}^{(1)}F^{-1} \;+\; O(F^{-2})\,,$$

where

$$\hat{\rho}^{(1)}$$

$$= \; -\mathcal{E}^{-1}\Bigg(\frac{\xi}{(\alpha+\xi)}\underbrace{E[\psi_f]}_{average\ passthrough}\Big(\frac{\alpha}{\alpha+\xi-1}(r_*(f)-\widetilde{\rho}^*_f)\underbrace{H(W)}_{repo\ concentration} \;-\; \underbrace{E[w^*_f\Lambda_f\Delta^*_f]}_{price\ impact\ internalization}\Big)$$

$$+\;\underbrace{\mathrm{Cov}\Big(\psi_f,\,\frac{\alpha w^*_f}{\alpha+\xi-1}(r_*(f)-\widetilde{\rho}^*_f)\;-\;w^*_f\Lambda_f\Delta^*_f\Big)}_{strategic\ interactions\ in\ the\ T-bill\ market}\Bigg)$$

4 Equilibrium RRP choice

Our derivations in the previous sections are based on the assumption of upward-sloping demand curves (4) for T-bills. In particular, we take the coefficients $a_f, \lambda(f)$ of these demand functions as given and cannot explain the funds' demand for RRP investments. In this section, we microfound this demand.

We assume that investing in RRP is associated with an implicit, non-monetary cost. For example, this cost might be driven by reputational concerns, whereby the fund fears its RRP investments might be interpreted as its inability to perform successful active investment choices. We assume that this cost is given by

$$\xi_f(\theta_f\Delta_f) \;+\; 0.5\beta_f(\theta_f\Delta_f)^2\,,$$

where θ_f is the fraction of residual cash invested into RRP. As a result, the objective of the fund is to maximize

$$\sum_b \frac{r_f(b)^{1-\alpha_b-\xi}-\widetilde{\rho}(\theta)r_f(b)^{-\alpha_b-\xi}}{r_f(b)^{-\alpha_b}w_f \;+\; F_{-f}(b)} \;+\; \widetilde{\rho}(\theta)d_f \;-\; (\xi_f(\theta_f\Delta_f) \;+\; 0.5\beta_f(\theta_f\Delta_f)^2)\,,$$

where

$$\widetilde{\rho}(\theta) \;=\; (\rho_* + (1-\theta)(\rho - \rho_*))\,.$$

Importantly, as above, funds are strategic in their trading decisions in the T-bill market and internalize their price impact: In equilibrium, investing $(1-\theta)\Delta_f$ of cash into T-bills moves the rate ρ by $\gamma_f(1-\theta)$, so that

$$\widetilde{\rho}(\theta) \;=\; (\rho_* + (1-\theta)(\rho - (1-\theta)\gamma_f - \rho_*))$$

As a result, the part of the objective that depends on θ can be rewritten as

$$(\rho_* + (1-\theta)(\rho - (1-\theta)\gamma_f - \rho_*))\Delta_f - (\xi_f(\theta_f\Delta_f) \;+\; 0.5\beta_f(\theta_f\Delta_f)^2)\,.$$

Optimizing over θ implies a demand function of (see, e.g., Malamud and Rostek (2017))

$$1-\theta_f(\hat{\rho}) \;=\; \frac{\hat{\rho} - \rho_* + \xi_f}{\gamma_f + \beta_f\Delta_f}\,,$$

and hence (4) takes the form

$$D_f^T(\rho) \;=\; \frac{\hat{\rho} - \rho_* + \xi_f}{\gamma_f + \beta_f\Delta_f}\Delta_f\,, \tag{11}$$

so that we recover the upward-sloping demand curves (4), but the coefficients $a_*(f),\ \lambda_*(f)$ are endogenous, determined in equilibrium through the strategic interaction of funds in the T-bill market. We will need the following characterization of this strategic interaction and equilibrium price impacts γ_f from Malamud and Rostek (2017).

Proposition 4.1 *We have*

$$\gamma_f \;=\; \frac{2\beta_f\Delta_f}{\beta_f\Delta_f(\lambda+b) - 2 + \sqrt{(\beta_f\Delta_f(\lambda+b))^2+4}}\,,$$

where $b > 0$ is the unique solution to

$$\sum_f \left(2 + \beta_f(\lambda + b) + \sqrt{(\beta_f \Delta_f(\lambda + b))^2 + 4}\right)^{-1} \; = \; 0.5\frac{b}{\lambda + b}\,.$$

When $F \to \infty$ and $\beta_f = O(1)$, this gives $b = b_0 + b_1 F + O(F^{-1})$ with

$$b_1 \; = \; \sum_f (\beta_f \Delta_f)^{-1}/F$$

and

$$b_0 \; = \; -\sum_f (\beta_f \Delta_f)^{-2}/(F b_1)\,.$$

The following result is a key testable implication of our theory.

Corollary 4.2 *T-bill liquidity negatively depends on Δ_f. A drop in λ leads to a drop in T-bill liquidity. As a result, γ_f go up and, through (11), lead to an increase in the share of residual cash invested in the RRP. Furthermore, funds' T-bill investments become less elastic with respect to changes in the T-bill rate. Finally, funds with larger Δ_f invest more into RRP, and more so when markets are illiquid.*

5 The aggregate impact of MMFs on the T-bill market

In this section, we focus on the aggregate impact of MMFs in T-bill markets. An important implication of our model is that MMFs have a price impact in the T-bill market, and on aggregate, they affect liquidity conditions. Here, we present empirical evidence consistent with these implications. In particular, we study the impact of a higher residual cash share[11] on two interest rate spreads – the 1-month T-bill minus RRP spread and the 1-month GC repo minus 1-month T-bill spread – as well as market liquidity. When MMFs have more residual cash left over from repo lending, the interest spread of T-bills over the RRP declines as they demand more T-bills. As predicted by the

[11] We measure the residual cash share as $\left(1 - \frac{\sum_f repo_{f,t}}{\sum_f repo_{f,t} + Tbill_{f,t} + RRP_{f,t}}\right) \times 100$

model, a higher residual cash share also deteriorates liquidity conditions in the T-bill markets. As a result, it drives up the liquidity premium on T-bills.

In our model, the variable of interest is the residual cash left over from repo lending to banks and not the actual cash amount invested into T-bills. Indeed, in equilibrium, the MMFs face the problem of optimally splitting the residual cash between the T-bills and the RRP, given the implicit costs of using the facility. The actual cash amount invested into T-bills is a complex, endogenous quantity depending on the elasticity of MMF's demand, market liquidity, and other strategic considerations. As a result, our model predicts no directly testable link between the share of total funds invested into T-bills and the outcome variables (see, Proposition 3.5). But instead, there is a direct link between the residual cash left over from repo lending and outcome variables.

The effects we estimate are causal, (to a large extent) statistically significant, and economically large. We establish causality through the use of instrumental variables guided by our theory. In OLS regressions, adding the average MMF share of residual cash left from repo lending significantly increases the explanatory power of the time series regressions to explain the 1-month T-bill-RRP spread and the 1-month General Collateral (GC) repo and 1-month T-bill spread (which is commonly used to gauge the liquidity premium of T-bills).[12] Quantitatively, using two-stage least squares regressions, we find that the partial impact of a standard deviation increase in the average MMF share of residual cash from repo lending on the spread between the 1-month T-bill rate and the RRP rate spread is equivalent to the partial effect of (i) around one percentage point increase in the federal funds rate, (ii) around a third of a percent increase in the T-bill supply normalized by GDP. Similarly, the partial impact of a standard deviation increase in the average MMF share of residual cash from repo lending on the spread between the 1-month GC repo rate and 1-month T-bill rate is equivalent to the partial effect of (i) around one percentage point rise in the federal funds rate, (ii) around one-fifth of a percent rise in the T-bill supply normalized by GDP. Finally, one standard deviation increase in the residual cash share increases our illiquidity measures by around half a standard deviation.

[12]In the literature, this differential is often used at three months. To further isolate any potential counterparty risks arising from long maturities of the repo transaction, we prefer using the 1-month spread.

Data description and summary statistics. To conduct our analysis, we collect monthly data averages on the 1-month T-bill rate, the effective federal funds rate, the rate on the RRP facility, the 1-month GC repo rate, and the VIX. We also collect data on T-bills outstanding that are held publicly and GDP (we interpolate monthly data from available quarterly data) to construct monthly data for bills to GDP and subtract the holdings of the Federal Reserve through its SOMA portfolio.

We also obtain the monthly average of the Bloomberg Liquidity Index from Bloomberg and obtain the weekly trading volume of T-bills from the Federal Reserve Bank of New York to construct an Amihud liquidity measure using the 1-month T-bill rate (see Amihud, 2002) and standardize it. For both measures, higher values correspond to lower liquidity.

We construct our key independent variable of interest, $residual\ cash\ share_t$, using our monthly MMF holdings data. We build this measure as the share of aggregate MMF investments on T-bills and the RRP divided by the aggregate investments on bank repos, T-bills, and the RRP.

We construct three instrumental variables. $HHI\ bank\ repo$ measures the HHI of funds in the repo market and lies between 0 and 10,000 (constructed by summing the squared market share of each fund in the repo market).[13] $HHI\ ex\ B*T\ FE$ is constructed by taking out the bank-time fixed effects (mean bank demand in a given month) from each original contract value and then constructing the HHI from the residual values stripped of average bank demand in a given month. It also lies between 0 and 10,000. Finally, we construct $post\ reform$, a dummy variable that takes on a value of 1 in the months after implementing the reform in October 2016.[14]

We present the summary statistics of variables used in this section in Table 2.

Residual cash from repo lending, market concentration, and the T-bill-RRP spread. To analyze the aggregate price impact of MMFs in the T-bill market, we estimate the following

[13] We define the Herfindahl-Hirschman index ('HHI') across fund market shares in the overall repo market as $HHI\ bank\ repo_t = \sum_{f=1}^{F} \left(\frac{bank\ repo_{f,t}}{bank\ repo_t} \times 100 \right)^2$.

[14] Some funds started adjusting to the reform a few months earlier than the introduction of the reform (see Aldasoro, Ehlers and Eren, 2022). Our results are robust to alternative ways this dummy variable is constructed around the implementation date.

Table 2: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max	P50
Tbill (1M)-RRP	140	4.07	7.89	-15	33	2
GC-Tbill (1M)	140	12.86	9.32	-2	43.5	10.94
residual cash share	140	38.3	20.84	7.31	86.23	34.94
HHI bank repo	140	259.41	73.38	160.12	384.69	277.67
HHI ex B*T FE	143	265.07	82.02	140.82	407.95	292.08
FFR	140	.61	.77	.05	2.56	.14
log(bills to GDP)	140	-2.23	.28	-2.67	-1.42	-2.31
VIX	140	18.23	6.8	10.13	57.74	16.29
BBG Liq.	140	1.24	.43	.6	2.73	1.17
Amihud Liq. (1MO)	140	0	1	-.77	8.21	-.27

Note: This table reports the summary statistics for the main variables used in regressions in this section. $Tbill\ (1M) - RRP$ and $GC - Tbill\ (1M)$ are the dependent variables in basis points. We use the 1-month T-bill and overnight RRP rates to calculate the $Tbill - RRP$ spread. Prior to the introduction of the RRP facility, we used the 1-month T-bill rate. $GC - Tbill$ spread is calculated using the 1-month GC repo rate minus the 1-month T-bill rate. *residual cash share* and ffr are measured in percentage points. HHI *bank repo* measures the HHI of funds in the repo market and is between 0 and 10,000 (constructed by summing the squared market share of each fund in the repo market). $HHI\ ex\ B*T\ FE$ is constructed by taking out the bank-time fixed effects (mean bank demand in a given month) from each original contract value and then constructing the HHI from the values stripped out of average bank demand on a given month. It is also between 0 and 10,000. $log(bills\ to\ GDP)$ is the log of total marketable bills held by the public minus bills held in the SOMA portfolio of the Federal Reserve. To construct monthly GDP data, we interpolate quarterly data into monthly data. VIX is the monthly average level of the index. P50 refers to the median. The sample is monthly time series between February 2011 and September 2022. *Data sources:* Crane Data, FRED, Bloomberg, US Treasury.

regression at the monthly level:

$$Tbill\ (1M) - RRP_t = \beta\ residual\ cash\ share_t + controls_t + \epsilon_t. \tag{12}$$

The dependent variable is the spread between the 1-month T-bill rate and the rate on the RRP facility in month t.[15] The variable $residual\ cash\ share_t$ captures the share of funds' AUM allocated to T-bills and the RRP facility. As controls, we include the Fed funds rate, the log of T-bill supply

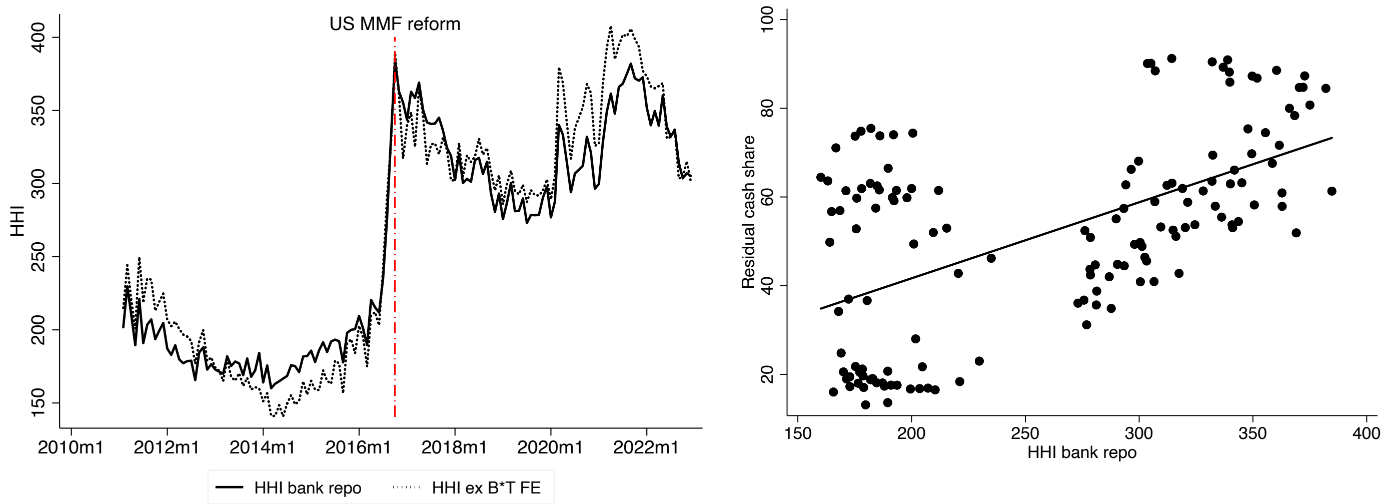
[15]The RRP rate was only operationalized towards the end of 2013 and paid very close to zero when the economy was at the zero lower bound. We used the T-bill rate prior to the introduction of RRP.

to GDP minus the holdings of the Federal Reserve in its SOMA portfolio, as well as the VIX, similar to Nagel (2016) and d'Avernas and Vandeweyer (2021). In all of our time series regressions, we report standard errors that are robust to arbitrary heteroskedasticity and autocorrelation. A key prediction of our model is that a higher share of residual cash from repo lending should have a negative effect on the spread, so $\beta < 0$. This follows directly from formula (10).

