Portfolio Choice and Settlement Frictions: a Theory of Endogenous Convenience Yields*

Javier Bianchi† and Saki Bigio‡

August 2025

Abstract

We study settlement frictions, stemming from the need to finance negative balances via an over-the-counter (OTC) market. We derive an endogenous convenience yield in closed form, and show how it can be embedded in a canonical portfolio problem. Using this framework, we examine how shifts in settlement frictions influence liquidity premia, the volume of overnight funding, the dispersion of market rates, and optimal portfolio allocations. On the normative front, we show that in the competitive equilibrium, investors may either over- or under-invest in liquid assets; moreover, higher risk aversion and tighter aggregate liquidity both increase the likelihood of under-accumulation.

Keywords: Asset Pricing, Portfolio Theory, OTC markets, convenience yields

JEL Classification: E51,E52, E58, G12, G21

*We thank Andy Atkeson, David Baqaee, Ricardo Lagos, and Pierre-Olivier Weill for helpful comments. Ali Haider Ismail, Jorge Mondragon, and Luis Yépez provided excellent research assistance. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

†Minneapolis Fed and NBER, email `javier.i.bianchi@gmail.com`

‡Department of Economics, UCLA and NBER, email `sbigio@econ.ucla.edu`

1. Introduction

The empirical finance literature has established that certain assets command substantial convenience yields—premia not accounted by their cash flows alone. The pattern appears across short-term assets such as U.S. Treasuries (Krishnamurthy and Vissing-Jørgensen, 2012), cash-like instruments (Nagel, 2016), and short-term dollar assets in international markets (Jiang, Krishnamurthy and Lustig, 2021; Engel and Wu, 2023). Convenience yields vary substantially across assets and over time, with important implications for the transmission of monetary policy, fiscal capacity, and international capital flows. While convenience yields have become central to understanding asset prices and macroeconomic policy, their theoretical foundations remain incomplete. This paper develops a tractable microfoundation for convenience yields arising from trading frictions in over-the-counter (OTC) financial markets and incorporates it into canonical portfolio theory.

In our theory, investors are subject to payoff risk—stemming from fundamental variations in returns—and liquidity risk. Liquidity risk emerges because cash flows are unpredictable, and financial imbalances must be settled in an OTC market. Examples of financial positions that expose investors to liquidity risk include deposit withdrawals for banks, margin calls for hedge funds, or claim payouts for insurance companies. To meet these obligations, investors hold buffers of liquid assets that can be deployed for settlement. When liquidity needs exceed available buffers, investors must borrow cash in a frictional OTC market. The interaction between portfolio-induced cash needs and OTC market frictions generates an endogenous convenience yield that lowers the returns on liquid assets. Moreover, to the extent that cash needs are correlated with asset returns, this creates an additional liquidity risk premium in excess of the conventional risk premium. Thus, convenience yields reflect both a first-order effect and the covariance between cash needs and asset payoffs. Crucially, because these yields endogenously depend on investors' portfolios and OTC market conditions, they are not invariant to policy and vary with market tightness and trading efficiency of the OTC market.

Our analysis yields a tractable derivation of convenience yields that can be easily embedded in standard portfolio problems. To do so, we build on the sequential OTC framework of Afonso and Lagos (2015b), with two key innovations. First, instead of taking settlement positions as given, we endogenize them by modeling the portfolio choices of investors. Second, we model portfolio managers as large institutions who delegate settlement trades to many small traders, following Shi (1997) and the OTC model of Atkeson, Eisfeldt and Weill (2015). Taking limits as trader size vanishes yields closed-form characterizations of the entire equilibrium path of trades and rates.

Settlement frictions manifest as a piecewise linear convenience yield function $\chi(s)$ that maps settlement positions s for each asset into an additional portfolio return. This function is kinked at zero, reflecting the asymmetry between borrowing costs and lending returns that is characteristic of the OTC market. Its slopes are precisely the expected lending return for investors with surplus and the borrowing costs for those with deficits arising from the OTC market. These convenience-yield coefficients depend on only three objects: a bargaining parameter, an initial market tightness (the ratio of aggregate deficits to surpluses) the depends on portfolio choices, and the terminal market tightness (which encodes the trading in the OTC market and, thereby, the matching technology). The asymmetry resulting from the kink induces risk-averse behavior even under risk-neutral preferences and implies that assets that generate volatile cash needs command higher convenience yields. Moreover, when settlements correlate with asset returns—as when margin calls intensify in downturns—an additional liquidity risk premium emerges. Thus, convenience yields reflect both pure OTC frictions and the interaction between liquidity risk and return risk.

Our analysis reveals several important properties linking OTC markets to convenience yields. First, the choice of matching technology fundamentally shapes market dynamics: under Cobb-Douglas matching, one side of the market can be depleted in finite time, causing trading to cease and convenience yields to reach their bounds; under Leontief matching, the short side always trades, preventing a complete market freeze. This distinction helps explain a puzzling observation in interbank markets where minimal trading persists even when rates reach the floor (e.g., see Lopez-Salido and Vissing-Jorgenson, 2023; Afonso, Giannone, Spada and Williams, 2023; Lagos and Navarro, 2023).[1] Second, we establish that convenience yields satisfy time dilation (the passage of time is mathematically equivalent to reduced matching efficiency) and symmetry (reversing market tightness and swapping bargaining powers yields identical surplus extraction). Third, our comparative statics yield sharp predictions: market tightness unambiguously increases convenience yields by steepening the convenience yield function, whereas matching efficiency has non-monotonic effects. The intuition behind this non-monotonicity—perhaps a surprising result—is that higher matching efficiency benefits the short side of the market, raising convenience yields when liquidity is scarce but potentially lowering them when it is abundant. Furthermore, our closed-form solutions facilitate the identification of structural OTC market parameters and shocks from observable data for trading volumes and spreads.

Finally, we show how the interaction between portfolio choice and settlement frictions generates a pecuniary externality that induces inefficient investment in liquid assets. Individual investors, when choosing portfolios, do not internalize how their cash needs affect aggregate market tightness and, thus, the liquidity yields faced by all market participants. This creates a wedge between private and social returns to holding liquid assets. We characterize conditions under which competitive equilibria feature over- or under-investment in liquidity. We show that under risk neutrality,

[1]Our analysis also shows that only certain matching functions, those with infinite rate of decay, can generate near-zero convenience yields with positive trading volumes.

investment in liquid assets is insufficient when the marginal impact of tightness on borrowing costs exceeds its impact on lending returns. This occurs in turn when aggregate surplus exceeds deficits. With risk aversion, the inefficiency is amplified as investors fail to account for how their portfolios affect others' liquidity risk. Strikingly, under balanced markets (equal settlement instrument deficits as surpluses) with symmetric shocks and Cobb-Douglas matching, the competitive equilibrium is constrained efficient under risk neutrality—but exhibits under-provision of liquidity under risk aversion. The findings have implications for liquidity regulation, suggesting that optimal policy depends critically on market conditions.

Literature Review. Our paper is related to a large literature on portfolio choice and asset pricing with liquidity frictions, such as portfolio constraints or transaction costs. Important examples in this literature include Constantinides (1986); Basak and Cuoco (1998); Vayanos and Vila (1999); Constantinides and Duffie (1996); Krueger and Lustig (2010); Holmstrom and Tirole (2001); Acharya and Pedersen (2005); Lagos (2010). Typically, these studies model liquidity frictions as exogenous. Our goal in this paper is to develop a tractable framework where liquidity premia emerge endogenously and explore the implications for asset portfolios.

Following Duffie, Garleanu and Pedersen (2005), a burgeoning literature on OTC markets has studied environments where assets are traded in the presence of search frictions.[2] This literature has identified how features of the trading environment, such as the speed of transactions and the heterogeneity in the motives for trade, affect trading volumes and impact liquidity premia. While the literature began with strong restrictions on portfolio holdings, namely binary holdings, work by Lagos and Rocheteau (2009) allows for arbitrary portfolio holdings, bringing this literature closer to standard portfolio theory. Hugonnier, Lester and Weill (2022) consider heterogeneity in asset selling speeds. Uslu (2019) introduce risk-averse behavior into this class of models. In these studies, trading speeds affect asset values because time is discounted, a feature that depresses the option value of selling assets when gains from trade emerge. He and Milbradt (2014) studies how liquidity in an OTC secondary market for bonds affects the issuer's default risk. Silva, Passadore and Kargar (2023) study portfolio problems that explicitly take into account trading times when agents want to modify their portfolios.[3] In the language of Hugonnier, Lester and Weill (2025), these papers study a semi-centralized setup, while we study a purely decentralized setup. Moreover, we focus on OTC frictions in the market for borrowing and lending as opposed to the sale of assets.

As stated above, the sequential OTC market is inherited from Afonso and Lagos (2015b). In a

[2]This literature has run in parallel to the money search literature, pioneered by Kiyotaki and Wright (1993). See Williamson and Wright (2010); Lagos, Rocheteau and Wright (2017) for recent surveys.

[3]In our model, time discounting plays no role, as trading occurs within a single period, although the sequence of trades matters because it affects the terms of trade when two investors match.

further application, Afonso and Lagos (2015a) obtains some closed-form formulas for volumes and OTC rates that mirror the ones found here. Afonso and Lagos (2015a) restrict positions to binary holdings and an increasing returns to scale matching function. Here, we show that, by working with a large number of traders within each investor—as in Shi (1997) and Atkeson, Eisfeldt and Weill (2015)—the OTC block can be readily embedded into portfolio theory. We furthermore characterize outcomes for a more general class of matching processes. In addition, we provide comparative statics with respect to market tightness and matching efficiency, and show how these outcomes can be mapped to observable empirical moments.

Our normative analysis is related to a broad literature on the welfare properties of competitive equilibrium with financial frictions, in particular, those studies analyzing the efficiency of risk-taking decisions. One branch of the literature that is more closely related focuses on over or under-investment in liquid assets (Jacklin, 1987; Bhattacharya and Gale, 1987; Farhi, Golosov and Tsyvinski, 2009; Yared, 2013; Geanakoplos and Walsh, 2017). These studies consider a Walrasian interbank market where the risk-free rate affects the degree of enforcement and risk sharing. In contrast, we consider a setting with an OTC market where a pecuniary externality emerges from congestion in the interbank market. The externality we study is related to other congestion externalities in the search and matching literature. In particular, Uslu (2019) identifies an externality where fast investors can capture a private transaction surplus larger than their contribution to surplus creation [4] In addition, Wong and Zhang (2023) shows that the degree of search and intermediation is inefficient in the OTC market where search intensity is endogenous. Different from these contributions, our approach compares the portfolio choices of individual investors vis-à-vis a social planner that takes as given the financial arrangements in the OTC market. In this respect, our analysis is closer to Arseneau and Rappoport (2017), who studies inefficient liquidity provision in a model where firms issue long-term bonds that are retraded by investors in an imperfectly liquid secondary market.[5]

Finally, we emphasize that while our earlier work (Bianchi and Bigio, 2022) employs the convenience yield function derived here to analyze the transmission and implementation of monetary policy through the banking system, the present paper develops its theoretical foundations.[6] Other applications include Arce, Nuño, Thaler and Thomas (2019) and Bigio and Sannikov (2019), who study optimal reserve policy; Bianchi, Bigio and Engel (2023), who examine exchange rate determination in an international context; Lopez-Salido and Vissing-Jorgenson (2023) who study quantita-

[4]In contrast, in a model with exogenous search effort, Afonso and Lagos (2015a,b) find that the competitive equilibrium coincides with the efficient allocation.

[5]A different strand of the literature analyzes pecuniary externalities that can result in excessive leverage (see e.g. Caballero and Krishnamurthy, 2001; Lorenzoni, 2008; Bianchi, 2011; Dávila and Korinek, 2018; Amador and Bianchi, 2024).

[6]An additional contribution relative to Bianchi and Bigio (2022) is the analysis of efficiency of portfolio choices.

tive tightening, and Bittner, Jamilov and Saidi (2025), who apply a similar framework with assortative matching to the German interbank market, De Fiore, Hoerova and Uhlig (2018) who study the role of collateral and interbank market freezes, and Ríos-Rull, Ritto, Takamura and Terajima (2024) who consider the role of capital requirements.[7]

Roadmap. The paper is organized as follows. Section 2 presents the environment. Section 3 presents the main theoretical results. Section 4 presents the applications and the normative analysis. Section 5 concludes.

2. Environment

We present an infinite-horizon model comprised of a unit mass of investors who take portfolio positions given exogenous returns. Assets and liabilities differ in their payoffs across states and, crucially, in their settlement risks, which gives rise to a liquidity-management problem. Investors are subject to idiosyncratic shocks, and trade overnight loans in an over-the-counter (OTC) market.

Asset structure. There is a collection of assets, $\{a_i\}$ where $i \in \mathbb{I}=\{1, 2, ..., I\}$, where an investor takes a positive position when long and a negative position when short. Assets are real and differ in their realized returns as well as their settlement properties. In addition, there is a special asset, m, that is used to settle all payoffs and represents "cash." Investors may hold a negative position in m during the period but must end each period with a strictly positive position.

2.1 The Portfolio and Balancing Stages

Each period consists of two stages: a *portfolio stage* and a *balancing stage*. In the portfolio stage, investors take portfolio positions, given possible asset returns and cash-flow shocks. In the balancing stage, investors are subject to cash-flow shocks and may face a shortage or surplus of cash. When the investor has a cash deficit, it can borrow cash in an OTC market or through a lender of last resort. In the case of banks, the lender of last resort can be a central bank, but more generally, it is a credit line from a cash provider. When it has a surplus, it can lend in the OTC market or hold cash and earn a short-term rate. We proceed next to describe in detail the two stages.

[7] In Piazzesi, Rogers and Schneider (2019) and Lenel, Piazzesi and Schneider (2019), settlement risks are used to explain the short-term rate puzzle and the determinacy of interest rate rules.

Portfolio Stage. In this first sub-stage, the individual investor makes portfolio holdings decisions. The investor enters the period with a given initial amount of wealth or equity e_t. Wealth is composed of a portfolio of assets, $\{a_t^i\}_{i\in\mathbb{I}}$, cash, m_t, loans borrowed from other investors (negative when loans are provided to other investors), f_t, and borrowing from the lender of last resort, w_t. That is,

$$e_t \equiv \sum_{i\in\mathbb{I}} R_t^i a_t^i + R_t^m m_t - \bar{R}_t^f f_t - R_t^w w_t, \tag{1}$$

where R denotes realized gross returns, and each return is indexed by X, the aggregate state. We assume that $R_t^w \geq R_t^m$, capturing a penalty for emergency borrowing.[8] The return $\bar{R}_t^f$ is a weighted average rate of OTC loans to be described below. This is the only return rate determined in equilibrium.

Starting from a given wealth level, the investor chooses the dividend—interpreted as an equity injection if negative—denoted by c_t, along with a new portfolio $\left\{\left\{\tilde{a}_{t+1}^i\right\}_{i\in\mathbb{I}}, \tilde{m}_{t+1}\right\}$. We use *'tilde'* to denote a portfolio variable chosen at the portfolio stage. The budget constraint is:

$$c_t + \sum_{i\in\mathbb{I}} \tilde{a}_{t+1}^i + \tilde{m}_{t+1} = e_t.$$

Balancing stage. Investors enter the balancing stage with the portfolio $\left\{\left\{\tilde{a}_{t+1}^i\right\}_{i\in\mathbb{I}}, \tilde{m}_{t+1}\right\}$. At the beginning of the balancing stage, the investor experiences cash-flow shocks $\{\omega_t^i\}$ to each of the assets in its portfolio

$$a_{t+1}^i = \tilde{a}_{t+1}^i(1+\omega_t^i). \tag{2}$$

where ω^i follows a C.D.F $\Phi^i(\cdot)$.

When $\tilde{a}_{t+1}^i\omega^i > 0$, this represents a *positive liquidity shock*; conversely, when $\tilde{a}_{t+1}^i\omega^i < 0$, it constitutes a *negative liquidity shock*. In the case where $\tilde{a}_t^i + 1 < 0$, a negative ω^i shock corresponds to a sudden withdrawal of liabilities, while a positive ω^i shock reflects a sudden inflow of liabilities. Conversely, when $\tilde{a}_t^i + 1 > 0$, a negative ω^i shock represents a decrease in assets (e.g., a loan prepayment or callable bond) and a positive ω^i shock as a need to invest more in a particular asset (e.g., a credit line). We cover some examples below in more detail.

Cash-flow shocks are settled with cash. Taking into account that the investor must end the period

[8] This assumption is consistent with monetary policy implementation, where the penalty rate is strictly above the interest on reserve balances.

with positive cash position, the investor's *cash surplus* after the realization of cash-flow shocks is:

$$s\left(\left\{\tilde{a}^i_{t+1}\right\}_{i\in\mathbb{I}}, \tilde{m}^i_{t+1}, \left\{\omega^i_t\right\}_{i\in\mathbb{I}}\right) = \underbrace{\tilde{m}^i_{t+1} - \sum_{i\in\mathbb{I}} \frac{R^i_{t+1}}{R^m_{t+1}} \omega^i_t \tilde{a}^i_{t+1}}_{\text{settlement balance}}. \tag{3}$$

The surplus s depends on the portfolio chosen in the first substage, $\left\{\left\{\tilde{a}^i_{t+1}\right\}_{i\in\mathbb{I}}, \tilde{m}_{t+1}\right\}$, and the shocks $\{\omega^i\}_{i\in\mathbb{I}}$. Notice that when $\tilde{a}^i_t + 1 > 0$, and $\omega^i_t > 0$ translates into an increase in assets settled with cash. If $\tilde{a}^i_t + 1 < 0$ the increase in liabilities corresponding to $\omega^i_t > 0$ is settled with an increase in cash. Equation (3) shows that if the investor faces sufficiently large negative cash-flow shocks, it may end up with a *cash deficit*. Implicit in this cash-flow accounting is the assumption that when settlement occurs, the return on each asset accrues to its holder, which is taken care of by the ratio of returns.

If an investor ends up with a cash deficit, it must either borrow in the OTC market or resort to the lender of last resort. Conversely, if it has a surplus, it may lend in the OTC market or retain the cash. At the end of the period, cash holdings are given by:

$$m_{t+1} = s\left(\left\{\tilde{a}^i_{t+1}\right\}_{i\in\mathbb{I}}, \tilde{m}_{t+1}, \left\{\omega^i_t\right\}_{i\in\mathbb{I}}\right) + f_{t+1} + w_{t+1} \geq 0. \tag{4}$$

The condition $m_{t+1} \geq 0$, combined with the assumption that $R^w \geq R^m$, implies that running a cash deficit is costly. Moreover, since investors may end up with different cash positions, there are gains from trade in the balancing stage. We assume that this money market is frictional, and operates as an *over-the-counter* (OTC) market—a natural assumption given that in practice, the money market is often decentralized: participants must search for counterparties with the opposite position.

Because this market is OTC, access to funding is not guaranteed. As a result, funding deficits may be only partially covered through the OTC market, with the remainder financed via the lender of last resort. Conversely, investors with a positive surplus may ultimately end up holding excess cash.

The outcome of the OTC market determines the amounts traded in the OTC market and borrowed from the lender of last resort, $\{f_t, w_t\}$, are determined by two endogenous probabilities $\left\{\Psi^+_t, \Psi^-_t\right\}$. In particular, given cash holdings $s_t \equiv s\left(\left\{\tilde{a}^i_{t+1}\right\}_{i\in\mathbb{I}}, \tilde{m}_{t+1}, \left\{\omega^i_t\right\}_{i\in\mathbb{I}}\right)$, we have that

$$f_t = \begin{cases} -\Psi^-_t s_t & \text{if } s_t \leq 0 \\ -\Psi^+_t s_t & \text{if } s_t > 0 \end{cases} \quad \text{and} \quad w_t = \begin{cases} -(1-\Psi^-_t) s_t & \text{if } s_t \leq 0 \\ 0 & \text{if } s_t > 0. \end{cases} \tag{5}$$

Substituting (5) into (4) the $t+1$ equity can be written as:

$$e_{t+1} \equiv \underbrace{\sum_{i\in\mathbb{I}} R^i_{t+1}\tilde{a}^i_{t+1} + R^m_{t+1}\tilde{m}_{t+1}}_{\text{portfolio return}} + \underbrace{\chi_{t+1}\left(s\left(\left\{\tilde{a}^i_{t+1}\right\}_{i\in\mathbb{I}}, \tilde{m}_{t+1}\right)\right)}_{\text{convenience yield}}, \tag{6}$$

where χ_{t+1} is the piecewise linear function

$$\chi_t(s) = \begin{cases} \chi^-_t s & \text{if } s \leq 0 \\ \chi^+_t s & \text{if } s > 0 \end{cases}, \tag{7}$$

where χ^- and χ^+ given by:

$$\chi^-_t = (\overline{R}^f_t - R^m_t)\Psi^-_t + (R^w_t - R^m_t)(1-\Psi^-_t), \qquad \chi^+_t = \Psi^+_t(\overline{R}^f_t - R^m_t). \tag{8}$$

The average interest rate of the OTC transactions $\overline{R}^f$ is also an outcome of the OTC market, as we will see below. Crucially, the convenience yield function $\chi(s)$ is asymmetric: settlement deficits generate steep marginal losses while surpluses generate lower marginal gains.

Equation (6) shows that the portfolio choice $\left\{\left\{\tilde{a}^i_{t+1}\right\}_{i\in\mathbb{I}}, \tilde{m}_{t+1}\right\}$ induces the direct portfolio returns, but also the indirect convenience yields associated with settlement frictions. The following section explains how the OTC market determines $\left\{\Psi^+_t, \Psi^-_t, \overline{R}^f_t\right\}$ and thus the convenience yields as an equilibrium outcome. Before we describe how the OTC market operates, we discuss several examples of assets that expose investors to liquidity risk in addition to return risk.

2.2 Interpretation and Asset-Class Examples

It is worth discussing some interpretations of the environment. If investors are banks, "cash" corresponds to central bank reserves, and emergency borrowing is the central bank's discount window. For non-bank financial institutions, however, cash often takes the form of bank deposits, and liquidity backstops are pre-negotiated arrangements, credit lines with banks rather than with the central bank. Similarly, the OTC market in our model may correspond to the secured interbank market in the case of banks, or to bilateral repo markets for non-bank institutions, abstracting in both cases from the collateral. In international settings, "cash" corresponds to deposits held in the U.S. correspondent banking system, liquidity support takes the form of contingent credit lines with global banks, and the OTC market is represented by short-term offshore funding markets such as the LIBOR market (see Bianchi et al., 2023).

We now illustrate how cash needs arise in specific institutional settings.

(i) *Funding Outflows:* A natural interpretation of our framework arises in the context of institutions exposed to unpredictable funding outflows. One example is a bank whose liabilities consist of demand deposits d, which are subject to withdrawal shocks and a reserve requirement ρ, as in Bianchi and Bigio (2022). The bank holds a portfolio composed of a fully liquid asset and a fully illiquid asset. This setting corresponds to a special case of our general framework, with $d = -\tilde{a}_t^1$, and the settlement position defined as:

$$s(d, m, \omega) \equiv m - d\left(\frac{R^d}{R^m}\omega - \rho(1+\omega)\right).$$

The same formulation applies to open-ended investment funds, where a redemption shock is the analog of a deposit withdrawal.

(ii) *Credit Lines:* A credit line specifies a multiple ω of an existing loan ℓ to be drawn upon the realization of a future state, at a pre-specified rate. The institution must then settle the exposure by delivering cash in the amount:

$$s(\ell, m, \omega) \equiv m - \omega\ell.$$

(iii) *Margin Calls:* A margin call requires that, in all states where the future asset return falls below a threshold k, the borrower must post additional collateral. If the loan amount is b, the required cash settlement is:

$$s(b, m) \equiv m - \max\left\{k - R^b_{t+1}, 0\right\} b.$$

Having provided context to the model, we move to explaining the workings of the OTC market.

3. OTC Market

In this section, we derive the equilibrium objects $\left\{\Psi_t^+, \Psi_t^-, \overline{R}_t^f\right\}$. We show how these objects are solely functions of market tightness, which we define below:

Definition 1. *Market tightness θ is defined as $\theta \equiv S^-/S^+$ where S^- and S^+ represent the sum of all negative and positive positions in the OTC market over the set of investors indexed by j:*

$$S^- = -\int_0^1 \min\left\{s^j, 0\right\} dj \quad \textit{and} \quad S^+ = \int_0^1 \max\left\{s^j, 0\right\} dj.$$

A higher tightness θ implies that there are relatively more investors with a cash deficit. The equilibrium outcome that $\left\{\Psi_t^+, \Psi_t^-, \overline{R}_t^f\right\}$ are only functions of θ results from the specific structure of the OTC market here. Notice that, unlike related OTC models such as Afonso and Lagos (2015b); Lagos and Navarro (2023), in our framework, trading probabilities and rates depend not on the full distribution of cash positions, but solely on a single statistic capturing market tightness.

OTC structure. The market structure consists of a sequence of over-the-counter (OTC) markets, following Afonso and Lagos (2015b). Unlike in Afonso and Lagos (2015b), where investors trade directly, we assume that each investor delegates its trade to a large number of traders, each handling an order of size Δ. Traders close cash deficits and lend surpluses on behalf of investors.

Traders are segmented into borrowers and lenders sides of the market, giving rise to a two-sided matching structure. If $\theta > 1$, the short side is the side with surplus. If $\theta < 1$, the short side is the side with deficit.

Trading occurs over N rounds, indexed by $n \in \mathcal{N} \equiv 1, 2, \ldots, N$. In each round, matches are formed according to a matching function. Once matched, borrowers and lenders bargain over the interest rate of a loan of cash of size Δ. Upon trade, the surplus trader transfers Δ cash to the deficit trader. The settlement assets plus the bargained rate are returned to the lender at the beginning of the next portfolio stage.

If a trader does not match in a given round, it may or may not match in the subsequent round. A trader in deficit who remains unmatched after round N borrows at the penalty rate R^w. Traders in surplus that are unmatched by round N, allocate only the cash rate R^m. We derive our analytical results in the limiting case where the order size Δ approaches zero and the number of trading rounds N tends to infinity.[9] This limit is technically convenient because it avoids the combinatorial challenges that arise when the identity of the matches matters. Here, the terms of trade vary by round, but not by the identity of counterparties.[10]

[9]The assumption of infinitesimal order size can also be found in Shi (1997); Atkeson et al. (2015). The assumption bears some realism to the extent that financial institutions often feature multiple traders operating their trading desks.

[10]The idea behind the modeling of fixed-sized order convention is that if a lender bank places lending orders that exceed its excess reserves, there is a chance it will not be able to have the funds to transfer to the borrowing bank. If a lending bank lacks the funds to transfer to the bank in deficit, the bank would violate a contract and face a large legal default cost. For sufficiently high costs, no bank will ever place lending orders above the amounts they hold. Assuming that orders are of a fixed size, investors cannot place more than $\eta(s, \Delta) \equiv \lfloor |s| / \Delta \rfloor$ orders. Here, $\lfloor x \rfloor$ is the floor function understood as the largest integer not greater than x. Because investors can only place integer numbers of orders, typically, there will be a remainder of reserves that cannot be lent or borrowed at the OTC market. These residuals can be borrowed (lent) at the penalty rate r^{DW} (at the cash rate r^m) directly. Mathematically, this residual is $\phi(x, \Delta) = s - \lfloor |s| / \Delta \rfloor \Delta$.

3.1 OTC market and its equilibrium

Denote $\{S_n^+, S_n^-\}$ the aggregate mass of surplus and deficit positions after the realization of matches in round n and where $S_0^+ = S^+$and $S_0^- = S^-$. The number of matches in round $n+1$, z_{n+1}, is given by

$$z_{n+1} \equiv \lambda_N G(S_n^+, S_n^-), \quad n \in \mathcal{N}.$$

where λ_N is a parameter capturing the matching efficiency, which we index by the total number of rounds N and $G : \mathbb{R}_+^2 \to \mathbb{R}_+$ is the matching function.[11]

We assume the matching function satisfies the following properties:

Assumption 1 (Matching Function)**.** *The matching function G satisfies:*

i. No disposal: $G(0, \cdot) = G(\cdot, 0) = 0$.

ii. Constant returns to scale. $G(\cdot, \cdot)$ is homogeneous of degree one.

iii. Symmetry: $G(a, b) = G(b, a)$, $\forall a, b$.

iv. Weak exhaustion: $\lambda_N G(S_n^+, S_n^-) \leq \min\{S_n^+, S_n^-\}$.

v. Strictly increasing: $G_a(a, b) \geq 0, G_b(a, b) \geq 0$.

vi. Weakly Concave: $G_{aa}(a, b) \leq 0, G_{bb}(a, b) \leq 0$.

No disposal implies that counterparts are needed for a match. Constant returns to scale imply that the number of matches scales proportionally with the size of the market. Symmetry assumes that the number of matches depends on the relative scarcity of borrowers or lenders, independently of which side is shorter. Week exhaustion means that there cannot be more matches than the shortest side of the market. Monotone increasing in each argument implies that adding more participants to either side leads to more matches. Finally, concavity implies diminishing returns in the number of matches as a function of one side of the market. We normalize $G(1, 1) = 1$, without loss of generality. Notice we can rescale the number of matches by λ_N.

Given, $\{S_0^+, S_0^-\}$, we construct the sequence of matches for each round z_n by tracking the evolution of surplus and deficit positions as follows:

$$S_n^+ \equiv S_{n-1}^+ - z_n \qquad \text{and} \qquad S_n^- \equiv S_{n-1}^- - z_n, \qquad \forall n \in \{1, 2, \dots, N+1\}.$$

[11]The subscript N is useful once we take the limit $N \to \infty$.

As matches take place, the surplus and deficit positions shrink in the following round. The recursion assumes implicitly that all matches result in trade.[12] Accordingly, we define market tightness in round n as

$$\theta_n \equiv \frac{S_n^-}{S_n^+}, \qquad n \in \{0, 1, 2, \ldots, N\}. \tag{9}$$

and matching probabilities $\{\psi_n^+, \psi_n^-\}$ for a trader in surplus and deficit respectively as

$$\psi_n^+ \equiv \frac{z_n}{S_{n-1}^+} = \lambda_N G\left(1, \theta_{n-1}\right), \quad \psi_n^- \equiv \frac{z_n}{S_{n-1}^-} = \lambda_N G\left(\frac{1}{\theta_{n-1}}, 1\right) \qquad n \in \mathcal{N}. \tag{10}$$

By convention, $\psi_{N+1}^+ = \psi_{N+1}^- = 0$, given that the last trading round with matches is N. By constant returns to scale, $\psi_n^+ = \theta_{n-1}\psi_n^-$.

Bargaining. At a given match, traders bargain over an OTC loan rate, R_n^f. Generically, each rate depends on the wealth of the counterparties and the round. We verify that as the size of trades vanishes, $\Delta \to 0$, rates only depend on the round.

