Regulation and Intermediation in Over-the-counter Markets*

Nicole Fleskes†

March 22, 2026

For the most recent version of this paper,
please visit this link.

Abstract

This paper examines the impact of banking regulation on the trading behavior of bank-affiliated dealers in over-the-counter (OTC) financial markets. I micro-found the regulatory costs of holding OTC traded assets using an equilibrium search model of a two-tiered OTC market in which risk-averse dealers facilitate trades for investors. Dealers choose portfolios of risk-free and risky OTC assets, and since they are bank-affiliated, are subject to a regulatory balance sheet constraint: a Basel-style risk- weighted asset ratio. Formally, I show that acquiring additional inventories tightens regulatory requirements and raises the shadow value of capital when regulatory slack is low. Regulatory balance sheet costs are heterogeneous and state-dependent, increasing with a dealer's distance-to-constraint and their inventories. In a calibration to the pre-2008 US corporate bond market, I show that as constraints bind, dealers reduce inventories at fire sale prices and only purchase assets at extreme discounts; however, trading costs rise. A policy experiment shows that stricter regulation increases trading costs and reduces intermediation quality in the investor market on average.

Keywords: Over-the-counter markets, regulation, inventory management
JEL Codes: G11, G18, G21, D53

*I am grateful for comments from my advisor Zachary Bethune, as well as from Kerry Back, Benjamin Lester, Shihan Shen, Isabelle Perrigne, and Julian Kozlowski. I also thank the participants of the 2025 Rice-LEMMA Monetary Conference and the 2025 St. Louis Federal Reserve Dissertation Fellows workshop for their feedback.

†Department of Economics, Rice University (nkf1@rice.edu)

1 Introduction

In the aftermath of the Great Financial Crisis (GFC), policymakers enacted a series of reforms to enhance financial stability and curb systemic risk in the banking sector. International standards such as Basel II.5 and Basel III, as well as the US Dodd-Frank Act, updated or introduced new rules and regulations to prevent excessive risk-taking by financial institutions. Central to these reforms were changes to balance sheet based regulation, such as higher capital requirements, stricter leverage ratios, and revised risk-weighted asset ratios.

An unintended consequence of these regulations has been their impact on dealer trading behavior in over-the-counter (OTC) financial markets, such as those for corporate bonds or treasuries. Many dealers in OTC markets are *bank-affiliated*, that is, they operate as subsidiaries of a large bank holding company (BHC).[1] Balance sheet regulatory ratios are calculated based on a BHC's consolidated balance sheet, which includes all subsidiary dealer trading assets used for market-making activities (Adrian et al., 2018; Cimon & Garriott, 2019). Accordingly, when making trading decisions, a bank-affiliated dealer anticipates the consequences for their parent BHC's regulatory position.

This paper studies how regulatory balance sheet ratios affect bank-affiliated dealer intermediation in OTC markets; specifically, the impact of risk-weighted asset ratios established by the Basel accords.[2] A risk-weighted asset ratio requires affected financial institutions to hold a minimum level of high-quality liquid capital relative to their risk-weighted assets such that the ratio is higher than a given percentage referred to as a capital requirement. The risk weights, the risk weight classifications of specific assets, and the capital requirement are determined by policy. There were two key changes to this ratio after the GFC. First, Basel II.5's market risk framework and Basel III revised risk-weighted asset classifications, moving some assets into higher risk weight categories. Second, Basel III increased the minimum capital requirement itself from 4% pre-crisis to 6%, phased in between 2013 and 2015 within the US. For large systemically important institutions, an additional

[1] A BHC may have many subsidiaries, including but not limited to: commercial banking, specialty lending, asset management, underwriting, and, of course, dealer trading desks for multiple OTC markets (Cetorelli & Stern, 2015).

[2] Other important ratios being the supplementary leverage ratio (SLR) and the liquidity coverage ratio (LCR).

cyclical capital buffer of 2.5% was phased in starting in 2015 and fully implemented by 2019 in the US. Basel III Endgame, to be implemented between 2025 and 2028, is expected to introduce additional revisions that further tighten these risk-weighted capital requirements.