We proceed as follows. We first provide the results of OLS regressions. As the quantity $residual\ cash\ share_t$ is an equilibrium outcome and endogenously determined, next, we leverage our model to devise instrumental variables and run 2SLS regressions to estimate the causal effect of $residual\ cash\ share_t$ on $Tbill\ (1M) - RRP_t$. To this end, we instrument $residual\ cash\ share_t$ with the market concentration of MMFs in the repo market, i.e., $HHI\ bank\ repo_t$. According to our model, greater concentration in the repo market means that funds charge, on average higher rates, thereby reducing banks' demand. In turn, funds have more residual cash that they must, at least in part, invest into T-bills – driving down the T-bill rate. As long as the market concentration in the repo market is not determined by the (unobservable) factors that also affect the T-bill rate, the exclusion condition is satisfied, and the HHI is a valid instrument. We show the time series of the instruments we use and the correlation with $residual\ cash\ share$ in Figure 3. To alleviate any remaining concerns about the exclusion restriction, we use alternative instrumental variables and show that our results are robust. The first one is constructed by stripping out the average bank-time level demand from contract values and reconstructing the HHI to strip it out of bank demand which might also be correlated with the T-bill rate. As the dashed line in Panel 3(a) shows, this leads to an alternative HHI measure highly correlated with the original one. Second, we use the $post\ reform$ dummy as an alternative instrument since the reform exogenously caused higher market concentration in the repo market. We report the results in Table 3. In all columns, we report standard errors that are robust to arbitrary heteroskedasticity and autocorrelation with bandwidths selected according to the automatical lag selection procedure in Newey and West (1994).

Our first result is that adding $residual\ cash\ share$ to the time series regression in addition to the variables that are commonly used in the literature, such as the federal funds rate, log(bills to GDP), and VIX, substantially improves the R^2 from 33% to 49% going from column (1) to column

Figure 3: HHI for funds in the repo market and residual cash share



(a) HHI of the repo market **(b)** HHI and residual cash share correlation

Notes: *HHI bank repo* measures the HHI of funds in the repo market and is between 0 and 10,000 (constructed by summing the squared market share of each fund in the repo market). *HHI ex* $B * T$ *FE* is constructed by taking out the bank-time fixed effects (mean bank demand in a given month) from each original contract value and then constructing the HHI from the values stripped out of average bank demand on a given month. it is also between 0 and 10,000. *residual cash share*$_t$ is constructed using the monthly MMF holdings data. For each fund on a given date, we subtract from one the share of repo lending to banks, which is one minus the total amount invested in repos with banks divided by the total amount invested in repos with banks, T-bills, and the RRP facility. We then average this across MMFs each month. Source: Crane Data.

(2). While this result is consistent with the theoretical expression for the equilibrium T-bill rate in Proposition 3.5, the OLS regression suffers from endogeneity which we address next.

We use instrumental variables to obtain a causal effect of *residual cash share* on *Tbill* $(1M) - RRP$. In column (3), we present the results of the first stage regression for *residual cash share* using *HHI bank repo* as the excluded instrument. As it was already alluded to in Panel 3(b), *HHI bank repo* has a highly statistically significant correlation with *residual cash share* suggesting it satisfies the relevance condition. In column (4), we estimate the second stage in which we show that the coefficient is more negative compared to its OLS counterpart. It remains statistically significant at the 1% level. Quantitatively, the partial impact of a one standard deviation increase in *residual cash share* on *Tbill* $(1M) - RRP_t$ is slightly more than five basis points (more than half a standard deviation). This effect is equivalent to the partial impact of around one percentage point increase in the federal funds rate and around one-third of a percent increase in the partial impact of bills-to-GDP.

Table 3: MMFs' residual cash share and the T-bill-RRP spread

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
			1st St.	2nd St.		1st St.	
	OLS	OLS	OLS	2SLS	2SLS	OLS	2SLS
VARIABLES	Tbill (1M)-RRP	Tbill (1M)-RRP	residual cash share	Tbill (1M)-RRP	Tbill (1M)-RRP	residual cash share	Tbill (1M)-RRP
residual cash share		-0.18***		-0.27***	-0.31***		-0.26***
		(0.03)		(0.05)	(0.07)		(0.05)
HHI bank repo			0.19***				
			(0.05)				
post reform						34.40***	
						(11.87)	
FFR	4.76***	4.94***	-5.83*	5.04***	5.08***	-11.93**	5.03***
	(1.78)	(1.21)	(3.07)	(1.23)	(1.27)	(5.02)	(1.21)
log(bills to GDP)	9.76***	16.20***	6.94	19.54***	21.06***	-3.33	19.25***
	(3.05)	(1.95)	(9.95)	(2.49)	(2.88)	(13.19)	(2.82)
VIX	-0.12	-0.11	0.21	-0.10	-0.10	0.26	-0.10
	(0.14)	(0.08)	(0.47)	(0.09)	(0.10)	(0.47)	(0.09)
Observations	140	140	140	140	140	140	140
R-squared	0.33	0.49	0.48			0.44	
Exc. Instrument				HHI	HHI ex B*T FE		post reform
Underidentification test (p-val)				0.04	0.07		0.04
Anderson-Rubin test (p-val)				0.00	0.00		0.00
Kleibergen-Paap rk Wald F stat				11.06	8.48		7.61

Note: This table reports results of various regressions with alternative specifications of the equation (12). Variable descriptions and summary statistics can be found in Table 2. Data are at a monthly frequency between February 2011 and September 2022. Columns (1) and (2) report the results of OLS regressions. Columns (3) and (4) report the first and second stages of a 2SLS regression, in which *HHI bank repo* instruments *residual cash share*. In column (5), we report the second stage in which we instrument *residual cash share* by $HHI\ ex\ B * T\ FE$. Columns (6) and (7) are the first and second stages of a 2SLS regression, in which *residual cash share* is instrumented by *post reform*. Wherever applicable, we report the p-values of the under-identification test, the Anderson-Rubin test, and the Kleibergen-Paap rk Wald F statistic of the first stage. Standard errors are robust to arbitrary heteroskedasticity and autocorrelation with the lag structure automatically selected using the Newey and West (1994) procedure. *** p<0.01, ** p<0.05, * p<0.1. Source: Crane Data, FRED, US Treasury.

In principle, changes in the HHI in the repo market could reflect movements in other (unobservable) factors that also affect the T-bill rate. In particular, the HHI could reflect banks' demand for repos with specific funds, which could be driven by factors that also affect the T-bill rate.[16] To address this concern, in column (5), we repeat the same second stage but use the $HHI\ ex\ B{*}T\ FE$ instead of the *HHI bank repo* as an instrument. The estimated coefficient is slightly more negative (from -0.27 to -0.31) and is statistically significant at the 1% level.

[16] Note that the exclusion restriction would only be violated if the unobservable factors systematically affect banks' demand for repos with large funds differently from smaller funds.

Next, we exploit the 2016 MMF reform, which was implemented in response to the repeated episodes of stress in this market during the GFC and the Eurozone crisis, and required prime institutional funds and municipal funds to switch to a floating net asset value (NAV) calculation. It also introduced the possibility of imposing redemption gates and fees at the discretion of the fund. Government and treasury funds, on the other hand, were allowed to operate with stable NAVs and without any redemption gates or fees. An unintended consequence of the reform was drastically increased concentration in the repo market (see Figure 3(a)). This increase in concentration was due to the reform and was exogenous to developments in the Treasury market and banks' demand. We use the reform to address the potential threat to identifying common factors driving the T-bill rate and the HHI (see also Aldasoro, Ehlers and Eren (2022) for details on the impact of the reform on the market structure). In columns (6) and (7), we show the first stage and the second stage, respectively, using the $post\ reform$ dummy as an alternative instrument. Finally, overall our instruments have desirable properties as they strongly correlate with the endogenous regressor, we reject the null hypothesis of under-identification, the Anderson-Rubin test yields very low p-values, and the first stage F-statistics are not very low (given the relatively small sample size).

MMFs' aggregate impact on Treasury market liquidity. In our model, on aggregate, MMFs' residual cash from repo lending also impacts liquidity conditions in the T-bill market. The model predicts a negative relationship between residual cash share in the aggregate and liquidity in the T-bill market. We test this implication using liquidity measures as an outcome variable in alternative specifications of equation (12). In particular, we use the *Bloomberg Liquidity Index* and the standardized *Amihud liquidity* measure based on weekly data. Since higher values of both measures correspond to lower liquidity, we expect a positive coefficient on *residual cash share*.

We report the results in Table 4. Overall, our results are similar for both measures. Therefore, for brevity, we report the results with Bloomberg Liquidity Index in columns (1) through (5) and show the 2SLS regression results for the benchmark case with the Amihud liquidity measure in column (6). Throughout the table, we report standard errors robust to arbitrary heteroskedasticity and autocorrelation with the lag structure automatically selected using the Newey and West (1994) procedure. For 2SLS regressions, we also report the p-values for the under-identification test, Anderson-Rubin test, and the first stage Kleibergen-Paap rk Wald F statistic.

Similar to the previous table, we first report the results of an OLS regression without including *residual cash share*. Adding *residual cash share* in column (2) improves the R^2 from 0.26 to 0.56. In column (3), we report the results from the 2SLS regressions in which we instrument *residual cash share* with *HHI bank repo*. The coefficient is significant at the 5% level. The partial causal impact of a one standard deviation increase in the *residual cash share* is around half a standard deviation increase in the Bloomberg Liquidity Index, which corresponds to lower liquidity as predicted by our model. In columns (4) and (5), we use $HHI\ ex\ B*T\ FE$ and $post\ reform$ as instruments, respectively. The results are similar, however, the coefficients are not statistically significant. In column (6), we repeat column (3) with the standardized measure of $Amihud\ Liq\ (1MO)$, which is the standardized Amihud liquidity measure for 1-month T-bills. The coefficient is statistically significant at the 10% level. The partial impact of *residual cash share* on $Amihud\ Liq\ (1MO)$ is similar to its effect on the Bloomberg Liquidity Index.

Residual cash and the liquidity premium. The liquidity premium that is the premium paid by investors for the liquidity services provided by US Treasuries, and in particular, T-bills, has fostered an important body of research both in macroeconomics (e.g. Krishnamurthy and Vissing-Jorgensen, 2012; Nagel, 2016) and in international finance (e.g. Jiang, Krishnamurthy and Lustig, 2021; Engel and Wu, 2021). Our results suggest that, beyond macroeconomic variables and liquidity preferences of agents, MMFs and the frictions they are facing are important determinants of the liquidity premium as well. This has consequences for the US and global markets through the important role of US Treasuries in global finance.

In this section, we investigate how MMFs affect the liquidity premium of T-bills, in particular through their price impact on the T-bill market. Our model predicts that the more funds allocated to T-bills, the more the T-bill rate would decline (see also results in Table 3). In turn, we expect that the liquidity premium, measured by the GC repo-T-bill spread, would increase.