The rate R_n^f is determined through Nash bargaining: The outside options for borrowers and lenders vary by round. In a match at the N-th round, the lender's outside option is R^m and the borrower's outside option is the outside borrowing rate, R^w. In earlier rounds, the outside option is the expected value of entering the next round with an unmatched position.

Now, consider an individual trader who bargains on behalf of investor j. The trader is responsible for closing the order of size Δ. When bargaining, traders must form an expectation of their investor's equity position in each round of the OTC market. These equity position depends on all other trades delegated by the investor. Constructing an expectation of what other traders do is a complex combinatorial problem. However, the limit as $\Delta \to 0$ circumvents that challenge. Before taking that limit, assume that when traders bargain, the trader uses the law of large numbers to estimate the equity position in each round and the average bargained rate $\overline{R}_n^f$ of other traders—a rate we have to solve for.

Under the law of large numbers assumption, if investor j has a cash deficit, $s^j < 0$, the trader forecasts that the fraction $\Psi^-\left(s^j + \Delta\right)$ will be borrowed across all rounds at an average rate $\overline{R}^f$ and $\left(1 - \Psi^-\right)\left(s^j + \Delta\right)$ will be borrowed at the discount rate. If the investor has a cash surplus, $s^j > 0$, the trader expects that $\Psi^+ \cdot \left(s^j - \Delta\right)$ funds will be lent at the OTC market and that $\left(1 - \Psi^+\right) \cdot \left(s^j - \Delta\right)$ of the funds will remain idle.

Recall that in (6) we determine that future equity is given by a convenience-yield function map-

[12] By no exhaustion and no disposal, if at any round either side of the market is exhausted—i.e., ends at zero—no further matches are formed.

ping an investor's position s to overall settlement costs through the function $\chi_{t+1}(s)$. In turn, this function depends on the trading probabilities $\{\Psi^-, \Psi^+\}$ and an average rate $\overline{R}^f$. Suppose the trader matches in round n. Thus, using $\left\{\Psi^-, \Psi^+, \overline{R}^f\right\}$, the trader forecasts the investor's equity to be:

$$e^j_\Delta = \underbrace{\sum_{i\in\mathbb{I}} a^i_{t+1}R^i_{t+1} + m_{t+1}R^m_{t+1} + \chi_{t+1}(s^j - \text{sign}\{s^j\}\Delta)}_{\equiv\mathcal{E}^j_{t+1}(\Delta)} + \text{sign}\{s^j\}\left(R^f_n - R^m_{t+1}\right)\Delta.$$

This is the analogue of e_{t+1} obtained in equation (6) with two modifications: First, the trader applies χ_{t+1} to the investor's cash position net of the amount it is delegated to trade, $s^j - \text{sign}\{s^j\}\Delta$. This is because it takes $\left\{\Psi^-, \Psi^+, \overline{R}^f\right\}$ as given for other trades. The term $\mathcal{E}^j_{t+1}(\Delta)$ defined above is the trader's estimate of equity, excluding its own trade. Second, the trader internalizes the effect of its bargained rate R^f_n. Thus, the payoff $\text{sign}\{s^j\}\left(R^f_n - R^m_{t+1}\right)\Delta$ depends on its bargain.

When bargaining, the trader internalizes its contribution to the investor's equity. We suppress the time subscript (and conditioning on the state) in what follows since the matching occurs intra-period. The value of the investor's future equity is $V(e')$. Let $J_M^{\text{sign}\left\{s^j\right\}}(n;\Delta) \in \left\{J^+_M(n;\Delta), J^-_M(n;\Delta)\right\}$ represent the round-n value functions of the surplus and deficit traders conditional on matching, respectively. These values satisfy:

$$J_M^{\text{sign}\{s^j\}}(n;\Delta) \;\;\equiv\;\; V(e') = V\left(\mathcal{E}^j(\Delta) + (\text{sign}\{s^j\})\left(R^f_n - R^m\right)\Delta\right).$$

Analogously, let $J_U^{\text{sign}\left\{s^j\right\}}(n;\Delta) \in \left\{J^+_U(n;\Delta), J^-_U(n;\Delta)\right\}$ represent the corresponding trader's value conditional on being unmatched in round n. These values are written recursively:

$$\begin{aligned} J_U^{\text{sign}\{s^j\}}(n;\Delta) &\equiv \mathbb{E}\left[V(e')\,|\text{unmatched at round } n\right] \\ &= \mathbb{E}\left[\psi_{n+1}^{\text{sign}\left\{s^j\right\}} J_M^{\text{sign}\{s^j\}}(n+1;\Delta) + \left(1-\psi_{n+1}^{\text{sign}\{s^j\}}\right) J_U^{\text{sign}\{s^j\}}(n+1;\Delta)\right]. \end{aligned}$$

This expression uses the trading probabilities obtained above. The value of an unmatched trader is the probability of matching in the next round times the value of matching its position, $\psi_{n+1}^{\text{sign}\{s^j\}} J_M^{\text{sign}\{s^j\}}(n+1;\Delta)$, plus the probability of not matching and going to the subsequent round with an unmatched position. If by round N the trader is unmatched, a surplus trader gains nothing (as if $R^f_n = R^m$), whereas a deficit trader pays the difference between the external borrowing rate (as if $R^f_n = R^w$):

$$J^+_U(N;\Delta) \equiv V\left(\mathcal{E}^j(\Delta)\right), \quad J^-_U(N;\Delta) \equiv V\left(\mathcal{E}^j(\Delta) - (R^w - R^m)\Delta\right).$$

With these value functions, we can describe the bargaining problem. We denote by $r \equiv R - 1$ the net rate of any gross return R. Let us refer to the trader associated with a deficit position by j and the trader associated with a surplus position by k. Upon a match, an OTC rate solves the following Nash-bargaining problem:

$$r_n^f(\Delta) = \arg\max_{r_n^f} \left\{ \mathcal{S}_n^-(\Delta)^\eta \times \mathcal{S}_n^+(\Delta)^{1-\eta} \right\} \tag{11}$$

where

$$\mathcal{S}_n^-(\Delta) = V\left(\mathcal{E}^j(\Delta) - \left(r_n^f - r^m\right)\Delta\right) - J_U^-(n;\Delta)$$
$$\mathcal{S}_n^+(\Delta) = V\left(\mathcal{E}^k(\Delta) + \left(r_n^f - r^m\right)\Delta\right) - J_U^+(n;\Delta).$$

For $\Delta > 0$, the combinatorial probability of who matches with whom induces risk, making the investor's equity a random variable. In the bargaining problem, the estimate $\mathcal{E}^j(\Delta)$ is treated as deterministic, so it is only an approximation that exploits the law of large numbers. Furthermore, because of the possible curvature of V, the solution to the bargaining problem also depends on the meeting counterparts. Once we take the limit $\Delta \to 0$, $\mathcal{E}^j(\Delta)$ is not an approximation but an exact estimate and, furthermore, the curvature of V plays no role. We work with that limit next.

Infinitesimal Orders. We now consider the limiting equilibrium as the order size vanishes, $\Delta \to 0$. When two traders meet during the different rounds, the interbank market rate satisfies the following:

Problem 1 (Infinitesimal Trade Bargaining Problem)**.** *The sequence of infinitesimal trade bargaining problems is:*

$$\max_{r_n^f \in \left\{r^m + \chi_n^+, r^m + \chi_n^-\right\}} \left(\chi_n^- - \left(r_n^f - r^m\right)\right)^\eta \left(\left(r_n^f - r^m\right) - \chi_n^+\right)^{1-\eta}$$

where $\{\chi_n^+, \chi_n^-\}$ *solve:*

$$\chi_n^+ = \left(r_{n+1}^f - r^m\right)\psi_{n+1}^+ + \chi_{n+1}^+\left(1 - \psi_{n+1}^+\right) \tag{12}$$

and

$$\chi_n^- = \left(r_{n+1}^f - r^m\right)\psi_{n+1}^- + \chi_{n+1}^-\left(1 - \psi_{n+1}^-\right) \tag{13}$$

for $n \in \mathcal{N}$*, with terminal conditions* $\chi_{N+1}^+ = 0$ *and* $\chi_{N+1}^- = r^w - r^m$*.*

This problem consists of a sequence of bargaining problems. The sequence $\{\psi_n^+, \psi_n^-\}_{n=1}^{N+1}$ is known. In round n, the bargained amount is r_n^f and the borrower's (lender's) outside options is the expected rate paid (received) conditional on going to the following round unmatched χ_n^- (χ_n^+).

Notice that χ_n^- and χ_n^+ inherit a recursive structure: Indeed, the expected cost of an unmatched deficit by round n is given by the cost of closing the deficit in the next round $\left(r_{n+1}^f - r^m\right)$ times the probability of matching in round $n+1$, ψ_{n+1}^-, plus the expected cost of heading to the subsequent round unmatched χ_{n+1}^- times the probability of not matching in round $n+1$, $\left(1-\psi_{n+1}^-\right)$; and an isomorphic recursion applies to χ_n^+.

Our first main result states that as $\Delta \to 0$, the solution to the bargaining problems (11) is indeed the solution to Problem 1. Thus, as long as trades are small, the identity of the counterparts is irrelevant to determining the OTC rate, and the only relevant information is the trading round n. This is a convenient property: we do not have to keep track of the evolution of the distribution of surplus positions across rounds, as do Afonso and Lagos (2015b). This feature simplifies the equilibrium dramatically.

As a result of the simplifications, r_n is obtained by solving a system of linear difference equations.

Proposition 1 (Limit of Bargaining Problems)**.** *As $\Delta \to 0$, any bargained rate at round n solves Problem 1. That is, $r_n^f = r^m + (1-\eta)\,\chi_n^- + \eta\chi_n^+$, for all $n \in \{1, 2, ..., N\}$.*

The proposition tells us that the spread between the OTC rate and the rate on cash, $r_n^f - r^m$, is given by a weighted average of the outside options, $(1-\eta)\,\chi_n^- + \eta\chi_n^+$.

A gist of the proof is as follows: we can divide the objective of the bargaining problem in (11) by any constant without changing the bargained rate. One such constant is the trade size Δ. Hence, we can write the objective as $[\mathcal{S}_n^-(\Delta)/\Delta]^\eta\,[\mathcal{S}_n^+(\Delta)/\Delta]^{1-\eta}$. Then, observe that

$$\lim_{\Delta\to 0^+} \frac{V\left(\mathcal{E}^j + \text{sign}\{s^j\}\left(r_n(\Delta) - r^m\right)\Delta\right) - V(\mathcal{E}^j + \chi_n^{\text{sign}\{s^j\}}\Delta)}{\Delta} =$$
$$V'\left(\mathcal{E}^j\right)\text{sign}\{s^j\}\left(r_n - r^m - \chi^{\text{sign}\{s^j\}}\right),$$

where the equality follows the definition of the derivative. Using this definition, the original bargaining problem simplifies,

$$\lim_{\Delta\to 0^+}\left\{\max_{r_n^f}\left[\mathcal{S}_n^-(\Delta)/\Delta\right]^\eta\left[\mathcal{S}_n^+(\Delta)/\Delta\right]^{1-\eta}\right\} =$$
$$V'\left(\mathcal{E}^j\right)^\eta V'\left(\mathcal{E}^k\right)^{1-\eta}\max_{r_n^f}\left[\chi_{n+1}^- - (r_n - r^m)\right]^\eta\left[(r_n - r^m) - \chi_{n+1}^+\right]^{1-\eta}.$$

Thus, as the transaction size shrinks, the rate is independent of the counterparties. This occurs because as the trade size vanishes, the trade the influence on the counterparts' wealth vanishes too. Thus, whether a trade takes place or not has a negligible effect on marginal utilities—the effect of the trade is of second order. Yet, traders want to maximize wealth, even though the size of the trade has a vanishing effect on wealth, a first-order effect. For that reason, each surplus is the difference in expected financing costs and is proportional to marginal utilities.

The difference between costs and benefits defines the total surplus in round n, which we define as $\Sigma_n \equiv \chi_n^- - \chi_n^+$. This surplus is the difference between the expected borrowing costs for lenders minus the lending benefit for lenders. The solution to the Nash bargaining problems renders the conventional surplus splitting rule:

$$\lim_{\Delta \to 0} \mathcal{S}_n^-(\Delta) = \chi_n^- - \left(r_n^f - r^m\right) = \eta \Sigma_n, \quad \lim_{\Delta \to 0} \mathcal{S}_n^+(\Delta) = \left(r_n^f - r^m\right) - \chi_n^+ = (1-\eta)\,\Sigma_n.$$

Algorithm. A byproduct of the proposition above is an algorithm to solve the OTC rate of each round: First, solve $\{\psi_n^+\}$ and $\{\psi_n^-\}$ forward, taking θ_0 as given. Then, solve $\{\chi_n^+\}$ and $\{\chi_n^-\}$ backwards and obtain r_n^f as the interest via (12) and (13).

The algorithm works because we can confirm that the solutions χ_0^+ and χ_0^-, are indeed the slopes of the liquidity yield function, χ, and furthermore, that $\bar{r}^f$, that defines the convenience-yield function $\chi\left(s\right)$.

Proposition 2 (Convenience-yield function and coefficients)**.** *The probability of closing deficit and surplus positions in the OTC market are given by:*

$$\Psi^- \equiv 1 - \prod_{n=1}^{N}\left(1-\psi_n^-\right) \textit{ and } \Psi^+ \equiv 1 - \prod_{n=1}^{N}\left(1-\psi_n^+\right).$$

The coefficients of the convenience yield function $\chi(s)$ are solutions to (12) and (13) that satisfy:

$$\begin{aligned}
\chi^- &= \Psi^-(\bar{r}^f - r^m) + (1-\Psi^-)(r^w - r^m) = \chi_0^-, \\
\chi^+ &= \Psi^+(\bar{r}^f - r^m) = \chi_0^+.
\end{aligned}$$

where $\bar{r}^f$ is the average of $\left\{r_n^f\right\}$ weighted by the trade volume.

The proposition confirms that the convenience-yield function delivers the exact cost or benefit of settling in the OTC market as trade size converges to zero. To construct the convenience-yield function, we equate $\{\chi^-, \chi^+\} = \left\{\chi_0^-, \chi_0^+\right\}$. This step guarantees that the overall matching probabilities $\{\Psi^-, \Psi^+\}$ and the average OTC rate determines the average settlement costs and benefits for borrowers and lenders, respectively.

We will show below in Section 3.3 that with additional structure on the matching technology, we can derive in closed form χ^+ and χ^-.

Infinite Rounds and Continuous-Time Limit. Let $N \to \infty$. As noted above, we define $\lambda_N = \bar{\lambda}/N$, so λ_N vanishes with N, but at a convergent rate $\bar{\lambda}$. This convergence rate allows us to convert the realization of a match into a Poisson process with time-varying intensity. We first derive the evolution of the masses of surpluses and deficits. The time interval between rounds is $1/N$, which will shrink to zero. Also, we index the normalized round by $\tau \in \{0, 1/N, 2/N, ..., 1\}$. Thus, as $N \to \infty$, we can associate a round with a point in time, $\tau \in [0,1]$. Therefore, from now on, we index all equilibrium variables by τ instead of by n.

Because the matching function is symmetric, we compress the notation by defining the normalized (the intensive form) matching function γ:

$$\gamma(\theta) \equiv G(\theta, 1).$$

As we increase the rounds, the probabilities of matching in any given instant τ vanish. However, given the tightness θ of a given round, matching probabilities turn into matching rates (intensities) defined as:

$$\psi^{+}(\theta) = \bar{\lambda}\gamma(\theta) \quad \psi^{-}(\theta) = \bar{\lambda}\gamma\left(\theta^{-1}\right). \tag{14}$$

These intensities satisfy (i) $\psi^{+}(\theta) = \psi^{-}(\theta^{-1})$ and (ii) $\psi^{+}(\theta) = \theta\psi(\theta^{-1})$, a matching consistency condition. The matching rates, in turn, determine the overall matching probabilities:

$$\Psi_{\tau}^{+} = 1 - \exp\left(-\int_{\tau}^{1} \psi_{x}^{+}(\theta)\, dx\right), \quad \Psi_{\tau}^{-} = 1 - \exp\left(-\int_{\tau}^{1} \psi_{x}^{-}(\theta)\, dx\right), \tag{15}$$

translating probability distributions into density functions. In the continuous-time limit, we can dispense with the non-exhaustion assumption as $\lambda_N \to 0$. However, we show below that for certain matching functions, the shortest side of the market may vanish in finite time. This density is distributed exponentially with time-varying intensities.

Lemma 1 (ODE for market tightness)**.** *Let the number of trading rounds $N \to \infty$. Then, index a round by $\tau = \frac{n}{N} \in [0,1]$. The ratio of deficit to surpluses θ_{τ} satisfies the following first-order homogeneous differential equation:*

$$\dot{\theta}_{\tau} = -\bar{\lambda}\theta_{\tau}\left[\gamma\left(\theta_{\tau}^{-1}\right) - \gamma(\theta_{\tau})\right], \quad \tau \in [0,1]. \tag{16}$$

where θ_0 given. The corresponding matching intensities are $\psi_{\tau}^{+} = \psi^{+}(\theta_{\tau})$ and $\psi_{\tau}^{-} = \psi^{-}(\theta_{\tau})$.

As the number of rounds increases, θ is translated from a sequence into a function of time, τ. This function satisfies the ordinary-differential equation (ODE), equation (16). Once we solve the ODE for tightness, we can solve for the matching intensities ψ_{τ}^{+} and ψ_{τ}^{-} in $\tau \in [0,1]$. Furthermore, with information of regarding the terminal market tightness, $\bar{\theta} \equiv \theta_1$, we recover the overall matching probabilities by using the matching consistency condition:

$$\Psi_0^+ = \frac{\bar{\theta} - \theta_0}{\bar{\theta}_1 - 1}, \quad \Psi_0^- = \frac{\bar{\theta} - \theta_0}{\bar{\theta}_1 - 1} \theta_0^{-1}.$$

Observe that the reciprocal tightness, θ^{-1}, satisfies the ODE:

$$\dot{\theta}^{-1} = \bar{\lambda}\theta^{-1}\left[\gamma(\theta) - \gamma\left(\theta^{-1}\right)\right] = -\bar{\lambda}\gamma(\theta_\tau)(1 - \theta_\tau), \tag{17}$$

where the second equality follows from symmetry and homogeneity, $\gamma(\theta) = \theta\gamma(\theta^{-1})$. Thus, the ODE for tightness or its reciprocal is the same.[13]

Proposition 3 (Properties of θ)**.** *The evolution of market tightness features the following properties:*

i. If $\theta_0 = 1$, then $\theta_\tau = 1$. If $\theta_0 > 1$ ($\theta_0 < 1$), then θ_τ is increasing (decreasing) with time.

ii. If $\theta_0 = 1$, matching rates are $\psi_\tau^+ = \psi_\tau^- = 1$. If $\theta_0 > 1$ ($\theta_0 < 1$), then ψ_τ^+ is increasing (decreasing) and ψ_τ^- is decreasing (increasing) with time.

iii. Take two matching functions, such that$\gamma(\theta) < \tilde{\gamma}(\theta), \quad \forall\theta$. Then, if $\theta > 1$ ($\theta < 1$), θ rises (falls) faster under $\tilde{\gamma}$.

The proposition shows that the short side of the market becomes relatively more scarce over the rounds, and this reflects on the trading probabilities. We also know that if a matching function creates more matches, it will result in a tighter market throughout, provided that the deficit side exceeds one initially.

Given ψ_τ^+ and ψ_τ^-, just as in the case with finite rounds, we can solve the differential equation for $\{\chi_\tau^+, \chi_\tau^-\}$ in closed form:

Proposition 4 (Liquidity-Yield Coefficients in the Continuous-Time Limit)**.** *Let the number of trading rounds $N \to \infty$. Then, the solution to $\{\chi_\tau^+, \chi_\tau^-\}$ is:*

$$\chi_\tau^+ = (r^w - r^m) \int_\tau^1 (1-\eta)\,\psi_y^+ \exp\left(\int_y^1 -\left((1-\eta)\,\psi_x^+ + \eta\psi_x^-\right) dx\right) dy$$

and

$$\chi_\tau^- = (r^w - r^m)\left(1 - \int_\tau^1 \eta\psi_y^- \exp\left(\int_y^1 -\left((1-\eta)\,\psi_x^+ + \eta\psi_x^-\right) dx\right)\right) dy,$$

[13] The Picard-Lindelöf Theorem guarantees the uniqueness of the solution, as long as γ is Lipschitz continuous. We note that for some matching functions, γ may fail to be Lipschitz continuous at $\theta = 0$ or ∞, but even in that case, uniqueness can be guaranteed by finding the path of θ backward in time to reach the initial condition θ_0 and using the fact that if $\theta_\tau = 0$ for any τit remains at zero for all $\tau' > \tau$.

for all $\tau = [0,1]$. In turn, the coefficients of the convenience-yield function are $\chi^{+} = \chi_0^{+}$ and $\chi^{-} = \chi_0^{-}$ and the OTC rate at time τ is given by $r_\tau^f = r^m + (1-\eta)\,\chi_\tau^{-} + \eta\chi_\tau^{+}$.

This proposition delivers $\{\chi_\tau^{+}, \chi_\tau^{-}\}$ in terms of the matching intensities. We observe several consistency properties: First, as $\tau \to 1$, the formula yields the terminal conditions $\chi_1^{+} = 0$ and χ_1^{-}. Second, if matching intensities approach zero, each side receives the terminal value $\chi_\tau^{+} = \chi_1^{+}$ or $\chi_\tau^{-} = \chi_1^{-}$. Third, χ_τ^{-} is increasing in time whereas χ_τ^{+} is decreasing: as the chances of trading fall with time, surplus positions have a lower expected value of lending benefits, whereas deficit positions expect a higher average rate. This tells us that regardless of the direction of rates as a function of time, matching rates dictate the direction of expected values of trading as a function of time.

The formula in the proposition has the interpretation of being equivalent to a setting where one side of the market extracts all the surplus with a constant probability. To see this, it is convenient to compute the surplus at time τ, which in this case is given by:

$$\Sigma_\tau \quad = (r^w - r^m)\,(1 - H_\tau^{+})\,(1 - H_\tau^{-})\,,$$

where $H_\tau^{+} \equiv 1 - \exp\left(\int_\tau^1 -\left((1-\eta)\,\psi_x^{+}\right)dx\right) \quad H_\tau^{-} \equiv 1 - \exp\left(\int_\tau^1 -\left(\eta\psi_x^{-}\right)dx\right)$.

This representation reveals two things: first, the surplus increases with time because, as time runs out, the option to trade in the future is falling. Second, the surplus at time τ is proportional to the trading between the surplus in the final round multiplied by the product of two probabilities, $(1 - H_\tau^{+})$ and $(1 - H_\tau^{-})$. The product $(1 - H_\tau^{+})\,(1 - H_\tau^{-})$ is analogous to the probability that neither side extracts any surplus in the future.[14]

We further simplify the formulas for χ_τ^{-} and χ_τ^{+} as time-varying fractions of the total surplus:

$$\chi_\tau^{+} = \int_\tau^1 (1-\eta)\,\psi_y^{+}\Sigma_y dy, \quad \chi_\tau^{-} = (r^w - r^m) - \int_\tau^1 \eta\psi_y^{-}\Sigma_y dy. \tag{18}$$

Thus, the outside options satisfy a dynamic surplus splitting rule and are obtained as the expected probability of extracting all the surplus in a future match.

Illustrations. Figure 1 presents an example of the OTC market using a Leontief matching function, for various initial conditions θ_0. The figure reports the movement of the OTC market rate

[14]Note $(r^w - r^m)\left(1 - \int_\tau^1 \frac{\partial}{\partial y}\exp\left(\int_y^1 -\left((1-\eta)\,\psi_x^{+} + \eta\psi_x^{-}\right)dx\right)\right)dy = (r^w - r^m)\exp\left(\int_\tau^1 -\left((1-\eta)\,\psi_x^{+} + \eta\psi_x^{-}\right)dx\right)$. These probabilities correspond to the probability that an (homogeneous) Poisson event is *not* realized between time τ and time 1. Indeed, H_τ^{+} and H_τ^{-} are the probabilities of two independent Poisson events occurring with respective intensities $(1-\eta)\,\psi_x^{+}$ and $\eta\psi_x^{-}$. Thus, H_τ^{+} and H_τ^{-} correspond to a compound process in which borrowers and lenders extract all the terminal surplus $(r^w - r^m)$ with probability $1-\eta$ and η respectively once matching.

through time, plotted together with the outside options. As the trading session ends, the outside options for the surplus (deficit) side = collapses to zero $(r^w - r^m)$. The outside options at the beginning of the trading session yield the convenience yield coefficients. Figure 2 presents analogous results for Cobb-Douglas matching. While the patterns within a trading day seem similar, comparative statics differ.

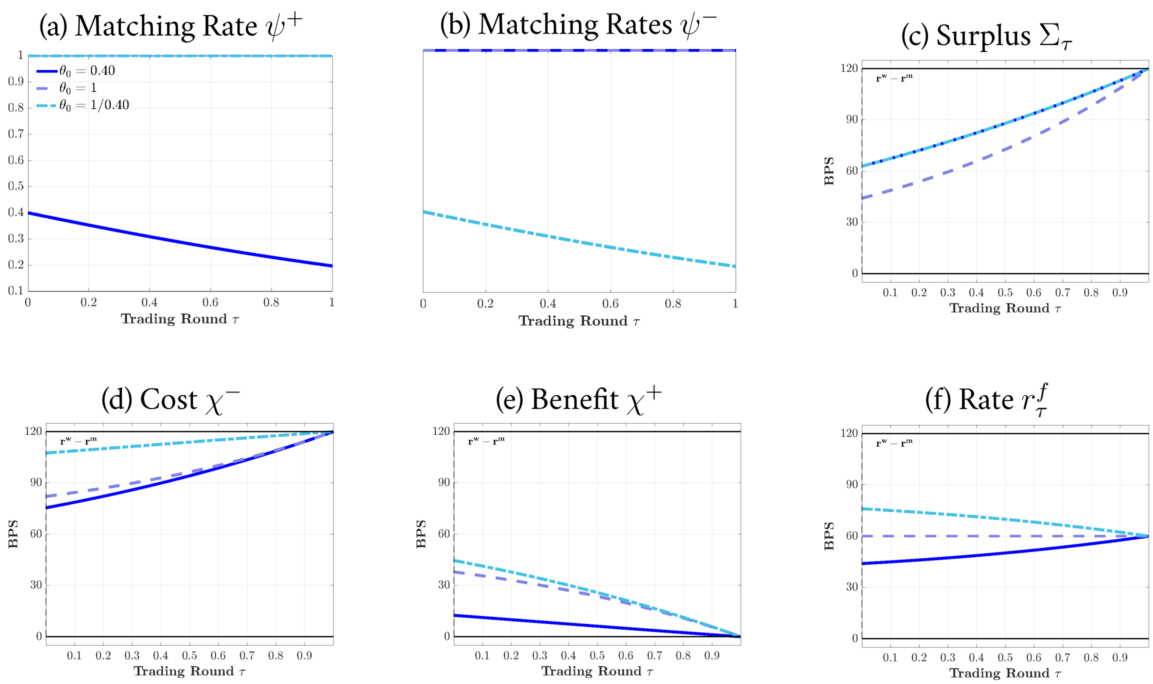

Figure 1: **Leontief Example:** Trading at various rounds.

Note: $\theta_0 = \theta$ is the initial market tightness, the ratio of the initial aggregate deficit and initial aggregate surplus. The example is calibrated with $\eta = 0.5$, $\bar{\lambda} = 1.2$, $r^w - r^m = 120$bps. See equation (9).

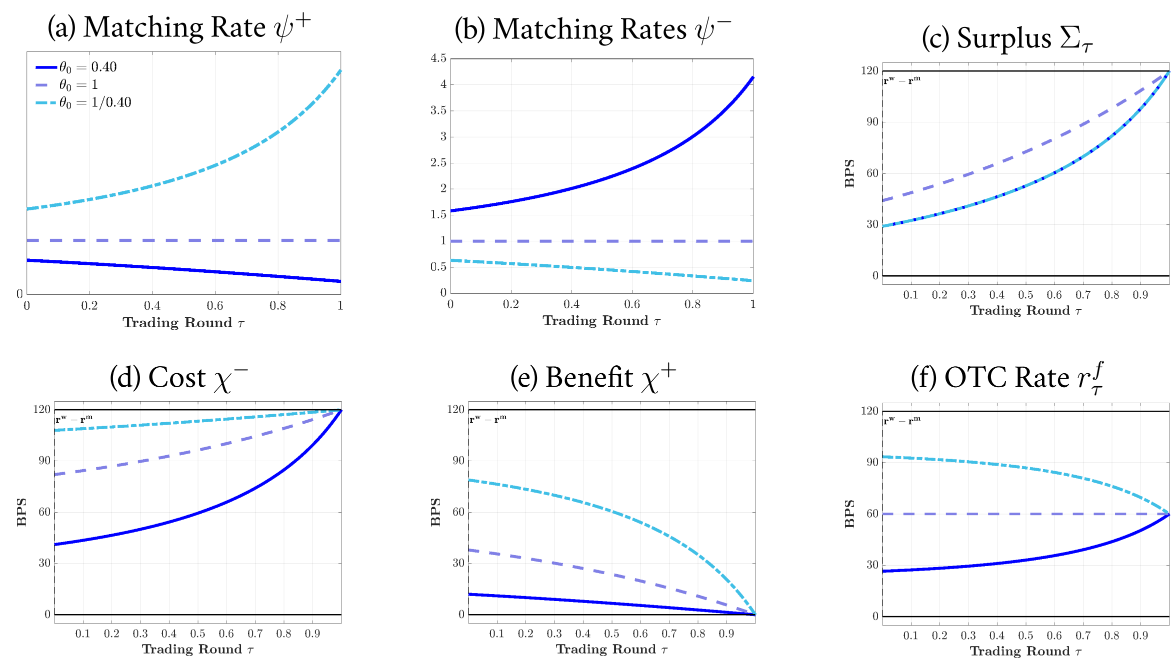

Figure 2: Cobb-Douglas Example: Trading at various rounds

Note: $\theta_0 = \theta$ is the initial market tightness, the initial aggregate deficit and the initial aggregate surplus. The example is calibrated with $\eta = 0.5$, $\bar{\lambda} = 1.2$, $r^w - r^m = 120$bps.

3.2 General Properties

In this section, we present general properties of the OTC market and liquidity-yield coefficients.