A large body of empirical evidence documents a few key shifts in OTC markets following the adoption of these regulatory changes: dealers now maintain smaller trading inventories and trading costs for investors have increased (Adrian et al., 2017; Bao et al., 2018; Bessembinder et al., 2018; Breckenfelder & Ivashina, 2021a; Choi et al., 2024; Cohen et al., 2024; Dick-Nielsen & Rossi, 2019; Duffie, 2023). The evidence on changes in various other liquidity measures (e.g., turnover, trade volumes, price impact) is heterogeneous, differing by market, time period, and study design. Although the empirical literature demonstrates the relevance of these regulations for OTC markets, the precise mechanism through which they affect dealer decision making to produce these outcomes is less well understood. Theoretical work addressing this question has thus far modeled regulatory balance sheet costs in a reduced-form or ad-hoc manner, for example, by either imposing a regulatory cap on OTC inventory stock, as in Duffie et al. (2023), or placing a holding cost per unit of inventory, as in Cohen et al. (2024).

In contrast, this paper aims to micro-found the regulatory balance sheet costs of OTC inventories.[3] To do so, I develop an equilibrium search model of a two-tiered OTC financial market in which bank-affiliated dealers intermediate trade for investors. As in Gârleanu (2009), all agents (dealers and investors) are risk averse and face idiosyncratic uncertainty regarding their endowment income and their private valuation of the OTC asset's return. Furthermore, these two processes may be correlated, reflecting the agent's liquidity preference. Investors' liquidity preferences change over time, whereas dealers' liquidity preferences remain constant. Dealers choose a portfolio of assets, including risky OTC traded assets, under a balance sheet level regulatory constraint: a risk-

[3]In this model, I focus exclusively on how regulation affects market making through inventory balance sheet costs, abstracting from funding costs. Usually, dealers finance their positions via short-term funding on the repo market; tighter regulation reduces a dealer's ability to obtain such funding (Adrian et al., 2018; Cimon & Garriott, 2019). I also abstract from competition from non bank-affiliated dealers and focus solely on the choices of bank-affiliated dealers. The rise of non-bank dealers to fill the immediacy gap left in recent years is thus an observed trend that this model cannot explain. Lastly, I abstract away from dealers' choice between risky-principal and agency trades, considering only risk-principal trades (An & Zheng, 2023; Cimon & Garriott, 2019; Saar et al., 2019).

weighted asset ratio. Investors, on the other hand, face only a non-negative equity requirement. Because endowment and dividend income are stochastic and trade in the OTC asset is frictional, both dealers and investors must satisfy a Value-at-Risk (VaR) condition to ensure compliance with their respective constraints. Dealers in this model facilitate trades in the OTC asset for investors in a decentralized investor market and can trade with other dealers in a more liquid yet still decentralized inter-dealer market. Each market is characterized by random pairwise meetings; when agents meet, they engage in proportional bilateral bargaining with equal bargaining power.

First, I show formally that the regulatory costs of OTC inventories borne by dealers arise endogenously from the balance sheet regulatory constraint and are heterogeneous, depending on a dealer's stock of OTC assets and how much regulatory slack they have (Proposition 1). Larger OTC positions directly raise minimum equity requirements, requiring dealers to commit more capital to remain compliant. The requirement induces an indirect precautionary savings motive: as regulatory slack shrinks, the shadow value of capital rises, making dealers more inclined to sell inventories as trading opportunities arise. Furthermore, because OTC assets are risky, the pressure to offload inventory when regulatory slack is low increases as a dealer's OTC assets grow. Search frictions are crucial to this mechanism, since infrequent trading opportunities prevent dealers from continuously re-balancing their portfolios.