We report the results in Table 4. As a baseline, in column (1), we run a similar regression to the ones in Nagel (2016) and d'Avernas and Vandeweyer (2021) for a monthly sample between

Table 4: MMFs' residual cash share and Treasury market liquidity

VARIABLES	(1) OLS BBG Liq.	(2) OLS BBG Liq.	(3) 2SLS BBG Liq.	(4) 2SLS BBG Liq.	(5) 2SLS BBG Liq.	(6) 2SLS Amihud Liq. (1MO)
residual cash share		0.01***	0.01**	0.01	0.01	0.03*
		(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
FFR	0.01	0.00	-0.00	0.00	0.00	0.38***
	(0.14)	(0.09)	(0.09)	(0.09)	(0.10)	(0.13)
log(bills to GDP)	-0.86***	-1.35***	-1.38***	-1.36***	-1.23***	-2.15**
	(0.23)	(0.19)	(0.25)	(0.33)	(0.29)	(0.97)
VIX	0.02**	0.02***	0.02***	0.02***	0.02***	0.07*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)
Observations	140	140	140	140	140	140
R-squared	0.26	0.56				
Exc. Instrument			HHI	HHI ex B*T FE	post reform	HHI
Underidentification test (p-val)			0.04	0.07	0.03	0.04
Anderson-Rubin test (p-val)			0.01	0.09	0.15	0.02
Kleibergen-Paap rk Wald F stat			11.06	8.48	7.95	11.06

Note: This table reports results of various regressions with alternative specifications of the equation (12) with the *Bloomberg Liquidity Index* in columns (1) through (5) as the dependent variable and the standardized *Amihud liquidity* measure for 1-month T-bills in column (6). For both variables, a higher value corresponds to lower liquidity. Variable descriptions and summary statistics can be found in Table 2. Data are at a monthly frequency between February 2011 and September 2022. Columns (1) and (2) report the results of OLS regressions. Columns (3), (4), and (5) report the second stage estimates of a 2SLS regression, in which HHI *bank repo*, HHI *ex* $B*T$ FE instrument *residual cash share*, and *post reform*, respectively. Wherever applicable, we report the p-values of the under-identification test and the Anderson-Rubin test and the Kleibergen-Paap rk Wald F statistic of the first stage. Standard errors are robust to arbitrary heteroskedasticity and autocorrelation with the lag structure automatically selected using the Newey and West (1994) procedure. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Source: Crane Data, FRED, US Treasury.

February 2011 and September 2022 as our MMF data started in February 2011. In column (2), we add *residual cash share* to the baseline regression and estimate an OLS regression. According to the estimates, the marginal effect of the added variable is economically sizable and statistically significant at the 5% level. Moreover, adding this variable improves the fit considerably as the R^2 goes from 36% to 42%. In columns (3) to (5), we repeat the two-stage least squares exercise as in Table 3 with three alternative instruments.[17] The point estimates are even larger when

[17] A word of caution is for the use of instrumental variables with the GC-T-bill spread on the left-hand side. MMFs account for a relatively smaller share of aggregate dollar repo markets, and GC repo is predominantly between

we use instrumental variables, with the estimates suggesting an equivalent partial effect of a one standard deviation change in *residual cash share* and around a 125 basis point increase in the federal funds rate and the partial impact of around one-fifth of a percent increase in bills-to-GDP. However, estimates in (3) and (4), when we use $HHI\ bank\ repo$ and $HHI\ ex\ B*T\ FE$, respectively, are not statistically significant. In column (5), when we use the $post\ reform$ dummy as the instrument, the result is statistically significant at 10%. It is important to note that as opposed to the literature, which only uses autocorrelation robust standard errors, we report both heteroskedasticity and autocorrelation robust standard errors, and combined with the small sample size, we lack statistical power.

6 Market power-price impact trade-off and portfolio allocations

In this section, we test the two key implications of the model using the contract-level data. The first prediction (Proposition 3.4) is that, on the one hand, higher market power allows MMFs to charge higher repo rates to banks. On the other hand, if funds have a large share in the T-bill market, they internalize their price impact in the T-bill market by charging lower repo rates. Our second key prediction (see Corollary 4.2) is that the allocation between T-bills and RRP favors RRP when the T-bill market is illiquid and therefore, the price impact in the T-bill market is greater.

6.1 Data description

We use a granular and rich dataset of US MMFs' portfolio holdings at month-ends obtained from Crane Data, which is based on the regulatory filings of US MMFs to the Securities and Exchange Commission (SEC N-MFP forms). The sample covers the universe of US MMF funds between February 2011 and November 2022. Holdings data are reported at each month's end.[18] For each holding, the dataset provides information on the face value in dollar amounts, the instrument, the

securities dealers. Therefore, we do not expect a direct effect of MMF market concentration on GC repo. However, it is hard to observe or rule out whether there might be indirect effects that might violate the exclusion restriction for the HHI in the MMF-bank repo lending market due to the very opaque nature of repo markets other than those between MMFs and banks.

[18]We use the same data cleaning procedure as in Aldasoro, Ehlers and Eren (2022). We refer the interested reader to that paper.

Table 5: MMFs' residual cash share and the GC-T-bill spread

	(1)	(2)	(3)	(4)	(5)
	OLS	OLS	2SLS	2SLS	2SLS
VARIABLES	GC-Tbill (1M)	GC-Tbill (1M)	GC-Tbill (1M)	GC-Tbill (1M)	GC-Tbill (1M)
residual cash share		0.13**	0.25	0.29	0.18*
		(0.06)	(0.16)	(0.22)	(0.10)
FFR	4.16***	4.03***	3.90***	3.86**	3.97***
	(1.17)	(1.24)	(1.43)	(1.56)	(1.05)
log(bills to GDP)	-18.64***	-23.29***	-27.88***	-29.32***	-25.29***
	(3.75)	(4.71)	(6.18)	(7.56)	(5.07)
VIX	0.29*	0.28**	0.27***	0.27**	0.28**
	(0.15)	(0.12)	(0.11)	(0.11)	(0.11)
Observations	140	140	140	140	140
R-squared	0.36	0.42			
Exc. Instrument			HHI	HHI ex B*T FE	post reform
Underidentification test (p-val)			0.02	0.04	0.04
Anderson-Rubin test (p-val)			0.02	0.05	0.00
Kleibergen-Paap rk Wald F stat			13.73	9.34	7.61

Note: This table reports results of various regressions with alternative specifications of the equation (12) with the $GC - Tbill\ (1M)$ as the dependent variable. Variable descriptions and summary statistics can be found in Table 2. Data are at a monthly frequency between February 2011 and September 2022. Columns (1) and (2) report the results of OLS regressions. Columns (3), (4), and (5) report the second stage estimates of a 2SLS regression, in which $HHI\ bank\ repo$, $HHI\ ex\ B*T\ FE$ instrument $residual\ cash\ share$, and $post\ reform$, respectively. Wherever applicable, we report the p-values of the under-identification test and the Anderson-Rubin test, and the Kleibergen-Paap rk Wald F statistic of the first stage. Standard errors are robust to arbitrary heteroskedasticity and autocorrelation with the lag structure automatically selected using the Newey and West (1994) procedure. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Source: Crane Data, FRED, US Treasury.

remaining maturity, and the (annualized) yield, among other contract characteristics. In addition, for repos, we observe whether the borrowing is backed by either Treasury, Government Agency, or Other collateral. By law, US MMFs are only allowed to invest in dollar-denominated instruments. Therefore, all transactions are denominated in dollars.

To measure funds' market share ('FMS') in repos with banks and the T-bill market, we define the following two metrics:

$$FMS\ bank\ repo_{f,t} = \frac{\sum_b bank\ repo_{f,b,t}}{\sum_f \sum_b bank\ repo_{f,b,t}} \times 100, \tag{13}$$

$$FMS\ treasury_{f,t} = \frac{amount\ treasury_{f,t}}{\sum_f amount\ treasury_{f,t}} \times 100, \tag{14}$$

where f denotes fund, b bank, and t the month. Higher values of $FMS\ bank\ repo$ ($FMS\ treasury$) proxy greater market power of a fund in the bank lending (T-bill) market.

Table 6: Summary statistics

Panel (a): Summary statistics for variables in Table 7 (contract level data)

Variable	Obs	Mean	Std. Dev.	Min	Max	P50
rate	266382	87.75	92.88	0	550	38
FMS bank repo	266382	1.84	2.12	0	11.89	.82
FMS treasury	266382	.81	1.15	0	15.72	.29

Panel (b): Summary statistics for variables in Table 8 (fund-time level data)

Variable	Obs	Mean	Std. Dev.	Min	Max	P50
RRP share	12997	16.94	30.75	0	99.87	0
FMS treasury	12997	.44	.87	0	10.53	.09
liqu tight (BBG index)	12997	1.3	.48	.67	2.76	1.25

Note: This table reports summary statistics for the key variables used in the empirical analysis. Using contract-level data, the upper panel reports the summary statistics of the variables used in Table 7. The sample period for the upper panel runs between February 2011 and November 2022, with holdings data reported at each month's end. *rate* refers to the repo rate and is in basis points. *FMS bank repo* is the market share of the fund in the repo market (see Eq. 13) *FMS treasury* is the market share of the fund in the T-bill market (see Eq. 14). Both of these variables are in percentage points. In the lower panel, we report the summary statistics of the variables used in Table 8 at the fund-time level. The sample period for the lower panel runs between October 2013 and November 2022 (i.e., after the introduction of the RRP facility). *RRP share* is the share a fund allocates between T-bills and the RRP facility. *FMS treasury* is the market share of the fund in the T-bill market (see Eq. 14). Both of these variables are in percentage points. *liqu tight* (*BBG index*) is the Bloomberg liquidity index which measures the average deviation of yields from a fair-value model. A higher number indicates lower liquidity. Source: Crane Data, Bloomberg.

6.2 Tests

We first test Proposition 3.4 and analyze the effects of funds' market shares ('FMS') in the repo market with banks and the T-bill market on repo rates charged to banks. We estimate variants of the following regression:

$$rate_{i(f,b),t} = \beta_1 \; FMS \; bank \; repo_{f,t} + \beta_2 \; FMS \; treasury_{f,t} + controls_{i,t} + \theta_t + \varepsilon_{i,t}. \tag{15}$$

The dependent variable $rate_{i(f,b),t}$ is the annualized interest rate in basis points on a contract i between fund f and bank b at time t. The explanatory variables $FMS \; bank \; repo_{f,t}$ and $FMS \; treasury_{f,t}$ denote fund f's market share in the bank repo and the T-bill markets in t (as defined in Equations (13) and (14)). To account for time-varying factors that affect different collateral types (US Treasury, government agency, or other collateral), the baseline regression includes time-varying fixed effects at the collateral type level (θ_t). Control variables are the size and the maturity of the contract. Standard errors are double clustered at both fund and time levels.

Proposition 3.4 states that both market power and funds' internalization of their price impact in the T-bill market should enter into their consideration when setting repo rates. In particular, funds with greater market power (proxied by fund market share or FMS bank repo) in the repo market charge higher rates, while a greater market share in the T-bill market lower the repo rates charged by the same fund due to the internalization of the price impact. We hence expect that $\beta_1 > 0$ and $\beta_2 < 0$.

Regression equation (15) faces the identification challenge that the observed rate could be determined by observable or unobservable time-varying factors that vary at the fund type (prime, government, or Treasury fund) or bank level. For example, if funds with a greater market share in the bank lending market lend to riskier banks, then any observed positive correlation between *FMS bank repo* and the rate reflects borrower characteristics, rather than market power. Moreover, prime funds might be subject to different shocks than government or treasury funds, which could influence the repo rate. As we will discuss in what follows, we address these identification challenges through the inclusion of granular time-varying fixed effects.

Table 7 shows that funds with a higher market share in the repo market charge higher repo rates, while, all else constant, funds with a higher market share in the T-bill market charge lower rates. Column (1) reports a positive coefficient on *FMS bank repo* ($\beta_1 > 0$) significant at the 5% level, and a negative coefficient on *FMS treasury* ($\beta_2 < 0$) significant at the 1% level. Column (2) shows that this pattern is robust to the inclusion of fund type*time fixed effects that account for any unobservable shocks that affect different fund types. Both coefficients are now significant at the 1% level.

To account for unobservable time-varying differences in bank characteristics, including changes in risk, size, or credit demand, column (3) includes bank*time fixed effects. Both coefficients keep their sign and significance and change only marginally in size. The stability of the coefficients suggests that unobservable borrower characteristics do not explain the correlation between fund market power and rates. In column (4) we include the log amount of a contract and finely-grained fixed effects for different maturities to control for the fact that rates could be correlated with the amount and maturity. Results show that even when we compare contracts of similar amounts and with comparable maturity, funds with a higher market share in the repo market still charge higher repo rates and a higher market share in the T-bill markets corresponds to lower repo rates. Finally, column (5) further allows any unobservable time-varying shock at the bank- or fund type-level to affect each collateral type differentially in each month. Yet our baseline result remains robust even when we include bank*collateral*time and fund type*collateral*time fixed effects.

In terms of magnitude, in column (5) the partial impact of a 2 percentage point increase in the fund market share in the repo market (roughly one standard deviation) corresponds to a 42 basis points higher repo rate. This is economically important given the overall mean of repo rates is 88 basis points. Similarly, a 1 percentage point increase in the share of a fund among other MMFs in the T-bill market (again roughly one standard deviation), leads to a 47 basis points lower repo rate. These results suggest that MMFs substantially internalize their price impact in the T-bill market when they are setting repo rates.