General Solution to the Liquidity-Yield Coefficients. While the solution for the path of market tightness depends on the matching function, the yield coefficients can be expressed solely as a function of the initial tightness, θ, and the terminal tightness.[15] This solution is valid even though the terminal values can be very different depending on the actual matching function that we use:

Proposition 5. *For any matching function, the convenience yield coefficients are:*

$$\chi^+_\tau = (r^w - r^m)\left(\frac{\bar{\theta} - \bar{\theta}^\eta \theta_\tau^{1-\eta}}{\bar{\theta} - 1}\right), \quad \chi^- = (r^w - r^m)\left(\frac{\bar{\theta} - \bar{\theta}^\eta \theta_\tau^{-\eta}}{\bar{\theta} - 1}\right).$$

Moreover, for $\theta = \theta_0$ the average OTC rate is given by:

$$\bar{r}^f = \phi(\theta)\, r^m + (1 - \phi(\theta)\,) r^w, \text{ where } \phi(\theta) = \frac{\left(\bar{\theta}/\theta\right)^\eta - \theta}{\bar{\theta}/\theta - 1}\phi(\theta) \in [0, 1]\,. \tag{19}$$

[15]Note, however, that $\bar{\theta}$ depends on the actual matching function.

The solution to $\{\chi^+, \chi^-\}$ is a continuous function of the tightness at some τ and the terminal tightness, θ and $\bar{\theta}$. $\{\chi^+, \chi^-\}$ are always positive and bounded by the terminal surplus, $r^w - r^m$, as expected. This property is convenient for analysis. There are many matching functions for which a solution to the (16) has no solution. Yet, we can establish various properties thanks to this calculation.[16] We furthermore obtain the average OTC rate as a weighted average of the cash rate and discount rate. The term $\phi(\theta)$ acts as an endogenous bargaining rate: unlike static bargaining, this bargaining rate depends on the evolution of outside options. $\phi(\theta)$ compactly captures the information from the matching function and the bargaining process.

As a corollary, we have that the time-evolution of variables in Figures 1 and 2, corresponding to Leontief and matching functions, hold across all matching functions.

Corollary 1 (Monotonicity Within Trading Session)**.** *For all $\tau \in (0,1)$, we have $\frac{d\chi_\tau^+}{d\tau} < 0$ and $\frac{d\chi_\tau^-}{d\tau} > 0$. Moreover, sgn $\left(\frac{dr_\tau^f}{d\tau}\right) = -sgn(\theta_0 - 1)$.*

Balanced Market Solution. Recall that when the OTC market is initially balanced, $\theta_0 = 1$, the market remains balanced throughout and, thus, matching rates are equalized: $\psi_\tau^+ = \psi_\tau^- = (1-\eta)\,\psi_\tau^+ + \eta\psi_\tau^- = \bar{\lambda}$. The following result follows from that observation:

Corollary 2 (Balanced Market)**.** *For $\theta_0 = 1$, $\Sigma_\tau = (r^w - r^m)\, e^{-\bar{\lambda}(1-\tau)}$, $r_\tau^f - r^m = (1-\eta)\,(r^w - r^m)$, and $\{\chi_\tau^+, \chi_\tau^-\} = \{(1-\eta)\,((r^w - r^m) - \Sigma_\tau)\,, \quad r^w - r^m - \eta\,((r^w - r^m) - \Sigma_\tau)\}$.*

When the market is balanced, the trading surplus is the terminal surplus, $(r^w - r^m)$, scaled by the probability of no matches in the remaining time $(1-\tau)$.[17] The yield coefficients are affine in the surplus, with the average cost of deficits increasing as time is running out—and the opposite for the surplus benefit. The yield coefficients balance exactly such that the average negotiated rates equal the one obtained under static bargaining (i.e., a one-round bargaining problem). We further verify that $\lim_{\theta \to 1} \phi(\theta) = \eta$, consistent with the balanced-market result.

Time Dilation. Another property of the OTC market equilibrium is time dilation: Let $\theta\left(\tau, \theta_0, \bar{\lambda}\right)$ denote the value of market tightness θ at time τ given an initial condition θ_0 and efficiency $\bar{\lambda}$: the function representing the solution (16)) as function of time, initial condition and parameters; with the same notation for $\left\{\gamma_\tau^+, \gamma_\tau^-, r_\tau^f, \chi_\tau^-, \chi_\tau^+, \Sigma_\tau\right\}$. Time dilation refers to the property that the passage of time is equivalent to a reduction in efficiency:

[16] This property is also convenient for quantitative analysis since it avoids the integrals in Proposition 4.
[17] This probability is an exponential with intensity $\bar{\lambda}$.

Proposition 6 (Time Dilation)**.** *Fix $\tau, \tau' \in [0,1]$ such that $\tau' > \tau$. Then,*

$$\theta\left(\tau', \theta_0, \bar{\lambda}\right) = \theta\left(\frac{\tau' - \tau}{1-\tau}, \theta\left(\tau, \theta_0, \bar{\lambda}\right), \bar{\lambda}\left(1-\tau\right)\right).$$

The same property holds for all the equilibrium objects $\left\{\gamma^+, \gamma^-, \chi^+, \chi^-, r^f\right\}$.

Time dilation allows us to obtain the value of market tightness at an instant τ', by computing first the value of market tightness at a prior instant τ: we can obtain the value at τ' by (i) solving the equilibrium renormalizing time by the remaining time $(1-\tau)$, (ii) scaling efficiency by the remaining time, and (iii) setting the initial condition to θ_τ. This property is useful because it tells us that any property (e.g., monotonicity, concavity, etc.) of the equilibrium functions at some moment in time generalizes to all times. It also reveals the recursive nature of the market and shows that the normalization of time to the interval $[0,1]$ is inconsequential.

Symmetry. We established symmetry for the trading intensities from the symmetry in the matching function. These property carries through to the market rates and the yield coefficients:

Proposition 7 (Symmetry)**.** *The OTC market satisfies:*

$$\Sigma\left(\tau, \theta, \eta, \bar{\lambda}\right) = \Sigma\left(\tau, \theta^{-1}, 1-\eta, \bar{\lambda}\right), \tag{20}$$

$$r^f\left(\tau, \theta, \eta, \bar{\lambda}\right) = \left(r^w - r^m\right) - r^f\left(\tau\theta^{-1}, 1-\eta, \bar{\lambda}\right), \tag{21}$$

$$\chi^-\left(\tau, \theta, \eta, \bar{\lambda}\right) = \left(r^w - r^m\right) - \chi^+\left(\tau, \theta^{-1}, 1-\eta, \bar{\lambda}\right) \tag{22}$$

$$\chi^+\left(\tau, \theta, \eta, \bar{\lambda}\right) = \left(r^w - r^m\right) - \chi^-\left(\tau, \theta^{-1}, 1-\eta, \bar{\lambda}\right). \tag{23}$$

This symmetry states that if we can reverse the market tightness and exchange the bargaining power of surplus and deficit sides, the surplus function is the same. For the expected cost of being in deficit, the benefit of a surplus, and the OTC rate, the same symmetry holds relative to $r^w - r^m$. Panel (a) of Figure 3 plots the liquidity-yield coefficients for various values of η and $\log\theta_0$ in the x axis. A rotation of 180 degrees around the center of the figure gives the same figure once we swap η for $(1-\eta)$. A similar pattern holds for the average OTC rate. This property reveals that any asymmetry in outcomes follows from asymmetries in assumed bargaining powers or the tightness. This symmetry property implies that we can solve the model for $\theta < 1$ and obtain solutions for $\theta > 1$, immediately. Furthermore, as we will see, many properties change exactly at $\theta = 1$, due to this property.

Bargaining Power. Next, we characterize how the borrower's bargaining affects outcomes.

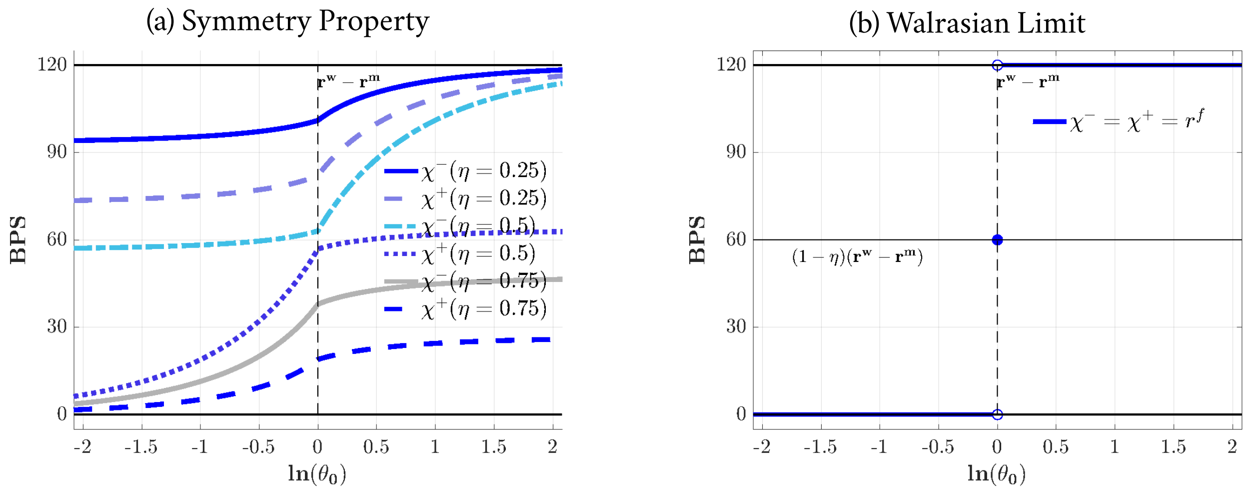

Figure 3: Symmetry and Walrasian Limit Properties

Note: OTC rate and convenience yield coefficients as functions of θ_0. Trading at various rounds. Note: $\theta_0 = \theta$ is the initial market tightness defined as the ratio of the initial aggregate deficit and initial aggregate surplus. The example in panel (a) is calibrated using $\bar{\lambda} = 0.8$, $r^w - r^m = 120$bps. Panel (a) is calibrated with $\eta = 0.5$, $r^w - r^m = 120$ bps.

Proposition 8 (Role of Bargaining Power)**.** *The equilibrium objects $\{\chi^+, \chi^-, \bar{r}^f\}$ are decreasing in η. In addition, at the extremes, we have that*

i) $\eta = 1$: $\bar{r}^f = r^m$, $\{\chi^+, \chi^-\} = \{0, (1 - \Psi^-)(r^w - r^m)\}$,

ii) $\eta = 0$: $\bar{r}^f = r^w$, $\{\chi^+, \chi^-\} = \{\Psi^+(r^w - r^m), (r^w - r^m)\}$.

Thus, as η increases, giving more bargaining power to borrowers, rates at every round fall. Thus, liquidity yields are decreasing. Regarding extrema, as borrowers extract all the matching surplus, rates go to the reserve rate, and the average cost of deficits is the penalty rate multiplied by the probability of matching in any round, $\chi^- = (1 - \Psi^-)(r^w - r^m)$; the opposite extreme follows by symmetry.

Limiting cases: matching efficiency. We also derive the limiting properties as the market efficiency is taken to its extreme values.

Proposition 9 (Efficiency Limits)**.** *The OTC market equilibrium satisfies:*

Walrasian: *As $\bar{\lambda} \to \infty$, the OTC market converges to its Walrasian market:*

i) If $\theta = 1$, we have

$$\Psi^+ = \Psi^- = 1, \quad \chi^+ = (r^w - r^m)(1 - \eta), \quad \chi^- = (r^w - r^m)(1 - \eta).$$

ii) If $\theta > 1$, we have

$$\Psi^{+} = 1, \Psi^{-} = \theta^{-1}, \quad \chi^{+} = \chi^{-} = (r^{w} - r^{m}), \quad \bar{r}^{f} = r^{w}.$$

iii) If $\theta < 1$:

$$\Psi^{+} = \theta, \quad \Psi^{-} = 1, \quad \chi^{+} = \chi^{-} = 0, \quad \bar{r}^{f} = r^{m}.$$

Static: *As $\bar{\lambda} \to 0$, the OTC market converges to a static bargaining: $\Psi^{+} = \Psi^{-} = 0$ and*

$$\chi^{+} = 0, \quad \chi^{-} = (r^{w} - r^{m}), \quad \bar{r}^{f} = r^{m} + (r^{w} - r^{m})(1 - \eta).$$

Proposition 9 shows that as efficiency increases, the OTC market rate approaches a Walrasian limit. In the Walrasian limit, if the market features an aggregate cash (scarcity of funds) deficit, $\theta > 1$, the average OTC market rate converges to the borrowers' outside option, r^{w}. In the opposite case, as $\theta < 1$, the rate converges to r^{m}. Likewise, the liquidity yields converge to the terminal trading surplus (in the case of an aggregate cash deficit) and zero (otherwise). Panel (b) of Figure 3 plots the rates and liquidity yields at the Walrasian limit.[18] On the other hand, when matching efficiency approaches zero, rates and yields converge to those of the static bargaining, and trade volume vanishes.

Market Tightness. Another property of interest regards how the convenience yield function varies with market tightness. We have the following corollary:

Corollary 3 (Monotonicity in market tightness)**.** *In any OTC market equilibrium, $\left\{\chi^{+}, \chi^{-}, \bar{r}^{f}\right\}$ are increasing in θ.*

The monotonicity of $\left\{\chi^{+}, \chi^{-}, \bar{r}^{f}\right\}$ is intuitive. Market tightness captures the relative size of settlement deficits. As deficits increase, lenders charge more for their funds and can match with greater probability. Borrowers pay higher interest rates on their OTC borrowings and are more likely to borrow from the last resource. The change from concavity to convexity around $\theta = 1$ follows immediately from the symmetry properties described above.

Next, we investigate the limit as market tightness approaches its extrema. By symmetry, it is sufficient to discuss the limit $\theta_0 \to 0$. Understanding this limit is important because it tells how the market behaves as the *aggregate* cash deficit vanishes. This limit highlights some subtle features of the

[18]In the knife-edge case where $\theta = 1$, the average rate is an average of outside options weighted by the bargaining power. In either case, the trading probability equals 1 for the shortest side of the market.

model. One might initially expect that as one side of the market vanishes—say, as $\theta_0 \to 0$—the short side (i.e., borrowers) would capture the full surplus, leading rates and yield coefficients to approach zero. However, the actual limiting behavior is more nuanced, and this intuition holds only under specific conditions.

Whether the shortest side of the market extracts all surplus depends on whether the intensive form of the matching function $\gamma(\cdot)$ is bounded above. Let $\bar{\gamma} \equiv \lim_{\theta \to 0} \gamma(\theta^{-1})$. Clearly, $\bar{\gamma}$ is bounded for some matching functions, e.g., for the harmonic mean, $G(a,b) = \left(\frac{1}{2}a^{-1} + \frac{1}{2}b^{-1}\right)^{-b}$ but not for others, e.g., the Cobb-Douglas, $G(a,b) = a^{1/2}b^{1/2}$. This bound is critical in determining the decay rate of θ as θ vanishes:

$$\frac{\dot{\theta}}{\theta} = -\bar{\lambda}\bar{\gamma} \quad \text{as} \quad \theta \to 0.$$

When the decay rate is finite, market tightness behaves as an exponentially decaying function for θ close to zero. That is, the deficit side of the market approaches zero, but is positive for all θ as $\theta \to 0$. By contrast, if $\bar{\gamma}$ is unbounded, the decay rate explodes, so the market tightness approaches zero extremely fast as $\theta \to 0$, leading to a singularity point in the ODE where θ actually reaches zero. Thus, θ_τ in that case may approach zero in finite time. This rate of decay, in turn, governs the behavior of the OTC rates at the extrema.

Proposition 10 (Tightness Limits)**.** *The convenience-yield function has the following limiting behavior:*

i. $\theta \to 0$*:*

$$\Psi^+ = 0, \qquad \Psi^- = 1 - e^{-\bar{\lambda}\bar{\gamma}},$$
$$\chi^+ = 0, \qquad \chi^- = (r^w - r^m)e^{-\bar{\lambda}\bar{\gamma}\eta}.$$

ii. $\theta \to \infty$*:*

$$\Psi^+ = 1 - e^{-\bar{\lambda}\bar{\gamma}}, \qquad \Psi^- = 0,$$
$$\chi^+ = (r^w - r^m)\left(1 - e^{-(1-\eta)\bar{\lambda}\bar{\gamma}}\right), \qquad \chi^- = r^w - r^m.$$

As tightness approaches zero, the lenders' matching probability approaches zero. Thus, $\chi^+ \to 0$. However, for the deficit side, the short side of the market, whether the overall probability of matching in the remaining rounds approaches one or remains less than one, depends on the asymptotic decay rate, $\bar{\gamma}$. If this decay rate is finite, although the deficit side is negligible relative to the surplus side, the matching probability converges, $1 - e^{-\bar{\lambda}\bar{\gamma}}$, a number strictly less than one. As a result χ^- does not vanish. If the asymptotic decay rate is unbounded, the matching probability does converge to one and, in that case, χ^{-} will vanish. By symmetry, the opposite occurs in the limit as $\theta \to \infty$. These properties reflect on the OTC market rate.

Notice that if the asymptotic decay rate is infinite, as market tightness vanishes, the average OTC rate approaches zero, which is akin to giving all the bargaining power to the deficit side. Instead, if the asymptotic decay rate is finite, a trader with surplus is able to extract some surplus in the improbable event that it gets to match. This occurs because the deficit side has a positive probability of not finding matches when $\bar{\gamma}$ is bounded. By symmetry, the opposite occurs as the market tightness explodes.

3.3 CES (Generalized Means) Matching Functions

The properties above apply to all matching functions that satisfy Assumption 1. A special case is the constant-elasticity of substitution (CES) class (also known as generalized means):[19]

$$G(a,b;p) = \left(\frac{1}{2}a^{p} + \frac{1}{2}b^{p}\right)^{\frac{1}{p}},$$

where concavity requires $p \leq 0$.[20] For $p = 0$, the matching function converges to a Cobb-Douglas (geometric mean), $G(a,b;p) = a^{1/2}b^{1/2}$, while for $p = -\infty$, it converges to the Leontief matching function, $\min\{a,b\}$.[21]

Figure 4 compares CES matching functions. Panel (a) plots $\dot{\theta}/\theta$ as a function of θ, showing that—except for the Cobb-Douglas case—the growth rate converges, indicating finite decay rates. Panel (b) traces the time path θ_τ from a common initial condition θ_0. Only under Cobb-Douglas does θ_τ reach zero; in all other cases, tightness decays toward zero but never vanishes in finite time.[22] The plots illustrate Proposition 3: tightness decays faster for higher values of p.

Panels (c) and (d) show liquidity yields as a function of $\log(\theta_0)$. Only for the Cobb-Douglas case the yield coefficients vanish (i.e., $r^w - r^m = 0$) for sufficiently small θ_0. This underscores that Cobb-Douglas is a knife-edge case within the CES family ($p \leq 0$), uniquely featuring infinite decay. Hence, it is the only specification where OTC rates and liquidity premia can be fully extinguished in finite time. In that sense, it delivers full cash-satiation: all surpluses are reallocated across trading rounds,

[19] A Theorem by Kolmogorov states that the only functions that satisfy symmetry and monotonicity satisfy that $G(a,b) = g^{-1}\left(\frac{1}{2}g(a) + \frac{1}{2}g(b)\right)$, for some monotone g. A special case of such functions is the CES class.

[20] The coefficient relates to the elasticity of substitution $\rho \geq 0$ in production via $p = 1/(1-\rho)$. For $p \leq 0$, we have $2^{1/p}G(a,b;p) \leq \min(a,b)$, so $\lambda_N < 2^{1/p}$ guarantees weak exhaustion in the finite rounds case.

[21] Other common special cases are, $p = -1$, so that the matching function becomes the harmonic mean: $G(a,b;p) = 2\left(\frac{1}{a^{-1}+b^{-1}}\right)$. It is also known that for $q > p$, $G(a,b;p) < G(a,b;q)$.

[22] Indeed, the solution to θ_τ may reach zero in some finite time $\tau < 1$ if and only if $p = 0$ (the matching function is Cobb-Douglas).

unlike a situation where there is no trade simply because no investors have a deficit.

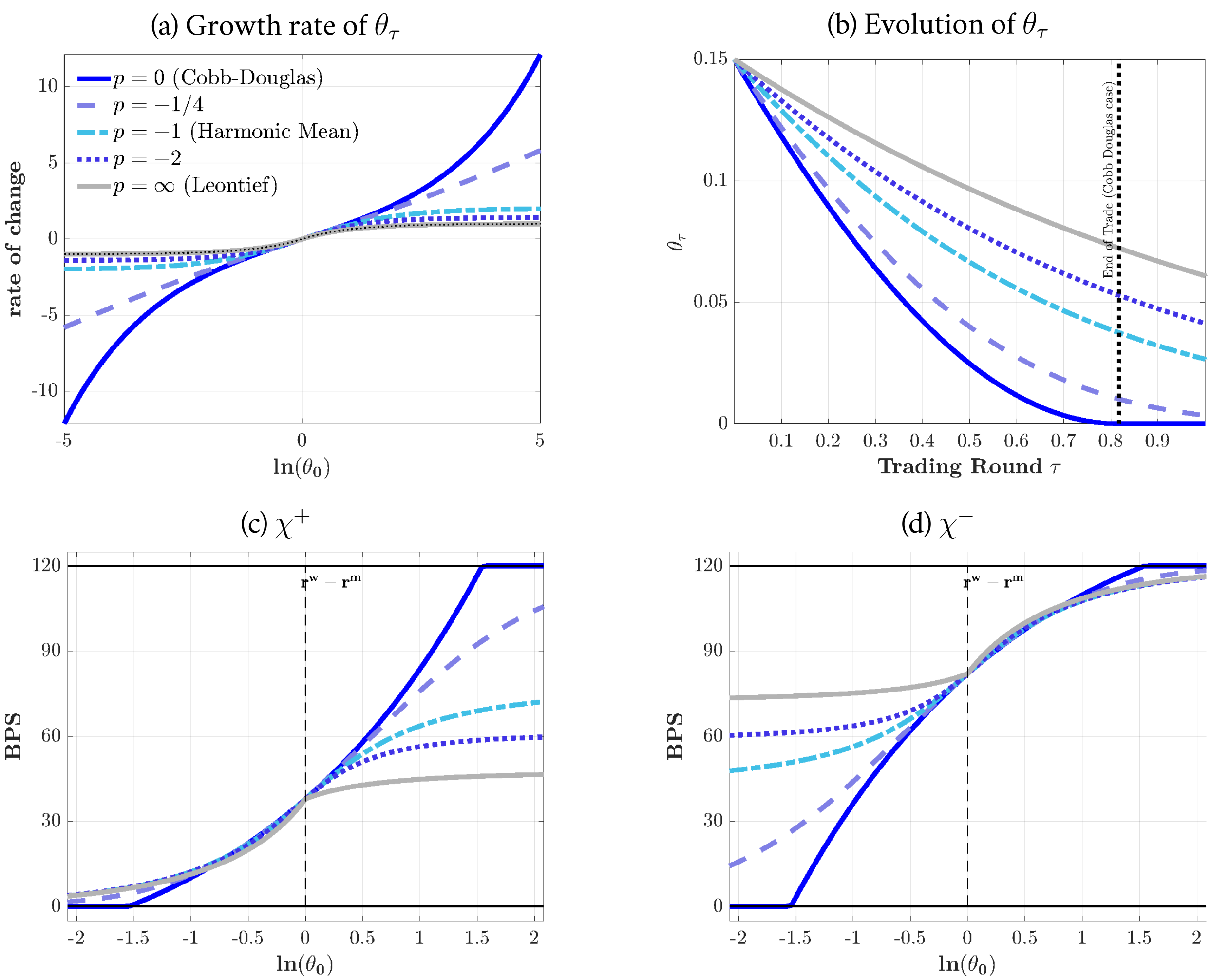

Figure 4: Comparison Across CES matching functions.

Note: $\theta = \theta_0$ is the initial market tightness. The example is calibrated with $\eta = 0.5$, $\bar{\lambda} = 1.2$, $r^w - r^m =$ 120bps in all cases.

Analytical Cases: Cobb-Douglas and Leontief. We are not aware of general solutions to the ODE for tightness (16). However, when the matching function is either Cobb-Douglas or Leontief, the evolution of market tightness can be characterized analytically.[23] Table 1 presents the corresponding solutions.

The Cobb-Douglas and Leontief cases display qualitatively distinct trading dynamics. Under Leontief matching, the short side's trading probabilities are exponentially distributed, independently

[23]A closed form can also be found for the Harmonic matching function.

of the initial tightness. Thus, θ_τ never vanishes fully throughout the trading session. Instead, θ_τ follows a logistic path. In the Cobb-Douglas case, the short side of the market may vanish before the trading time is over, if the stop time T in Table 1 is $T < 1$; evaluating the formula for θ_τ at the stopping time, $\tau = T$, yields zero or infinity, which is consistent with the end of trade.[24] In particular, under Cobb-Douglas matching, market tightness features the following *Walrasian thresholds*:

$$\theta^{+}(\bar{\lambda}) \equiv \left(\frac{e^{\bar{\lambda}}-1}{e^{\bar{\lambda}}+1}\right)^{2} \quad \text{and} \quad \theta^{-}(\bar{\lambda}) \equiv \frac{1}{\theta^{-}(\bar{\lambda})} = \left(\frac{e^{\bar{\lambda}}+1}{e^{\bar{\lambda}}-1}\right)^{2}.$$

If the initial tightness is sufficiently unbalanced:

$$\theta_0 \notin \left[\theta^{-}(\bar{\lambda}), \theta^{+}(\bar{\lambda})\right],$$

trade will stop before the final trading round and the OTC rates will coincide with those of the Walrasian limit. Equivalently, the Cobb-Douglas case exhibits *a Walrasian efficiency threshold*:

$$\bar{\lambda}^{\star}(\theta_0) \equiv \frac{1}{2}\log\left(\left|\frac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}}\right|\right),$$

such that trade vanishes in finite time if $\bar{\lambda} < \bar{\lambda}^{\star}(\theta_0)$.

The behavior of the yield coefficients and the average rate as a function of $\log(\theta_0)$ is depicted in Figure 5. Panels (a) and (b) depict the outcomes as functions of the initial condition for the Leontief and Cobb-Douglas cases. Both figures have sigmoid-like patterns, but the differences regarding concavity and the limits as θ moves to its extreme are clear. In the Cobb-Douglas case, as market tightness becomes sufficiently unbalanced, falling outside of the interval of the Walrasian limits, the short side of the market trades is fully matched. In such cases, the outcome coincides with the Walrasian limit, and the OTC rate corresponds to the outside option of the short side. That is, if investors in surplus exceed those in deficit, the OTC rate equals r^w. Conversely, if investors in deficit exceed those in surplus, the OTC rate equals r^m.

[24] The formula for θ_τ is actually associated with the hyperbolic tangent function, thereby featuring singularities.

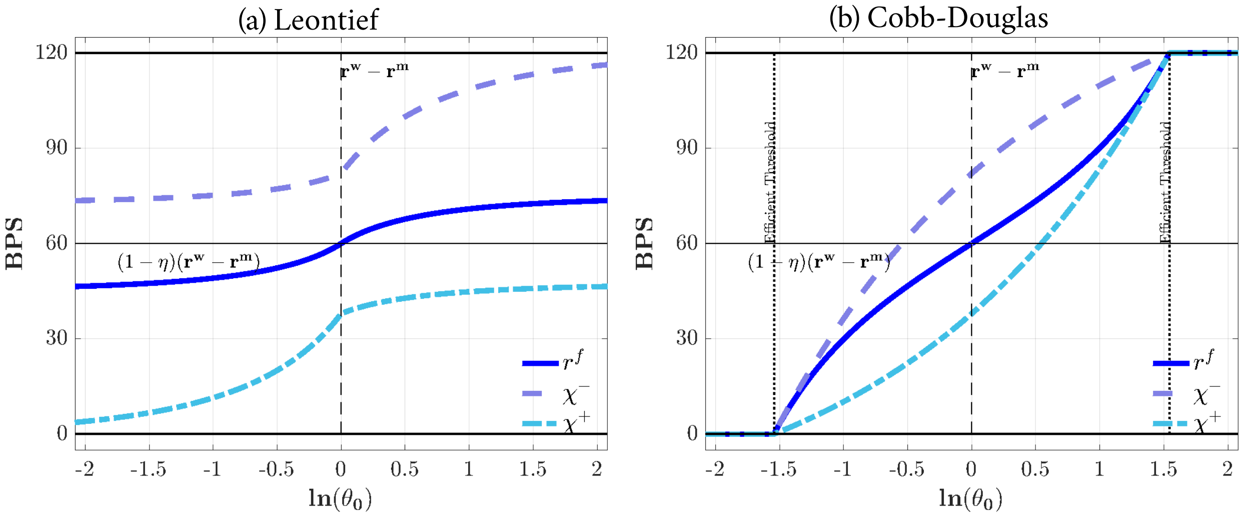

Figure 5: Analytic Solutions: OTC rates and yield coefficients for Leontief and Cobb-Douglas Matching Functions.

Note: The OTC rate and convenience yield coefficients are plotted as functions of θ_0. $\theta_0 = \theta$ is the initial market tightness, defined as the ratio of the initial aggregate deficit and initial aggregate surplus. Both panels are calibrated using $\eta = 0.5$, $\bar{\lambda} = 1.2$, $r^w - r^m = 120$bps.

Table 1: Analytical Formulas:

Matching	Cobb-Douglas ($p = 0$)	Leontief ($p = -\infty$)
$\boldsymbol{\theta(\tau),\ \tau \in [0, T]}$	$\left(\dfrac{\left(1+\sqrt{\theta_0}\right)e^{-\bar{\lambda}t} - \left(1-\sqrt{\theta_0}\right)}{\left(1+\sqrt{\theta_0}\right)e^{-\bar{\lambda}t} + \left(1-\sqrt{\theta_0}\right)}\right)^2$	$\begin{cases} 1+(\theta-1)\,e^{\bar{\lambda}\tau}, & \theta > 1 \\ \dfrac{\theta_0}{\theta_0+(1-\theta_0)e^{\bar{\lambda}\tau}}, & \theta < 1 \end{cases}$
$\boldsymbol{T}$	$\min\left\{\dfrac{1}{\bar{\lambda}}\log\left(\left\vert\dfrac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}}\right\vert\right), 1\right\}$	∞
$\boldsymbol{\Psi^+}$	$1-e^{-\bar{\lambda}T}\left(\dfrac{\left(1+\sqrt{\theta_0}\right)+\left(1-\sqrt{\theta_0}\right)e^{\bar{\lambda}T}}{\left(1+\sqrt{\theta_0}\right)+\left(1-\sqrt{\theta_0}\right)}\right)^2$	$\begin{cases} 1-e^{-\bar{\lambda}}, & \theta_0 \geq 1 \\ \theta_0(1-e^{-\bar{\lambda}}), & \theta_0 < 1 \end{cases}$
$\boldsymbol{\Psi^-}$	$1-e^{-\bar{\lambda}T}\left(\dfrac{\left(1+\sqrt{\theta_0}\right)-\left(1-\sqrt{\theta_0}\right)e^{\bar{\lambda}T}}{\left(1+\sqrt{\theta_0}\right)-\left(1-\sqrt{\theta_0}\right)}\right)^2$	$\begin{cases} (1-e^{-\bar{\lambda}})\theta_0^{-1}, & \theta_0 > 1 \\ 1-e^{-\bar{\lambda}}, & \theta_0 \leq 1 \end{cases}$

Note: Closed-form solutions for $\theta(t)$, stopping time T, and matching probabilities Ψ^+,Ψ^-.