Given the model's complexity, I solve the steady state numerically using a novel iterative procedure that addresses the high-dimensional fixed-point problem involving agent value functions, trading rules, and distributions. Specifically, I extend the finite difference algorithm developed by Achdou et al. (2022) to account for what is essentially a state-based endogenous borrowing constraint, as well as transition intensities across agent states that are determined through bilateral bargaining. I calibrate the baseline model to the pre-GFC period using BHC level quarterly balance sheet data (FR-Y9C filings) and corporate bond transaction level data, focusing on risky-principal trades in investment-grade bonds (Academic TRACE). I use a capital requirement of 4% as a baseline and compare it to an economy in which dealers face no risk-based capital requirement.

I find that the presence of a risk-weighted asset ratio creates a regulatory motive for trade:

dealers trade not only to earn intermediation profits, but also to manage their distance from the regulatory constraint. Because the regulatory constraint causes the marginal value of an additional unit to decline sharply as dealers become more constrained, dealer buy and sell prices in both investor and inter-dealer markets also fall: dealers are willing to sell at low prices, and to buy only at deep discounts in the region near the constraint. At the same time, the expected spread, or the wedge between the two, rises as dealers demand higher compensation for round-trip trades. Using their balance sheets to facilitate trades bears a larger regulatory cost when dealers are less well-capitalized, consistent with the empirical evidence that more constrained bank-affiliated dealers charge higher trading costs (Adrian et al., 2017; Breckenfelder & Ivashina, 2021a). Furthermore, this effect scales with dealer OTC holdings.

I also find that OTC inventories are mean-reverting, consistent with both theoretical (Üslü, 2019) and empirical (Dick-Nielsen & Rossi, 2019) evidence. However, I also show that target inventories decline as agents become less well-capitalized. When capital is tight, agents reduce their target risky OTC holdings to relax balance sheet constraints and lower portfolio volatility. Finally, I show that constrained dealers are worse intermediaries, in that they are less willing to both buy and sell in high trade volumes. This result also matches the empirical evidence that bonds handled by constrained dealers are less liquid, with reduced turnover and volumes (Adrian et al., 2017; Bao et al., 2018).

I use this model to study the impact of stricter regulation by increasing the capital requirement from 4% to 8.5%. Under stricter regulation, investor spreads rise in the cross-section: more constrained dealers charge higher spreads for the same amount of regulatory slack. Consistent with post-GFC patterns, average investor spreads rise and dealers reduce their inventories. Furthermore, investors adopt less extreme positions due to these higher trading costs. Stricter regulation also leads dealers to intermediate less in the investor market, reducing investor turnover, while relying more on the inter-dealer market to reallocate inventories among themselves to manage their balance sheet constraints. Under stricter regulation, dealers lean away from using their balance sheets to facilitate investor trades, preferring instead to trade in the more liquid inter-dealer market for regulatory reasons.

This paper is organized as follows. Section 2 reviews the related literature. Section 3 outlines the model and establishes the micro-foundations of regulatory costs. Section 4 discusses the method used to solve the model, the calibration, and baseline model outcomes. Section 5 explores the case when capital requirements increase due to stricter regulation. Section 6 concludes.

2 Related Literature

Many search-theoretic approaches to studying OTC markets follow the seminal works of Duffie et al. (2005) and Lagos and Rocheteau (2009), in which dealers can intermediate trades in an illiquid asset for investors. However, in this setting the study of dealer inventory management is limited: dealers have no private valuation of the OTC traded asset, are risk-neutral, and have access to a centralized inter-dealer market. Consequently, dealers do not hold steady state inventories and engage only in matchmaking activities for investors. Weill (2007), Rocheteau and Weill (2011), and Fleskes (2025) extend this framework to study the out of steady state dynamics of dealer inventory, identifying under what conditions dealers use their inventory capacity temporarily to smooth imbalances in customer trading needs due to unexpected liquidity crises. In Cohen et al. (2024), the authors impose an asset-in-advance constraint: dealers must hold assets to fulfill customer orders, and so have incentive to hold inventories in the steady state. The regulatory cost of inventories is modeled in their work as an ad-hoc per-unit cost; the authors show that when this cost increases, dealer intermediation falls. In this paper, I micro-found regulatory inventory costs through a constrained portfolio problem.