Next, we turn to Prediction 2 (Corollary 4.2), which states that the share of residual cash left

Table 7: Funds have market power in the repo market, but also internalize their price impact in the T-bill market

	(1)	(2)	(3)	(4)	(5)
VARIABLES	rate	rate	rate	rate	rate
FMS bank repo	0.209**	0.357***	0.296***	0.272***	0.209**
	(0.105)	(0.109)	(0.091)	(0.098)	(0.090)
FMS treasury	-0.495***	-0.590***	-0.596***	-0.567***	-0.477***
	(0.150)	(0.152)	(0.143)	(0.137)	(0.141)
Observations	266,382	266,382	266,195	266,193	265,500
R-squared	0.957	0.958	0.963	0.964	0.968
collateral*time FE	✓	✓	✓	✓	-
fund type*time FE	-	✓	✓	✓	-
bank*time FE	-	-	✓	✓	-
fund type*collateral*time FE	-	-	-	-	✓
bank*collateral*time	-	-	-	-	✓
controls	-	-	-	✓	✓

Note: This table reports the results of the regressions for alternative specifications of equation (15). Variable descriptions and summary statistics can be found in Table 6. The unit of observation is a contract between a fund and a bank reported as part of the disclosure of MMFs' portfolio holdings snapshot at month ends between February 2011 and November 2022. In column (1), we include collateral*time fixed effects. Collateral categories are US Treasury, government agency, and other collateral. In column (2), we add fund type*time fixed effects. Fund type categories are government, Treasury, and prime funds catered to retail or institutional investors. In column (3), we add bank*time fixed effects. In column (4), we add control variables which are the log contract size and fixed effects of finely-grained maturity buckets. In column (5), we include fund type*collateral*time and bank*collateral*time fixed effects as well as control variables. Standard errors, double clustered at fund and time level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Source: Crane Data.

over from repo lending allocated to overnight RRP is higher for funds with a higher share in the T-bill market, and in particular when Treasury market liquidity is low.

To test this prediction, we estimate regressions at the fund f-month t level:

$$RRP\ share_{f,t} = \delta_1\ FMS\ treasury_{f,t} + \delta_2\ liquidity\ tightness_t \qquad (16)$$
$$+ \delta_3\ FMS\ treasury_{f,t} \times liquidity\ tightness_t + controls_{f,t} + \theta_t + \varepsilon_{f,t}.$$

The dependent variable $RRP\ share_{f,t}$ is the share of cash left over from repo lending allocated to RRP as opposed to T-bills for fund f at time t. The explanatory variable $FMS\ treasury_{f,t}$ denotes fund f's market share in the T-bill market in t (as defined in Equation (14)). The variable $liquidity\ tightness_t$ measures liquidity conditions in the overall Treasury market, measured by the average deviation of yields from a fair-value model, with higher values indicating lower liquidity (proxied by the Bloomberg Liquidity Index). To account for time-varying factors that affect all funds, the baseline regression includes time-fixed effects (θ_t) as we progressively saturate the regressions with more fixed effects and control variables. Standard errors are double clustered at the level of both fund and time. Prediction 2 states that funds with greater market power in the T-bill market allocate a greater share of their assets to RRP, especially when liquidity conditions in the Treasury market are worse ($\delta_3 > 0$).

Table 8 shows results consistent with Prediction 2. Column (1) includes only $FMS\ treasury$ as the dependent variable and shows a highly significant positive relationship between a fund's market share in the T-bill market and the share of assets allocated to RRP. This correlation remains unaffected when we include time-fixed effects in column (2). In terms of magnitude, a 1 percentage point higher T-bill market share is associated with a 4.3 percentage point higher RRP share of the total residual cash left over from repo lending and invested between T-bills and RRP (corresponding to around a fourth of the mean RRP share). Column (3) introduces the interaction term of $FMS\ treasury$ with $liquidity$ and shows a positive coefficient on the interaction term, significant at the 1% level. Column (4) adds fund-fixed effects to control for any unobservable fund-specific characteristics. Column (5) adds fund type*time fixed effects to control for any time-varying differences across different fund types as well as control variables such as the log change in the assets under management and interaction of $FMS\ treasury$ with the 1-month T-bill rate and the federal funds rate separately. In column (6) we hence replicate column (5) but use the *change* in the liquidity indicator as the explanatory variable. We obtain similar results.

Across specifications, the sign and significance of our coefficients of interest remain similar, suggesting that the predicted relationship between fund market share, liquidity, and the RRP share is not due to unobservable fund characteristics, nor time-varying shocks that affect different fund types, nor changes in the assets under management, nor reflecting changes in the Fed funds rate or T-bill rate that could affect funds with different market shares differentially.

Table 8: Treasury market liquidity and the share of residual cash going to RRP

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	RRP share	RRP share	RRP share	RRP share	RRP share	RRP share
FMS treasury	3.758***	4.355***	-3.721**	-11.582***	-10.665***	-4.700***
	(1.189)	(1.130)	(1.843)	(2.421)	(2.090)	(1.186)
FMS treasury × liqu tight			6.620***	5.412***	5.312***	
			(2.067)	(1.272)	(1.397)	
FMS treasury × Δ liqu tight						8.973***
						(3.052)
Observations	12,997	12,997	12,997	12,980	11,865	11,774
R-squared	0.011	0.243	0.251	0.606	0.653	0.652
time FE	-	✓	✓	✓	-	-
fund FE	-	-	-	✓	✓	✓
fund type*time FE	-	-	-	-	✓	✓
controls	-	-	-	-	✓	✓

Note: This table reports the results of the regressions for alternative specifications of equation (16). Variable descriptions and summary statistics can be found in Table 6. Observations are at the fund-time level constructed from the holding level data reported as part of the disclosure of MMFs' portfolio holdings snapshot at month ends between the introduction of the RRP facility in October 2013 and November 2022. In column (1), we run a simple bivariate OLS regression. In column (2), we include time-fixed effects. In column (3), we include the interaction between $FMS\ treasury$, which measures the share of an individual fund among all MMFs in the T-bill market, and *liqu tight*, which is the Bloomberg Liquidity Index (higher values correspond to lower liquidity in the entire Treasury market). In column (4), we further add fund fixed effects. In column (5), we further add fund type*time fixed effects and several control variables at the fund-time level (log change in AUM, the interaction between $FMS\ treasury$ and the federal funds rate as well as the 1-month T-bill rate). In column (6), we repeat column (5) but replace *liqu tight* with Δ*liqu tight*, which measures the changes in the Bloomberg Liquidity Index. Standard errors, double clustered at fund and time level, are in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Source: Crane Data.

7 Policy implications

Our results have policy implications for the transmission of monetary policy and benchmark rates.

First, MMFs typically receive inflows when the federal funds rate is higher (e.g. Duffie and Krishnamurthy, 2016; Drechsler, Savov and Schnabl, 2017; Xiao, 2020). Therefore, the market power-price impact trade-off could lead to downward pressure on T-bill rates during these episodes, weakening the effectiveness of interest rate policy. This effect might also transmit to the long end of the yield curve due to the high sensitivity of long-term rates to short-term rates (Hanson, Lucca and Wright, 2021) and might lead to negative externalities if it incentivizes private money creation (Greenwood, Hanson and Stein, 2015).

Second, a smaller central bank balance sheet would mean a smaller role for the RRP facility. This would intensify the price impact of MMFs on T-bill rates, and arguably especially so during flight-to-quality episodes in which MMFs receive large inflows. For example, during the Covid-19 crisis, government and Treasury MMFs received close to $1 trillion inflows within weeks (e.g. Eren, Schrimpf and Sushko, 2020b).

Third, as the transition from credit-sensitive benchmark rates to risk-free repo-based benchmark rates is underway (Huang and Todorov, 2022), an often-mentioned reason for this transition is to have "a more resilient rate than LIBOR because of [...] the depth and liquidity of the markets that underlie it" (ARCC, 2021).[19] We show that new benchmark rates could also be affected by the market structure and liquidity conditions in repo and Treasury markets which could have large spillover effects to many other markets in which these benchmark rates are used.

Finally, our results imply that market concentration in the MMF sector has broader consequences for the macroeconomy through its effect on short-term money market interest rates and spreads. As a result, policy reforms that have impacted market concentration might have had unintended consequences such as the reduction of T-bill rates and an increase in the liquidity premium.

[19]The Alternative Reference Rates Committee (ARRC) is a group of private-market participants convened by the Federal Reserve Board and the New York Fed to help ensure a smooth transition from LIBOR to Secured Overnight Financing Rate (SOFR).

8 Conclusion

In this paper, we show that repo markets and T-bill markets are connected through the large presence of MMFs in both markets. MMFs have market power in the repo market, but they are also large in the T-bill market. As a result, they also have a price impact on the T-bill market. They set repo rates with banks and interact with other MMFs in the T-bill market. Strategic and optimal price setting and portfolio choice determine how key interest rates and spreads are determined in equilibrium and how MMFs allocate their portfolio between repos with banks, the Federal Reserve, and T-bills. The key drivers of these choices are frictions due to market concentration in the repo market and liquidity conditions in the T-bill market. Our results suggest that monetary policy transmission can weaken due to these frictions, central bank balance sheet size plays an important role in affecting MMFs' trade-offs and macroeconomic outcomes, and these market frictions can affect benchmark rates. Our future work will focus on the broader macroeconomic consequences of our findings.

References

Afonso, Gara, Lorie Logan, Antoine Martin, William Riordan, and Patricia Zobel, "How the Fed's overnight reverse repo facility works," *Liberty Street Economics Blog*, 2022.

___ **, Marco Cipriani, Adam M Copeland, Anna Kovner, Gabriele La Spada, and Antoine Martin**, "The market events of mid-september 2019," *FRB of New York Staff Report*, 2020, (918).

___ **,** ___ **, and Gabriele La Spada**, "Banks' Balance-Sheet Costs, Monetary Policy, and the ON RRP," *FRB of New York Staff Report*, 2022, (1041).

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

Amihud, Yakov, "Illiquidity and stock returns: cross-section and time-series effects," *Journal of Financial Markets*, 2002, *5* (1), 31–56.

___ **and Haim Mendelson**, "Liquidity, maturity, and the yields on US Treasury securities," *The Journal of Finance*, 1991, *46* (4), 1411–1425.

Anderson, Alyssa, Wenxin Du, and Bernd Schlusche, "Arbitrage Capital of Global Banks," *working paper*, 2022.

ARCC, "Transition from LIBOR," https://www.newyorkfed.org/arrc/sofr-transition 2021. Accessed: 2022-12-10.

Avalos, Fernando, Torsten Ehlers, and Egemen Eren, "September stress in dollar repo markets: passing or structural?," *BIS Quarterly Review*, 2019.

CGFS, "Repo market functioning report," *Committee on the Global Financial System Papers*, 2017, (59).

Chernenko, Sergey and Adi Sunderam, "Frictions in Shadow Banking: Evidence from the Lending Behavior of Money Market Mutual Funds," *Review of Financial Studies*, 2014, *27*, 1717–1750.

Cipriani, Marco and Gabriele La Spada, "Investors' appetite for money-like assets: The MMF industry after the 2014 regulatory reform," *Journal of Financial Economics*, 2021, *140* (1), 250–269.

Copeland, Adam, Antoine Martin, and Michael Walker, "Repo runs: Evidence from the tri-party repo market," *Journal of Finance*, 2014, *69* (6), 2343–2380.

___ , Darrell Duffie, and Yilin Yang, "Reserves were not so ample after all," Technical Report, National Bureau of Economic Research 2021.

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

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

Du, Wenxin, Benjamin M Hébert, and Wenhao Li, "Intermediary balance sheets and the treasury yield curve," Technical Report, National Bureau of Economic Research 2022.

Duffee, Gregory R, "Idiosyncratic variation of Treasury bill yields," *The Journal of Finance*, 1996, *51* (2), 527–551.

Duffie, Darrell, "Still the World's Safe Haven?," *Redesigning the US Treasury market after the COVID-19 crisis', Hutchins Center on Fiscal and Monetary Policy at Brookings, available online at https://www. brookings. edu/research/still-the-worlds-safe-haven*, 2020.

___ and Arvind Krishnamurthy, "Passthrough efficiency in the Fed's new monetary policy setting," in "Designing Resilient Monetary Policy Frameworks for the Future. Federal Reserve Bank of Kansas City, Jackson Hole Symposium" 2016, pp. 1815–1847.

d'Avernas, Adrien and Quentin Vandeweyer, "Intraday liquidity and money market dislocations," *Available at SSRN*, 2020.

___ and ___ , "Treasury bill shortages and the pricing of short-term assets," Technical Report, Working Paper 2021.

Engel, Charles and Steve Pak Yeung Wu, "Liquidity and exchange rates: An empirical investigation," Technical Report, National Bureau of Economic Research 2021.

Eren, Egemen and Philip D Wooldridge, "Non-bank financial institutions and the functioning of government bond markets," *BIS Papers*, 2021, (119).

___ , Andreas Schrimpf, and Vladyslav Sushko, "US dollar funding markets during the Covid-19 crisis - the international dimension," *BIS Bulletin no.15*, 2020.

——, ——, and ——, "US dollar funding markets during the Covid-19 crisis - the money market fund turmoil," *BIS Bulletin no.14*, 2020.

Goldreich, David, Bernd Hanke, and Purnendu Nath, "The price of future liquidity: Time-varying liquidity in the US Treasury market," *Review of Finance*, 2005, *9* (1), 1–32.

Greenwood, Robin and Dimitri Vayanos, "Bond supply and excess bond returns," *The Review of Financial Studies*, 2014, *27* (3), 663–713.

——, **Samuel G Hanson, and Jeremy C Stein**, "A comparative-advantage approach to government debt maturity," *The Journal of Finance*, 2015, *70* (4), 1683–1722.

Han, Song and Kleopatra Nikolaou, "Trading Relationships in the OTC Market for Secured Claims: Evidence from Triparty Repos," FEDS Working Paper 2016-64, Federal Reserve Board 2016.

Hanson, Samuel G, David O Lucca, and Jonathan H Wright, "Rate-amplifying demand and the excess sensitivity of long-term rates," *The Quarterly Journal of Economics*, 2021, *136* (3), 1719–1781.