4. Applications

4.1 Portfolio choices and convenience yields

We now study optimal portfolios in the presence of settlement risk. We maintain the assumption that all returns are exogenous, except for the endogenous OTC rate $\bar{R}^f$. Investor preferences are represented by

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_t), \tag{24}$$

where $\beta < 1$ is the time discount factor, and $u(c) = \frac{c^{1-\gamma}-1}{1-\gamma}$ is the utility function over the consumption good with $\gamma \geq 0$. As described in Section 2, each period, investors start with an initial wealth e, which depends on previous portfolio choices and the realization of returns, and choose their portfolio decisions subject to the budget constraint (1) We present below the problem of the investor in recursive form. We use X to denote an aggregate state.

Problem 2 (Investor's Problem)**.** *The savings-portfolio problem is:*

$$V_t(e, X) = \max_{\left\{c, \left\{\tilde{a}^i_{t+1}\right\}_{i\in\mathbb{I}}, m\tilde{_{t+1}}\right\}} u(c) + \beta\mathbb{E}\left[V_{t+1}(e_{t+1}, X')\right], \tag{25}$$

subject to: (1), (3), *and* (6).

In the investor's problem, $\mathbb{E}$ denotes the expectation operator with respect to the idiosyncratic liquidity shock in the settlement stage in the current period and the next period's realization of returns. Because preferences are homothetic and constraints linear, the problem admits portfolio separation: consumption (or dividends) can be chosen independently of portfolio weights, which themselves depend only on returns and settlement risks, not on total equity.[25]

The investor chooses portfolio weights on assets $\{a^i\}$ and the cash asset m to maximize the expected utility of returns, incorporating both standard asset returns and convenience yields arising from settlement frictions.[26]

$$\max_{m, \left\{a^i\right\}_{i\in\mathbb{I}}} \left(\mathbb{E}\left[\sum_{i\in\mathbb{I}} R^i_{t+1}\left(X'\right) a^i + R^m_{t+1}(X')m + \chi_{t+1}\left(s\left(\left\{a^i\right\}_{i\in\mathbb{I}}, m, \left\{\omega^i_{t+1}\right\}_{i\in\mathbb{I}}\right); \theta_t\right)\right]^{1-\gamma}\right)^{\frac{1}{1-\gamma}} \tag{26}$$

[25]See Proposition 3 of Bianchi and Bigio (2022) for a full characterization.

[26]The weights are defined as portfolio holdings relative to equity. With abuse of notation, we denote the weights with the same notation as the holdings.

subject to:

$$\sum_{i \in \mathbb{I}} a^i + m = 1$$

where we make explicit the dependence of the convenience yield on the market tightness, with a slight abuse of notation.

Figure 6 illustrates the forces underlying the portfolio problem. Panel (a) of Figure 6 shows how portfolio choice affects settlement risk through the convenience yield function $\chi(s;\theta)$. The purple kinked function represents $\chi(s;\theta)$, which maps settlement positions s into additional payoffs. When $s < 0$ (settlement deficit), investors face average borrowing costs χ^-; when $s > 0$ (surplus), they earn lending returns with slope χ^+. The asymmetry in the yield coefficients reflects higher penalty rates for emergency borrowing than returns on surplus lending. The panel shows two settlement distributions arising from different portfolio choices. The high-risk portfolio (red, dashed) generates a wide distribution of settlement needs, with significant probability mass in deficit regions. The low-risk portfolio (blue, solid) concentrates settlement positions at higher values. Due to the kink at zero, the high-settlement risk portfolio incurs expected losses even when distributions have zero mean. This illustrates a key insight: assets that generate volatile settlement needs command convenience yields to compensate for these expected losses, even for risk-neutral investors. Panel (b) of Figure 6 illustrates shows how individual portfolio choices affects the liquidity yield function$\chi(s;\theta)$ of others.

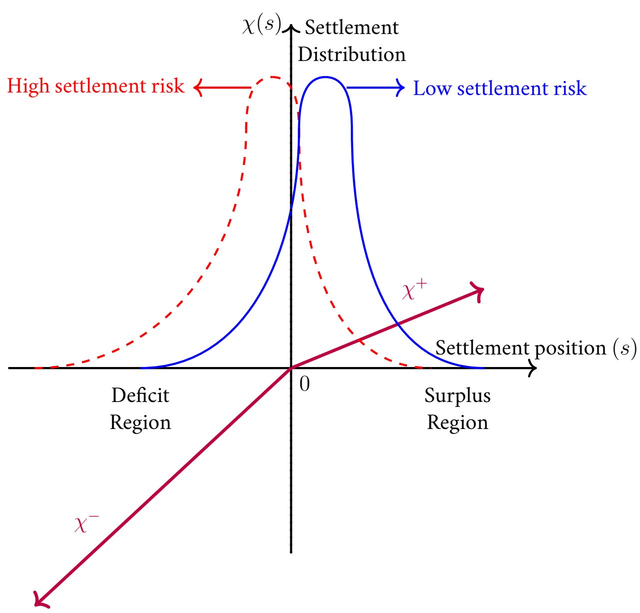

Figure 6: Convenience yield function $\chi(s)$ and settlement risk.

Convenience yields. Let us define $\chi_s \equiv \frac{\partial \chi}{\partial s}$ which is equal to χ^+ if $s \geq 0$ and χ^-otherwise. Taking first-order conditions in the portfolio problem (26), and assuming strictly interior portfolio positions, we obtain

$$\underbrace{\mathbb{E}_X\left[R^i\right]-R^m}_{\text{premium}} = \underbrace{\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial m}-\chi_s\frac{\partial s}{\partial a^i}\right]}_{\text{first-order liquidity premium}} - \underbrace{\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma},\chi_s\frac{\partial s}{\partial m}-\chi_s\frac{\partial s}{\partial a^i}\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]}}_{\text{liquidity risk premium}}$$
$$-\underbrace{\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma},R^i\left(X\right)\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]}}_{\text{conventional risk premium}}. \tag{27}$$

The optimal portfolio equalizes the marginal utility-weighted returns across assets, accounting for both price risk and liquidity risk due to settlement. The left-hand side is the difference in the expected return of asset i relative to cash. At an optimal solution, the representation in (27) shows that this asset's premium equals the sum of a convenience yield (liquidity premium) and a conventional risk premium. Thus, the equation captures a trade-off between expected return differentials against settlement risk and a conventional risk premium. The former emerges because asset a^i exposes the investor to settlement risk, whereas a larger m provides liquidity. In particular, in the case of a negative settlement shock, the investor obtains a higher return on cash compared to a less liquid asset and vice-versa. Assets that induce greater liquidity risk command a greater premium.

The liquidity premium term can be further unpacked into two terms: The first term captures how changes in the portfolio affect the expected settlement costs.[27] This term is present under risk neutrality. The terms are negative even when the cash position is zero in expectation because the liquidity yield function is concave. In turn, the second term captures the covariance term between liquidity payoffs and the discount factor: when the investor is risk-averse, states with negative returns penalize additional settlement costs.

There are some important lessons here: First, settlement risks induce determinate portfolios even among risk-neutral investors. Unlike standard portfolio problems where the riskiness of assets is given, here, by choosing its cash position, investors control the amount of risk. Given the concavity of χ, investors must be compensated with returns in order to hold portfolios that are more exposed to settlement risk. Second, the standard risk-adjustment insufficient to price the risk associated with an asset. Crucially, without considering the cash needs induced by an asset, the return in payoffs

[27]Note that χ is differentiable almost everywhere. While changing the portfolio affects, at the margin, the probability of being short or long, this does not enter into the optimality condition because the bank obtains the same marginal payoff (i.e., zero) evaluated at the threshold ω^*. Using Leibniz's, we can show that the effect disappears given that $s=0$ when the shock is ω^*.

is not enough to capture the full extent of risk. Conversely, convenience yields cannot be treated as pricing factors independent of risk, because the correlation between the riskiness of the asset and the liquidity needs must be considered. Another possible application of the theory is to investigate how the equity premium puzzle (Merha and Prescott, 1985) is impacted by the presence of convenience yields.

4.2 Parameter Identification

An important empirical question is how to identify the underlying OTC market parameters and settlement risk from observable data? Our analytical formulas reveal that the average OTC rate $\bar{r}^f$ and the liquidity yield coefficients $\{\chi^+, \chi^-\}$ depend on three key elements: market tightness θ (which, given portfolios, reflects the distribution of unobservable settlement shocks), matching efficiency $\bar{\lambda}$, and bargaining power η. However, these parameters are not directly observable. This section develops an approach to infer them from two potential sources of data: (i) convenience yields on liquid assets, which reveal the shadow value of liquidity, and (ii) OTC market outcomes, particularly rate dispersion and trading volumes.

An estimation of these structural parameters is important because it helps quantify the sources of variation in convenience yields over time and across assets. Furthermore, this estimation enables counterfactual analyses. In what follows, we focus for the most part on the Leontief matching function, which yields particularly sharp comparative statics that facilitate identification.

Market tightness. It is worth recalling a key result from our earlier analysis. As established in Corollary 3, the convenience yield coefficients and the average OTC rate $\left\{\chi^+, \chi^-, \bar{r}^f\right\}$ are all monotonically increasing in market tightness θ. This monotonicity enables identification of θ from observable data. For instance, the convenience yield on illiquid bonds (29) depends on a weighted sum of the liquidity coefficients, which is also monotonic in θ given portfolios. Similarly, the average OTC rate increases monotonically with θ. Once we identify θ from these observables, we can back out parameters of the unobserved settlement shock distribution, such as the variance of withdrawal risk or the frequency of margin calls. Note that the net interest margin in equation (30) depends on a weighted difference of the yield coefficients, which need not be monotonic in θ.[28]

Matching efficiency. Let us now turn to the comparative static with respect to the efficiency parameter $\bar{\lambda}$.

[28]The curvature properties of the yield coefficients—whether convex or concave—may provide additional information.

Proposition 11 (Comparative Statics in Leontief Case)**.** *If the matching function is Leontief, we have that: i) Let $\theta < 1$, then χ^- is decreasing in $\bar{\lambda}$, $\overline{r}^f$ is decreasing in $\bar{\lambda}$, and χ^+ is non-montonic in $\bar{\lambda}$. ii) Let $\theta > 1$, then χ^+ is increasing in $\bar{\lambda}$, $\overline{r}^f$ is increasing in $\bar{\lambda}$, and χ^- is non-montonic in $\bar{\lambda}$. iii) Let $\theta = 1$, then*

$$\frac{\partial \chi^+}{\partial \bar{\lambda}} = (r^w - r^m)(1-\eta)\frac{\partial \Psi^+}{\partial \bar{\lambda}} > 0, \qquad \frac{\partial \chi^-}{\partial \bar{\lambda}} = -(r^w - r^m)\eta \frac{\partial \Psi^-}{\partial \bar{\lambda}} < 0.$$

The comparative static reveals a subtle but important insight: convenience yields may be non-monotonic in the matching efficiency. When markets are tight ($\theta > 1$), improving matching efficiency has two opposing effects: it increases the probability that deficit investors find lenders (reducing their borrowing costs), but it also strengthens lenders' bargaining position, raising equilibrium rates. The net effect on χ^- depends on which force dominates. This non-monotonicity means that higher convenience yields need not signal market dysfunction—they could reflect improved matching that benefits the scarce side of the market. The same non-monotonic pattern characterizes the average OTC rate $\overline{r}^f$. Figure 7 illustrates these relationships: Panel (a) confirms our theoretical results for the Leontief case, while Panel (b) demonstrates that this pattern extends to the Cobb-Douglas matching function, suggesting it is a robust feature of OTC markets outside of the Leontief case.

The non-monotonicity of rates and liquidity premia with respect to efficiency implies that convenience yields cannot be interpreted mechanically as evidence of worsening market efficiency. To see this, consider the first-order liquidity premium in equation (27) for an asset that does not expose investors to settlement risk, i.e., $\frac{\partial s}{\partial a^i} = 0$. In that case, the premium reduces to $\mathbb{E}X, \omega\left[\chi s\right]$. An increase in this premium could be consistent with two opposite interpretations. On the one hand, it might signal weaker OTC efficiency, raising investors' demand for cash as insurance against settlement needs. On the other hand, it could reflect greater efficiency, which enhances the option value of holding cash to exploit lending opportunities. Thus, higher observed convenience yields do not unambiguously point to deteriorating trading efficiency.

(a) Leontief (b) Cobb-Douglas

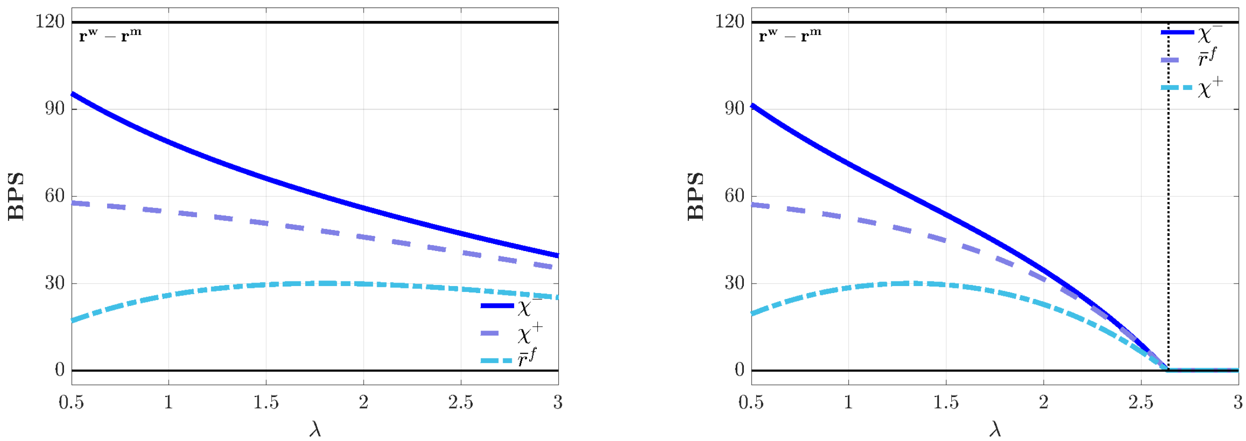

Figure 7: Effects of Matching Efficiency

Note: Rates and Yield Coefficients as functions of λ for Leontief and Cobb-Douglas Matching Functions. Note: The OTC rate and liquidity yield coefficients are plotted as functions of $\bar{\lambda}$. Both panels are calibrated using $\eta = 0.5$, $\theta_0 = 0.75$, $r^w - r^m = 120$bps.

OTC Rate Dispersion. The dispersion of OTC rates within each trading day—the difference between rates at the beginning and end of the trading session—provides another observable moment for parameter identification. Recall from Proposition **??** that the equilibrium rate varies during the trading session according to market evolution of the tightness θ_τ. Since rates change monotonically throughout the day, increasing (decreasing) with time if $\theta > 1$ ($\theta < 1$), we can measure dispersion in rates:

$$Q \equiv \max_\tau r^f_\tau - \min_\tau r^f_\tau = \left| r^f_1 - r^f_0 \right|.$$

We obtain the following relationship between parameters and the dispersion:

Corollary 4 (Comparative Statics: Rate Dispersion)**.** *Rate dispersion features the following comparative statics:*

i) If $\theta < 0$, then $\frac{\partial Q}{\partial \theta} < 0$;

ii) If $\theta > 1$, then $\frac{\partial Q}{\partial \theta} > 0$;

iii) If $\theta = 1$, then $\frac{\partial Q}{\partial \theta} = 0$;

iv) $\frac{\partial Q}{\partial \lambda} \geq 0$ with equality if and only if $\theta = 1$.

The intuition is straightforward. Rate dispersion reflects how much the market "unwinds" during the trading session. When the market is unbalanced ($\theta \neq 1$), the scarce side gets progressively better terms as trading proceeds, creating larger rate movements. The more unbalanced the market—in either direction—the greater the dispersion. Similarly, higher matching efficiency $\bar{\lambda}$ accelerates this unwinding process, amplifying dispersion. At $\theta = 1$, the market is balanced and rates remain constant throughout the session, yielding $Q = 0$.

Figure 8 illustrates these patterns. Panel (a) shows the symmetric U-shape of dispersion around $\theta = 1$ for the, while panel (b) demonstrates how dispersion increases with matching efficiency. Panels (c) and (d) show the corresponding mappings for the Cobb-Douglas case, demonstrating that these patterns are robust across matching technologies. These relationships provide additional moments for identification: observing both the level of OTC rates (which pins down θ) and their intraday dispersion Q (which helps identify $\bar{\lambda}$) can jointly identify OTC market parameters.

Relative Volumes of Market and Outside Funding. Define the relative volume as the ratio of lender-of-last-resort borrowing to OTC market borrowing:

$$I(\theta) \equiv \frac{1 - \Psi^{-}(\theta)}{\Psi^{-}(\theta)} = \begin{cases} \frac{e^{-\bar{\lambda}}}{1-e^{-\bar{\lambda}}} & \theta \leq 1 \\ \\ \frac{1-\left(1-e^{-\bar{\lambda}}\right)\theta^{-1}}{\left(1-e^{-\bar{\lambda}}\right)\theta^{-1}} & \theta > 1. \end{cases}$$

This ratio captures market efficiency in reallocating liquidity: if $I(\theta) = 0$ indicates perfect reallocation through the OTC market whereas $I(\theta) = \infty$ indicates no OTC volume at all. We call $I(\theta)$ the relative volume.

These results reveal a key property for identification: Given Proposition 1, the relative volume decreases monotonically with matching efficiency $\bar{\lambda}$ regardless of market tightness. This contrasts sharply with convenience yields, which can be non-monotonic in $\bar{\lambda}$. Moreover, when markets have excess liquidity ($\theta < 1$), relative volume depends only on $\bar{\lambda}$ and is independent of θ in the Leontief case. Combined with our earlier results, this suggests that the relative volume is convenient to pin down $\bar{\lambda}$.

Figure 9 illustrates these patterns for the Leontief case (panels a and b). Panels (c and d) demonstrate a similar pattern for the Cobb-Douglas case, demonstrating again that patterns are robust across matching technologies. Notice that in panel c, as the tightness crosses the threshold of efficient trade, indicated by the vertical lines, the relative volume is either zero (when deficits are the short side) or, otherwise, $\frac{S^{-}-S^{+}}{S^{+}}$. Likewise, in panel d, as efficiency crosses a threshold, the relative volume is zero.

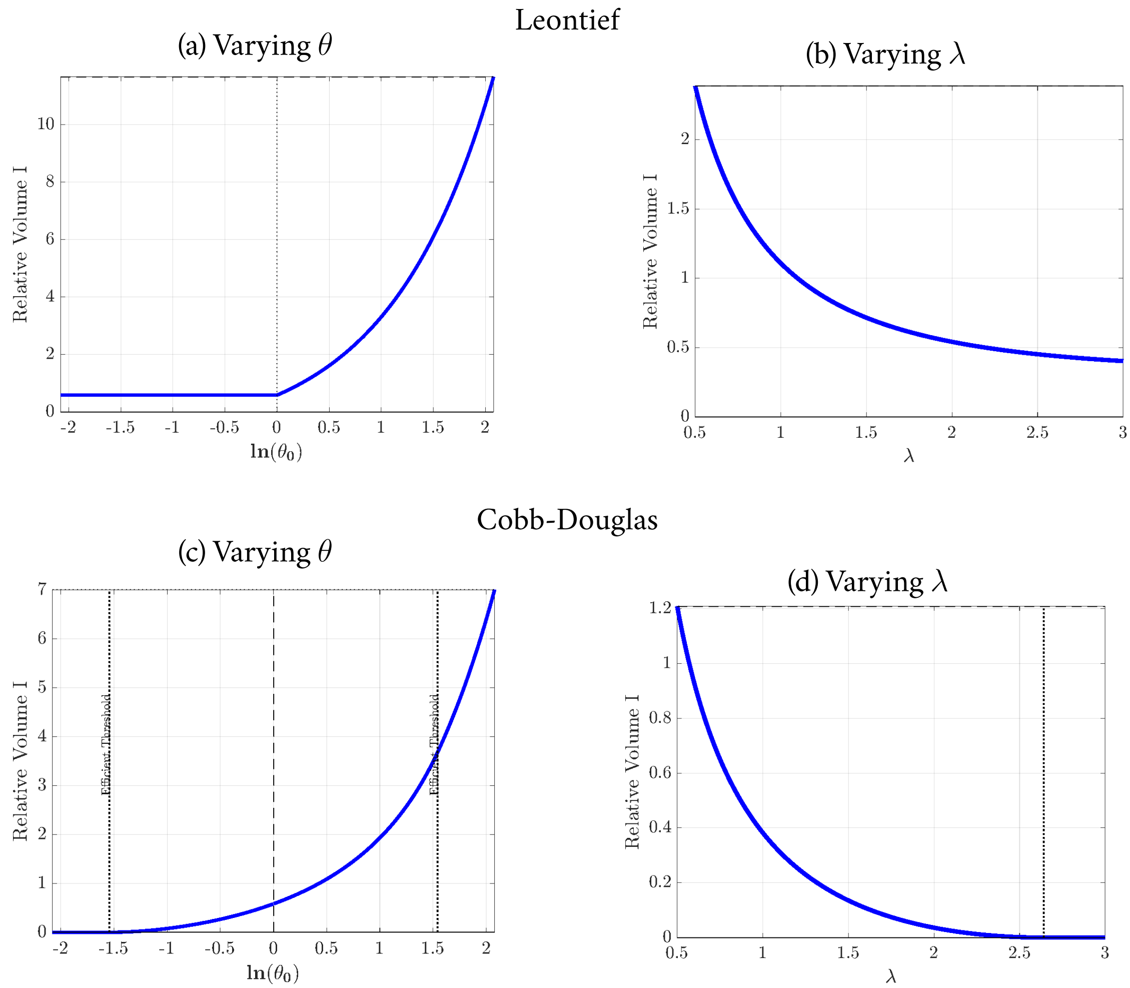

Figure 9: Relative Volume as Function of $\{\lambda, \theta_0\}$.

Note: The relative trading volume is plotted as a function of $\bar{\lambda}$ and θ_0 for the Leontief and Cobb-Douglas matching functions. Both panels are calibrated using $\eta = 0.5$, $r^w - r^m = 120$bps. When we vary $\bar{\lambda}$, we set $\theta_0 = 0.75$. When we vary θ_0 we set $\bar{\lambda} = 1.2$.

Summing up. Our comparative statics results suggest practical approaches for identifying the three key parameters of the model $\{\theta, \bar{\lambda}, \eta\}$ from observable data. The key insight is to exploit the different monotonicity properties to achieve robust identification: relative volume $I(\theta)$ appears to provide a clean first-stage identification of $\bar{\lambda}$ because it decreases monotonically with matching efficiency regardless of market conditions. Once $\bar{\lambda}$ is pinned down, the monotonic relationship between convenience yields and θ allows for second-stage identification. Crucially, since θ depends on both portfolios and the distribution of settlement shocks, observing portfolios allows us to infer the unobservable shock distribution Φ.

For empirical work, this framework suggests focusing data collection on: (i) portfolio compositions across institutions, (ii) spreads between liquid and illiquid assets, (iii) the level and intraday range of OTC rates, and (iv) the relative use of emergency lending facilities, if available. Together, these moments may sufficiently identify both the market microstructure driving convenience yields and the underlying distribution of liquidity shocks. We provide a sketch of an identification strategy next.

Algorithm 1 Suggested Identification Approach

Given possible observable data:

- Portfolio holdings across investors, $\{m, a^i\}$, convenience yields on liquid assets: $R^b - R^m$, an average OTC rate: $\bar{r}^f$, rate dispersion: $Q = |r_1^f - r_0^f|$, relative volume: I

Step 1: Identify matching efficiency $\bar{\lambda}$

-Use $I(\theta)$ to infer $\bar{\lambda}$ (monotonic for all θ)

Step 2: Identify market tightness θ and shock distribution Φ

-Given $\bar{\lambda}$ from Step 1:

-Use convenience yields or $\bar{r}^f$ to infer θ

-Given observed portfolios and implied θ, back out the distribution Φ of shocks (e.g., withdrawal risk parameters)

Step 3: Identify bargaining power η

-Consider the historical average of $\bar{r}^f$ relative to r^m and r^w
or examine the ratio χ^+/χ^- near balanced markets

Step 4: Validation

- Check implied $Q(\theta, \bar{\lambda})$ aligns with observed dispersion
- Assess consistency across different liquid assets

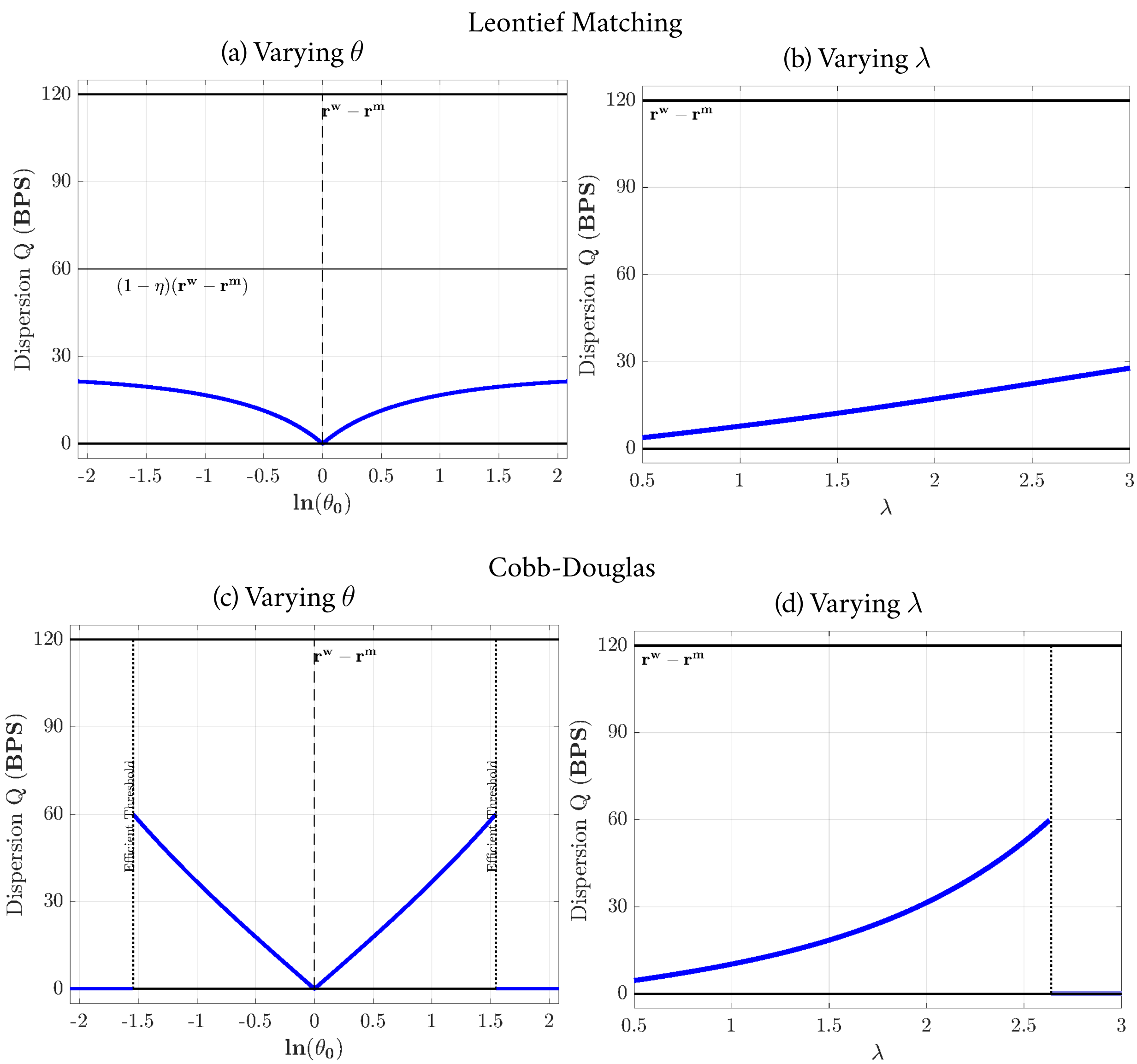

Figure 8: Dispersion of OTC Rates as function of $\{\lambda, \theta_0\}$.

Note: The dispersion of the OTC rate is plotted as a function of $\bar{\lambda}$ and θ_0 for the Leontief and Cobb-Douglas matching functions. Both panels are calibrated using $\eta = 0.5$, $r^w - r^m = 120$bps. When we vary $\bar{\lambda}$, we set $\theta_0 = 0.75$. When we vary θ_0 we set $\bar{\lambda} = 1.2$.

4.3 Normative Analysis

In this section, we study the implications of the theory for liquidity regulation. We do so by comparing the decentralized portfolio choices vis-à-vis a social planner's portfolio choices. We will show that by affecting the tightness in the OTC market and the degree of congestion, the portfolio decisions will be socially inefficient.

To illustrate the externality, we consider a simple example. We specialize the model to the case where investors' portfolios are composed of a liquid asset m, an illiquid asset b, and deposits d. In particular, deposits are subject to withdrawal risk such that (3) is given by

$$s\left(\{b,d\}, m\right) \equiv m + \left(\frac{R^d}{R^m}\omega - \rho(1+\omega)\right)d, \tag{28}$$

and their budget constraint is: $b+m=1+d$ where ω is distributed F with mean zero.

We study the individual investor's problem next and then consider the planner's counterpart.