Dealer inventories play a more natural role when markets are fully decentralized.[4] Hugonnier

[4]Another parallel literature that studies dealer inventory management is based on a class of search models in the tradition of Rubinstein and Wolinsky (1987), such as in Gu et al. (2024), Johri and Leach (2002), Masters (2007), Nosal et al. (2015, 2019), Shevchenko (2004), Watanabe (2010, 2020), and Wright and Wong (2014) and Gong and Wright (2024). These types of inventory models differ from the previously discussed works in that they primarily study three-sided goods markets: there are usually distinct sellers, middlemen, and buyers who all occupy different roles in the economy and face different frictional markets. The middleman's inventory management problem is therefore summarized by their pricing strategies on both sides of the market, wholesale and investor. Carrasco and Harrison (2023) and Carrasco and Smith (2017) focus on the pricing strategies of a dealer liquidating their assets.

et al. (2020) allow for fully decentralized two-tier markets for trade in an OTC asset. Risk-neutral dealers and investors are heterogeneous in their private valuations of the asset and engage in bilateral trade with binary asset holdings. A key result is that, in the steady state, only dealers with middle-of-the-road private valuations provide intermediation services, and that dealer intermediation is characterized by intermediation chains. Yang and Zeng (2021) use a similar two-tier market framework in which risk-neutral dealers can now hold multiple units of the asset. By relaxing the binary asset holding assumption, the authors show that there are multiple equilibria in which dealers coordinate their liquidity provision depending on asset fundamentals.

In the above models, dealers have a special exogenous ability to intermediate in an investor market. Other works in this literature focus instead on one-tier fully decentralized markets to study endogenous intermediation: which agents decide to become intermediaries. Rather than designating a special set of agents who exogenously have the ability to intermediate, heterogeneous agents all can trade with each other. These markets are characterized by a core-periphery structure and intermediation chains: at the core of the trading network are agents who engage heavily in intermediation, and at the periphery are those who behave more as value investors. The agents that arise as intermediaries tend to be those whose type allows them to extract better terms of trade: they have faster meeting technologies, better information, or more bargaining power (Bethune et al., 2022; Donaldson et al., 2018; Farboodi et al., 2023; Hugonnier et al., 2021; Jarosch et al., 2016; Üslü, 2019). In contrast to these works, I show that it is certain agent states rather than fixed agent types which drive the heterogeneity in intermediation behavior attributable to balance sheet regulation. In this model, dealers are all identical in that they face the same risky income processes. However, realizations for each dealer generate heterogeneity in the first dimension, equity, leading dealers to have differing levels of regulatory slack and desire to intermediate.

In the vast majority of studies of OTC markets in this literature, dealers are risk-neutral. Duffie et al. (2007) and Gârleanu (2009) are the foundational works that depart from this assumption, relying on CARA utility preferences to micro-found quadratic utility over multiple units of the OTC asset.[5]

[5] The same method is used to micro-found the reduced form utility specifications in Vayanos and Weill (2008), Praz (2014), and Üslü (2019).

Kargar et al. (2023) is the exception, explicitly considering risk-averse investors and aggregate risk. However, their work focuses on the pricing of the OTC asset and the portfolio choice problem of investors; dealers act as simple risk-neutral profit-maximizing matchmakers. In contrast, this paper investigates the portfolio problem of risk-averse dealers subject to a regulatory balance sheet constraint. Liquidity is therefore a consequence of dealers' constrained portfolio decisions.

There is also a literature on inventory management by dealers in OTC markets in the finance tradition based on the seminal works of Amihud and Mendelson (1980) and T. Ho and Stoll (1981) and T. S. Y. Ho and Stoll (1983).[6] Works in this tradition primarily focus on the determination of bid-ask spreads set by a (usually) monopolistic risk-neutral dealer, with investors arriving at the OTC market with fixed trading demands. Duffie (2023) extends Amihud and Mendelson (1980) by imposing a maximum level of inventory that the dealer can hold: the higher the balance sheet costs of inventories, the lower the inventory cap. As the level of inventories approaches the regulatory implied limit, bid-ask spreads widen to account for the increasing marginal cost of dealer balance sheet space. Similarly, Wang and Zhong (2022) study the impact of the Basel III risk weight changes by varying the maximum level of inventories a dealer can hold: they find that the increased capital requirements result in an increase in order rejection rates by dealers. Adrian et al. (2020) study the impact of increased overnight inventory costs and show that dealers adjust their pricing strategies to offload inventories as the end of the trading day approaches.