He, Zhiguo, Stefan Nagel, and Zhaogang Song, "Treasury inconvenience yields during the covid-19 crisis," *Journal of Financial Economics*, 2022, *143* (1), 57–79.

Hu, Grace Xing, Jun Pan, and Jiang Wang, "Tri-party repo pricing," *Journal of Financial and Quantitative Analysis*, 2021, *56* (1), 337–371.

Huang, Wenqian and Karamfil Todorov, "The post-Libor world: a global view from the BIS derivatives statistics," 2022.

Huber, Amy Wang, "Market Power in Wholesale Funding: A Structural Perspective from the Triparty Repo Market," Technical Report, Working Paper 2022.

Jiang, Zhengyang, Arvind Krishnamurthy, and Hanno Lustig, "Foreign safe asset demand and the dollar exchange rate," *The Journal of Finance*, 2021, *76* (3), 1049–1089.

Kacperczyk, Marcin and Philipp Schnabl, "How Safe Are Money Market Funds?," *The Quarterly Journal of Economics*, 2013, *128* (3), 1413–42.

Klingler, Sven and Suresh M Sundaresan, "Diminishing Treasury Convenience Premiums: Effects of Dealers' Excess Demand in Auctions," *Available at SSRN 3556502*, 2020.

Krishnamurthy, Arvind and Annette Vissing-Jorgensen, "The aggregate demand for treasury debt," *Journal of Political Economy*, 2012, *120* (2), 233–267.

___ **and** ___ , "The impact of treasury supply on financial sector lending and stability," *Journal of Financial Economics*, 2015, *118* (3), 571–600.

___ **, Stefan Nagel, and Dmitry Orlov**, "Sizing up repo," *Journal of Finance*, 2014, *69* (6), 2381–2417.

Lenel, Moritz, "Safe assets, collateralized lending and monetary policy," *Stanford Institute for Economic Policy Research Discussion Paper*, 2017, pp. 17–010.

___ **, Monika Piazzesi, and Martin Schneider**, "The short rate disconnect in a monetary economy," *Journal of Monetary Economics*, 2019, *106*, 59–77.

Li, Wenhao, Yiming Ma, and Yang Zhao, "The passthrough of treasury supply to bank deposit funding," Technical Report 2019.

Li, Yi, "Reciprocal lending relationships in shadow banking," *Journal of Financial Economics*, 2021, *141* (2), 600–619.

Longstaff, Francis A, "The term structure of very short-term rates: New evidence for the expectations hypothesis," *Journal of Financial Economics*, 2000, *58* (3), 397–415.

___ , "The flight-to-liquidity premium in US Treasury bond prices," 2004.

Malamud, Semyon and Marzena Rostek, "Decentralized exchange," *American Economic Review*, 2017, *107* (11), 3320–62.

Nagel, Stefan, "The liquidity premium of near-money assets," *The Quarterly Journal of Economics*, 2016, *131* (4), 1927–1971.

Newey, Whitney K and Kenneth D West, "Automatic lag selection in covariance matrix estimation," *The Review of Economic Studies*, 1994, *61* (4), 631–653.

Schmidt, Lawrence, Allan Timmermann, and Russ Wermers, "Runs on Money Market Mutual Funds," *American Economic Review*, September 2016, *106* (9), 2625–57.

Schrimpf, Andreas, Hyun Song Shin, and Vladyslav Sushko, "Leverage and margin spirals in fixed income markets during the Covid-19 crisis," *Working Paper*, 2020.

Sunderam, Adi, "Money creation and the shadow banking system," *The Review of Financial Studies*, 2015, *28* (4), 939–977.

Xiao, Kairong, "Monetary transmission through shadow banks," *The Review of Financial Studies*, 2020, *33* (6), 2379–2420.

A Proofs

Proof of Lemma 3. In the presence of internalization, we get that MMF is maximizing

$$\frac{r_f(b)^{-\alpha_b} w_f}{r_f(b)^{-\alpha_b} w_f \ + \ F_{-f}(b)} (R_*^{\xi} r_f(b)^{1-\xi} - \rho R_*^{\xi} r_f(b)^{-\xi})$$

which is equivalent to maximizing

$$\frac{r_f(b)^{1-\alpha_b-\xi} - \rho r_f(b)^{-\alpha_b-\xi}}{r_f(b)^{-\alpha_b} w_f \ + \ F_{-f}(b)} ,$$

and the first order condition is

$$\begin{aligned}
&((1-\alpha_b-\xi) r_f(b)^{-\alpha_b-\xi} + (\alpha_b+\xi)\rho r_f(b)^{-\alpha_b-\xi-1})(r_f(b)^{-\alpha_b} w_f \ + \ F_{-f}(b)) \\
&+ \alpha_b r_f(b)^{-\alpha_b-1} w_f (r_f(b)^{1-\alpha_b-\xi} - \rho r_f(b)^{-\alpha_b-\xi})) \ = \ 0
\end{aligned}$$

which is equivalent to

$$\begin{aligned}
&((1-\alpha_b-\xi) r_f(b) + (\alpha_b+\xi)\rho) \\
&+ \frac{\alpha_b r_f(b)^{-\alpha_b} w_f (r_f(b) - \rho)}{r_f(b)^{-\alpha_b} w_f \ + \ F_{-f}(b)} \ = \ 0
\end{aligned}$$

so that

$$r_f(b) \ = \ r_*(b) \ + \ \frac{1}{\alpha_b+\xi-1} \frac{\alpha_b r_f(b)^{-\alpha_b} w_f (r_f(b)-\rho)}{\Gamma_*(b)} .$$

Q.E.D.

Proof of Proposition 3.2. We have

$$r_f(b) \ = \ r_*(b) \ + \ F^{-1} r_f^{(1)}(b) \ + \ F^{-2} r_f^{(2)}(b) \ + \ O(F^{-3}) .$$

Our goal is to find $r_f^{(1)}(b),\ r_f^{(2)}(b)$. Substituting, we get

$$
\begin{aligned}
\Gamma_*(b) \;&=\; F^{-1}\sum_f r_f(b)^{-\alpha_b} w_f^* \\
&=\; F^{-1}\sum_f w_f^*\left(r_*(b) \;+\; F^{-1}r_f^{(1)}(b) \;+\; F^{-2}r_f^{(2)}(b) \;+\; O(F^{-3})\right)^{-\alpha_b} \\
&=\; F^{-1}\sum_f w_f^*\Bigg(r_*(b)^{-\alpha_b} \;-\alpha_b r_*(b)^{-1}\Big(F^{-1}r_f^{(1)}(b) \;+\; F^{-2}r_f^{(2)}(b) \;+\; O(F^{-3})\Big) \\
&+\; 0.5\alpha_b(\alpha_b+1)r_*(b)^{-2}F^{-2}(r_f^{(1)}(b))^2\Bigg) \\
&=\; r_*(b)^{-\alpha_b} \;-\; F^{-1}\alpha_b r_*(b)^{-1-\alpha_b}E[r_f^{(1)}(b)] \\
&+\; F^{-2}\alpha_b r_*(b)^{-1-\alpha_b}\Big(0.5(\alpha_b+1)r_*(b)^{-1}E[(r_f^{(1)}(b))^2] - E[r_f^{(2)}(b)]\Big) \;+\; O(F^{-3}) \\
&=\; r_*(b)^{-\alpha_b} \;+\; \Gamma_*(b)^{(1)}F^{-1} \;+\; \Gamma_*(b)^{(2)}F^{-2} \;+\; O(F^{-3})\,.
\end{aligned}
$$

Substituting this gives

$$
\begin{aligned}
&O(F^{-3}) + r_f(b) \;=\; r_*(b) \\
&+\; \frac{1}{\alpha_b+\xi-1}\frac{\alpha_b\Big(r_*(b) \;+\; F^{-1}r_f^{(1)}(b) \;+\; F^{-2}r_f^{(2)}(b)\Big)^{-\alpha_b} w_f\Big(\Big(r_*(b) \;+\; F^{-1}r_f^{(1)}(b) \;+\; F^{-2}r_f^{(2)}(b)\Big) - \rho\Big)}{r_*(b)^{-\alpha_b} \;+\; \Gamma_*(b)^{(1)}F^{-1} \;+\; \Gamma_*(b)^{(2)}F^{-2}} \\
&=\; r_*(b) \;+\; F^{-1}\frac{1}{\alpha_b+\xi-1}\alpha_b r_*(b)^{-\alpha_b}\Big(1 \;-\; F^{-1}r_f^{(1)}(b)r_*(b)^{-1}\alpha_b\Big)w_f^*\Big(\Big(r_*(b) \;+\; F^{-1}r_f^{(1)}(b)\Big) - \rho\Big) \\
&\times r_*(b)^{\alpha_b}\left(1 \;-\; r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)}F^{-1}\right) \\
&=\; r_*(b) \;+\; F^{-1}\frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\Big(1 \;-\; F^{-1}r_f^{(1)}(b)r_*(b)^{-1}\alpha_b - r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)}F^{-1}\Big)\Big(\Big(r_*(b) \;+\; F^{-1}r_f^{(1)}(b)\Big) - \rho\Big) \\
&=\; r_*(b) \;+\; F^{-1}\frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\Big((r_*(b)-\rho) \;+\; F^{-1}\Big(r_f^{(1)}(b) - (r_f^{(1)}(b)r_*(b)^{-1}\alpha_b + r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)})(r_*(b)-\rho)\Big)\Big) \\
&=\; r_*(b) \;+\; F^{-1}\frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\Big((r_*(b)-\rho) \;+\; F^{-1}\Big(r_f^{(1)}(b) - (r_f^{(1)}(b)r_*(b)^{-1}\alpha_b + r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)})(r_*(b)-\rho)\Big)\Big)
\end{aligned}
$$

Thus,

$$
r_f^{(1)}(b) \;=\; \frac{w_f^*\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)
$$

and

$$r_f^{(2)}(b) \;=\; \frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\Big(r_f^{(1)}(b)-(r_f^{(1)}(b)r_*(b)^{-1}\alpha_b+r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)})(r_*(b)-\rho)\Big)$$

and

$$\Gamma_*(b)^{(1)} \;=\; -\alpha_b r_*(b)^{-1-\alpha_b}E[r_f^{(1)}(b)] \;=\; -\alpha_b r_*(b)^{-1-\alpha_b}E[\frac{w_f^*\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)]$$

$$=\; -\,\alpha_b r_*(b)^{-1-\alpha_b}\frac{\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)H(W)$$

where

$$H(W) \;=\; F^{-1}\sum_f (w_f^*)^2\,.$$

Thus,

$$r_f^{(2)}(b) \;=\; \frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\Big(r_f^{(1)}(b)-\Big(r_f^{(1)}(b)r_*(b)^{-1}\alpha_b-\alpha_b r_*(b)^{-1}\frac{\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)H(W)\Big)(r_*(b)-\rho)\Big)$$

$$=\; \frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\Big(r_f^{(1)}(b)-\alpha_b r_*(b)^{-1}\Big(r_f^{(1)}(b)-\frac{\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)H(W)\Big)(r_*(b)-\rho)\Big)$$

$$=\; \frac{w_f^*\alpha_b}{\alpha_b+\xi-1}\Big(\frac{w_f^*\alpha_b}{\alpha_b+\xi-1}-\alpha_b r_*(b)^{-1}\Big(\frac{w_f^*\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)-\frac{\alpha_b}{\alpha_b+\xi-1}(r_*(b)-\rho)H(W)\Big)\Big)(r_*(b)-\rho)$$

Q.E.D.

Proof of Proposition 3.3. We have

$$
\begin{aligned}
\hat{\rho} \;&=\; \rho_* \;+\; \frac{U_{-f} - a_*(f)\Delta_f((R(b))_{b\in B})}{V_{-f} + \lambda_*(f)\Delta_f((R(b))_{b\in B})} \\
&=\; \rho_* \;+\; \frac{U_{-f} - a_*(f)\left(d_f - \sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)}\right)}{V_{-f} + \lambda_*(f)\left(d_f - \sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)}\right)} \\
&=\; \rho_* \;+\; \frac{U_{-f} - a_*(f)\left(d_f - \sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)}\right)}{V_{-f} + \lambda_*(f)d_f} \\
&\times \left(1 + \frac{1}{V_{-f}+\lambda_*(f)d_f}\sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)} + \Big(\frac{1}{V_{-f}+\lambda_*(f)d_f}\sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)}\Big)^2\right) \;+\; O(F^{-3}) \\
&=\; \rho_* \;+\; \frac{U_{-f} - a_*(f)d_f}{V_{-f} + \lambda_*(f)d_f} \\
&+\; \left(\frac{a_*(f)}{V_{-f}+\lambda_*(f)d_f} \;+\; \frac{U_{-f} - a_*(f)d_f}{(V_{-f}+\lambda_*(f)d_f)^2}\right)\sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)} \\
&+ \frac{U_{-f} - a_*(f)d_f}{V_{-f} + \lambda_*(f)d_f}\Big(\frac{1}{V_{-f}+\lambda_*(f)d_f}\sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)}\Big)^2 + \;O(F^{-3}) \\
&=\; \rho_* \;+\; \frac{U_{-f} - a_*(f)d_f}{V_{-f} + \lambda_*(f)d_f} \\
&+\; \Xi_{-f}\sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)} + \frac{U_{-f} - a_*(f)d_f}{V_{-f} + \lambda_*(f)d_f}\Big(\frac{1}{V_{-f}+\lambda_*(f)d_f}\sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b}w_f}{\Gamma_*(b)}\Big)^2 + \;O(F^{-3})
\end{aligned}
$$

where we have defined

$$
\Xi_{-f} \;=\; \left(\frac{a_*(f)}{V_{-f}+\lambda_*(f)d_f} \;+\; \frac{U_{-f} - a_*(f)d_f}{(V_{-f}+\lambda_*(f)d_f)^2}\right).
$$