Investor's problem The investor's problem, after substituting its budget constraint, thus becomes:

Problem 3 (Withdrawal Risk Problem). *The investor's problem in example is:*

$$\max_{m\geq 0, d\geq 0} \left\{\mathbb{E}_{\omega}\left[R^b + R^b(d-m) + \ R^m m - \ R^d d + \chi\left(s\left(d,m\right);\theta\right)\right]^{1-\gamma}\right\}^{\frac{1}{1-\gamma}},$$

In this simple example, the optimality condition (27), is simple and links the portfolio to a convenience yield, a spread, between illiquid assets and cash:

$$\underbrace{R^b - R^m}_{\text{liquidity premium}} = \chi^{+}\left(1-\tilde{\Phi}\left(\omega^{*}\right)\right) + \chi^{-}\tilde{\Phi}\left(\omega^{*}\right) = \chi^{+} + \Sigma\left(\theta\right)\tilde{\Phi}\left(\omega^{*}\right)\cdot \tag{29}$$

where

$$\tilde{\Phi}\left(\omega^{*}\right) = \underbrace{\underbrace{\Phi\left(\omega^{*}\right)}_{\text{deficit prob}} \cdot \underbrace{\frac{\mathbb{E}_{\omega}\left[R^e(\omega)^{-\gamma}|\omega<\omega^{*}\right]}{\mathbb{E}_{\omega}\left[R^e(\omega)^{-\gamma}\right]}}_{\text{risk-aversion correction}}}_{\text{risk-adjusted deficit probability}}, \quad \omega^{*} \equiv \frac{\rho - \dfrac{m}{d}}{\dfrac{R^d}{R^m} - \rho}.$$

Thus, the liquidity premium is a weighted average of the liquidity yield coefficients, $\chi^{+}\left(1-\tilde{\Phi}\left(\omega^{*}\right)\right) + \chi^{-}\tilde{\Phi}\left(\omega^{*}\right)$. In the expression, $\tilde{\Phi}\left(\omega^{*}\right)$ is the risk-adjusted probability of falling in a cash deficit.[29] The probability of a deficit is $\Phi\left(\omega^{*}\right)$ where ω^{*} is the threshold shock that puts banks in deficit—recall

[29]The risk adjustment scales probabilities by marginal utilities of wealth.

that lower values of ω represent an outflow of funds. Thus, the liquidity premium of cash over illiquid assets is their risk-adjusted expected return of coming into the OTC market with an extra unit of cash.

In the special case of risk-neutrality, we have that (29) becomes:

$$\underbrace{R^b - R^m}_{\text{liquidity premium}} = \chi^+ + (R^w - R^m)\left(1 - \Psi^-(\theta)\right)\Phi(\omega^*).$$

The corresponding difference between illiquid assets and deposits is an interest margin associated with the liquidity risk of deposits:

$$\underbrace{R^b - R^d}_{\text{liquidity premium}} = \chi^+ + \left(\chi^- - \chi^+\right)\tilde{\Phi}(\omega^*)\left(\left(\frac{R^d}{R^m} - \rho\right)\mathbb{E}_\omega\left[\omega R^e(\omega)^{-\gamma}|\omega < \omega^*\right] - \rho\right)\cdot \tag{30}$$

When deposits are riskier, as captured by the left tail of the distribution, this requires a higher return on loans.

Next, we discuss how features of the OTC market affect the liquidity yield coefficients and, consequently, the return premia. In turn, we discuss how OTC market frictions can be inferred from convenience yields or OTC market data to discipline the theory.

Constrained efficient allocation. We assume the planner chooses the portfolio shares (d, b, m) on behalf of investors while investors retain the choice of how much to save overall and how much to consume. The planner takes as given all portfolio returns and the structure of the OTC market, as encoded in the convenience yield function. Because investors are ex-ante identical, the problem can be expressed as follows.

Problem 4. *The planner's problem is:*

$$\max_{m\geq 0, d\geq 0}\left\{\mathbb{E}_\omega\left[R^b(d-m) + R^m m - R^d d + \chi\left(s\left(\omega, m, d\right), \theta(m,d)\right)\right]^{1-\gamma}\right\}^{\frac{1}{1-\gamma}},$$

subject to (28) *and the market tightness*

$$\theta(m,d) \equiv \frac{\int_{-1}^{\omega^*} s\left(\omega, m, d\right)\Phi(d\omega)}{\int_{\omega^*}^{\infty} s\left(\omega, m, d\right)\Phi(d\omega)}, \quad \omega^* \equiv \frac{\rho - \dfrac{m}{d}}{\dfrac{R^d}{R^m} - \rho}.$$

The key distinction is that the planner considers how the portfolio determines θ and how in turn this affects the convenience yield function χ. Notice that because all investors are identical ex ante, there are no redistribution considerations.

The planner's first-order conditions with respect to m and b yields the following first-order condition:

$$\underbrace{R^b - R^m}_{\text{asset premium}} = \chi^+ + \left(\chi^- - \chi^+\right) \cdot \tilde{\Phi}\left(\omega^*\right) + \frac{\partial\theta}{\partial m}\frac{\partial\chi^+(\theta)}{\partial\theta} \cdot (1 - \tilde{\Phi}(\omega^*)) \cdot \mathbb{E}\left[s \cdot \frac{R^e(\omega)^{-\gamma}}{\mathbb{E}\left[R^e(\omega)^{-\gamma} | \ \omega > \omega^*\right]} | \ \omega > \omega^*\right] + \frac{\partial\theta}{\partial m}\frac{\partial\chi^-(\theta)}{\partial\theta} \cdot \tilde{\Phi}(\omega^*) \cdot \mathbb{E}\left[s \cdot \frac{R^e(\omega)^{-\gamma}}{\mathbb{E}\left[R^e(\omega)^{-\gamma} | \ \omega < \omega^*\right]} | \ \omega < \omega^*\right]. \quad (31)$$

Just like individual investors, the planner trades off the higher return on loans with the liquidity benefits of cash, considering the uses of cash in the OTC market. However, the planner internalizes the pecuniary externality that emerges because it understands how portfolio choices will affect trading probabilities, as encoded in the convenience yield χ.

A key insight is that the sign of the externality is ambiguous: it depends on the derivatives of the convenience-yield coefficients. To understand whether investors over- or under-invest in liquid assets, it is useful to consider first the limiting case with risk-neutral investors. With $\gamma \to 0$, given that $\frac{\partial\theta}{\partial m} < 0, \frac{\partial\chi^+(\theta)}{\partial\theta} > 0, \frac{\partial\chi^-(\theta)}{\partial\theta} > 0$, we have that the *planner values cash more* than individual investors if and only if

$$\frac{\partial\chi^+(\theta)}{\partial\theta} \cdot S^+ > \frac{\partial\chi^-(\theta)}{\partial\theta} \cdot S^- \quad (32)$$

This inequality underscores that there is an under-accumulation of liquid assets when the planner perceives that higher cash holdings (lower market tightness) raise the marginal return on liquid assets more when in deficit, compared to the case in surplus. That is, when market tightness goes up, this favors investors that are in deficit (by allowing them to borrow at a lower rate and by raising the probability of a match) relative to investors that are in surplus (as they now must lend at a lower rate and face a lower matching probability).

Suppose that the planner picks a portfolio with $\theta = 1$. Under the case with Cobb-Douglas matching, we know that $\frac{\partial\chi^+(\theta)}{\partial\theta} = \frac{\partial\chi^-(\theta)}{\partial\theta}$. If, in addition, the shock is symmetric $F(\omega^*) = 0.5$ and $\mathbb{E}\left[-s \mid \omega < \omega^*\right] = \mathbb{E}\left[s \mid \omega \geq \omega^*\right]$.[30] In this case, it follows that there is neither over- nor under-accumulation of liquid assets. However, if investors were risk-averse, the risk adjustment correction would imply that by accumulating more liquid assets, the planner would effectively provide more insurance. Because individual investors do not internalize these benefits, the competitive equilibrium would feature under-accumulation of liquid assets. In addition, an allocation with a higher

[30]To see that $\frac{\partial\chi^+(\theta)}{\partial\theta}|_{\theta=1} = \frac{\partial\chi^-(\theta)}{\partial\theta}|_{\theta=1}$ under Cobb-Douglas we can exploit the symmetry property of derivatives.

probability of being in deficit $\Phi(\omega^*)$ or with a more sensitive convenience yield $\frac{\partial \chi^-(\theta)}{\partial \theta}$ implies that the planner perceives a higher value from higher liquid holdings.

It is also interesting to discuss the case with Leontief matching. In this case, while matching probabilities do not change with market tightness on the short-side of the market at the margin, aggregate cash holdings affect endogenously the outside options of investors, the OTC rate does change, and so does χ^+ and χ^-. Notice that if the OTC market were static, the OTC rate would be fixed, and so in this case, it would suffice to know whether the market features excess surplus or deficit to trace the sign of the inefficiency, under risk neutrality. In particular, if the market had, on average, excess surplus, the planner would value less cash than individual investors at the margin. Conversely, if the market had, on average, excess deficits, the planner would value more cash than individual investors at the margin. This follows because in the former case $\frac{\partial \chi^-(\theta)}{\partial \theta} = 0$ and $\frac{\partial \chi^+(\theta)}{\partial \theta} > 0$ while in the latter $\frac{\partial \chi^+(\theta)}{\partial \theta} = 0$ and $\frac{\partial \chi^-(\theta)}{\partial \theta} > 0$.

This externality has implications for prudential policy and liquidity regulation. Our results highlight that the optimal liquidity regulation does not necessarily involve minimum requirements on liquidity holdings. In fact, depending on market conditions, the competitive equilibrium may feature over or under investment in liquid assets.[31]

5. Conclusions

We develop a tractable microfoundation for convenience yields arising from trading frictions in OTC markets for settlement instruments and show how it can be readily introduced into a canonical portfolio problem. We further characterize how the convenience yield function depends on market tightness, bargaining power, and matching efficiency, and show that convenience yields reflect both direct OTC frictions and the interaction between liquidity and return risk. The framework generates closed-form expressions for rates and spreads that facilitate comparative statics and quantitative analysis. Finally, we show that individual investors fail to internalize how their portfolio choices influence aggregate market tightness, which leads to over- or under-investment in liquid assets depending on the level of tightness and the degree of risk aversion.

Our framework abstracts from several important features of real-world markets that offer fertile ground for future work. First, we do not model large market makers or networks (see e.g., Bech and Atalay, 2010), or the role of collateral, which may shape both the terms and scope of OTC trades. In

[31]We highlight that in our environment, the planner maximizes the welfare of the banks and the net proceeds from the trades at r^m and r^w are not rebated back to the banks. Ismail and Zuniga (2025) study the converse problem, where the planner internalizes the revenues from discount window borrowing and rebates them back to banks. Their analysis focuses on the welfare implications from an asset-demand perspective, i.e., the wedge between loan and deposit rates.

addition, we abstract from the multi-layered nature of real-world liquidity provision, where banks provide credit lines and deposits that serve as settlement instruments for non-bank financial institutions. Equipped with suitable data, our framework can contribute to a deeper understanding of the determinants of convenience yields and of asset prices more broadly.

References

Acharya, Viral V and Lasse Heje Pedersen, "Asset pricing with liquidity risk," *Journal of financial Economics*, 2005, *77* (2), 375–410.

Afonso, Gara and Ricardo Lagos, "The Over-the-Counter Theory of the Fed Funds Market: A Primer," *Journal of Money, Credit and Banking*, June 2015, *47* (2), 127–154.

— **and** — , "Trade Dynamics in the Market for Federal Funds," *Econometrica*, 2015, *83* (1), 263–313.

— **, Domenico Giannone, Gabriele La Spada, and John C. Williams**, "Scarce, Abundant, or Ample? A Time-Varying Model of the Reserve Demand Curve," June 2023. Federal Reserve Bank of New York Staff Reports 1019.

Amador, Manuel and Javier Bianchi, "Bank runs, fragility, and regulation," 2024. NBER Working Paper No. 32341.

Arce, Óscar, Galo Nuño, Dominik Thaler, and Carlos Thomas, "A Large Central Bank Balance Sheet? Floor vs Corridor Systems in a New Keynesian Environment," *Journal of Monetary Economics*, 2019.

Arseneau, David M. and David E. Rappoport, "Private and Public Liquidity Provision in Over-the-Counter Markets," Technical Report 2017-033 2017.

Atkeson, Andrew G., Andrea L. Eisfeldt, and Pierre-Olivier Weill, "Entry and Exit in OTC Derivatives Markets," *Econometrica*, 2015, *83* (6), 2231–2292.

Basak, Suleyman and Domenico Cuoco, "An Equilibrium Model with Restricted Stock Market Participation," *The Review of Financial Studies*, 1998, *11* (2), 309–341.

Bech, Morten L. and Enghin Atalay, "The topology of the federal funds market," *Physica A: Statistical Mechanics and its Applications*, 2010, *389* (22), 5223–5246.

Bhattacharya, Sudipto and Douglas Gale, "Preference Shocks, Liquidity, and Central Bank Policy," in William A. Barnett and Kenneth J. Singleton, eds., *New Approaches to Monetary Economics*, Cambridge: Cambridge University Press, 1987, pp. 69–88.

Bianchi, Javier, "Overborrowing and systemic externalities in the business cycle," *American Economic Review*, 2011, *101* (7), 3400–3426.

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

— **,** — **, and Charles Engel**, "Scrambling for Dollars: International Liquidity, Banks and Exchange Rates," 2023. Mimeo, Minneapolis Fed.

Bigio, Saki and Yuliy Sannikov, "A Model of Intermediation, Money, Interest, and Prices," 2019. Unpublished Manuscript.

Bittner, Christian, Rustam Jamilov, and Farzad Saidi, "Assortative matching, interbank markets, and monetary policy," Technical Report, ECONtribute Discussion Paper 2025.

Caballero, Ricardo J and Arvind Krishnamurthy, "International and domestic collateral constraints in a model of emerging market crises," *Journal of monetary Economics*, 2001, *48* (3), 513–548.

Constantinides, George M., "Capital Market Equilibrium with Transaction Costs," *Journal of Political Economy*, 1986, *94*, 842–862.

_ **and Darrell Duffie**, "Asset Pricing with Heterogeneous Consumers," *Journal of Political Economy*, 1996, *104* (2), 219–240.

Dávila, Eduardo and Anton Korinek, "Pecuniary externalities in economies with financial frictions," *The Review of Economic Studies*, 2018, *85* (1), 352–395.

Duffie, Darrell, Nicolae Garleanu, and Lasse Heje Pedersen, "Over-the-Counter Markets," *Econometrica*, 2005, *73* (6), 1815–1847.

Engel, Charles and Steve Pak Yeung Wu, "Liquidity and Exchange Rates: An Empirical Investigation," *The Review of Economic Studies*, 2023, *90* (5), 2395–2438.

Farhi, Emmanuel, Mikhail Golosov, and Aleh Tsyvinski, "A Theory of Liquidity and Regulation of Financial Intermediation," *The Review of Economic Studies*, 2009, *76* (3), 973–992.

Fiore, Fiorella De, Marie Hoerova, and Harald Uhlig, "Money Markets, Collateral and Monetary Policy," 2018. NBER Working Paper 253190.

Geanakoplos, John and Kieran James Walsh, "Inefficient Liquidity Provision," *Economic Theory*, 2017. Advance online publication.

He, Zhiguo and Konstantin Milbradt, "Endogenous liquidity and defaultable bonds," *Econometrica*, 2014, *82* (4), 1443–1508.

Holmstrom, Bengt and Jean Tirole, "LAPM: A Liquidity Based Asset Pricing Model," *Journal of Finance*, October 2001, *56* (5), 1837–1867.

Hugonnier, Julien, Benjamin Lester, and Pierre-Olivier Weill, "Heterogeneity in decentralized asset markets," *Theoretical Economics*, 2022, *17* (3), 1313–1356.

_ **,** _ **, and** _ , *The Economics of Over-the-Counter Markets: A Toolkit for the Analysis of Decentralized Exchange*, Princeton University Press, 2025.

Ismail, Ali Haider and Diego Zuniga, "Welfare Cost of Convenience Yield," 2025. Working Paper.

Jacklin, Charles J., "Demand Deposits, Trading Restrictions, and Risk Sharing," in Edward C. Prescott and Neil Wallace, eds., *Contractual Arrangements for Intertemporal Trade*, Minneapolis: University of Minnesota Press, 1987, pp. 26–47.

Jiang, Zhengyang, Arvind Krishnamurthy, and Hanno Lustig, "Foreign Safe Asset Demand and the Dollar Exchange Rate," *Journal of Finance*, June 2021, *76* (3), 1049–1089.

Kiyotaki, Nobuhiro and Randall Wright, "A search-theoretic approach to monetary economics," *The American Economic Review*, 1993, pp. 63–77.

Krishnamurthy, Arvind and Annette Vissing-Jørgensen, "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 2012, *120* (2), 233–267.

Krueger, Dirk and Hanno Lustig, "When is market incompleteness irrelevant for the price of aggregate risk (and when is it not)?," *Journal of Economic Theory*, 2010, *145* (1), 1–41.

Lagos, Ricardo, "Asset prices and liquidity in an exchange economy," *Journal of Monetary Economics*, 2010, *57* (8), 913–930.

_ and Gaston Navarro, "Monetary Policy Operations: Theory, Evidence, and Tools for Quantitative Analysis," June 2023. NBER working paper 31370.

_ and Guillaume Rocheteau, "Liquidity in Asset Markets with Search Frictions," *Econometrica*, 2009, *77*, 403–426.

_ , _ , and Randall Wright, "Liquidity: A New Monetarist Perspective," *Journal of Economic Literature*, 2017, *55* (2), 371–440.

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

Lopez-Salido, David and Annette Vissing-Jorgenson, "Reserve Demand, Interest Rate Control, and Quantitative Tightening," 2023. Working Paper.

Lorenzoni, Guido, "Inneficient Credit Booms," *Review of Economics Studies*, July 2008, *75* (2), 809–833.

Merha, Rajnish and Edward C. Prescott, "The Equity Premium: a Puzzle," *Journal of Monetary Economics*, 1985, *10*, 335–359.

Nagel, Stefan, "The Liquidity Premium of Near-Money Assets," *The Quarterly Journal of Economics*, 2016, *131* (4), 1927–1971.

Piazzesi, Monika, Ciaran Rogers, and Martin Schneider, "Money and Banking in a New Keynesian Model," 2019. Working Paper, Stanford.

Ríos-Rull, José-Víctor, Joao Ritto, Tamon Takamura, and Yaz Terajima, "Real Monetary Policy," October 2024. Very Preliminary.

Shi, Shouyong, "A Divisible Search Model of Fiat Money," *Econometrica*, 1997, *65*, 75–102.

Silva, Dejanir, Juan Passadore, and Mahyar Kargar, "Liquidity and Risk in OTC Markets: A Theory of Asset Pricing and Portfolio Flows," 2023. Working Paper.

Uslu, Semih, "Pricing and Liquidity in Decentralized Asset Markets," *Econometrica*, 2019, *87* (6), 2079–2140.

Vayanos, Dimitri and Jean-Luc Vila, "Equilibrium Interest Rate and Liquidity Premium with Transaction Costs," *Economic Theory*, 1999, *13*, 509–539.

Williamson, Stephen D. and Randall Wright, "New Monetarist Economics: Models," in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3, Elsevier, 2010, chapter 2, p. 25–96.

Wong, Russell and Mengbo Zhang, "Disintermediating the Federal funds market," 2023. Working Paper.

Yared, Pierre, "Public Debt Under Limited Private Credit," *Journal of the European Economic Association*, 2013, *11* (2), 229–245.

Appendix (not intended for publication)

Appendix Table of Contents

A. Proof of Proposition 1

We begin with an auxiliary Lemma showing that market tightness follows a difference equation. With the market tightness, we obtain the matching probabilities at each round:

Lemma 2. *Let θ_0 be the initial market tightness. Then, the ratio $\{\theta_n\}$ features the following law of motion:*

$$\theta_n = \theta_{n-1} \frac{(1 - \lambda_N G(1/\theta_{n-1}, 1))}{(1 - \lambda_N G(1, \theta_{n-1}))} \qquad \forall n \in \{1, 2, \ldots, N\}.$$

and the matching probabilities can be expressed in terms of the ratio via:

$$\psi_n^+ = \lambda_N G(1, \theta_{n-1}) \text{ and } \psi_n^- = \lambda_N G(1/\theta_{n-1}, 1).$$

Proof. By definition and homogeneity:

$$\theta_n = \frac{S_n^-}{S_n^+} = \frac{S_{n-1}^- - z_n}{S_{n-1}^+ - z_n} = \theta_{n-1} \frac{(1 - \lambda_N G(1/\theta_{n-1}, 1))}{(1 - \lambda_N G(1, \theta_{n-1}))}, \qquad \forall n \in \{1, 2, \ldots, N\}.$$

where the second equality follow from the definition of z_n and uses its homogeneity property. □

The lemma shows that we can track matching probabilities in terms of the initial market tightness, without reference to the terms of trade. It also shows that the these probabilities are scale invariant. We use these observations in what follows, as the Lemma permits us to treat trading probabilities as exogenous series.

The proof also makes use of the following standard result in probability, which we include for ease of completeness:

Lemma 3 (Conditional Probability Decomposition)**.** *For any starting round $n \geq 0$, the conditional probabilities of matching satisfy:*

$$\sum_{k=n+1}^{N} \psi_k^{\pm} \prod_{m=n+1}^{k-1} (1 - \psi_m^{\pm}) + \prod_{m=n+1}^{N} (1 - \psi_m^{\pm}) = 1$$

That is, starting from round n, the probability of matching in some future round plus the probability of never matching equals 1.

Proof. Let $P_{n,j} = \prod_{m=n+1}^{j} (1 - \psi_m^{\pm})$ be the probability of not matching from round $n+1$ through round j, with the convention that $P_{n,n} = 1$.

Then:

$$\begin{aligned}\sum_{k=n+1}^{N}\psi_k^{\pm}\prod_{m=n+1}^{k-1}(1-\psi_m^{\pm}) &= \sum_{k=n+1}^{N}\psi_k^{\pm}P_{n,k-1}\\ &= \sum_{k=n+1}^{N}[P_{n,k-1}-P_{n,k}] \quad (\text{since } P_{n,k}=P_{n,k-1}(1-\psi_k^{\pm}))\\ &= (P_{n,n}-P_{n,n+1})+(P_{n,n+1}-P_{n,n+2})+\cdots+(P_{n,N-1}-P_{n,N})\\ &= P_{n,n}-P_{n,N}\\ &= 1-\prod_{m=n+1}^{N}(1-\psi_m^{\pm})\end{aligned}$$

where we used $P_{n,n}=1$ and the telescoping sum. The result follows by rearranging. □

Next, we describe the limit of the bargaining problem as $\Delta \to 0$. Recall that the trader's estimate of equity, excluding its own trade, is:

$$\mathcal{E}^j(\Delta) \equiv \sum_{i\in\mathbb{I}} a_{t+1}^i R_{t+1}^i + m_{t+1}R_{t+1}^m + \chi_{t+1}(s^j - \text{sign}\{s^j\}\Delta).$$

The proof of Proposition 1 is as follows:

Proof. The Nash bargaining problem at round n with trade size Δ has surplus for trader with position s^j:

$$\mathcal{S}_n^{\text{sign}\{s^j\}}(\Delta) = V(\mathcal{E}^j(\Delta) + \text{sign}\{s^j\}(r_n^f - r^m)\Delta) - J_U^{\text{sign}\{s^j\}}(n;\Delta).$$

Step 1: Outside option recursion. The unmatched value satisfies:

$$J_U^{\text{sign}\{s^j\}}(n;\Delta) = \psi_{n+1}^{\text{sign}\{s^j\}}V(\mathcal{E}^j(\Delta)+\text{sign}\{s^j\}(r_{n+1}^f-r^m)\Delta) + (1-\psi_{n+1}^{\text{sign}\{s^j\}})J_U^{\text{sign}\{s^j\}}(n+1;\Delta),$$

with terminal round, N, values given by:

$$\begin{aligned}J_U^{+}(N;\Delta) &= V(\mathcal{E}^j(\Delta)) \quad (\text{surplus traders hold cash at rate } r^m)\\ J_U^{-}(N;\Delta) &= V(\mathcal{E}^j(\Delta)-(r^w-r^m)\Delta) \quad (\text{deficit traders borrow at rate } r^w).\end{aligned}$$

Expanding the recursion forward from round n:

$$J_U^{\text{sign}\{s^j\}}(n;\Delta) = \sum_{k=n+1}^{N}\left[\prod_{m=n+1}^{k-1}(1-\psi_m^{\text{sign}\{s^j\}})\right]\psi_k^{\text{sign}\{s^j\}}V(\mathcal{E}^j(\Delta)+\text{sign}\{s^j\}(r_k^f-r^m)\Delta)$$
$$+\left[\prod_{m=n+1}^{N}(1-\psi_m^{\text{sign}\{s^j\}})\right]V(\mathcal{E}^j(\Delta)+\text{sign}\{s^j\}\chi_{N+1}^{\text{sign}\{s^j\}}\Delta).$$

By Lemma 3, this corresponding to the value of $\mathbb{E}\left[V(\mathcal{E}^j(\Delta)+\text{sign}\{s^j\}(r_n^f-r^m)\Delta)|\text{unmatched by n}\right]$. Importantly, notice that the expectation assumes that the negotiated rated in future rounds r_n^f only depends on future rounds, an assumption that we have to verify below.

Now, consider the normalized difference::

$$\frac{J_U^{\text{sign}\{s^j\}}(n;\Delta)-V(\mathcal{E}^j(\Delta))}{\Delta} = \sum_{k=n+1}^{N}\left[\prod_{m=n+1}^{k-1}(1-\psi_m^{\text{sign}\{s^j\}})\right]\psi_k^{\text{sign}\{s^j\}}\frac{V(\mathcal{E}^j+\text{sign}\{s^j\}(r_k^f-r^m)\Delta)-V(\mathcal{E}^j)}{\Delta}$$
$$+\left[\prod_{m=n+1}^{N}(1-\psi_m^{\text{sign}\{s^j\}})\right]\frac{V(\mathcal{E}^j+\text{sign}\{s^j\}\chi_{N+1}^{\text{sign}\{s^j\}}\Delta)-V(\mathcal{E}^j)}{\Delta}.$$

As $\Delta\to 0$, by the definition of derivative:

$$\lim_{\Delta\to 0}\frac{J_U^{\text{sign}\{s^j\}}(n;\Delta)-V(\mathcal{E}^j(\Delta))}{\Delta} = \text{sign}\{s^j\}\cdot V'(\mathcal{E}^j)\cdot\chi_n^{\text{sign}\{s^j\}}$$

where:

$$\chi_n^{\text{sign}\{s^j\}} = \sum_{k=n+1}^{N}(r_k^f-r^m)\left[\prod_{m=n+1}^{k-1}(1-\psi_m^{\text{sign}\{s^j\}})\right]\psi_k^{\text{sign}\{s^j\}}+\chi_{N+1}^{\text{sign}\{s^j\}}\prod_{m=n+1}^{N}(1-\psi_m^{\text{sign}\{s^j\}}). \quad \text{(A.1)}$$

The variable $\chi_n^{\text{sign}\{s^j\}}$ represents the expected financing cost/benefit conditional on being unmatched at round n, for $+$ and $-$ positions. The variables satisfy the following recursion:

$$\chi_n^{\text{sign}\{s^j\}} = \psi_{n+1}^{\text{sign}\{s^j\}}(r_{n+1}^f-r^m)+(1-\psi_{n+1}^{\text{sign}\{s^j\}})\chi_{n+1}^{\text{sign}\{s^j\}} \quad \text{(A.2)}$$

with terminal conditions $\chi_{N+1}^{+}=0$ and $\chi_{N+1}^{-}=r^w-r^m$.

Step 2: Limiting surplus. Consider the following term:

$$\frac{S_n^{\text{sign}\{s^j\}}(\Delta)}{\Delta} = \frac{\left(V(\mathcal{E}^j(\Delta)+\text{sign}\{s^j\}(r_n^f-r^m)\Delta)-V(\mathcal{E}^j(\Delta)\right)-\left(J_U^{\text{sign}\{s^j\}}(n;\Delta)-V(\mathcal{E}^j(\Delta))\right)}{\Delta}.$$

Taking the $\Delta \to 0$ limit and substituting the result in Step 1:

$$\lim_{\Delta \to 0} \frac{\mathcal{S}_n^{\text{sign}\{s^j\}}(\Delta)}{\Delta} = \text{sign}\{s^j\} V'(\mathcal{E}^j(0))[(r_n^f - r^m) - \chi_n^{\text{sign}\{s^j\}}].$$

Step 3: Nash bargaining. The bargained rate solves:

$$r_n^f(\Delta) = \arg\max_{r_n} [S_n^-(\Delta)]^\eta [S_n^+(\Delta)]^{1-\eta}.$$

Since multiplying the objective by a positive constant Δ^{-1} doesn't change the maximizer:

$$r_n^f(\Delta) = \arg\max_{r_n} \frac{[S_n^-(\Delta)]^\eta [S_n^+(\Delta)]^{1-\eta}}{\Delta} = \arg\max_{r_n} \left[\frac{S_n^-(\Delta)}{\Delta}\right]^\eta \left[\frac{S_n^+(\Delta)}{\Delta}\right]^{1-\eta}.$$

Taking the limit as $\Delta \to 0$:

$$r_n^f = \lim_{\Delta \to 0} r_n^f(\Delta) = \lim_{\Delta \to 0} \left\{ \arg\max_{r_n} \left[\frac{S_n^-(\Delta)}{\Delta}\right]^\eta \left[\frac{S_n^+(\Delta)}{\Delta}\right]^{1-\eta} \right\}.$$

By the Theorem of the Maximum, since the objective function is continuous in both r_n and Δ, and the constraint set $[r^m, r^w]$ we can pass limits inside the maximum operator:

$$\lim_{\Delta \to 0} r_n^f = \arg\max_{r_n} \left\{ \lim_{\Delta \to 0} \left[\frac{S_n^-(\Delta)}{\Delta}\right]^\eta \left[\frac{S_n^+(\Delta)}{\Delta}\right]^{1-\eta} \right\}.$$

Substituting the limits from Step 2:

$$r_n^f = \arg\max_{r_n} \left\{ [V'(\mathcal{E}^j(0))]^\eta [V'(\mathcal{E}^k(0))]^{1-\eta} \cdot [\chi_n^- - (r_n - r^m)]^\eta [(r_n - r^m) - \chi_n^+]^{1-\eta} \right\}.$$

Since $V'(\mathcal{E}^j(0))$ and $V'(\mathcal{E}^k(0))$ are positive constants independent of r_n^f, this reduces to Problem 1:

$$r_n^f = \arg\max_{r_n} [\chi_n^- - (r_n - r^m)]^\eta [(r_n - r^m) - \chi_n^+]^{1-\eta}.$$

where χ_n^+ and χ_n^- satisfy A.1.

Step 4: First-order condition. We now solve the maximization problem. The FOC yields: $\eta[(r_n^f - r^m) - \chi_n^+] = (1-\eta)[\chi_n^- - (r_n^f - r^m)]$. Therefore:

$$r_n^f = r^m + (1-\eta)\chi_n^- + \eta\chi_n^+,$$

as stated by the proposition. This justifies the assumption that outside options only consider rates that depend on rates without the need to consider counterparties. □

B. Proof of Proposition 2

Proof. The matching probabilities over N rounds are:

$$\Psi^{+} = 1 - \prod_{n=1}^{N}(1-\psi_n^{+}) \quad \text{and} \quad \Psi^{-} = 1 - \prod_{n=1}^{N}(1-\psi_n^{-}).$$

These represent the probability of matching at least once during the OTC stage.