In contrast to this model style, in this work all agents engage in bilateral trade and I focus on an equilibrium in which investor distributions and investor demands are endogenous equilibrium objects. Furthermore, inventory costs are endogenous, derived via a balance sheet constraint rather than through an exogenous inventory cap. This paper is, to the best of my knowledge, the first to study dealer inventory management from a joint portfolio-choice and OTC search perspective.

[6] Both this literature and the middlemen literature a la Rubinstein and Wolinsky (1987) have origins in Garman (1976).

3 A portfolio choice model with decentralized trade

Time is continuous and infinite. All agents, dealers and investors, are infinitely lived and discount time at a constant rate $r \in (0, 1)$. There is a numeraire good called consumption c, a risk-free asset ω in perfectly elastic supply, and a risky over-the-counter (OTC) traded asset n in some positive supply in shares $S > 0$. The risk-free asset can be invested in an outside account that pays the discount rate and can be costlessly converted into the numeraire. Alternatively, agents can borrow in the risk-free asset at the discount rate: ω denotes the stock of *net* risk-free assets an agent holds. The risky OTC asset earns a risky dividend dD_t and all agents earn income from a risky endowment $d\eta_t$, both paid out in units of the risk-free asset. Agents can hold multiple divisible units of the risk-free asset ω and the risky OTC asset $n \geq 0$.

I frame the preceding setup as the problem of a trading desk manager operating as a subsidiary of a larger institution (i.e., a bank holding company). The risk-free asset is interpreted as cash, or another equally liquid and safe asset such as reserves, which serves as the unit of account. The trading desk manager uses their cash holdings to buy and sell the OTC asset, and all income (from dividends or endowments) is realized in this unit of account. The outside account paying the discount rate r is interpreted as a reserve account if positive; alternatively, agents can borrow from an outside market at the same rate. The manager controls their trading strategies and portfolio decisions: their holdings of cash (the risk-free asset) and OTC asset inventories (the risky OTC asset). However, they do not control the activities of the rest of the institution. The endowment is interpreted as any non-trading income the parent institution may periodically receive from their other lines of business. *Dealers* are subsidiaries of a BHC. Therefore, their outside income is that from other BHC subsidiaries (e.g., commercial banking, investment banking, etc.). *Investors* are institutional investors such as mutual funds, pension funds, etc. Their outside income is derived from other investments they make or other business activities.

All agents have constant absolute risk aversion (CARA) utility over consumption of the numeraire good with risk aversion parameter κ: $u(c) = -e^{-\kappa c}$. As such, consumption of the numeraire good can be negative. Consumption can be interpreted as profits remitted from the trading desk to the shareholders of the larger institution, with negative consumption indicating realized losses. Agents have idiosyncratic uncertainty over their endowment income and their private valuation of the OTC traded asset's dividend. Let $Z_t = [Z_{\eta,t}, Z_{D,t}]'$ be a two-dimensional standard Brownian motion for $t \geq 0$ defined on a filtered probability space $(\Omega, F, \mathbb{P})$ where $Z_{\eta,t}$ and $Z_{D,t}$ are independent. Following Duffie et al. (2007), Gârleanu (2009), and Praz (2014), the stochastic cumulative dividend process per unit of risky OTC asset n follows for constants μ_D and σ_D:

$$dD_t = \mu_D dt + \sigma_D dZ_{D,t}. \tag{1}$$

The stochastic endowment process follows for constants μ_η and σ_η:

$$d\eta_t^i = \mu_\eta dt + \sigma_\eta Z_{\eta,t}^i. \tag{2}$$

$Z_{\eta,t}^i$ is defined as:

$$Z_{\eta,t}^i = \rho_t^i dZ_{D,t} + \sqrt{1 - (\rho_t^i)^2} dZ_{\eta,t}. \tag{3}$$

An agent's endowment process is stochastic with drift μ_η and with an element of uncertainty $\sigma_\eta dZ_{\eta,t}^i$. An agent's type i determines how correlated their endowment is with their private valuation of the OTC asset's dividend: ρ_t^i is the instantaneous correlation for an agent of type i between the OTC asset dividend and their endowment. If an agent's OTC asset dividend income is negatively correlated with their endowment income, the OTC asset dividend acts as a hedge.