Therefore,

$$
\begin{aligned}
&\frac{\partial \hat{\rho}}{\partial r_f(u)} \\
&= \ \Xi_{-f} w_f \Bigg(R_*^\xi \frac{-(\xi+\alpha_u) r_f(u)^{-\xi-\alpha_u-1}\Gamma_*(u) + \alpha_u w_f r_f(u)^{-\alpha_u-1} r_f(u)^{-\xi-\alpha_u}}{\Gamma_*(u)^2} \Bigg) \\
&+ 2w_f^2 \frac{U_{-f} - a_*(f)d_f}{(V_{-f} + \lambda_*(f)d_f)^3} \Big(\sum_b (R_*/r_*(b))^\xi \Big) \Bigg(R_*^\xi \frac{-(\xi+\alpha_u) r_*(u)^{-\xi-\alpha_u-1}\Gamma_*(u)}{\Gamma_*(u)^2} \Bigg) + \ O(F^{-3}) \\
&= \ \Xi_{-f} w_f R_*^\xi \Bigg(-(\xi+\alpha_u) r_f(u)^{-\xi-\alpha_u-1}\Gamma_*(u)^{-1} + \alpha_u w_f r_*(u)^{-\alpha_u-1} r_*(u)^{-\xi-\alpha_u}\Gamma_*(u)^{-2} \Bigg) \\
&+ 2w_f^2 \frac{U_{-f} - a_*(f)d_f}{(V_{-f} + \lambda_*(f)d_f)^3} \Big(\sum_b (R_*/r_*(b))^\xi \Big) \Bigg(R_*^\xi \frac{-(\xi+\alpha_u) r_*(u)^{-\xi-\alpha_u-1}\Gamma_*(u)}{\Gamma_*(u)^2} \Bigg) + \ O(F^{-3}) \\
&= \ \Xi_{-f} w_f^* F^{-1} R_*^\xi \Bigg(-(\xi+\alpha_u) r_*(u)^{-\xi-\alpha_u-1}(1-(\xi+\alpha_u+1) r_*(u)^{-1} r_f^{(1)} F^{-1}) r_*(u)^{\alpha_u} (1 - r_*(u)^{\alpha_u}\Gamma_*(u)^{(1)} F^{-1}) \\
&+ \alpha_u w_f^* F^{-1} r_*(u)^{-\alpha_u-1} r_*(u)^{-\xi-\alpha_u} r_*(u)^{2\alpha_u} \Bigg) \\
&+ 2(w_f^*)^2 F^{-2} \frac{U_{-f} - a_*(f)d_f}{(V_{-f} + \lambda_*(f)d_f)^3} \Big(\sum_b (R_*/r_*(b))^\xi \Big) \Bigg(R_*^\xi \frac{-(\xi+\alpha_u) r_*(u)^{-\xi-\alpha_u-1}\Gamma_*(u)}{\Gamma_*(u)^2} \Bigg) + \ O(F^{-3}) \\
&= \ -\Xi_{-f} w_f^* F^{-1} R_*^\xi (\xi+\alpha_u) r_*(u)^{-\xi-1} \\
&+ \ \Xi_{-f} w_f^* F^{-2} R_*^\xi r_*(u)^{-\xi-1} \Bigg((\xi+\alpha_u)\Big((\xi+\alpha_u+1) r_*(u)^{-1} r_f^{(1)} + r_*(u)^{\alpha_u}\Gamma_*(u)^{(1)} \Big) \ + \ \alpha_u w_f^* \Bigg) - \ Q_f^*(u) F^{-2}
\end{aligned}
\tag{17}
$$

where we have defined

$$
Q_f^*(u) \ = \ 2(w_f^*)^2 F^{-2} \frac{U_{-f} - a_*(f)d_f}{(V_{-f} + \lambda_*(f)d_f)^3} \Big(\sum_b (R_*/r_*(b))^\xi \Big) \Bigg(R_*^\xi (\xi+\alpha_u) r_*(u)^{-\xi-1} \Bigg)
$$

Thus,

$$
\frac{\partial \hat{\rho}}{\partial r_f(u)} \ = \ \hat{\rho}^{(1)}_{r_f(u)} F^{-1} \ + \ \hat{\rho}^{(2)}_{r_f(u)} F^{-2} \ + \ O(F^{-3}), \tag{18}
$$

where, using that

$$
\Gamma_*(b) \ = \ r_*(b)^{-\alpha_b} \ + \ \Gamma_*(b)^{(1)} F^{-1},
$$

with

$$\Gamma_*(b)^{(1)} \;=\; -\alpha_b r_*(b)^{-1-\alpha_b} E[r_f^{(1)}(b)]\,, \tag{19}$$

and *where* $r_f^{(1)}(b)$ *is to be determined later in general equilibrium.* Q.E.D.

Proof of Proposition 3.4. The first order condition with respect to a particular bank u is

$$\begin{aligned}
0 \;&=\; \frac{\partial}{\partial r_f(u)} \sum_b \frac{r_f(b)^{1-\alpha_b-\xi} - \widetilde{\rho}_f r_f(b)^{-\alpha_b-\xi}}{r_f(b)^{-\alpha_b} w_f \;+\; F_{-f}(b)} + \frac{\partial}{\partial r_f(u)} \widetilde{\rho}_f d_f / w_f \\
&= \Big(((1-\alpha_u-\xi) r_f(u)^{-\alpha_u-\xi} + (\alpha_u+\xi)\widetilde{\rho}_f r_f(u)^{-\alpha_u-\xi-1})(r_f(u)^{-\alpha_u} w_f \;+\; F_{-f}(u)) \\
&+ \alpha_u r_f(u)^{-\alpha_u-1} w_f (r_f(u)^{1-\alpha_u-\xi} - \widetilde{\rho}_f r_f(u)^{-\alpha_u-\xi}) \Big) \Big(r_f(u)^{-\alpha_u} w_f \;+\; F_{-f}(u) \Big)^{-2} \\
&+ \frac{\partial \widetilde{\rho}_f}{\partial r_f(u)} w_f^{-1} \Delta_f
\end{aligned}$$

where we have defined

$$\widetilde{\rho}_f \;=\; \Big((\hat{\rho}(r_f) - \rho_*)(a_*(f) + \lambda_*(f)(\hat{\rho}(r_f) - \rho_*)) + \rho_* \Big)$$

to be the effective T-bill rate. Now, by (17), we have

$$\begin{aligned}
&\frac{\partial}{\partial r_f(u)} \widetilde{\rho}_f \\
&= \frac{\partial}{\partial r_f(u)} \Big((\hat{\rho}(r_f) - \rho_*)(a_*(f) + \lambda_*(f)(\hat{\rho}(r_f) - \rho_*)) + \rho_* \Big) \\
&= \frac{\partial \hat{\rho}_u}{\partial r_f(u)} (a_*(f) + 2\lambda_*(f)(\hat{\rho}(r_f) - \rho_*))\,.
\end{aligned}$$

Thus,

$$\begin{aligned}
0 \;&=\; \Big(((1-\alpha_u-\xi) r_f(u)^{-\alpha_u-\xi} + (\alpha_u+\xi)\widetilde{\rho}_f r_f(u)^{-\alpha_u-\xi-1}) \Gamma_*(b) \\
&+ \alpha_u r_f(u)^{-\alpha_u-1} w_f (r_f(u)^{1-\alpha_u-\xi} - \widetilde{\rho}_f r_f(u)^{-\alpha_u-\xi}) \Big) \\
&+ \; \Gamma_*(b)^2 \frac{\partial \widetilde{\rho}_f}{\partial r_f(u)} \Delta(f) w_f^{-1} \;+\; O(F^{-3})\,.
\end{aligned}$$

Diving this identity by $r_f(u)^{-\alpha_u-\xi-1}(\alpha_u+\xi-1)\Gamma_*(b)$, we get

$$0 \;=\; O(F^{-3}) \;+\; \Big((-r_f(u)+\frac{\alpha_u+\xi}{\alpha_u+\xi-1}\widetilde{\rho}_f))$$

$$+\,\alpha_u r_f(u)^{-\alpha_u-1}w_f(r_f(u)^{1-\alpha_u-\xi}-\widetilde{\rho}_f r_f(u)^{-\alpha_u-\xi})(\alpha_u+\xi-1)^{-1}\Gamma_*(u)^{-1}r_f(u)^{\alpha_u+\xi+1}\Big)$$

$$(r_f(u)^{-\alpha_u-\xi-1}(\alpha_u+\xi-1)\Gamma_*(b))^{-1}\frac{\partial\widetilde{\rho}_f}{\partial r_f(u)}\Gamma_*(b)^2 w_f^{-1}\Delta_f$$

so that, using (17) and (18), we get

$$
\begin{aligned}
r_f(u) &= \frac{\alpha_u+\xi}{\alpha_u+\xi-1}\widetilde{\rho}_f \\
&+ \frac{\alpha_u}{\alpha_u+\xi-1}\frac{r_f(u)^{-\alpha_u} w_f(r_f(u)-\widetilde{\rho}_f)}{\Gamma_*(u)} \\
&+ (r_f(u)^{-\alpha_u-\xi-1}(\alpha_u+\xi-1))^{-1}\frac{\partial\widetilde{\rho}_f}{\partial r_f(u)}\Gamma_*(b) w_f^{-1}\Delta_f \;+\; O(F^{-3}) \\
&= \frac{\alpha_u+\xi}{\alpha_u+\xi-1}\widetilde{\rho}_f \\
&+ \frac{\alpha_u}{\alpha_u+\xi-1}\frac{w_f(r_f(u)^{1-\alpha_u} - r_f(u)^{-\alpha_u}\widetilde{\rho}_f)}{\Gamma_*(u)} \\
&+ (r_f(u)^{-\alpha_u-\xi-1}(\alpha_u+\xi-1))^{-1}\frac{\partial\widetilde{\rho}_f}{\partial r_f(u)}\Gamma_*(b) w_f^{-1}\Delta_f \;+\; O(F^{-3}) \\
&= \frac{\alpha_u+\xi}{\alpha_u+\xi-1}\widetilde{\rho}_f \\
&+ \frac{\alpha_u}{\alpha_u+\xi-1}\frac{w_f\Big(r_*(u)^{1-\alpha_u}(1+(1-\alpha_u)r_*(u)^{-1}r_f^{(1)}(u)F^{-1}) - r_*(u)^{-\alpha_u}(1-\alpha_u r_*(u)^{-1}r_f^{(1)}(u)F^{-1})\widetilde{\rho}_f\Big)}{r_*(u)^{-\alpha_u} \;+\; \Gamma_*(u)^{(1)}F^{-1}} \\
&+ ((\alpha_u+\xi-1))^{-1} r_*(u)^{\alpha_u+\xi+1}\big(1+(\alpha_u+\xi+1)r_*(u)^{-1}r_f^{(1)}(u)F^{-1}\big) \\
&\times \Big(\hat{\rho}^{(1)}_{r_f(u)}F^{-1} \;+\; \hat{\rho}^{(2)}_{r_f(u)}F^{-2}\Big)(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*)) \\
&\times (r_*(u)^{-\alpha_u} \;+\; \Gamma_*(u)^{(1)}F^{-1})w_f^{-1}\Delta_f \;+\; O(F^{-3}) \\
&= \frac{\alpha_u+\xi}{\alpha_u+\xi-1}\widetilde{\rho}_f \\
&+ \frac{\alpha_u w_f}{\alpha_u+\xi-1}\Big((r_*(u)^{1-\alpha_u} - r_*(u)^{-\alpha_u}\widetilde{\rho}_f) + \;((1-\alpha_u)r_*(u)^{-\alpha_u} + \alpha_u r_*(u)^{-\alpha_u-1}\widetilde{\rho}_f)r_f^{(1)}(u)F^{-1}\Big) \\
&\times r_*(u)^{\alpha_u}(r_*(u)^{\alpha_u} \;-\; r_*(u)^{\alpha_u}\Gamma_*(u)^{(1)}F^{-1}) \\
&+ ((\alpha_u+\xi-1))^{-1} r_*(u)^{\alpha_u+\xi+1}\big(1+(\alpha_u+\xi+1)r_*(u)^{-1}r_f^{(1)}(u)F^{-1}\big) \\
&\times F^{-1}\Big(\hat{\rho}^{(1)}_{r_f(u)} \;+\; \hat{\rho}^{(2)}_{r_f(u)}F^{-1}\Big)(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*)) \\
&\times (r_*(u)^{-\alpha_u} \;+\; \Gamma_*(u)^{(1)}F^{-1})w_f^{-1}\Delta_f \;+\; O(F^{-3})
\end{aligned}
$$