Verification of χ^+ and χ^-. From Proposition 1 we obtain the recursion (A.1). Solving that recursion forward, from round 0 to round N:

$$\chi_0^{\pm} = \sum_{n=1}^{N}(r_n^f - r^m)\psi_n^{\pm}\prod_{k=1}^{n-1}(1-\psi_k^{\pm}) + \chi_{N+1}^{\pm}\prod_{n=1}^{N}(1-\psi_n^{\pm}). \tag{A.3}$$

where we used $\chi_{N+1}^{+} = 0$ and $\chi_{N+1}^{-} = r^w - r^m$. The first term is the OTC rates weighted by the unconditional matching probabilities, and the second term is the terminal value weighted by the probability of never matching $(1-\Psi^{\pm})$.

Recall that probabilities in the expansion add up to 1: apply Lemma 3 to $n = 1$. Thus, we can define the volume-weighted average rate as:

$$\overline{r}^f - r^m = \sum_{n=1}^{N}(r_n^f - r^m)\frac{\psi_n^{\pm}\prod_{k=1}^{n-1}(1-\psi_k^{\pm})}{\Psi^{\pm}},$$

where the weights in the sum add up to 1.

Substituting the definition of weighted rates in (A.3), we obtain:

$$\begin{aligned}
\chi_0^{+} &= \Psi^{+}(\overline{r}^f - r^m) + 0\cdot(1-\Psi^{+}) = \Psi^{+}(\overline{r}^f - r^m) = \chi^{+}\\
\chi_0^{-} &= \Psi^{-}(\overline{r}^f - r^m) + (r^w - r^m)(1-\Psi^{-}) = \chi^{-}.
\end{aligned}$$

Consistency of weights. Probabilities are proportional to traded amounts since $z_n = \psi_n^{\pm}S_{n-1}^{\pm}$ and $S_{n-1}^{\pm} = \prod_{k=1}^{n-1}(1-\psi_k^{\pm})S_0^{\pm}$ and $\sum_{n=1}^{N} z_n = \Psi^{\pm}S_0^{\pm}$. Thus, the weights $\varkappa_n^{\pm} = \frac{\psi_n^{\pm}\prod_{k=1}^{n-1}(1-\psi_k^{\pm})}{\Psi^{\pm}}$ are identical for surplus and deficit sides due to the homogeneity of the matching function (see Lemma 2), ensuring a unique $\overline{r}^f$. $\square$

C. Proof of Lemma 1

Consider the continuous-time limit as $N \to \infty$. Let $\Delta = 1/N$ denote the time between rounds, and index continuous time by $\tau \in [0,1]$. We derive the ODE in the statement of the Lemma.

Proof. **Step 1: Evolution of masses.** In discrete time, the masses of surplus and deficit positions evolve as:

$$S_{n+1}^{\pm} = S_n^{\pm} - \lambda_N G(S_n^+, S_n^-).$$

With $\lambda_N = \bar{\lambda}/N$ and continuous time $\tau = n/N$, taking $\Delta \to 0$:

$$\dot{S}_\tau^{\pm} = \lim_{\Delta \to 0} \frac{S_{\tau+\Delta}^{\pm} - S_\tau^{\pm}}{\Delta} = -\bar{\lambda} G(S_\tau^+, S_\tau^-).$$

Both sides shrink at the same absolute rate since matches clear equal amounts from each side.

Step 2: Dynamics of tightness. Define $\theta_\tau = S_\tau^-/S_\tau^+$. Taking logarithms and differentiating:

$$\begin{aligned}
\frac{\dot{\theta}_\tau}{\theta_\tau} &= \frac{d}{d\tau}\ln(\theta_\tau) = \frac{d}{d\tau}[\ln(S_\tau^-) - \ln(S_\tau^+)] \\
&= \frac{\dot{S}_\tau^-}{S_\tau^-} - \frac{\dot{S}_\tau^+}{S_\tau^+} \\
&= -\bar{\lambda}\frac{G(S_\tau^+, S_\tau^-)}{S_\tau^-} + \bar{\lambda}\frac{G(S_\tau^+, S_\tau^-)}{S_\tau^+}.
\end{aligned}$$

Using homogeneity of degree one: $G(S^+, S^-) = S^+ G(1, \theta) = S^- G(\theta^{-1}, 1)$. Therefore:

$$\frac{\dot{\theta}_\tau}{\theta_\tau} = -\bar{\lambda} G(\theta_\tau^{-1}, 1) + \bar{\lambda} G(1, \theta_\tau).$$

Using the normalized matching function, $\gamma(\theta) \equiv G(1, \theta)$, and noting that $G(\theta^{-1}, 1) = \gamma(\theta^{-1})$ by definition:

$$\dot{\theta}_\tau = -\bar{\lambda}\theta_\tau[\gamma(\theta_\tau^{-1}) - \gamma(\theta_\tau)].$$

Step 3: Matching intensities. The instantaneous matching rates per unit mass are the limit as $\Delta \to 0$ of $\psi_n^{\pm}/\Delta = \frac{G(S_n^+, S_n^-)}{S_n^{\pm}\Delta}$. By the calculations above:

$$\psi_\tau^+ = \bar{\lambda}\gamma(\theta_\tau) \quad \text{and} \quad \psi_\tau^- = \bar{\lambda}\gamma(\theta_\tau^{-1}).$$

Note that $\psi_\tau^+ = \theta_\tau \psi_\tau^-$ follows from $\gamma(\theta) = \theta\gamma(\theta^{-1})$ (by symmetry). □

D. Proof of Proposition 3

Proof. From Lemma 1, the ODE for market tightness is:

$$\dot{\theta}_\tau = -\bar{\lambda}\theta_\tau[\gamma(\theta_\tau^{-1}) - \gamma(\theta_\tau)]. \tag{A.4}$$

1. Steady state. If $\theta_0 = 1$, then $\gamma(\theta_0^{-1}) = \gamma(1) = \gamma(\theta_0)$, so $\dot{\theta}_0 = 0$. Then, by uniqueness of solutions to ODEs, $\theta_\tau = 1$ for all $\tau \in [0, 1]$.

2. Monotonicity. The sign of $\dot{\theta}_\tau$ depends on $\gamma(\theta_\tau^{-1}) - \gamma(\theta_\tau)$. By symmetry of G:

$$\gamma(\theta^{-1}) = G(1, \theta^{-1}) = G(\theta^{-1}, 1) = \frac{G(1,\theta)}{\theta} = \frac{\gamma(\theta)}{\theta}.$$

Therefore:

$$\dot{\theta}_\tau = -\bar{\lambda}\theta_\tau \left[\frac{\gamma(\theta_\tau)}{\theta_\tau} - \gamma(\theta_\tau)\right] = \bar{\lambda}\gamma(\theta_\tau)(\theta_\tau - 1)$$

Since $\gamma(\theta) > 0$ for all $\theta > 0$: (i) If $\theta_0 > 1$, then $\dot{\theta}_0 > 0$, so θ_τ increases. (ii) If $\theta_0 < 1$, then $\dot{\theta}_0 < 0$, so θ_τ decreases.

3. Effect of matching functions. Consider two matching functions $\gamma, \tilde{\gamma}$, such that $\gamma(\theta) < \tilde{\gamma}(\theta)$ for all . From the ODE (A.4): (i) If $\theta > 1$: $\dot{\theta}$ is larger under $\tilde{\gamma}$, so θ rises faster. (ii) If $\theta < 1$: $\dot{\theta}$ is smaller under $\tilde{\gamma}$, so θ falls faster. □

E. Proof of Proposition 4

We proceed to derive the integral form in the proposition.

Proof. **Step 1: Continuous-time limit of the recursions.** From the discrete-rounds recursion:

$$\chi_n^{\pm} = \psi_{n+1}^{\pm}(r_{n+1}^f - r^m) + (1 - \psi_{n+1}^{\pm})\chi_{n+1}^{\pm}$$

we can rearrange terms to obtain: $\chi_{n+1}^{\pm} - \chi_n^{\pm} = \psi_{n+1}^{\pm}[\chi_{n+1}^{\pm} - (r_{n+1}^f - r^m)]$. Dividing both sides by Δ, substituting the discrete round probabilities for the corresponding intensities, and the change of variables $\Delta = 1/N$ and $\tau = n/N$, we have:

$$\frac{\chi^{\pm}_{\tau+\Delta} - \chi^{\pm}_{\tau}}{\Delta} = \frac{\psi^{\pm}_{\tau+\Delta}}{\Delta}[\chi^{\pm}_{\tau+\Delta} - (r^f_{\tau+\Delta} - r^m)].$$

Now, recall that:

$$\lim_{\Delta \to 0} \frac{\psi^{\pm}_{\tau+\Delta}}{\Delta} = \lim_{\Delta \to 0} \frac{\lambda_N(\Delta)}{\Delta} \cdot G(S^+_n, S^-_n) = \bar{\lambda} G(S^+_\tau, S^-_\tau).$$

Thus, taking $\Delta \to 0$ we obtain an ODE for the outside options:

$$\dot{\chi}^{\pm}_{\tau} = \psi^{\pm}_{\tau}[\chi^{\pm}_{\tau} - (r^f_\tau - r^m)].$$

Step 2: Substituting the bargained rate. From Proposition 1, after the change of index from rounds to time

$$r^f_\tau = r^m + (1-\eta)\chi^-_\tau + \eta\chi^+_\tau.$$

Hence, the ODE for the outside options becomes a coupled system:

$$\dot{\chi}^+_\tau = \psi^+_\tau[\chi^+_\tau - (1-\eta)\chi^-_\tau - \eta\chi^+_\tau] = -(1-\eta)\psi^+_\tau(\chi^-_\tau - \chi^+_\tau) \tag{A.5}$$
$$\dot{\chi}^-_\tau = \psi^-_\tau[\chi^-_\tau - (1-\eta)\chi^-_\tau - \eta\chi^+_\tau] = \eta\psi^-_\tau(\chi^-_\tau - \chi^+_\tau). \tag{A.6}$$

Step 3: Solving for the joint surplus. Define $\Sigma_\tau \equiv \chi^-_\tau - \chi^+_\tau$ as the joint surplus. Then, the joint surplus must satisfy the following ODE:

$$\dot{\Sigma}_\tau = \dot{\chi}^-_\tau - \dot{\chi}^+_\tau = [\eta\psi^-_\tau + (1-\eta)\psi^+_\tau]\Sigma_\tau.$$

This is a linear ODE with backward solution:

$$\Sigma_\tau = \Sigma_1 \exp\left(-\int_\tau^1 [\eta\psi^-_s + (1-\eta)\psi^+_s]ds\right).$$

Since the terminal conditions for the outside options are $\chi^+_1 = 0$ and $\chi^-_1 = r^w - r^m$, the terminal condition for the ODE is given by $\Sigma_1 = r^w - r^m$. Thus,

$$\Sigma_\tau = (r^w - r^m) \exp\left(-\int_\tau^1 [\eta\psi^-_s + (1-\eta)\psi^+_s]ds\right). \tag{A.7}$$

Step 4: Solving for individual components. Substituting the integral in (A.7) back into the

ODE system (A.5)-(A.6):

$$\dot{\chi}_\tau^+ = -(1-\eta)\psi_\tau^+\Sigma_\tau, \qquad \dot{\chi}_\tau^- = \eta\psi_\tau^-\Sigma_\tau.$$

Integrating from τ to 1 and using the terminal conditions:

$$\chi_\tau^+ - 0 = \int_\tau^1 (1-\eta)\psi_y^+\Sigma_y dy, \quad \chi_\tau^- - (r^w - r^m) = -\int_\tau^1 \eta\psi_y^-\Sigma_y dy.$$

Substituting (A.7), the expression for Σ_y:

$$\begin{aligned}
\chi_\tau^+ &= (r^w - r^m)\int_\tau^1 (1-\eta)\psi_y^+ \exp\left(-\int_y^1 [\eta\psi_x^- + (1-\eta)\psi_x^+]dx\right) dy \\
\chi_\tau^- &= (r^w - r^m)\left[1 - \int_\tau^1 \eta\psi_y^- \exp\left(-\int_y^1 [\eta\psi_x^- + (1-\eta)\psi_x^+]dx\right) dy\right].
\end{aligned}$$

The convenience-yield coefficients are defined as the $\tau = 0$ values of the integrals above: $\chi^\pm = \chi_0^\pm$, and the OTC rate is $r_\tau^f = r^m + (1-\eta)\chi_\tau^- + \eta\chi_\tau^+$. □

To verify that our solution is consistent with the discrete case in that weighted-average rates can be obtaine from $\chi_0^\pm$, we trace the relationship between matching intensities and probabilities in continuous time. The matching intensities $\psi_\tau^\pm$ represent the instantaneous matching rate per unit time. These are intensities, but cannot be interpreted directly as PDFs. The probability of not matching from time 0 to τ is therefore:

$$1 - F^\pm(\tau) = \exp\left(-\int_0^\tau \psi_s^\pm ds\right),$$

where the CDF of matching by time τ is $F^\pm(\tau) \equiv 1 - \exp\left(-\int_0^\tau \psi_s^\pm ds\right)$. Hence, the PDF (probability density of matching at time τ) is:

$$f^\pm(\tau) = \frac{dF^\pm}{d\tau} = \psi_\tau^\pm \exp\left(-\int_0^\tau \psi_s^\pm ds\right) = \psi_\tau^\pm(1 - F^\pm(\tau)).$$

This confirms the hazard rate relationship: $\psi_\tau^\pm = f^\pm(\tau)/(1 - F^\pm(\tau))$. The volume-weighted average rate is $\overline{r}^f = \int_0^1 \varkappa_\tau^\pm r_\tau^f d\tau$ where the weights $\varkappa_\tau^\pm = f^\pm(\tau)/\Psi^\pm$ represent the fraction of total volume traded at time τ.

Verification: Integrating the ODE $\dot{\chi}_\tau^- = \psi_\tau^-[\chi_\tau^- - (r_\tau^f - r^m)]$ multiplied by $(1 - F^-(\tau))$ and using integration by parts:

$$\Psi^-(\bar{r}^f - r^m) = \chi_0^- - (1-\Psi^-)(r^w - r^m)$$

which confirms $\chi_0^- = \Psi^-(\bar{r}^f - r^m) + (1-\Psi^-)(r^w - r^m) = \chi^-$. Similarly for $\chi_0^+ = \chi^+$. This is the same consistency condition as in the discrete rounds.

F. Proof of Proposition 5

Proof. We begin establishing the result for χ^+. From Proposition 4, we obtained (18):

$$\chi_\tau^+ = \int_\tau^1 (1-\eta)\psi_t^+ \Sigma_t dt, \tag{A.8}$$

where the surplus satisfies:

$$\Sigma_\tau = (r^w - r^m)\exp\left(-\int_\tau^1 [\eta\psi_x^- + (1-\eta)\psi_x^+]dx\right). \tag{A.9}$$

Recall that overall matching probabilities, given in (15), satisfy:

$$e^{-\int_\tau^1 \psi_s^+ ds} = 1 - \Psi_\tau^+, \quad e^{-\int_\tau^1 \psi_s^- ds} = 1 - \Psi_\tau^-.$$

Thus, we have that:

$$\Sigma_\tau = (r^w - r^m)(1-\Psi_\tau^+)^{1-\eta}(1-\Psi_\tau^-)^\eta. \tag{A.10}$$

Consider the following proposed solution χ_τ^+ as a function of θ_τ:

$$\chi^+(\theta_\tau) = (r^w - r^m)\frac{\bar{\theta} - \bar{\theta}^\eta \theta_\tau^{1-\eta}}{\bar{\theta} - 1}. \tag{A.11}$$

If the proposal is correct, its time derivative must coincide with the term inside the integral, and we must verify that it satisfies the terminal condition. Taking the derivative with respect to τ:

$$\frac{\partial \chi_\tau^+}{\partial \tau} = (r^w - r^m)\frac{-(1-\eta)\bar{\theta}^\eta \theta^{-\eta}}{\bar{\theta} - 1}\dot{\theta}. \tag{A.12}$$

By market clearing:

$$\bar{\theta} = \frac{1-\Psi_\tau^-}{1-\Psi_\tau^+}\theta_\tau, \tag{A.13}$$

a condition we can use to substitute out $\bar{\theta}$ from (A.12). We obtain:

$$\frac{\partial \chi_\tau^+}{\partial \tau} = -(r^w - r^m)(1-\eta)\frac{\left(\frac{1-\Psi_\tau^-}{1-\Psi_\tau^+}\right)^\eta}{\frac{1-\Psi_\tau^-}{1-\Psi_\tau^+}\theta - 1}\dot{\theta}. \tag{A.14}$$

Using the market clearing condition $\Psi_\tau^- \theta_\tau = \Psi_\tau^+$, the denominator in the fraction becomes:

$$\frac{\theta - \Psi_\tau^- \theta - 1 + \Psi_\tau^+}{1 - \Psi_\tau^+} = \frac{\theta - 1}{1 - \Psi_\tau^+}.$$

We then have that:

$$\frac{\partial \chi_\tau^+}{\partial \tau} = -(r^w - r^m)(1-\eta)\frac{(1-\Psi_\tau^-)^\eta (1-\Psi_\tau^+)^{1-\eta}}{\theta - 1}\dot{\theta}. \tag{A.15}$$

Next, we substitute out $\dot{\theta}$. Recall from equation (17) that market tightness follows:

$$\dot{\theta} = -\bar{\lambda}\theta(\gamma(\theta^{-1}) - \gamma(\theta)).$$

Replacing the matching intensities given in (14); $\psi_\tau^+ = \bar{\lambda}\gamma(\theta_\tau)$ and $\psi_\tau^- = \bar{\lambda}\gamma(\theta_\tau^{-1})$ combined with the market clearing condition $\psi_\tau^- = \theta_\tau \psi_\tau^+$ we arrive at:

$$\dot{\theta} = -(\bar{\lambda}\theta\gamma(\theta^{-1}) - \bar{\lambda}\theta\gamma(\theta)) = -\psi_\tau^+(\theta - 1). \tag{A.16}$$

Substituting this step into (A.15), we obtain:

$$\frac{\partial \chi_\tau^+}{\partial \tau} = (r^w - r^m)(1-\eta)(1-\Psi_\tau^-)^\eta (1-\Psi_\tau^+)^{1-\eta}\psi_\tau^+ = (1-\eta)\psi_\tau^+ \Sigma_\tau. \tag{A.17}$$

This matches the integrand exactly. Since the terminal condition is also satisfied (at $\tau = 1$, $\theta_1 = \bar{\theta}$ giving $\chi_1^+ = 0$), the formula is verified. Applying the integral to this term yields χ^+. The result is valid for any matching function.

To verify the formula,

$$\chi^- = (r^w - r^m)\frac{\bar{\theta} - \bar{\theta}^\eta \theta_0^{-\eta}}{\bar{\theta} - 1},$$

we use the property that $\Sigma_\tau = \chi_\tau^- - \chi_\tau^+$. Thus, we obtain that if the formula is correct:

$$\Sigma_\tau = (r^w - r^m)\frac{\bar{\theta}^\eta \theta_\tau^{1-\eta} - \bar{\theta}^\eta \theta_\tau^{-\eta}}{\bar{\theta} - 1} = (r^w - r^m)\bar{\theta}^\eta \theta_\tau^{1-\eta}\frac{\theta_\tau - 1}{\theta_\tau(\bar{\theta} - 1)}. \tag{A.18}$$

We now verify this formula using (A.10):

$$\Sigma_\tau = (r^w - r^m)(1 - \Psi_\tau^+) \left(\frac{1 - \Psi_\tau^-}{1 - \Psi_\tau^+} \right)^\eta .$$

Equation (A.13) implies:

$$\bar{\theta} - 1 = \frac{1 - \Psi_\tau^-}{1 - \Psi_\tau^+} \theta - 1 = \frac{\theta - 1}{1 - \Psi_\tau^+} \Rightarrow 1 - \Psi_\tau^+ = \frac{\theta - 1}{\bar{\theta} - 1}. \tag{A.19}$$

where the second line uses the clearing condition. Back into the equation above, we obtain:

$$\Sigma_\tau = (r^w - r^m) \left(\frac{\theta - 1}{\bar{\theta} - 1} \right) \left(\frac{\bar{\theta}}{\theta} \right)^\eta .$$

Multiplying and dividing by θ yields the target (A.18). □

G. Proof of Corollary 1

Proof. **Part 1: Monotonicity of χ^+ and χ^-.** From Theorem 1, the outside option for surplus positions is:

$$\chi_\tau^+ = (r^w - r^m) \left[\frac{\bar{\theta} - \bar{\theta}^\eta \theta_\tau^{1-\eta}}{\bar{\theta} - 1} \right] .$$

Differentiating with respect to τ (noting that $\bar{\theta}$ is fixed):

$$\frac{d\chi_\tau^+}{d\tau} = - \frac{(r^w - r^m)(1 - \eta)\bar{\theta}^\eta}{(\bar{\theta} - 1)} \cdot \frac{\dot{\theta}_\tau}{\theta_\tau^\eta}.$$

For $\theta_0 > 1$: all terms are positive, yielding $\frac{d\chi_\tau^+}{d\tau} < 0$. For $\theta_0 < 1$: the numerator has $\bar{\theta} - 1 < 0$ and $\dot{\theta}_\tau < 0$, so their ratio is positive, again yielding $\frac{d\chi_\tau^+}{d\tau} < 0$.

Part 2: Monotonicity of . From Theorem 1, the outside option for deficit positions is:

$$\chi_\tau^- = (r^w - r^m) \left[\frac{\bar{\theta} - \bar{\theta}^\eta \theta_\tau^{-\eta}}{\bar{\theta} - 1} \right] .$$

Differentiating:

$$\frac{d\chi_\tau^-}{d\tau} = \frac{(r^w - r^m)\eta\bar{\theta}^\eta}{(\bar{\theta} - 1)} \cdot \frac{\dot{\theta}_\tau}{\theta_\tau^{\eta+1}}.$$

For both $\theta_0 > 1$ and $\theta_0 < 1$, the ratio $\frac{\dot{\theta}_\tau}{(\bar{\theta}-1)}$ is positive, yielding $\frac{d\chi_\tau^-}{d\tau} > 0$.

Part 3: Monotonicity of r^f_τ. From Proposition 1, the OTC rate satisfies $r^f_\tau = r^m + \eta\chi^+_\tau + (1-\eta)\chi^-_\tau$. Differentiating and substituting Parts 1 and 2:

$$\frac{dr^f_\tau}{d\tau} = \frac{(r^w - r^m)\eta(1-\eta)\bar{\theta}^\eta\dot{\theta}_\tau}{(\bar{\theta}-1)\theta_\tau^{\eta+1}}(1-\theta_\tau).$$

The sign is determined by $(1-\theta_\tau)\cdot\frac{\dot{\theta}_\tau}{(\bar{\theta}-1)}$:

- If $\theta_0 > 1$: then $(1-\theta_\tau) < 0$ and $\frac{\dot{\theta}_\tau}{(\bar{\theta}-1)} > 0$, so $\frac{dr^f_\tau}{d\tau} < 0$.
- If $\theta_0 < 1$: then $(1-\theta_\tau) > 0$ and $\frac{\dot{\theta}_\tau}{(\bar{\theta}-1)} > 0$, so $\frac{dr^f_\tau}{d\tau} > 0$.
- If $\theta_0 = 1$: then $(1-\theta_\tau) = 0$, so $\frac{dr^f_\tau}{d\tau} = 0$.

This completes the proof. □

As the OTC stage progresses, the value of surplus declines while the cost of deficit rises. The direction of rate movement depends critically on market imbalance: rates decline in deficit markets as surplus becomes scarce, and rise in surplus markets as deficits become scarce.

H. Proof of Corollary 2

Proof. When $\theta_0 = 1$, market tightness remains constant ($\theta_\tau = 1$ for all τ)— by Proposition 3. Thus, $\psi^+_\tau = \psi^-_\tau = \bar{\lambda}$ for all τ. From Proposition 4, the joint surplus is:

$$\Sigma_\tau = (r^w - r^m)\exp\left(-\int_\tau^1 [\eta\bar{\lambda} + (1-\eta)\bar{\lambda}]ds\right) = (r^w - r^m)e^{-\bar{\lambda}(1-\tau)}.$$

Hence, the outside options simplify to:

$$\chi^+_\tau = (r^w - r^m)\int_\tau^1 (1-\eta)\bar{\lambda}e^{-\bar{\lambda}(1-y)}dy = (1-\eta)(r^w - r^m)(1 - e^{-\bar{\lambda}(1-\tau)})$$
$$\chi^-_\tau = (r^w - r^m)\left[1 - \int_\tau^1 \eta\bar{\lambda}e^{-\bar{\lambda}(1-y)}dy\right] = (r^w - r^m) - \eta(r^w - r^m)(1 - e^{-\bar{\lambda}(1-\tau)}).$$

Noting that $(r^w - r^m)(1 - e^{-\bar{\lambda}(1-\tau)}) = (r^w - r^m) - \Sigma_\tau$ we have:

$$\chi^+_\tau = (1-\eta)[(r^w - r^m) - \Sigma_\tau], \quad \chi^-_\tau = (r^w - r^m) - \eta[(r^w - r^m) - \Sigma_\tau].$$

Thus, the OTC rates across rounds must be constant: $r^f_\tau = r^m + (1-\eta)\chi^-_\tau + \eta\chi^+_\tau = r^m + (1-\eta)(r^w - r^m)$. Since the calculations are valid for any $\tau \in [0,1]$, including zero $\tau = 0$, the corollary follows from these calculations. □

I. Proof of Proposition 6

Proof. **Step 1: Time rescaling for θ.** From Lemma 1, market tightness satisfies:

$$\dot{\theta}_\tau = -\bar{\lambda}\theta_\tau[\gamma(\theta_\tau^{-1}) - \gamma(\theta_\tau)].$$

Assume time τ has elapsed. Consider the solution of market tightness from time τ to τ' with initial condition θ_τ. Define the rescaled time $s = \frac{t-\tau}{1-\tau}$ for $t \in [\tau, 1]$, so that $s \in [0,1]$. Then $dt = (1-\tau)ds$.

By change of variables, the ODE for tightness becomes:

$$\frac{d\theta}{ds} = (1-\tau)\frac{d\theta}{dt} = -(1-\tau)\bar{\lambda}\theta[\gamma(\theta^{-1}) - \gamma(\theta)]$$

This is exactly the same ODE, but with efficiency parameter $\bar{\lambda}(1-\tau)$ and initial condition $\theta(0) = \theta_\tau$. At rescaled time $s' = \frac{\tau'-\tau}{1-\tau}$, we have:

$$\theta(\tau', \theta_0, \bar{\lambda}) = \theta\left(\frac{\tau'-\tau}{1-\tau}, \theta(\tau, \theta_0, \bar{\lambda}), \bar{\lambda}(1-\tau)\right).$$

For $\Psi^\pm$ the property follows from the market clearing condition $\bar{\theta} = \frac{1-\Psi^-_\tau}{1-\Psi^+_\tau}\theta_\tau$ and (A.19). For $\chi^\pm$ the property follows from Proposition 5. All of these objects can be expressed as a function of θ_t. Thus, the property is established immediately.

□

J. Proof of Proposition 7

Proof. **Step 1: Symmetry of market tightness.** From the ODE in Lemma 1:

$$\dot{\theta}_\tau = -\bar{\lambda}\theta_\tau[\gamma(\theta_\tau^{-1}) - \gamma(\theta_\tau)]$$

Define, $\phi_\tau \equiv \theta_\tau^{-1}$, we have $\dot{\phi}_\tau = -\theta_\tau^{-2}\dot{\theta}_\tau$. Substituting:

$$\dot{\phi}_\tau = -\phi_\tau^2 \cdot (-\bar{\lambda}\phi_\tau^{-1})[\gamma(\phi_\tau) - \gamma(\phi_\tau^{-1})] = -\bar{\lambda}\phi_\tau[\gamma(\phi_\tau^{-1}) - \gamma(\phi_\tau)].$$

This is the same ODE as for θ. Hence, if θ_τ solves it with initial condition θ_0, then θ_τ^{-1} solves it with initial condition θ_0^{-1}.

Step 2: Symmetry of matching intensities. By definition:

$$\psi^+(\theta) = \bar{\lambda}\gamma(\theta) = \bar{\lambda}\theta\gamma(\theta^{-1}) = \theta\gamma(\theta^{-1}) = \psi^-(\theta^{-1}).$$

Step 3: Symmetry of the spread. From Proposition 4:

$$\Sigma_\tau(\theta, \eta) = (r^w - r^m) \exp\left(-\int_\tau^1 [\eta\psi_s^-(\theta_s) + (1-\eta)\psi_s^+(\theta_s)]ds\right).$$

Using $\psi^+(\theta) = \psi^-(\theta^{-1})$ and $\psi^-(\theta) = \psi^+(\theta^{-1})$:

$$\begin{aligned}
\Sigma_\tau(\theta^{-1}, 1-\eta) &= (r^w - r^m) \exp\left(-\int_\tau^1 [(1-\eta)\psi_s^-(\theta_s^{-1}) + \eta\psi_s^+(\theta_s^{-1})]ds\right) \\
&= (r^w - r^m) \exp\left(-\int_\tau^1 [(1-\eta)\psi_s^+(\theta_s) + \eta\psi_s^-(\theta_s)]ds\right) \\
&= \Sigma_\tau(\theta, \eta).
\end{aligned}$$

Step 4: Symmetry of convenience yields. From the expressions in Proposition 4:

$$\begin{aligned}
\chi_\tau^+(\theta^{-1}, 1-\eta) &= (r^w - r^m) \int_\tau^1 \eta\psi_y^+(\theta_y^{-1})\Sigma_y(\theta^{-1}, 1-\eta)dy \\
&= (r^w - r^m) \int_\tau^1 \eta\psi_y^-(\theta_y)\Sigma_y(\theta, \eta)dy \\
&= (r^w - r^m) - \chi_\tau^-(\theta, \eta).
\end{aligned}$$

Thus, $\chi_\tau^-(\theta^{-1}, 1-\eta) = (r^w - r^m) - \chi_\tau^+(\theta, \eta)$.