Up to this point, dealers and investors are identical. I will now discuss the first of two fundamental ways in which they differ. Investors are labeled so because they have a fundamental trading need that dealers do not: their liquidity preferences, captured by ρ^i, vary over time. Specifically, they receive idiosyncratic preference shocks that arrive according to a Poisson process with intensity $\gamma > 0$, and with probability $\pi(i') > 0$ they switch to type i' with correlation parameter $\rho^{i'}$. This is what drives investor trading in the OTC market. For example, an investor of type i that holds their perfect amount of OTC assets n may receive a liquidity preference shock and become type i', at which time they would like to change their OTC asset holdings to n'. To adjust their holdings, they take advantage of the periodic trading opportunities in the OTC market, to be described shortly. In contrast, dealers have fixed liquidity preferences, maintaining the same correlation ρ over time.

Each agent records on the asset side of their balance sheet their marked-to-market net holdings of the risk-free asset ω at a reference price normalized to 1, and the OTC asset n at reference price $\tilde{p}$, to be defined precisely shortly. An agent's equity E is the sum of the market value of their net assets:

$$E = \omega + \tilde{p}n.$$

There is a unit measure of a continuum of both dealers and investors. Denote the distributions of dealers and investors to be $F(E, n)$ and $H(E, n, i)$ respectively, where i denotes the investor's type. Let the relative prevalence of dealers in the economy be $p_d \in (0, 1)$. In reality, there are very few dealers (a few hundred) and a large mass of investors (a few million), and so p_d is quite small. However, the identity of dealers and where they can be found is generally known; therefore, the relative prevalence of dealers in the economy matters only for market clearing and not for finding a trading partner in the OTC market.

Trade in the risky OTC asset is fully bilateral. Investors are only permitted to trade the OTC asset through dealers, whereas dealers can trade both with investors and with other dealers. Agents meet according to a Poisson process with intensity $\lambda_c > 0$ between investors and dealers, and with intensity $\lambda_d > \lambda_c > 0$ between dealers: the inter-dealer market is more liquid than the investor market. Upon receiving a trade opportunity, an agent draws a counter-party from the relevant distribution, either $F(E, n)$ or $H(E, n, i)$. Once a trading partner is drawn, agents engage in proportional bargaining, otherwise known as Kalai (1977) bargaining, over the terms of trade of the OTC asset where agents have equal bargaining power. Consequently, the OTC asset trades at different prices depending on which agents meet.

The market value of the OTC asset, captured by reference price $\tilde{p}$, is defined as the *average* volume-weighted price of the risky OTC asset based on realized trades. The average price is used to value agent asset holdings as it best reflects current accounting standards in the United States. According to Financial Accounting Standards Board (FASB) regulation[7], assets for which there are active (even if decentralized) markets must be marked-to-market using recently quoted prices for that particular asset or a similarly classed asset.

The second fundamental difference between dealers and investors concerns regulation. Investors are subject to a simple non-negative equity requirement to prevent default: $E \geq 0$. However, dealers are subject to a risk-based capital requirement in the spirit of Basel II and Basel III, which stipulates that they must maintain a level of high-quality liquid capital relative to their risk-weighted assets of at least $\Lambda\%$. In this model, this liquid capital is simply agent equity, E, and risk-weighted assets are calculated as the sum of the market value of dealer risk-free and risky OTC asset holdings, where risk-free assets have a 0% risk weight and risky OTC asset inventories have some positive risk weight set by policy, $\nu\%$. The risk-weight $\nu\%$ typically ranges from 0% to 250% in increments of 50%, with cash and US treasuries using a risk weight of 0%. The risk-based capital requirement in this model is in practice a risk-weighted assets ratio:

$$\frac{E}{0 \times \omega + \nu \times \tilde{p}n} \geq \Lambda. \tag{4}$$

Define the minimum level of equity required by this regulation for a dealer with a stock of risky OTC assets n to be:

$$\bar{E}(n) = \Lambda \nu \tilde{p} n. \tag{5}$$

The regulatory capital requirement faced by dealers is stricter than the non-negative equity condition faced by investors. If a dealer holds non-zero OTC inventories, their regulatory minimum level of equity is strictly positive if $\Lambda > 0$ and $\nu > 0$. When dealers take on OTC inventories, their minimum required equity level rises, all else equal.

Due to the definitions of the above stochastic processes, income from agent endowments $d\eta_t$ and dividends dD_t are random variables with unbounded support and are not known ex-ante of agent

[7]FASB ASC 820-10 (Fair Value Measurements and Disclosures)

consumption and portfolio decisions. This differs from other standard continuous-time incomplete-markets models with borrowing constraints, as in Achdou et al. (2022), in which income at each instant is known, although it may evolve over time via a diffusion or jump process. In these standard models, there is a finite level of maximum consumption implied by the savings function that ensures that agents never violate their borrowing constraints. Because time is continuous, agents continuously approach their borrowing constraints in small increments; unlike discrete time, there are no jumps. As a result, there is a level of small but finite consumption such that at the borrowing constraint, the agent's wealth exclusively drifts back into the unconstrained region. This is possible exclusively because income is known when agents make their consumption decisions, so they will never violate the constraint with certainty.

Because in this model income is random and on an unbounded domain, the level of consumption needed to ensure that agents never violate their constraint is $-\infty$. Agents do not know their income realizations before making their consumption decisions, and these income realizations can be negative infinity. This is permissible because of CARA utility, but less than ideal for obtaining theoretical and computational results. Therefore, I make an assumption regarding how agents can satisfy their constraints: either the non-negative equity condition for investors, or the regulatory capital requirement for dealers.

Value-at-Risk (VaR) constraint. The Value-at-Risk (VaR) constraint requires that agents must act in such a manner that the chance of violating their relevant constraint, even under more extreme negative income shocks, is very small. Specifically, this means that, given a stock of OTC assets n, agents must manage their risk-free assets such that within the next Δt period of time the probability of violating their constraint is no more than ϵ%. Formally, the regulatory VaR is the ϵ-quantile $\forall n$ of the distribution of an agent's regulatory compliance:

$$VaR_\epsilon(R_{\Delta t}(n)) = \text{inf}\{q \in \mathcal{R} : Pr\left(R_{\Delta t}(n) \leq q\right) \geq \epsilon\}.$$

Where the degree of an agent's compliance for each type of agent is defined $\forall n$ as:

$$R_{\Delta t}(n) = \begin{cases} E_{t+\Delta t} - \bar{E}(n) & \text{for dealers;} \\ E_{t+\Delta t} - 0 & \text{for investors.} \end{cases}$$

If $R_{\Delta t}(n)$ is negative, this indicates that an agent is in violation of their regulatory minimum level of equity. The regulatory requirement is then $\forall n$:

$$VaR_{\epsilon}(R_{\Delta t}(n)) \geq 0 \tag{6}$$

which ensures that the probability at time t of violating the regulatory constraint over the horizon $t + \Delta t$ is less than ϵ%. The VaR constraint (6) guarantees that the agent remains compliant with probability of at least $(1 - \epsilon)$%. Smaller values of ϵ or longer time horizons $t + \Delta t$ tighten the constraint, making dealers more conservative and leading to more severe portfolio adjustments.