Thus,

$$
\begin{aligned}
& r_f(u)^{(1)} \\
= \;& \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} \frac{r_*(u)^{1-\alpha_u} - r_*(u)^{-\alpha_u} \widetilde{\rho}_f}{r_*(u)^{-\alpha_u}} \\
+ \;& ((\alpha_u + \xi - 1))^{-1} r_*(u)^{\alpha_u + \xi + 1} \hat{\rho}^{(1)}_{r_f(u)} r_*(u)^{-\alpha_u} w_f^{-1} \Delta_f \\
= \;& \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} (r_*(u) - \widetilde{\rho}_f) \\
- \;& ((\alpha_u + \xi - 1))^{-1} r_*(u)^{\xi+1} \Big(w_f^* \Xi_{-f} \left(R_*^{\xi} \frac{(\xi + \alpha_u) r_*(u)^{-\xi - \alpha_u - 1}}{r_*(u)^{-\alpha_u}} \right) \Big) \Delta_f^* (a_*(f) + 2\lambda_*(f)(\hat{\rho}(r_f) - \rho_*)) \\
= \;& \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} (r_*(u) - \widetilde{\rho}_f) - \frac{\xi + \alpha_u}{\xi + \alpha_u - 1} w_f^* \Xi_{-f} R_*^{\xi} \Delta_f^* (a_*(f) + 2\lambda_*(f)(\hat{\rho}(r_f) - \rho_*))
\end{aligned}
$$

Hence, by (19), we have

$$
\Gamma_*(b)^{(1)} \;=\; -\alpha_b r_*(b)^{-1-\alpha_b} E[r_f^{(1)}(b)]\,.
$$

Furthermore, defining

$$
\Delta_f^* \;=\; \Delta_f / w_f,
$$

we get

$$
\begin{aligned}
&r_f(u)^{(2)}\\
&=\ \hat{\rho}^{(2)}_{r_f(u)}((\alpha_u+\xi-1))^{-1}r_*(u)^{\alpha_u+\xi+1}r_*(u)^{-\alpha_u}\Delta^*_f(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*))\\
&+\ \frac{\alpha_u w_f}{\alpha_u+\xi-1}\Big(-(r_*(u)^{1-\alpha_u}-r_*(u)^{-\alpha_u}\widetilde{\rho}_f)r_*(u)^{2\alpha_u}\Gamma_*(u)^{(1)}\\
&+\ ((1-\alpha_u)r_*(u)^{-\alpha_u}+\alpha_u r_*(u)^{-\alpha_u-1}\widetilde{\rho}_f)r_f^{(1)}(u)r_*(u)^{\alpha_u}\Big)\\
&+\ ((\alpha_u+\xi-1))^{-1}r_*(u)^{\alpha_u+\xi+1}\hat{\rho}^{(1)}_{r_f(u)}\Big((\alpha_u+\xi+1)r_*(u)^{-1}r_f^{(1)}(u)r_*(u)^{-\alpha_u}\ +\ \Gamma_*(u)^{(1)}\Big)\\
&\times\Delta^*_f(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*))\\
&=\ ((\alpha_u+\xi-1))^{-1}r_*(u)^{\xi+1}\Delta^*_f\\
&\times\left(\Xi_{-f}w^*_f R^\xi_* r_*(u)^{-\xi-1}\left((\xi+\alpha_u)\Big((\xi+\alpha_u+1)r_*(u)^{-1}r_f^{(1)}+r_*(u)^{\alpha_u}\Gamma_*(u)^{(1)}\Big)\ +\ \alpha_u w^*_f\right)-\ Q^*_f(u)\right)\\
&\times(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*))\\
&+\ \frac{\alpha_u w^*_f}{\alpha_u+\xi-1}\Big(-(r_*(u)^{1-\alpha_u}-r_*(u)^{-\alpha_u}\widetilde{\rho}_f)r_*(u)^{2\alpha_u}\Gamma_*(u)^{(1)}\\
&+\ ((1-\alpha_u)r_*(u)^{-\alpha_u}+\alpha_u r_*(u)^{-\alpha_u-1}\widetilde{\rho}_f)r_f^{(1)}(u)r_*(u)^{\alpha_u}\Big)\\
&+\ ((\alpha_u+\xi-1))^{-1}r_*(u)^{\alpha_u+\xi+1}\\
&\times\Big(-\Xi_{-f}w^*_f R^\xi_*(\xi+\alpha_u)r_*(u)^{-\xi-1}\Big)\Big((\alpha_u+\xi+1)r_*(u)^{-1}r_f^{(1)}(u)r_*(u)^{-\alpha_u}\ +\ \Gamma_*(u)^{(1)}\Big)\\
&\times\Delta^*_f(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*))\\
&=\ -((\alpha_u+\xi-1))^{-1}r_*(u)^{\xi+1}\Delta^*_f Q^*_f(u)\ +\ (\alpha_u+\xi-1))^{-1}r_*(u)^{\xi+1}\Delta^*_f\Xi_{-f}w^*_f R^\xi_* r_*(u)^{-\xi-1}\alpha_u w^*_f\\
&+\ \Omega_1 r_f^{(1)}\ +\ \Omega_2\Gamma_*(u)^{(1)}\,.
\end{aligned}
$$

Here, we have defined

$$
\begin{aligned}
\Omega_1 \;=\; & ((\alpha_u+\xi-1))^{-1} r_*(u)^{\xi+1}\Delta_f^* \\
& \times \left(\Xi_{-f} w_f^* R_*^{\xi} r_*(u)^{-\xi-1}\left((\xi+\alpha_u)\Big((\xi+\alpha_u+1) r_*(u)^{-1}\Big)\right)\right) \\
& \times \left(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*)\right) \\
& + \; \frac{\alpha_u w_f^*}{\alpha_u+\xi-1}\Big(((1-\alpha_u) r_*(u)^{-\alpha_u}+\alpha_u r_*(u)^{-\alpha_u-1}\widetilde{\rho}_f) r_*(u)^{\alpha_u}\Big) \\
& + \; ((\alpha_u+\xi-1))^{-1} r_*(u)^{\alpha_u+\xi+1} \\
& \times \Big(-\Xi_{-f} w_f^* R_*^{\xi}(\xi+\alpha_u) r_*(u)^{-\xi-1}\Big)\Big((\alpha_u+\xi+1) r_*(u)^{-1} r_*(u)^{-\alpha_u}\Big)\Delta_f^*(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*)) \\
& = \; \Xi_{-f} w_f^* \Delta_f^*(a_*(f)+2\lambda_*(f)(\hat{\rho}(r_f)-\rho_*)) R_*^{\xi} r_*(u)^{-1}\Bigg(((\alpha_u+\xi-1))^{-1}\Bigg(\\
& \left((\xi+\alpha_u)\Big((\xi+\alpha_u+1)\Big)\right)\Bigg) \\
& + \; ((\alpha_u+\xi-1))^{-1} \\
& \times \Big(-(\xi+\alpha_u)\Big)\Big((\alpha_u+\xi+1)\Big)\Bigg) \\
& + \; \frac{\alpha_u w_f^*}{\alpha_u+\xi-1}\Big(((1-\alpha_u) r_*(u)^{-\alpha_u}+\alpha_u r_*(u)^{-\alpha_u-1}\widetilde{\rho}_f) r_*(u)^{\alpha_u}\Big) \\
& = \; \frac{\alpha_u w_f^*}{\alpha_u+\xi-1}\Big(((1-\alpha_u) r_*(u)^{-\alpha_u}+\alpha_u r_*(u)^{-\alpha_u-1}\widetilde{\rho}_f) r_*(u)^{\alpha_u}\Big)
\end{aligned}
$$

and

$$
\begin{aligned}
\Omega_2 \ &= \ ((\alpha_u + \xi - 1))^{-1} r_*(u)^{\xi+1} \Delta_f^* \\
&\times \left(\Xi_{-f} w_f^* R_*^\xi r_*(u)^{-\xi-1} \left((\xi + \alpha_u) \Big(r_*(u)^{\alpha_u} \Big) \right) \right) \\
&\times (a_*(f) + 2\lambda_*(f)(\hat{\rho}(r_f) - \rho_*)) \\
&+ \ \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} \Big(- (r_*(u)^{1-\alpha_u} - r_*(u)^{-\alpha_u} \widetilde{\rho}_f) r_*(u)^{2\alpha_u} \Big) \\
&+ \ ((\alpha_u + \xi - 1))^{-1} r_*(u)^{\alpha_u + \xi + 1} \\
&\times \Big(- \Xi_{-f} w_f^* R_*^\xi (\xi + \alpha_u) r_*(u)^{-\xi-1} \Big) \\
&\times \Delta_f^* (a_*(f) + 2\lambda_*(f)(\hat{\rho}(r_f) - \rho_*)) \\
&= \ \Delta_f^* (a_*(f) + 2\lambda_*(f)(\hat{\rho}(r_f) - \rho_*)) \left(((\alpha_u + \xi - 1))^{-1} r_*(u)^{\xi+1} \left(\Xi_{-f} w_f^* R_*^\xi r_*(u)^{-\xi-1} \left((\xi + \alpha_u) \Big(r_*(u)^{\alpha_u} \Big) \right) \right) \right. \\
&+ ((\alpha_u + \xi - 1))^{-1} r_*(u)^{\alpha_u + \xi + 1} \\
&\times \left. \Big(- \Xi_{-f} w_f^* R_*^\xi (\xi + \alpha_u) r_*(u)^{-\xi-1} \Big) \right) \\
&+ \ \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} \Big(- (r_*(u)^{1-\alpha_u} - r_*(u)^{-\alpha_u} \widetilde{\rho}_f) r_*(u)^{2\alpha_u} \Big) \\
&= \ \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} \Big(- (r_*(u)^{1} - \widetilde{\rho}_f) r_*(u)^{\alpha_u} \Big) .
\end{aligned}
$$

Summarizing, we get

$$
r_f(u) \ = \ \frac{\alpha_u + \xi}{\alpha_u + \xi - 1} \widetilde{\rho}_f \ + \ F^{-1} r_f(u)^{(1)} \ + \ F^{-2} r_f(u)^{(2)} \ + \ O(F^{-2})
$$

with

$$
r_f(u)^{(1)} \ = \ \frac{\alpha_u w_f^*}{\alpha_u + \xi - 1} (r_*(u) - \widetilde{\rho}_f) - \frac{\xi + \alpha_u}{\xi + \alpha_u - 1} w_f^* \Xi_{-f} R_*^\xi \Delta_f^* (a_*(f) + 2\lambda_*(f)(\hat{\rho}(r_f) - \rho_*))
$$

and

$$
\Gamma_*(b)^{(1)} \ = \ -\alpha_b r_*(b)^{-1-\alpha_b} E[r_f^{(1)}(b)]
$$

and

$$
\begin{aligned}
&r_f(u)^{(2)} \\
&= \; -((\alpha_u+\xi-1))^{-1} r_*(u)^{\xi+1} \Delta_f^* Q_f^*(u) \; + \; (\alpha_u+\xi-1))^{-1} r_*(u)^{\xi+1} \Delta_f^* \Xi_{-f} w_f^* R_*^{\xi} r_*(u)^{-\xi-1} \alpha_u w_f^* \\
&+ \; \Omega_1 r_f^{(1)} \; + \; \Omega_2 \Gamma_*(u)^{(1)} \, .
\end{aligned}
$$

Q.E.D.

Proof of Proposition 3.5. In equilibrium,

$$
\hat\rho \; = \; \hat\rho^* \; + \; \hat\rho^{(1)} F^{-1} \; + \; \hat\rho^{(2)} F^{-2} \; + \; O(F^{-3}) \, ,
$$

where

$$
\hat\rho^* \; = \; \rho_* \; + \; \frac{S - a - \sum_f a_*(f) \Delta_f(0)}{\lambda + \sum_f \lambda_*(f) \Delta(0)}
$$

is the level of rates absent market power, where

$$
\Delta_f(0) \; = \; \left(d_f - \sum_b (R_*/r_f(b))^{\xi} \frac{r_f(b)^{-\alpha_b} w_f}{\Gamma_*(b)} \right) \, .
$$

Similarly,

$$
\widetilde\rho_f \; = \; \widetilde\rho_f^* \; + \; \widetilde\rho_f^{(1)} F^{-1} \; + \; \widetilde\rho_f^{(2)} F^{-1} \; + \; O(F^{-3})
$$

where

$$
\widetilde\rho_f^* \; = \; \Big((\hat\rho^* - \rho_*)(a_*(f) + \lambda_*(f)(\hat\rho^* - \rho_*)) + \rho_* \Big)
$$

and

$$
\begin{aligned}
\widetilde{\rho}_f \;&=\; \Big((\hat{\rho}^* \;+\; \hat{\rho}^{(1)}F^{-1} \;+\; \hat{\rho}^{(2)}F^{-2} - \rho_*)(a_*(f) + \lambda_*(f)(\hat{\rho}^* \;+\; \hat{\rho}^{(1)}F^{-1} \;+\; \hat{\rho}^{(2)}F^{-2} - \rho_*)) + \rho_*\Big) \\
&=\; \rho_* \;+\; (\hat{\rho}^* \;+\; \hat{\rho}^{(1)}F^{-1} \;+\; \hat{\rho}^{(2)}F^{-2} - \rho_*)a_*(f) \\
&+\; (\hat{\rho}^* \;+\; \hat{\rho}^{(1)}F^{-1} \;+\; \hat{\rho}^{(2)}F^{-2} - \rho_*)^2\lambda_*(f) \\
&=\; \rho_* \;+\; (\hat{\rho}^* \;+\; \hat{\rho}^{(1)}F^{-1} \;+\; \hat{\rho}^{(2)}F^{-2} - \rho_*)a_*(f) \\
&+\; \Big((\hat{\rho}^* - \rho_*)^2 \;+\; 2(\hat{\rho}^* - \rho_*)\hat{\rho}^{(1)}F^{-1} \;+\; 2(\hat{\rho}^* - \rho_*)\hat{\rho}^{(2)}F^{-2} + (\hat{\rho}^{(1)})^2F^{-2}\Big)\lambda_*(f) \\
&=\; \widetilde{\rho}^*_f \;+\; 2(\hat{\rho}^* - \rho_*)\hat{\rho}^{(1)}\lambda_*(f)F^{-1} \;+\; (2(\hat{\rho}^* - \rho_*)\hat{\rho}^{(2)} + (\hat{\rho}^{(1)})^2)\lambda_*(f)F^{-2} \;+\; O(F^{-3})
\end{aligned} \tag{20}
$$