Step 5: Symmetry of the OTC rate. Using $r^f = r^m + (1-\eta)\chi^- + \eta\chi^+$ the symmetry follows from the symmetry of outside options. $\square$

K. Proof of Proposition 8

Proof. **Part 1: Monotonicity in η.** We establish that χ^+, χ^-, and $\overline{r}^f$ are all decreasing in η. From the general formulas:

$$\chi^+ = (r^w - r^m)\frac{\bar\theta - \bar\theta^\eta\theta_\tau^{1-\eta}}{\bar\theta - 1}, \quad \chi^- = (r^w - r^m)\frac{\bar\theta - \bar\theta^\eta\theta_\tau^{-\eta}}{\bar\theta - 1},$$

we establish the result by taking derivatives with respect to η (holding θ_τ and $\bar\theta$ fixed):

$$\frac{\partial\chi^+}{\partial\eta} = -\frac{(r^w - r^m)}{\bar\theta - 1}\cdot\bar\theta^\eta\theta_\tau^{1-\eta}\ln\left(\frac{\bar\theta}{\theta_\tau}\right), \quad \frac{\partial\chi^-}{\partial\eta} = -\frac{(r^w - r^m)}{\bar\theta - 1}\cdot\bar\theta^\eta\theta_\tau^{-\eta}\ln\left(\frac{\bar\theta}{\theta_\tau}\right).$$

We establish that $\ln\left(\frac{\bar\theta}{\theta_\tau}\right)$ and $(\bar\theta - 1)$ always have the same sign: If (i) $\theta_0 > 1$: Market tightness increases over time, so $\bar\theta > \theta_\tau > 1$. Thus $\ln\left(\frac{\bar\theta}{\theta_\tau}\right) > 0$ and $(\bar\theta - 1) > 0$. Otherwise, if (ii) $\theta_0 < 1$: Market tightness decreases over time, so $\bar\theta < \theta_\tau < 1$. Thus $\ln\left(\frac{\bar\theta}{\theta_\tau}\right) < 0$ and $(\bar\theta - 1) < 0$. At $\theta_0 = 1$: $\bar\theta = \theta_\tau = 1$, so both expressions equal zero.[32] Therefore, $\frac{\ln\left(\frac{\bar\theta}{\theta_\tau}\right)}{(\bar\theta - 1)} > 0$ whenever $\theta_0 \neq 1$, which implies:

$$\frac{\partial\chi^+}{\partial\eta} < 0, \quad \frac{\partial\chi^-}{\partial\eta} < 0$$

The negotiated rate at time τ is: $r_\tau^f = r^m + (1-\eta)\chi_\tau^- + \eta\chi_\tau^+$. Differentiating with respect to η yields a negative value:

$$\frac{\partial r_\tau^f}{\partial\eta} = \underbrace{(\chi_\tau^+ - \chi_\tau^-)}_{<0} + \underbrace{(1-\eta)\frac{\partial\chi_\tau^-}{\partial\eta}}_{<0} + \underbrace{\eta\frac{\partial\chi_\tau^+}{\partial\eta}}_{<0} < 0.$$

Finally, we note that the average OTC rate $\overline{r}^f$ inherits this property since it is a weighted average of rates across all trading rounds: $\frac{\partial\overline{r}^f}{\partial\eta} < 0$.

Part 2: Limits. When $\eta = 1$:

$$\begin{aligned}
\chi_0^+ &= (r^w - r^m)\int_0^1 0\cdot\psi_y^+\exp\left(-\int_y^1\psi_x^- dx\right)dy = 0 \\
\chi_0^- &= (r^w - r^m)\left[1 - \int_0^1\psi_y^-\exp\left(-\int_y^1\psi_x^- dx\right)dy\right] = (r^w - r^m)(1 - \Psi^-).
\end{aligned}$$

[32] We can evaluate the derivative at $\theta_0 = 1$ using L'Hospital's rule.

Therefore: $\overline{r}^f = r^m + 0 \cdot \chi^- + 1 \cdot \chi^+ = r^m$. When $\eta = 0$:

$$\chi_0^+ = (r^w - r^m) \int_0^1 \psi_y^+ \exp\left(-\int_y^1 \psi_x^+ dx\right) dy = (r^w - r^m)\Psi^+$$
$$\chi_0^- = (r^w - r^m) \left[1 - \int_0^1 0 \cdot \psi_y^- \exp\left(-\int_y^1 \psi_x^+ dx\right) dy\right] = r^w - r^m.$$

Therefore: $\overline{r}^f = r^m + 1 \cdot \chi^- + 0 \cdot \chi^+ = r^m + (r^w - r^m) = r^w$. $\square$

L. Proof of Proposition 9

Proof. **Part 1: Static limit ($\bar{\lambda} \to 0$).** As $\bar{\lambda} \to 0$, both matching intensities vanish: $\psi_\tau^\pm \to 0$. From Proposition 4:

$$\chi^+ = (r^w - r^m) \int_0^1 (1-\eta)\psi_y^+ \exp\left(-\int_y^1 [\eta\psi_x^- + (1-\eta)\psi_x^+]dx\right) dy \to 0$$
$$\chi^- = (r^w - r^m) \left[1 - \int_0^1 \eta\psi_y^- \exp\left(-\int_y^1 [\eta\psi_x^- + (1-\eta)\psi_x^+]dx\right) dy\right] \to r^w - r^m.$$

Therefore, by the dilation property, all rates must equal: $\overline{r}^f \to r^m + (1-\eta)(r^w - r^m)$.

Part 2: Walrasian limit ($\bar{\lambda} \to \infty$). First, consider a balance market where $\theta = 1$. As $\bar{\lambda} \to \infty$, $\Sigma_0 \to 0$ as $\Psi^\pm \to 0$. From Corollary 2, it follows immediately that:

$$\chi^\pm \to (r^w - r^m)(1-\eta), \quad \overline{r}^f = r^m + (1-\eta)(r^w - r^m).$$

When the market is unbalanced, we must be more careful. Start with $\theta > 1$. As $\bar{\lambda} \to \infty$, the short (surplus) side matches instantly while the long (deficit) side faces rationing. This property follows because $\dot{\theta}_\tau = \bar{\lambda}\gamma(\theta_\tau)(\theta_\tau - 1) > 0$, implying that tightness explodes: $\theta_\tau \to \infty$. Thus, surplus traders match instantly: $\Psi^+ \to 1$ whereas $\Psi^- = S^+/S^- \to \theta_0^{-1}$.

Hence, using the consistency property (Proposition 2), we have:

$$\chi_0^+ \to \overline{r}^f - r^m, \quad \chi_0^- \to (\overline{r}^f - r^m)\theta_0^{-1} + (r^w - r^m)(1 - \theta_0^{-1}).$$

Recall from Proposition 4 that the bargained rate at $\tau = 0$ satisfies:

$$r_0 - r^m = (1-\eta)\chi_0^- + \eta\chi_0^+.$$

Thus, as efficiency increases to infinity:

$$r_0^f - r^m \to (1-\eta)((\bar{r}^f - r^m)\theta_0^{-1} + (r^w - r^m)(1-\theta_0^{-1})) + \eta(\bar{r}^f - r^m).$$

We further know that as $\bar{\lambda} \to \infty$, trades occur faster, i.e., matching intensities explode. Hence, the distribution of trade concentrates toward the earliest rounds and, thus, the volume-weighted average rate converges to the first-trading round rate: $\bar{r}^f \to r_0^f$. Therefore, the limit above requires:

$$(r_0^f - r^m)(1-\eta)(1-\theta_0^{-1}) \to (r^w - r^m)(1-\eta)(1-\theta_0^{-1}) \Leftrightarrow r_0^f \to r^w.$$

By the dilation property, $r_\tau^f \to r^w, \forall \tau$. Thus, we conclude that when $\theta > 1$, as $\bar{\lambda} \to \infty$, we have $\chi^{\pm} \to r^w - r^m$. When $\theta < 1$, by the symmetry property (Proposition 7), $\chi^{\pm} = 0, \bar{r}^f = r^m$. $\square$

M. Proof of Corollary 3

Proof. The proof proceeds in two stages: forward induction to establish monotonicity of market tightness, followed by backward induction to establish monotonicity of outside options and rates.

Part 1: Forward induction on θ_n.

Assume $\frac{\partial \theta_k}{\partial \theta_0} > 0$ for all $k \leq n$. From the market tightness recursion:

$$\theta_{n+1} = \theta_n \frac{1-\psi_{n+1}^-}{1-\psi_{n+1}^+},$$

where $\psi_{n+1}^+ = \lambda(N)G(1,\theta_n)$ and $\psi_{n+1}^- = \lambda(N)G(\theta_n^{-1},1)$. Since G is increasing in each argument:

$$\frac{\partial \psi_{n+1}^+}{\partial \theta_n} = \lambda(N)G_2(1,\theta_n) > 0, \qquad \frac{\partial \psi_{n+1}^-}{\partial \theta_n} = \lambda(N)G_1(\theta_n^{-1},1)\cdot(-\theta_n^{-2}) < 0.$$

Taking the derivative of θ_{n+1} with respect to θ_0:

$$\frac{\partial \theta_{n+1}}{\partial \theta_0} = \frac{\partial \theta_n}{\partial \theta_0}\cdot \frac{1-\psi_{n+1}^-}{1-\psi_{n+1}^+} + \theta_n \frac{\partial}{\partial \theta_0}\left[\frac{1-\psi_{n+1}^-}{1-\psi_{n+1}^+}\right].$$

For $\theta_0 > 1$, we have $\theta_n > 1$ for all n. The first term is positive by the inductive hypothesis. The second term is also positive because as θ_n increases, the numerator $(1-\psi_{n+1}^-)$ increases (since $\frac{\partial \psi_{n+1}^-}{\partial \theta_n} < 0$) and the denominator $(1-\psi_{n+1}^+)$ decreases (since $\frac{\partial \psi_{n+1}^+}{\partial \theta_n} > 0$). Therefore, $\frac{\partial \theta_{n+1}}{\partial \theta_0} > 0$, completing the induction. This establishes Part 1 and implies:

$$\frac{\partial \psi_n^+}{\partial \theta_0} > 0 \quad \text{and} \quad \frac{\partial \psi_n^-}{\partial \theta_0} < 0 \quad \text{for all } n.$$

Part 2: Backward induction on outside options and rates. The terminal values are constants:

$$\chi_N^+ = 0 \implies \frac{\partial \chi_N^+}{\partial \theta_0} = 0, \quad \chi_N^- = r^w - r^m \implies \frac{\partial \chi_N^-}{\partial \theta_0} = 0.$$

Proceeding by backward induction: Assume for round $n+1$ that $\frac{\partial \chi_{n+1}^+}{\partial \theta_0} \geq 0$, $\frac{\partial \chi_{n+1}^-}{\partial \theta_0} \geq 0$, and $\frac{\partial r_{n+1}^f}{\partial \theta_0} \geq 0$. From the outside option recursions:

$$\begin{aligned} \chi_n^+ &= \psi_{n+1}^+(r_{n+1}^f - r^m) + (1 - \psi_{n+1}^+)\chi_{n+1}^+, \\ \chi_n^- &= \psi_{n+1}^-(r_{n+1}^f - r^m) + (1 - \psi_{n+1}^-)\chi_{n+1}^-. \end{aligned}$$

Analysis for χ_n^+: Differentiating with respect to θ_0:

$$\frac{\partial \chi_n^+}{\partial \theta_0} = \frac{\partial \psi_{n+1}^+}{\partial \theta_0}(r_{n+1}^f - r^m - \chi_{n+1}^+) + \psi_{n+1}^+ \frac{\partial r_{n+1}^f}{\partial \theta_0} + (1 - \psi_{n+1}^+)\frac{\partial \chi_{n+1}^+}{\partial \theta_0}.$$

From the bargaining solution, $r_{n+1}^f = r^m + \eta\chi_{n+1}^+ + (1-\eta)\chi_{n+1}^-$, which implies:

$$r_{n+1}^f - r^m - \chi_{n+1}^+ = (1-\eta)(\chi_{n+1}^- - \chi_{n+1}^+) \equiv (1-\eta)\Sigma_{n+1} > 0,$$

where $\Sigma_{n+1} > 0$ is the total surplus. Examining each term:

1. $\frac{\partial \psi_{n+1}^+}{\partial \theta_0} \cdot (1-\eta)\Sigma_{n+1} = (+) \cdot (+) > 0$,
2. $\psi_{n+1}^+ \cdot \frac{\partial r_{n+1}^f}{\partial \theta_0} \geq 0$ (by inductive hypothesis),
3. $(1 - \psi_{n+1}^+) \cdot \frac{\partial \chi_{n+1}^+}{\partial \theta_0} \geq 0$ (by inductive hypothesis).

Therefore, $\frac{\partial \chi_n^+}{\partial \theta_0} > 0$.

Similarly, differentiating:

$$\frac{\partial \chi_n^-}{\partial \theta_0} = \frac{\partial \psi_{n+1}^-}{\partial \theta_0}(r_{n+1}^f - r^m - \chi_{n+1}^-) + \psi_{n+1}^- \frac{\partial r_{n+1}^f}{\partial \theta_0} + (1 - \psi_{n+1}^-)\frac{\partial \chi_{n+1}^-}{\partial \theta_0}.$$

Now, $r_{n+1}^f - r^m - \chi_{n+1}^- = \eta(\chi_{n+1}^+ - \chi_{n+1}^-) = -\eta\Sigma_{n+1} < 0$. Examining the terms:

1. $\frac{\partial \psi_{n+1}^-}{\partial \theta_0} \cdot (-\eta \Sigma_{n+1}) = (-) \cdot (-) > 0$,

2. $\psi_{n+1}^- \cdot \frac{\partial r_{n+1}^f}{\partial \theta_0} \geq 0$ (by inductive hypothesis),

3. $(1 - \psi_{n+1}^-) \cdot \frac{\partial \chi_{n+1}^-}{\partial \theta_0} \geq 0$ (by inductive hypothesis).

Therefore, $\frac{\partial \chi_n^-}{\partial \theta_0} > 0$.

Since $r_n^f = r^m + \eta \chi_n^+ + (1-\eta)\chi_n^-$:

$$\frac{\partial r_n^f}{\partial \theta_0} = \eta \frac{\partial \chi_n^+}{\partial \theta_0} + (1-\eta)\frac{\partial \chi_n^-}{\partial \theta_0} > 0.$$

This completes the backward induction for discrete rounds. □

Corollary 5 (Continuous-Time Limit)**.** *The monotonicity properties extend to the continuous-time limit. For $\theta_0 > 1$ and all $\tau \in (0,1)$:*

$$\frac{\partial \chi_\tau^+}{\partial \theta_0} > 0, \quad \frac{\partial \chi_\tau^-}{\partial \theta_0} > 0, \quad \frac{\partial r_\tau^f}{\partial \theta_0} > 0.$$

Proof. As $N \to \infty$, the discrete-time equilibrium converges uniformly to the continuous-time equilibrium on $[0,1]$. Since strict inequalities are preserved under uniform convergence, the monotonicity properties carry over to the continuous-time limit. □

N. Proof of Proposition 10

Proof. We first prove the result for $\theta \to 0$ and then use symmetry to show the opposite limit. As $\theta \to 0$:

$$\frac{\dot{\theta}}{\theta} = -\bar{\lambda}[\gamma(\theta^{-1}) - \gamma(\theta)] \approx -\bar{\lambda}\gamma(\theta^{-1}) \to -\bar{\lambda}\bar{\gamma},$$

where we used the no-disposal property of the matching function to establish that: $\gamma(\theta) \to 0$ as $\theta \to 0$. If $\bar{\gamma} < \infty$, matching intensities are finite, so trade is active throughout all rounds. Moreover, θ_τ approximately satisfied exponential decay: $\theta_\tau \sim \theta_0 e^{-\bar{\lambda}\bar{\gamma}\tau}$.

For surplus traders:

$$\Psi^+ = 1 - \exp\left(-\int_0^1 \bar{\lambda}\gamma(\theta_\tau)d\tau\right) \to 0$$

since $\gamma(\theta) \to 0$ as $\theta \to 0$.

For deficit traders, if $\bar{\gamma} < \infty$:

$$\Psi^- = 1 - \exp\left(-\int_0^1 \bar{\lambda}\gamma(\theta_\tau^{-1})d\tau\right) \approx 1 - \exp(-\bar{\lambda}\bar{\gamma}).$$

If $\bar{\gamma} = \infty$, then $\Psi^- \to 1$ as $\theta \to 0$.

From Proposition 4, for finite $\bar{\gamma}$:

$$\chi^+ \to (r^w - r^m)\int_0^1 (1-\eta)\psi_y^+ \Sigma_y dy \to 0,$$

given that ψ_y^+ converges uniformly to zero as $\theta \to 0$ and the surplus is always finite. For χ^- we have that:

$$\chi^- \to (r^w - r^m)\left[1 - \int_0^1 \eta\bar{\lambda}\bar{\gamma}e^{-\eta\bar{\lambda}\bar{\gamma}(1-y)}dy\right] = (r^w - r^m)e^{-\bar{\lambda}\bar{\gamma}\eta}.$$

The first limit passes limits inside the integral and uses the continuity of $\exp(\cdot)$.

For $\bar{\gamma}$ infinite, we bound χ^- above by $(r^w - r^m)e^{-\bar{\lambda}\gamma(\theta_0)\eta}$, noting that $\gamma(\theta_\tau)$ is increasing in time. Then apply the limit as $\theta_0 \to \infty$ to find that $\chi^- = 0$. Hence, the solution above remains valid.

Since $\chi^+ \to 0$, the average interbank rate is:

$$\bar{r}^f = r^m + \eta(r^w - r^m)e^{-\bar{\lambda}\bar{\gamma}\eta}.$$

Finally, we obtain the corresponding formulas for $\theta \to \infty$. By symmetry. Using Proposition 7, this is case (i) with $\theta \leftrightarrow \theta^{-1}$, $\psi^+ \leftrightarrow \psi^-$, and $\eta \leftrightarrow 1-\eta$. Thus:

$$\Psi^+ = 1 - e^{-\bar{\lambda}\bar{\gamma}}, \quad \Psi^- = 0, \quad \chi^+ = (r^w - r^m)(1 - e^{-(1-\eta)\bar{\lambda}\bar{\gamma}}), \quad \chi^- = r^w - r^m.$$

□

O. Analytic Solutions: Calculations for Table 1

O.1 Leontief Case

For the Leontief matching function $G(a,b) = \min\{a,b\}$, we have $\gamma(\theta) = \min\{1,\theta\}$.

Case 1: $\theta_0 > 1$. By Lemma 1, ODE for market tightness becomes:

$$\dot{\theta}_\tau = -\bar{\lambda}\theta_\tau[\theta_\tau^{-1} - 1] = \bar{\lambda}(\theta_\tau - 1).$$

Solving this linear ODE with initial condition θ_0:

$$\theta_\tau = 1 + (\theta_0 - 1)e^{\bar{\lambda}\tau}$$

The matching intensities are:

$$\begin{aligned}
\psi_\tau^+ &= \bar{\lambda}\gamma(\theta_\tau) = \bar{\lambda} \\
\psi_\tau^- &= \bar{\lambda}\gamma(\theta_\tau^{-1}) = \frac{\bar{\lambda}}{\theta_\tau} = \frac{\bar{\lambda}}{1 + (\theta_0 - 1)e^{\bar{\lambda}\tau}}.
\end{aligned}$$

The matching probability, Ψ^+, is thus:

$$\Psi^+ = 1 - \exp\left(-\int_0^1 \bar{\lambda} d\tau\right) = 1 - e^{-\bar{\lambda}}.$$

For Ψ^-, we use that $\Psi^- = \Psi^+/\theta_0$ to conclude that:

$$\Psi^- = \frac{1 - e^{-\bar{\lambda}}}{\theta_0}.$$

We can verify that the same conclusion is reached by evaluating the integral: $\Psi^- = 1 - \exp(-\int_0^1 \psi_\tau^- d\tau)$.

Case 2: $\theta_0 < 1$. By symmetry (Proposition 7), we have that:

$$\theta_\tau^{-1} = 1 + (\theta_0^{-1} - 1)e^{\bar{\lambda}\tau} \Leftrightarrow \theta_\tau = \frac{\theta_0}{\theta_0 + (1 - \theta_0)e^{\bar{\lambda}\tau}}.$$

And, also, by symmetry:

$$\Psi^+ = \theta_0(1 - e^{-\bar{\lambda}}), \quad \Psi^- = 1 - e^{-\bar{\lambda}}.$$

O.2 Cobb-Douglas Case

For the Cobb-Douglas matching function $G(a, b) = a^{1/2}b^{1/2}$, we have $\gamma(\theta) = \theta^{1/2}$. The ODE for tightness specializes to:

$$\dot{\theta}_\tau = -\bar{\lambda}\theta_\tau[\theta_\tau^{-1/2} - \theta_\tau^{1/2}] = -\bar{\lambda}\theta_\tau^{1/2}(1 - \theta_\tau).$$

We solve the ODE by change of variables: $u \equiv \theta_\tau^{1/2}$. Then $\dot{\theta}_\tau = 2u\dot{u}$, giving:

$$2u\dot{u} = -\bar{\lambda}u(1-u^2) \Rightarrow \dot{u} = -\frac{\bar{\lambda}}{2}(1-u^2)$$

Separating variables:

$$\frac{du}{1-u^2} = -\frac{\bar{\lambda}}{2}d\tau.$$

Using partial fractions, we write:

$$\frac{1}{1-u^2} = \frac{1}{2}\left(\frac{1}{1+u} + \frac{1}{1-u}\right),$$

Thus, integrating both sides of the relationship above:

$$\frac{1}{2}\ln\left(\frac{1+u}{1-u}\right) = -\frac{\bar{\lambda}\tau}{2} + C.$$

Since, the initial condition $u(0) = \sqrt{\theta_0}$, we obtain:

$$\ln\left(\frac{1+u}{1-u}\right) = -\bar{\lambda}\tau + \ln\left(\frac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}}\right)$$

We clear out u taking the exponential on both sides:

$$\frac{1+u_\tau}{1-u_\tau} = \frac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}}e^{-\bar{\lambda}\tau}.$$

Therefore:

$$u_\tau = \frac{(1+\sqrt{\theta_0})e^{-\bar{\lambda}\tau} - (1-\sqrt{\theta_0})}{(1+\sqrt{\theta_0})e^{-\bar{\lambda}\tau} + (1-\sqrt{\theta_0})} \Rightarrow \theta_\tau = u_\tau^2 = \left(\frac{(1+\sqrt{\theta_0})e^{-\bar{\lambda}\tau} - (1-\sqrt{\theta_0})}{(1+\sqrt{\theta_0})e^{-\bar{\lambda}\tau} + (1-\sqrt{\theta_0})}\right)^2.$$

We can verify that for $\tau = 0$ the initial condition holds.

Stopping time: For the Cobb-Douglas case, the asymptotic limit $\bar{\gamma}$ is unbounded. Thus, trade can vanish in finite time. The ODE is valid locally as long as trade occurs. Next, we obtain the value of time at which the trade stops.

Trade stops when $\theta_\tau = 0$ which can be reached if $\theta_0 < 1$. Likewise, trade stops at the time where $\theta_\tau = 0$ asymptotes to infinity, (a case only relevant when $\theta_0 > 1$).

When $\theta_0 < 1$, note that the denominator in the formula θ_τ is always positive. Thus, the numerator reaches zero at: $T = \frac{1}{\bar{\lambda}}\ln\left(\frac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}}\right)$. When $\theta_0 > 1$, the numerator is always positive, but the denominator reaches zero at $T = \frac{1}{\bar{\lambda}} \ln \left(\frac{1+\sqrt{\theta_0}}{\sqrt{\theta_0}-1} \right)$. Thus, the stopping time in either case occurs when:

$$T = \min \left\{ \frac{1}{\bar{\lambda}} \ln \left| \frac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}} \right|, 1 \right\}.$$

The effective stopping time is $\min\{T, 1\}$.

Matching probabilities: Matching probabilities follow from:

$$\Psi^{\pm} = 1 - \exp\left(- \int_0^T \psi_s^{\pm} ds \right).$$

where $\psi_\tau^+ = \bar{\lambda}\sqrt{\theta_\tau} = \bar{\lambda} u_\tau$ and $\psi_\tau^- = \bar{\lambda}/\sqrt{\theta_\tau} = \bar{\lambda} u_\tau^{-1}$. We proceed with some calculations:

$$\int_0^T \psi^{\pm} d\tau = \int_0^T \bar{\lambda} u_s ds.$$

Let $\alpha = \frac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}}$. Then:

$$u_s = \frac{\alpha e^{-\bar{\lambda} s} - 1}{\alpha e^{-\bar{\lambda} s} + 1}.$$

To integrate, note that: $u_s = \frac{1}{\bar{\lambda}} \frac{\partial[-\bar{\lambda} s - 2\ln(\alpha e^{-\bar{\lambda} s}+1)]}{\partial s} = \frac{1}{\bar{\lambda}}(-\bar{\lambda} + \frac{2\bar{\lambda} e^{-\bar{\lambda} s}}{\alpha e^{-\bar{\lambda} s}+1})$. Therefore:

$$\int_0^T -\bar{\lambda} u_s ds = \int_0^T \frac{\partial[\bar{\lambda} s + 2\ln(\alpha e^{-\bar{\lambda} s} + 1)]}{\partial s} ds = \bar{\lambda} s + 2\ln(\alpha e^{-\bar{\lambda} s} + 1)|_{s=0}^{s=T}.$$

Thus,

$$\exp\left(- \int_0^T \psi_s^- ds \right) = \exp\left(\bar{\lambda} \min\{T, 1\} + 2\ln\left(\frac{\alpha e^{-\bar{\lambda} s} + 1}{\alpha + 1} \right) \right).$$

Thus, we have that:

$$\Psi^+ = 1 - \exp(-\bar{\lambda} T) \left(\frac{\alpha + e^{-\bar{\lambda} T}}{\alpha + 1} \right)^2.$$

Substituting α, we get the formula in the table.

The formula for Ψ^- is obtained using $\Psi^- = \Psi^+/\theta_0 = \Psi^+/u_0^2$:

$$\Psi^- = \frac{\Psi^+}{\theta_0} = \left(\frac{\alpha+1}{\alpha-1} \right)^2, \quad \Psi^+ = \left(\frac{\alpha+1}{\alpha-1} \right)^2 - \exp(-\bar{\lambda} T) \left(\frac{\alpha e^{-\bar{\lambda} T} + 1}{\alpha - 1} \right)^2.$$

Substituting back $\alpha = \frac{1+\sqrt{\theta_0}}{1-\sqrt{\theta_0}}$:

$$\begin{aligned}
\Psi^- &= \left(\frac{2\sqrt{\theta_0}}{2}\right)^2 \left(\frac{1}{\theta_0}\right) - \exp(-\bar{\lambda}T) \left(\frac{(1+\sqrt{\theta_0})e^{-\bar{\lambda}T} + (1-\sqrt{\theta_0})}{2\sqrt{\theta_0}}\right)^2 \left(\frac{1}{\theta_0}\right) \\
&= 1 - \exp(-\bar{\lambda}T) \left(\frac{(1+\sqrt{\theta_0})e^{-\bar{\lambda}T} - (1-\sqrt{\theta_0})e^{\bar{\lambda}T}}{2\sqrt{\theta_0}}\right)^2 \\
&= 1 - \exp(-\bar{\lambda}T) \left(\frac{(1+\sqrt{\theta_0}) - (1-\sqrt{\theta_0})e^{\bar{\lambda}T}}{(1+\sqrt{\theta_0}) - (1-\sqrt{\theta_0})}\right)^2 .
\end{aligned}$$

This completes the derivation of both Ψ^+ and Ψ^- for the Cobb-Douglas case.

P. Verification of Special Cases

We derive the convenience yield coefficients for both the Leontief and Cobb-Douglas matching functions by computing their integrals.

P.1 Verification of the Leontief Case ($\theta_0 > 1$)

Step 1: Specialize market dynamics. With the Leontief matching function, from Table 1 we have:

$$\theta_\tau = 1 + (\theta_0 - 1)e^{\bar{\lambda}\tau}, \quad \psi_\tau^+ = \bar{\lambda}, \quad \psi_\tau^- = \frac{\bar{\lambda}}{\theta_\tau} = \frac{\bar{\lambda}}{1 + (\theta_0 - 1)e^{\bar{\lambda}\tau}}$$

The terminal tightness is $\bar{\theta} = \theta_1 = 1 + (\theta_0 - 1)e^{\bar{\lambda}}$. Thus, we compute the following integrals:

$$\int_t^1 \psi_\tau^+ \, d\tau = \int_t^1 \bar{\lambda} \, d\tau = \bar{\lambda}(1-t),$$

and

$$\begin{aligned}
\int_t^1 \psi_\tau^- \, d\tau = \int_t^1 \frac{\bar{\lambda}}{1 + (\theta_0 - 1)e^{\bar{\lambda}\tau}} \, d\tau &= \left[\bar{\lambda}\tau - \ln\left(1 + (\theta_0 - 1)e^{\bar{\lambda}\tau}\right)\right]_t^1 \\
&= \bar{\lambda}(1-t) - \ln\left(\frac{1 + (\theta_0 - 1)e^{\bar{\lambda}}}{1 + (\theta_0 - 1)e^{\bar{\lambda}t}}\right).
\end{aligned}$$

The convenience yield coefficients (outside options) take the following forms::

$$\chi_\tau^+ = (1-\eta)(r^w - r^m)\mathbb{P}_\tau^+ \quad \text{and} \quad \chi_\tau^- = (r^w - r^m)(1-\eta\mathbb{P}_\tau^-), \tag{A.20}$$

where

$$\begin{aligned}
\mathbb{P}_t^+ &= \int_t^1 \psi_\tau^+ e^{-[\eta\Psi_\tau^- + (1-\eta)\Psi_\tau^+]}\, d\tau \\
&= \int_t^1 \bar{\lambda} e^{-\bar{\lambda}(1-\tau)+\eta\ln\left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right)}\, d\tau \\
&= \int_t^1 \left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right)^\eta \bar{\lambda} e^{-\bar{\lambda}(1-\tau)}\, d\tau \\
&= \left[\frac{1+(\theta_0-1)e^{\bar{\lambda}\tau}}{(1-\eta)(\theta_0-1)e^{\bar{\lambda}}}\left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right)^\eta\right]_t^1 \\
&= \frac{1+(\theta_0-1)e^{\bar{\lambda}}}{(1-\eta)(\theta_0-1)e^{\bar{\lambda}}} - \frac{1+(\theta_0-1)e^{\bar{\lambda}t}}{(1-\eta)(\theta_0-1)e^{\bar{\lambda}}}\left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}t}}\right)^\eta \\
&= \frac{\theta_1\left(1-\left(\frac{\theta_t}{\theta_1}\right)^{1-\eta}\right)}{(1-\eta)(\theta_1-1)} = \frac{1}{1-\eta}\cdot\frac{\theta_1 - \theta_t^{1-\eta}\theta_1^\eta}{\theta_1 - 1} \\
&= \frac{1}{1-\eta}\left(\frac{\bar{\theta}}{\theta}\right)^\eta \left(\frac{\theta^\eta\bar{\theta}^{1-\eta}-\theta}{\bar{\theta}-1}\right).
\end{aligned}$$

and where

$$\begin{aligned}
\mathbb{P}_t^- &= \int_t^1 \psi_\tau^- e^{-[\eta\Psi_\tau^- + (1-\eta)\Psi_\tau^+]}\, d\tau \\
&= \int_t^1 \left(\frac{\bar{\lambda}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right) e^{-\bar{\lambda}(1-\tau)+\eta\ln\left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right)}\, d\tau \\
&= \int_t^1 \left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right)^\eta \left(\frac{\bar{\lambda}e^{-\bar{\lambda}(1-\tau)}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right) d\tau \\
&= \left[\left(\frac{-1}{\eta(\theta_0-1)e^{\bar{\lambda}}}\right)\left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}\tau}}\right)^\eta\right]_t^1 \\
&= \frac{1}{\eta(\theta_0-1)e^{\bar{\lambda}}}\left(\frac{1+(\theta_0-1)e^{\bar{\lambda}}}{1+(\theta_0-1)e^{\bar{\lambda}t}}\right)^\eta - \frac{1}{\eta(\theta_0-1)e^{\bar{\lambda}}} \\
&= \frac{\left(\frac{\theta_1}{\theta_t}\right)^\eta - 1}{\eta(\theta_1-1)}.
\end{aligned}$$

Thus, we obtain the following formulas:

$$\chi_\tau^+ = (r^w - r^m)\left(\frac{\theta_1}{\theta_\tau}\right)^\eta \left(\frac{\theta_\tau^\eta \theta_1^{1-\eta} - \theta_\tau}{\theta_1 - 1}\right) \quad \chi_\tau^- = (r^w - r^m)\left(\frac{\theta_1}{\theta_\tau}\right)^\eta \left(\frac{\theta_\tau^\eta \theta_1^{1-\eta} - 1}{\theta_1 - 1}\right).$$

From here, substitute for $\tau = 0$ and obtain $\{\chi_0^+, \chi_0^-\}$.