Therefore,

$$\sum_f a_*(f)\Delta_f \;=\; E[a_f d_f^*] - \sum_b E\left[a_f (R_*/r_f(b))^{\xi} \frac{r_f(b)^{-\alpha_b}}{\Gamma_*(b)}\right]$$

$$= \; E[a_f d_f^*] - \sum_b E\Bigg[a_f R_*^{\xi}$$

$$\times \frac{r_*(b)^{-\alpha_b-\xi} - (\alpha_b+\xi) r_*(b)^{-\alpha_b-\xi-1}(r_f(b)^{(1)}F^{-1} + r_f(b)^{(2)}F^{-2}) + 0.5(\alpha_b+\xi)(\alpha_b+\xi+1)(r_f(b)^{(1)})^2 F^{-2}}{r_*(b)^{-\alpha_b} + \Gamma_*(b)^{(1)}F^{-1} + \Gamma_*(b)^{(2)}F^{-2}}\Bigg]$$

$$= \; E[a_f d_f^*] - \sum_b E\Bigg[a_f R_*^{\xi}$$

$$\times \Bigg(r_*(b)^{-\alpha_b-\xi} - (\alpha_b+\xi) r_*(b)^{-\alpha_b-\xi-1}(r_f(b)^{(1)}F^{-1} + r_f(b)^{(2)}F^{-2})$$

$$+ 0.5(\alpha_b+\xi)(\alpha_b+\xi+1) r_*(b)^{-\alpha_b-\xi-1}(r_f(b)^{(1)})^2 F^{-2}\Bigg)$$

$$\times r_*(b)^{\alpha_b}\Bigg(1 - r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)}F^{-1} - r_*(b)^{\alpha_b}\Gamma_*(b)^{(2)}F^{-2} + (r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)}F^{-1})^2\Bigg)\Bigg]$$

$$= \; E[a_f d_f^*] - \sum_b E\Bigg[a_f R_*^{\xi}$$

$$\Bigg(r_*(b)^{-\xi} \;+\; F^{-1}\Big(- r_*(b)^{-\xi} r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)} - (\alpha_b+\xi) r_*(b)^{-\xi-1} r_f(b)^{(1)}\Big)$$

$$+ \; F^{-2}\Big(r_*(b)^{-\xi}(-r_*(b)^{\alpha_b}\Gamma_*(b)^{(2)} + (r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)})^2) + (\alpha_b+\xi) r_*(b)^{-\xi-1} r_f(b)^{(1)} r_*(b)^{\alpha_b}\Gamma_*(b)^{(1)}$$

$$+ 0.5(\alpha_b+\xi)(\alpha_b+\xi+1) r_*(b)^{-\xi-2}(r_f(b)^{(1)})^2 - (\alpha_b+\xi) r_*(b)^{-\xi-1} r_f(b)^{(2)}\Big)\Bigg)\Bigg]$$

By assumption, all banks are identical, and hence we can rewrite it as

$$\sum_f a_*(f)\Delta_f \;=\; A \;+\; (A_{1,1} r_f^{(1)} \;+\; A_{1,2}\Gamma_*^{(1)})F^{-1}$$

$$+ \;(B_{2,0}(r_f^{(1)})^2 \;+\; B_{1,1} r_f^{(1)}\Gamma_*^{(1)} \;+\; B_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2} \;+\; O(F^{-3}),$$

and similarly

$$\sum_f \lambda_*(f)\Delta_f \;=\; C \;+\; (C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})F^{-1}$$

$$+\;(D_{2,0}(r_f^{(1)})^2 \;+\; D_{1,1}r_f^{(1)}\Gamma_*^{(1)} \;+\; D_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2} \;+\; O(F^{-3})\,,$$

so that

$$\left(\lambda + \sum_f \lambda_*(f)\Delta_f\right)^{-1} \;=\; \Bigg(\lambda + \bar{\lambda} \;+\; (C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})F^{-1}$$

$$+\;(D_{2,0}(r_f^{(1)})^2 \;+\; D_{1,1}r_f^{(1)}\Gamma_*^{(1)} \;+\; D_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2} \;+\; O(F^{-3})\Bigg)^{-1}$$

$$=\;(\lambda+\bar{\lambda})^{-1}\Bigg(1 - (\lambda+\bar{\lambda})^{-1}(C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})F^{-1}$$

$$-\,(\lambda+\bar{\lambda})^{-1}\Big((D_{2,0}(r_f^{(1)})^2 \;+\; D_{1,1}r_f^{(1)}\Gamma_*^{(1)} \;+\; D_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2}\Big)$$

$$+\;(\lambda+\bar{\lambda})^{-2}(C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})^2F^{-2}\Bigg)$$

To solve for $\hat{\rho}^{(1)}, \hat{\rho}^{(2)}$ we proceed to solving the market clearing equation

$$
\begin{aligned}
&\hat{\rho}^* \;+\; \hat{\rho}^{(1)}F^{-1} \;+\; \hat{\rho}^{(2)}F^{-1} \;+\; O(F^{-3}) \\
&= \hat{\rho} \;=\; \rho_* \;+\; \frac{S - a - \sum_f a_*(f)\Delta_f}{\lambda + \sum_f \lambda_*(f)\Delta_f} \\
&=\; \rho_* \;+\; \frac{S - a - \sum_f a_*(f)\left(d_f - \sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b} w_f}{\Gamma_*(b)}\right)}{\lambda + \sum_f \lambda_*(f)\left(d_f - \sum_b (R_*/r_f(b))^{\xi}\frac{r_f(b)^{-\alpha_b} w_f}{\Gamma_*(b)}\right)} \\
&=\; \Bigg(S - a - A \;-\; (A_{1,1}r_f^{(1)} \;+\; A_{1,2}\Gamma_*^{(1)})F^{-1} \\
&-\; (B_{2,0}(r_f^{(1)})^2 \;+\; B_{1,1}r_f^{(1)}\Gamma_*^{(1)} \;+\; B_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2}\Bigg) \\
&\times (\lambda + \bar{\lambda})^{-1}\Bigg(1 - (\lambda + \bar{\lambda})^{-1}(C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})F^{-1} \\
&- (\lambda + \bar{\lambda})^{-1}\Big((D_{2,0}(r_f^{(1)})^2 \;+\; D_{1,1}r_f^{(1)}\Gamma_*^{(1)} \;+\; D_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2}\Big) \\
&+\; (\lambda + \bar{\lambda})^{-2}(C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})^2 F^{-2}\Bigg) \\
&=\; (S - a - A)(\lambda + \bar{\lambda})^{-1} \;-\; (\lambda + \bar{\lambda})^{-1}(A_{1,1}r_f^{(1)} \;+\; A_{1,2}\Gamma_*^{(1)})F^{-1} \\
&-\; (S - a - A)(\lambda + \bar{\lambda})^{-2}(C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})F^{-1} \\
&+\; (\lambda + \bar{\lambda})^{-2}(A_{1,1}r_f^{(1)} \;+\; A_{1,2}\Gamma_*^{(1)})(C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})F^{-2} \\
&- (S - a - A)(\lambda + \bar{\lambda})^{-2}\Big((D_{2,0}(r_f^{(1)})^2 \;+\; D_{1,1}r_f^{(1)}\Gamma_*^{(1)} \;+\; D_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2}\Big) \\
&+\; (S - a - A)(\lambda + \bar{\lambda})^{-2}(C_{1,1}r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})^2 F^{-2} \\
&- (\lambda + \bar{\lambda})^{-1}(B_{2,0}(r_f^{(1)})^2 \;+\; B_{1,1}r_f^{(1)}\Gamma_*^{(1)} \;+\; B_{0,2}(\Gamma_*^{(1)})^2 + a_1 r_f^{(2)} + a_2\Gamma_*^{(2)})F^{-2} \;+\; O(F^{-3})\,.
\end{aligned}
$$

We now summarize these equations for the first-order approximation:

$$\hat{\rho}^{(1)} \;=\; -(\lambda+\bar{\lambda})^{-1}(A_{1,1} r_f^{(1)} \;+\; A_{1,2}\Gamma_*^{(1)})$$
$$-\;(S-a-A)(\lambda+\bar{\lambda})^{-2}(C_{1,1} r_f^{(1)} \;+\; C_{1,2}\Gamma_*^{(1)})$$
$$r_f(u)^{(1)} \;=\; \frac{\alpha_u+\xi}{\alpha_u+\xi-1}\widetilde{\rho}_f^{(1)} \;+\; \frac{\alpha_u w_f^*}{\alpha_u+\xi-1}\left(1+\frac{\alpha_u w_f^*}{\alpha_u+\xi-1}F^{-1}\right)(r_*(u)-\widetilde{\rho}_f) \;-\; w_f^*\Lambda_f\Delta_f^*$$
$$\widetilde{\rho}_f \;=\; \widetilde{\rho}_f^* \;+\; 2(\hat{\rho}^*-\rho_*)\hat{\rho}^{(1)}\lambda_*(f)F^{-1} \;+\; (2(\hat{\rho}^*-\rho_*)\hat{\rho}^{(2)}+(\hat{\rho}^{(1)})^2)\lambda_*(f)F^{-2} \;+\; O(F^{-3})$$
$$\Gamma_*(b)^{(1)} \;=\; -\alpha_b r_*(b)^{-1-\alpha_b} E[r_f^{(1)}(b)]$$
$$A_{1,1} \;=\; (\alpha+\xi) r_*^{-\xi-1} R_*^{\xi} E[a_f]$$
$$A_{1,2} \;=\; R_*^{\xi} r_*^{-\xi} r_*^{\alpha} E[a_f]$$
$$C_{1,1} \;=\; (\alpha+\xi) r_*^{-\xi-1} R_*^{\xi} E[\lambda_f]$$
$$C_{1,2} \;=\; R_*^{\xi} r_*^{-\xi} r_*^{\alpha} E[\lambda_f]$$

where we have used (20) and (9).

Since we assume that all banks are homogeneous, we can omit the dependence on u, b.

Let also

$$\psi_f \;=\; (\alpha+\xi) r_*^{-\xi-1} R_*^{\xi}((\lambda+\bar{\lambda})^{-1} a_f + (S-a-A)(\lambda+\bar{\lambda})^{-2}\lambda_f)$$

Thus, we end up with the first point system

$$\hat{\rho}^{(1)} \;=\; -E\left[\left(\psi_f-\frac{\alpha}{\alpha+\xi}E[\psi_f]\right)\left(\frac{\alpha+\xi}{\alpha+\xi-1}\Big(2(\hat{\rho}^*-\rho_*)\hat{\rho}^{(1)}\lambda_*(f)\Big) \;+\; \frac{\alpha w_f^*}{\alpha+\xi-1}(r_*-\widetilde{\rho}_f) \;-\; w_f^*\Lambda_f\Delta_f^*\right)\right]$$

so that

$$\hat{\rho}^{(1)} \;=\; \frac{-E\left[\left(\psi_f-\frac{\alpha}{\alpha+\xi}E[\psi_f]\right)\left(\frac{\alpha w_f^*}{\alpha+\xi-1}(r_*-\widetilde{\rho}_f) \;-\; w_f^*\Lambda_f\Delta_f^*\right)\right]}{1+E\left[\left(\psi_f-\frac{\alpha}{\alpha+\xi}E[\psi_f]\right)\left(\frac{\alpha+\xi}{\alpha+\xi-1}\Big(2(\hat{\rho}^*-\rho_*)\lambda_*(f)\Big)\right)\right]}$$

Q.E.D.

Proof of Proposition 4.1. Equilibrium price impacts satisfy

$$\gamma_f \;=\; \frac{1}{\lambda + b - (\gamma_f + \beta_f)^{-1}}$$

where

$$b \;=\; \sum_f (\gamma_f + b_f)^{-1}\,,$$

and the first claims follow by a direct calculation. To prove asymptotics, we note that, with $b = b_0 + Fb_1 + O(F^{-1})$, we get

$$\begin{aligned}
& 2 + \beta_f(\lambda + b) + \sqrt{(\beta_f(\lambda + b))^2 + 4} \\
&\approx\; 2 + \beta_f(\lambda + b_0 + b_1 F) + \beta_f b_1 F(1 + (\lambda + b_0)/(b_1 F)) \\
&=\; 2\beta_f b_1 F(1 + \frac{1 + \beta_f(\lambda + b_0)}{b_1 \beta_f F})
\end{aligned}$$

so that

$$\begin{aligned}
& \sum_f (2\beta_f b_1 F)^{-1}(1 - \frac{1 + \beta_f(\lambda + b_0)}{\beta_f b_1 F}) \\
&=\; 0.5\frac{b_0 + b_1 F}{\lambda + b_0 + b_1 F} \;=\; 0.5(b_1 F)^{-1}(b_0 + b_1 F)(1 - \frac{(\lambda + b_0)}{b_1 F}) \\
&=\; 0.5(1 - \lambda/(b_1 F))
\end{aligned}$$

Equating the coefficients gives

$$b_1 \;=\; E[\beta_f^{-1}/w_f^*]$$

and

$$b_0 \;=\; -\sum_f \beta_f^{-2}/(Fb_1)$$

Q.E.D.