P.2 Leontief Case ($\theta_0 < 1$) via Symmetry

For $\theta_0 < 1$, we can apply the symmetry property from Proposition 7 rather than repeat the calculations.

$$\chi^+(\theta_0, \eta) = (r^w - r^m) - \chi^-(\theta_0^{-1}, 1-\eta).$$

Since $\theta_0^{-1} > 1$, we can use our formula:

$$\chi^-(\theta_0^{-1}, 1-\eta) = (r^w - r^m)\frac{\bar{\theta}^{-1} - (\bar{\theta}^{-1})^{1-\eta}(\theta_0^{-1})^{-(1-\eta)}}{\bar{\theta}^{-1} - 1} = (r^w - r^m)\frac{1 - \bar{\theta}^\eta \theta_0^{(1-\eta)}}{1 - \bar{\theta}}.$$

Thus, we have that:

$$\chi^+(\theta_0, \eta) = (r^w - r^m)\left(1 - \frac{1 - \bar{\theta}^\eta \theta_0^{(1-\eta)}}{1 - \bar{\theta}}\right) = (r^w - r^m)\left(\frac{\bar{\theta} - \bar{\theta}^\eta \theta_0^{(1-\eta)}}{\bar{\theta} - 1}\right).$$

Thus, the same formula is valid for $\chi^+(\theta_0, \eta)$ regardless of whether $\theta_0 > 1$ or $\theta_0 < 1$. Since the formula for χ^+ holds when $\theta_0 < 1$, the formula for χ^- must also hold in both cases.

P.3 Cobb-Douglas Case

TBA

Step 1: Specialize market dynamics. With the Cobb-Douglas matching function:

$$\theta_\tau = \left(\frac{(1+\sqrt{\theta_0})e^{-\bar{\lambda}\tau} - (1-\sqrt{\theta_0})}{(1+\sqrt{\theta_0})e^{-\bar{\lambda}\tau} + (1-\sqrt{\theta_0})}\right)^2, \quad \psi_\tau^+ = \bar{\lambda}\sqrt{\theta_\tau}, \quad \psi_\tau^- = \frac{\bar{\lambda}}{\sqrt{\theta_\tau}}.$$

Define $u = e^{-\bar{\lambda}\tau}$. Then:

$$\theta_\tau = \left(\frac{\alpha u - 1}{\alpha u + 1}\right)^2.$$

Step 2: Computing the surplus The surplus at time τ is:

$$\Sigma_\tau = (r^w - r^m)\exp\left(-\int_\tau^T [\eta\psi_s^- + (1-\eta)\psi_s^+]ds\right)$$

From previous calculations:

$$\int_\tau^T \psi_s^- ds = \ln\left(\frac{\alpha - e^{\bar{\lambda}T}}{\alpha + e^{\bar{\lambda}T}}\right) - \ln\left(\frac{\alpha - e^{\bar{\lambda}\tau}}{\alpha + e^{\bar{\lambda}\tau}}\right) \tag{A.21}$$

$$\int_\tau^T \psi_s^+ ds = \ln\left(\frac{\alpha + e^{\bar{\lambda}\tau}}{\alpha - e^{\bar{\lambda}\tau}}\right) - \ln\left(\frac{\alpha + e^{\bar{\lambda}T}}{\alpha - e^{\bar{\lambda}T}}\right) \tag{A.22}$$

For stopping time T (when exhaustion occurs): $\alpha e^{-\bar{\lambda}T} = 1 \implies T = \frac{1}{\bar{\lambda}}\ln(\alpha)$

Step 4: Computing $\mathbb{P}_t^+$ Starting with:

$$\mathbb{P}_t^+ = \int_t^T \psi_\tau^+ e^{-[\eta\Psi_\tau^- + (1-\eta)\Psi_\tau^+]} d\tau$$

Substituting the exponent terms:

$$\mathbb{P}_t^+ = \int_t^T \bar{\lambda}\sqrt{\theta_\tau}\exp\left[-\eta\ln\left(\frac{\alpha - e^{\bar{\lambda}T}}{\alpha + e^{\bar{\lambda}T}} \cdot \frac{\alpha + e^{\bar{\lambda}\tau}}{\alpha - e^{\bar{\lambda}\tau}}\right)\right. \tag{A.23}$$

$$\left. -(1-\eta)\ln\left(\frac{\alpha + e^{\bar{\lambda}\tau}}{\alpha - e^{\bar{\lambda}\tau}} \cdot \frac{\alpha - e^{\bar{\lambda}T}}{\alpha + e^{\bar{\lambda}T}}\right)\right] d\tau \tag{A.24}$$

Simplifying the exponential:

$$\mathbb{P}_t^+ = \int_t^T \bar{\lambda}\left(\frac{\alpha e^{-\bar{\lambda}\tau} - 1}{\alpha e^{-\bar{\lambda}\tau} + 1}\right)\left(\frac{\alpha - e^{\bar{\lambda}T}}{\alpha + e^{\bar{\lambda}T}}\right)^{-\eta}\left(\frac{\alpha + e^{\bar{\lambda}\tau}}{\alpha - e^{\bar{\lambda}\tau}}\right)^{-\eta}\left(\frac{\alpha + e^{\bar{\lambda}\tau}}{\alpha - e^{\bar{\lambda}\tau}}\right)^{-(1-\eta)} d\tau$$

This simplifies to:

$$\mathbb{P}_t^+ = \int_t^T \bar{\lambda}\left(\frac{\alpha e^{-\bar{\lambda}\tau} - 1}{\alpha e^{-\bar{\lambda}\tau} + 1}\right)\left(\frac{\alpha - e^{\bar{\lambda}T}}{\alpha + e^{\bar{\lambda}T}}\right)^{-\eta}\left(\frac{\alpha - e^{\bar{\lambda}\tau}}{\alpha + e^{\bar{\lambda}\tau}}\right) d\tau$$

Using substitution $u = e^{-\bar{\lambda}\tau}$, $du = -\bar{\lambda}e^{-\bar{\lambda}\tau}d\tau = -\bar{\lambda}u\,d\tau$:

$$\mathbb{P}_t^+ = \int_{e^{-\bar{\lambda}t}}^{1/\alpha} \left(\frac{\alpha u - 1}{\alpha u + 1}\right)\left(\frac{\alpha - 1}{\alpha + 1}\right)^{-\eta}\left(\frac{\alpha - 1/u}{\alpha + 1/u}\right)\left(-\frac{du}{u}\right)$$

After integration:

$$\mathbb{P}_t^+ = \frac{1}{2\alpha(1-\eta)}\left[\frac{\alpha e^{-\bar{\lambda}t} + 1}{\alpha e^{-\bar{\lambda}t} - 1}\left(\frac{\alpha - e^{\bar{\lambda}t}}{\alpha + e^{\bar{\lambda}t}}\right)^{\eta} - \frac{2\alpha}{1-\eta}\right]$$

Step 5: Computing $\mathbb{P}_t^-$ Similarly:

$$\mathbb{P}_t^- = \int_t^T \psi_\tau^- e^{-[\eta\Psi_\tau^- + (1-\eta)\Psi_\tau^+]}d\tau$$

With full exponent terms:

$$\mathbb{P}_t^- = \int_t^T \frac{\bar{\lambda}}{\sqrt{\theta_\tau}}\left(\frac{\alpha - e^{\bar{\lambda}T}}{\alpha + e^{\bar{\lambda}T}}\right)^{-\eta}\left(\frac{\alpha + e^{\bar{\lambda}\tau}}{\alpha - e^{\bar{\lambda}\tau}}\right)^{-\eta}\left(\frac{\alpha + e^{\bar{\lambda}\tau}}{\alpha - e^{\bar{\lambda}\tau}}\right)^{-(1-\eta)} d\tau$$

After simplification and integration:

$$\mathbb{P}_t^- = \frac{1}{\eta}\left[\left(\frac{\alpha - e^{\bar{\lambda}t}}{\alpha + e^{\bar{\lambda}t}}\right)^{\eta} - \left(\frac{\alpha - 1}{\alpha + 1}\right)^{\eta}\right]$$

Step 6: Final expressions for χ^+ and χ^- Setting $t = 0$ and using the relationships:

$$\chi^+ = (r^w - r^m)\frac{\bar{\theta} - \bar{\theta}^\eta\theta_0^{1-\eta}}{\bar{\theta} - 1}, \quad \chi^- = (r^w - r^m)\frac{\bar{\theta} - \bar{\theta}^\eta\theta_0^{-\eta}}{\bar{\theta} - 1}$$

Q. Proof of Proposition 11

Proof. Recall that:

$$\chi^+ = (r^w - r^m)\frac{\bar{\theta} - \bar{\theta}^\eta\theta^{1-\eta}}{\bar{\theta} - 1}, \quad \chi^- = (r^w - r^m)\frac{\bar{\theta} - \bar{\theta}^\eta\theta^{-\eta}}{\bar{\theta} - 1},$$

where $\bar{\theta} = \bar{\theta}(\theta, \bar{\lambda})$ is the terminal market tightness given in Table 1. Define:

$$h^+(\bar{\theta}) \equiv \frac{\bar{\theta} - \bar{\theta}^\eta \theta^{1-\eta}}{\bar{\theta} - 1}, \quad h^-(\bar{\theta}) \equiv \frac{\bar{\theta} - \bar{\theta}^\eta \theta^{-\eta}}{\bar{\theta} - 1}.$$

Step 1: Derivatives with respect to $\bar{\theta}$. For $h^+(\bar{\theta})$, using the quotient rule:

$$\begin{aligned} h^+_{\bar{\theta}}(\bar{\theta}) &= \frac{[1 - \eta\bar{\theta}^{\eta-1}\theta^{1-\eta}](\bar{\theta} - 1) - [\bar{\theta} - \bar{\theta}^\eta\theta^{1-\eta}]}{(\bar{\theta} - 1)^2} \\ &= \frac{\eta\bar{\theta}^{\eta-1}\theta^{1-\eta} + (1-\eta)\bar{\theta}^\eta\theta^{1-\eta} - 1}{(\bar{\theta} - 1)^2} \\ &= \frac{\bar{\theta}^\eta\theta^{1-\eta}[(1-\eta)\bar{\theta} + \eta] - \bar{\theta}}{(\bar{\theta} - 1)^2 \cdot \bar{\theta}}. \end{aligned}$$

Similarly, for $h^-(\bar{\theta})$:

$$h^-_{\bar{\theta}}(\bar{\theta}) = \frac{\bar{\theta}^\eta\theta^{-\eta}[(1-\eta)\bar{\theta} + \eta] - \bar{\theta}}{(\bar{\theta} - 1)^2 \cdot \bar{\theta}}.$$

Since $(\bar{\theta} - 1)^2 > 0$ and $\bar{\theta} > 0$, the sign of $h^\pm_{\bar{\theta}}$ is determined by:

$$\text{sign}(h^\pm_{\bar{\theta}}) = \text{sign}\left(\bar{\theta}^\eta\theta^{\mp\eta}[(1-\eta)\bar{\theta} + \eta] - \bar{\theta}\right). \tag{A.25}$$

Equivalently, $h^\pm_{\bar{\theta}} > 0$ if and only if:

$$\theta^{\mp\eta}[(1-\eta)\bar{\theta} + \eta] > \bar{\theta}^{1-\eta}. \tag{A.26}$$

Step 2: Case $\theta > 1$. For Leontief matching: $\bar{\theta} = 1 + (\theta - 1)e^{\bar{\lambda}} \in [\theta, \infty)$ and $\frac{\partial\bar{\theta}}{\partial\bar{\lambda}} = (\theta - 1)e^{\bar{\lambda}} > 0$. We test the monotonicity of χ^+. We verify condition (A.26) with the minus sign: $\theta^{1-\eta}[(1-\eta)\bar{\theta} + \eta] > \bar{\theta}^{1-\eta}$. At $\bar{\theta} = \theta$:

$$\theta^{1-\eta}[(1-\eta)\theta + \eta] = \theta^{1-\eta}[\theta - \eta(\theta - 1)] = \theta^{2-\eta} - \eta\theta^{1-\eta}(\theta - 1) > \theta^{1-\eta},$$

since $\theta > 1$ implies $\theta^{2-\eta} > \theta^{1-\eta}$ and $\eta\theta^{1-\eta}(\theta - 1) > 0$.

To verify the inequality holds throughout $[\theta, \infty)$, consider:

$$D(\bar{\theta}) = \theta^{1-\eta}[(1-\eta)\bar{\theta} + \eta] - \bar{\theta}^{1-\eta}.$$

Taking the derivative:

$$D'(\bar{\theta}) = (1-\eta)\theta^{1-\eta} - (1-\eta)\bar{\theta}^{-\eta} = (1-\eta)(\theta^{1-\eta} - \bar{\theta}^{-\eta}).$$

The term in parentheses is positive since $\theta, \bar{\theta} > 1$. Since $D(\theta) > 0$ and $D'(\bar{\theta}) > 0$ for $\bar{\theta} > \theta$, we have $D(\bar{\theta}) > 0$ for all $\bar{\theta} \geq \theta$. Therefore $h^+_{\bar{\theta}} > 0$ for all $\bar{\theta} \geq \theta$, and:

$$\frac{\partial \chi^+}{\partial \bar{\lambda}} = (r^w - r^m) h^+_{\bar{\theta}} \cdot \frac{\partial \bar{\theta}}{\partial \bar{\lambda}} > 0.$$

Next, we test the non-monotonicity of χ^-. We check condition (A.26) with the plus sign: $\theta^{-\eta}[(1-\eta)\bar{\theta} + \eta] > \bar{\theta}^{1-\eta}$. At $\bar{\theta} = \theta$:

$$\begin{aligned} \theta^{-\eta}[(1-\eta)\theta + \eta] &= \theta^{-\eta}[\theta - \eta(\theta - 1)] \\ &= \theta^{1-\eta} - \eta\theta^{-\eta}(\theta - 1) < \theta^{1-\eta}. \end{aligned}$$

So the condition is violated at $\bar{\theta} = \theta$. As $\bar{\theta} \to \infty$:

$$\theta^{-\eta}[(1-\eta)\bar{\theta} + \eta] \sim (1-\eta)\theta^{-\eta}\bar{\theta} \to \infty,$$

while $\bar{\theta}^{1-\eta} \to \infty$ at a slower rate (since $1 - \eta < 1$). Thus, the condition is eventually satisfied. By continuity of $\bar{\theta}$ in $\bar{\lambda}$, $h^-_{\bar{\theta}}$ changes sign on $[\theta, \infty)$, so χ^- is non-monotonic in $\bar{\lambda}$ when $\theta > 1$.

Finally, we check the monotonicity of $\bar{r}^f$. The average interest rate can be written as:

$$\bar{r}^f = \phi(\theta) r^m + (1 - \phi(\theta)) r^w,$$

where the endogenous bargaining power is:

$$\phi(\theta) = \begin{cases} 1 - \dfrac{\bar{\theta} - \theta^{1-\eta}\bar{\theta}^{\eta}}{\bar{\theta} - \theta} & \text{if } \theta > 1, \\ \eta & \text{if } \theta = 1, \\ 1 - \dfrac{\theta^{-1} - \theta^{-\eta}\bar{\theta}^{\eta-1}}{\bar{\theta}^{-1} - \theta^{-1}} & \text{if } \theta < 1. \end{cases}$$

For $\theta > 1$, note that $\phi(\theta) = 1 - h^+(\bar{\theta})$. Since we have shown that $h^+_{\bar{\theta}} > 0$:

$$\frac{\partial \phi}{\partial \bar{\lambda}} = -\frac{\partial h^+}{\partial \bar{\theta}} \cdot \frac{\partial \bar{\theta}}{\partial \bar{\lambda}} < 0.$$

Therefore:

$$\frac{\partial \bar{r}^f}{\partial \bar{\lambda}} = \frac{\partial \phi}{\partial \bar{\lambda}}(r^m - r^w) = -\frac{\partial \phi}{\partial \bar{\lambda}}(r^w - r^m) > 0.$$

Thus $\bar{r}^f$ is increasing in $\bar{\lambda}$ when $\theta > 1$.

Step 3: Case $\theta < 1$. he results for $\theta < 1$ follow by symmetry (Proposition 7).

Step 4: Case $\theta = 1$. When $\theta = 1$, $\bar{\theta} = 1$ for all $\bar{\lambda}$, so $\bar{r}^f = r^m + (1-\eta)(r^w - r^m)$ is constant. Using the explicit formulas:

$$\chi^+ = \Psi^+(\bar{r}^f - r^m), \quad \chi^- = (r^w - r^m) - \Psi^-(r^w - \bar{r}^f),$$

where $\Psi^\pm = 1 - e^{-\bar{\lambda}}$. Therefore:

$$\frac{\partial \chi^+}{\partial \bar{\lambda}} = (1-\eta)(r^w - r^m)e^{-\bar{\lambda}} > 0, \quad \frac{\partial \chi^-}{\partial \bar{\lambda}} = -\eta(r^w - r^m)e^{-\bar{\lambda}} < 0.$$

□

R. Proof of Corollary 4

Proof. When $\theta > 1$, the OTC rate is decreasing over time (Corollary 1), so $Q = r_0^f - r_1^f$. When $\theta < 1$, the rate is increasing, so $Q = r_1^f - r_0^f$. Since the terminal rate r_1^f is constant, all variation in Q comes from the initial rate r_0^f. From Corollary 3, $\bar{r}^f$ is increasing in both θ and $\bar{\lambda}$ when $\theta > 1$, and decreasing in both when $\theta < 1$. Therefore, for $\theta > 1$, both $\frac{\partial Q}{\partial \theta} > 0$ and $\frac{\partial Q}{\partial \bar{\lambda}} > 0$. For $\theta < 1$, the sign of Q reverses, so $\frac{\partial Q}{\partial \theta} = -\frac{\partial \bar{r}^f}{\partial \theta} < 0$ while $\frac{\partial Q}{\partial \bar{\lambda}} = -\frac{\partial \bar{r}^f}{\partial \bar{\lambda}} > 0$. When $\theta = 1$, the rate is constant over time, so $Q = 0$ and both derivatives vanish. □

S. Derivation of Portfolio Conditions

S.1 Pricing Conditions

Risk Aversion Case. We have the following problem:

$$\max_A \left(\mathbb{E}_{X,\omega} \left[\left(R^m e + \sum_{i\in\mathbb{I}} \underbrace{\left(R^i\left(X_t\right) - R^m\right)}_{\text{Liquidity Premium}} \bar{a}^i_t + \underbrace{\chi_{t+1} \left(s \left(\{\bar{a}\}_{i\in\mathbb{I}}, e - \sum_{i\in\mathbb{I}} \bar{a}^i_t \right) \right)}_{\text{Liquidity Yield}} \right)^{1-\gamma} \right] \right)^{\frac{1}{1-\gamma}}$$

subject to $\Gamma_t \cdot A_t \geq 0$. The first-order condition is:

$$\begin{aligned} a^i: \quad & \mathbb{E}_{X,\omega} \left[(1-\gamma)\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \left(\mathbb{E}_X \left[R^i\left(X_t\right) - R^m \right] + \mathbb{E}_{X,\omega} \left[\chi_s \frac{\partial s}{\partial a^i} \right] \right) \right] = 0 \\ & \mathbb{E}_{X,\omega} \left[\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \left(\mathbb{E}_X \left[R^i\left(X_t\right) - R^m \right] + \mathbb{E}_{X,\omega} \left[\chi_s \frac{\partial s}{\partial a^i} \right] \right) \right] = 0 \end{aligned}$$

Taking the second term of the expression in parentheses to the right-hand side:

$$\mathbb{E}_{X,\omega} \left[\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \mathbb{E}_X \left[R^i\left(X_t\right) - R^m \right] \right] = -\mathbb{E}_{X,\omega} \left[\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \mathbb{E}_{X,\omega} \left[\chi_s \tfrac{\partial s}{\partial a^i} \right] \right]$$

If we take into account the covariance formula:

$$\begin{aligned} & \mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \mathbb{E}_X \left[R^i\left(X_t\right) - R^m \right] + \mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, R^i\left(X\right) - R^m \right] \\ & \quad = -\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \mathbb{E}_{X,\omega} \left[\chi_s \frac{\partial s}{\partial a^i} \right] - \mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, \chi_s \frac{\partial s}{\partial a^i} \right] \end{aligned}$$

Rearranging:

$$\begin{aligned} & \mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \mathbb{E}_X \left[R^i\left(X_t\right) - R^m \right] \\ & \quad = -\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \mathbb{E}_{X,\omega} \left[\chi_s \frac{\partial s}{\partial a^i} \right] - \mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, \chi_s \frac{\partial s}{\partial a^i} \right] \\ & \qquad - \mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, R^i\left(X\right) - R^m \right] \end{aligned}$$

Now, we can obtain the asset premium:

$$\begin{aligned} \mathbb{E}_X \left[R^i\left(X_t\right) - R^m \right] = & -\frac{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right] \mathbb{E}_{X,\omega} \left[\chi_s \frac{\partial s}{\partial a^i} \right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right]} - \frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, \chi_s \frac{\partial s}{\partial a^i} \right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right]} \\ & - \frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, R^i\left(X\right) - R^m \right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma} \right]} \end{aligned}$$

Simplifying:

$$\mathbb{E}_X\left[R^i\left(X_t\right)\right]-R^m=-\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i}\right]-\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma},\chi_s\frac{\partial s}{\partial a^i}\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]}$$
$$-\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma},R^i\left(X\right)\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]}$$

Therefore:

$$\mathbb{E}_X\left[R^i\left(X_t\right)\right]-R^m=-\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i}\right]-\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma},R^i\left(X\right)+\chi_s\frac{\partial s}{\partial a^i}\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]} \tag{A.27}$$

S.1.1 Risk Neutral Case

We have the same problem as the previous case, but considering $\gamma=0$:

$$\max_A\left(\mathbb{E}_{X,\omega}\left[R^m e+\sum_{i\in\mathbb{I}}\underbrace{\left(R^i\left(X_t\right)-R^m\right)}_{\text{Liquidity Premium}}\bar{a}_t^i+\underbrace{\chi_{t+1}\left(s\left(\{\bar{a}\}_{i\in\mathbb{I}},e-\sum_{i\in\mathbb{I}}\bar{a}_t^i\right)\right)}_{\text{Liquidity Yield}}\right]\right)$$

subject to $\Gamma_t\cdot A_t\geq 0$. The first-order condition is:

$$a^i:\quad \mathbb{E}_{X,\omega}\left[\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right]+\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i}\right]\right]=0$$
$$\mathbb{E}_{X,\omega}\left[\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right]\right]+\mathbb{E}_{X,\omega}\left[\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i}\right]\right]=0$$

Thus, we have the asset premium:

$$\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right]=-\mathbb{E}_{X,\omega}\left[\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i}\right]\right]$$
$$\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right]=-\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i}\right]$$
$$\mathbb{E}_X\left[R^i\left(X_t\right)\right]-R^m=-\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i}\right]$$

S.2 Efficiency

S.2.1 Risk Aversion Case

We have the same problem as the pricing conditions case, so we will follow the same steps to solve it:

$$\max_A \left(\mathbb{E}_{X,\omega} \left[\left(R^m e + \sum_{i\in\mathbb{I}} \underbrace{\left(R^i\left(X_t\right) - R^m\right)}_{\text{Liquidity Premium}} \bar{a}^i_t + \underbrace{\chi_{t+1} \left(s\left(\left\{\bar{a}\right\}_{i\in\mathbb{I}}, e - \sum_{i\in\mathbb{I}} \bar{a}^i_t \right), \theta(\{\bar{a}\}_{i\in\mathbb{I}}) \right)}_{\text{Liquidity Yield}} \right)^{1-\gamma} \right] \right)^{\frac{1}{1-\gamma}}$$

subject to $\Gamma_t \cdot A_t \geq 0$. The first-order condition is:

$$a^i: \quad \mathbb{E}_{X,\omega} \left[(1-\gamma)\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\right] \left(\mathbb{E}_X \left[R^i\left(X_t\right) - R^m\right] + \mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i} + \chi_\theta \frac{\partial \theta}{\partial a^i} \right] \right) \right] = 0$$

$$\mathbb{E}_{X,\omega} \left[\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\right] \left(\mathbb{E}_X \left[R^i\left(X_t\right) - R^m\right] + \mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i} + \chi_\theta \frac{\partial \theta}{\partial a^i} \right] \right) \right] = 0$$

Reordering the last expression:

$$\mathbb{E}_{X,\omega}\left[\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\right]\mathbb{E}_X\left[R^i\left(X_t\right) - R^m\right]\right] = -\mathbb{E}_{X,\omega}\left[\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\left(\chi_s\tfrac{\partial s}{\partial a^i} + \chi_\theta\tfrac{\partial \theta}{\partial a^i}\right)\right]\right]$$

Using the covariance formula:

$$\begin{aligned}
&\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\right]\mathbb{E}_X\left[R^i\left(X_t\right) - R^m\right] + \mathbb{COV}_{X,\omega}\left[(R^e)^{-\gamma}, R^i\left(X\right) - R^m\right] \\
&\quad = -\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\right]\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i} + \chi_\theta\frac{\partial \theta}{\partial a^i}\right] \\
&\quad\quad - \mathbb{COV}_{X,\omega}\left[(R^e)^{-\gamma}, \chi_s\frac{\partial s}{\partial a^i} + \chi_\theta\frac{\partial \theta}{\partial a^i}\right]
\end{aligned}$$

Rearranging:

$$\begin{aligned}
&\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\right]\mathbb{E}_X\left[R^i\left(X_t\right) - R^m\right] \\
&\quad = -\mathbb{E}_{X,\omega}\left[(R^e)^{-\gamma}\right]\mathbb{E}_{X,\omega}\left[\chi_s\frac{\partial s}{\partial a^i} + \chi_\theta\frac{\partial \theta}{\partial a^i}\right] \\
&\quad\quad - \mathbb{COV}_{X,\omega}\left[(R^e)^{-\gamma}, \chi_s\frac{\partial s}{\partial a^i} + \chi_\theta\frac{\partial \theta}{\partial a^i}\right] \\
&\quad\quad - \mathbb{COV}_{X,\omega}\left[(R^e)^{-\gamma}, R^i\left(X\right) - R^m\right]
\end{aligned}$$

Finally, we can obtain the asset premium:

$$
\begin{aligned}
\mathbb{E}_X\left[R^i\left(X_t\right)\right]-R^m = & -\mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i}\right]-\mathbb{E}_{X,\omega}\left[\chi_\theta \frac{\partial \theta}{\partial a^i}\right] \\
& -\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, R^i(X)+\chi_s \frac{\partial s}{\partial a^i}\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]} \\
& -\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, \chi_\theta \frac{\partial \theta}{\partial a^i}\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]}
\end{aligned}
$$

Therefore:

$$
\mathbb{E}_X\left[R^i\left(X_t\right)\right]-R^m=-\left(\mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i}\right]+\frac{\mathbb{COV}_{X,\omega}\left[\left(R^e\right)^{-\gamma}, R^i(X)+\chi_s \frac{\partial s}{\partial a^i}\right]}{\mathbb{E}_{X,\omega}\left[\left(R^e\right)^{-\gamma}\right]}\right)-\left(\mathbb{E}_{X,\omega}\left[\chi_\theta \frac{\partial \theta}{\partial a^i}\right]+\frac{\mathbb{COV}_{X,\omega}}{\mathbb{E}_{X,}} \tag{A.28}
$$

S.2.2 Risk Neutral Case

Now we are going to solve the last problem considering $\gamma=0$:

$$
\max_A\left(\mathbb{E}_{X,\omega}\left[R^m e+\textstyle\sum_{i\in\mathbb{I}} \underbrace{\left(R^i\left(X_t\right)-R^m\right)}_{\text{Liquidity Premium}} \bar{a}_t^i+\underbrace{\chi_{t+1}\left(s\left(\{\bar{a}\}_{i\in\mathbb{I}}, e-\sum_{i\in\mathbb{I}} \bar{a}_t^i\right), \theta(\{\bar{a}\}_{i\in\mathbb{I}})\right)}_{\text{Liquidity Yield}}\right]\right)
$$

subject to $\Gamma_t \cdot A_t \geq 0$. The first-order condition is:

$$
\begin{aligned}
a^i: \quad & \mathbb{E}_{X,\omega}\left[\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right]+\mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i}+\chi_\theta \frac{\partial \theta}{\partial a^i}\right]\right]=0 \\
& \mathbb{E}_{X,\omega}\left[\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right]\right]+\mathbb{E}_{X,\omega}\left[\mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i}+\chi_\theta \frac{\partial \theta}{\partial a^i}\right]\right]=0
\end{aligned}
$$

Finally, we have the asset premium:

$$
\begin{aligned}
\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right] & =-\mathbb{E}_{X,\omega}\left[\mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i}+\chi_\theta \frac{\partial \theta}{\partial a^i}\right]\right] \\
\mathbb{E}_X\left[R^i\left(X_t\right)-R^m\right] & =-\mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i}+\chi_\theta \frac{\partial \theta}{\partial a^i}\right] \\
\mathbb{E}_X\left[R^i\left(X_t\right)\right]-R^m & =-\mathbb{E}_{X,\omega}\left[\chi_s \frac{\partial s}{\partial a^i}\right]-\mathbb{E}_{X,\omega}\left[\chi_\theta \frac{\partial \theta}{\partial a^i}\right]
\end{aligned}
$$