What this assumption provides is a very low but finite closed-form maximum level of consumption at the constraint point, $c^{VaR}(E, n)$ for dealers and $c^{VaR}(E, n, i)$ for investors, such that the agent's equity drifts back into the unconstrained region with a very large probability. It also allows me to fine-tune through my choices of Δt and ϵ how risk-averse dealers are with respect to regulation, in addition to their general risk aversion κ. In practice, when solving this model computationally using finite difference methods following Achdou et al. (2022), the effective probability of agents violating the regulatory constraint is zero. What this VaR assumption determines is the shape of the value function near the constraint and thus the degree to which the balance sheet regulation shapes agent trading behavior.

Furthermore, this assumption reflects how institutional and dealer investors manage risk in the real world. Internal risk limits determined by VaR constraints calculated using a variety of different methods or simulations are actively used by financial institutions to manage risk and have been shown to directly affect dealer risk-appetites when making the market (Anderson et al., 2023).

3.1 Agent problems

All agents are born with an initial level of the risk-free asset, $\omega_0 > 0$. Agents receive income from their endowment $d\eta_t$, their holdings of the risk-free asset $r\omega_t dt$, and their holdings of the risky OTC asset $n_t dD_t$. Agents consume the numeraire good $c_t dt$, and can trade a_t units at price P_t determined by bargaining when trading opportunities arrive (described in the next section). An agent's equity therefore evolves according to:

$$\begin{aligned} dE_t =& d\eta_t^i - c_t dt + rE_t dt + n_t \left(dD_t + d\tilde{p}_t - r\tilde{p}_t dt\right) \\ &+ (\tilde{p}_t - P_t) dn_t. \end{aligned} \tag{7}$$

An agent's OTC asset holdings evolve according to:

$$dn_t = 1\{\text{OTC trade opportunity}\} a_t. \tag{8}$$

Higher endowment income, $d\eta_t$, and higher direct income from the risk-free and risky OTC assets, $rE_t dt$ and $n_t dD_t$, increase an agent's equity. When an agent consumes, $c_t dt$, equity falls. $d\tilde{p}_t$ captures the change in equity per unit of risky OTC asset due to changes in the reference price (capital gains). The term $r\tilde{p}_t dt$ captures the per-unit opportunity cost of holding a risky OTC asset, the value $\tilde{p}_t$ of which otherwise could be earning the discount rate r. Finally, when an agent trades, they balance their effective trade price, P_t, relative to the reference price at which the asset is recorded on their balance sheet, $\tilde{p}_t$. The term $\tilde{p}_t - P_t$ is called *slippage* by practitioners and captures the net gain or loss to an agent's equity when trading the illiquid asset. In centralized markets, slippage is always zero as the trade price coincides with the reference price.

At time t, a dealer of state (E, n) maximizes their expected discounted value of consumption subject to equations (7), (8), as well as the regulatory VaR constraint (6):

$$W_t(E, n) = \sup_{\{c(s)\}_{s=t}^{\infty}} E_t \left[\int_t^{\infty} -e^{r(s-t)} e^{-\kappa c(s)} ds \mid E_t = E, n_t = n \right]. \tag{9}$$

Similarly, an investor of state (E, n, i) solves:

$$V_t(E, n, i) = \sup_{\{c(s)\}_{s=t}^{\infty}} E_t \left[\int_t^{\infty} -e^{r(s-t)} e^{-\kappa c(s)} ds \mid E_t = E, n_t = n, i_t = i \right]. \tag{10}$$

3.2 Trade in the OTC asset

When agents meet in the OTC market, they engage in proportional bilateral bargaining (Kalai, 1977) with equal bargaining power over the terms of trade (a, P) where a is the trade size.

Inter-dealer trade. First, consider inter-dealer trade. When $a > 0$, this indicates a "home" dealer (E, n) buy and an "away" dealer (E', n') sell. The home dealer maximizes their post-trade surplus with respect to (a, P),

$$\max_{P \geq 0, a} \left[W(E + (\tilde{p} - P)a, n + a) - W(E, n) \right], \tag{11}$$

subject to the following constraints:

$$W(E + (\tilde{p} - P)a, n + a) - W(E, n) = W(E' - (\tilde{p} - P)a, n' - a) - W(E', n'),$$
$$-n \leq a \leq n',$$
$$E + (\tilde{p} - P)a \geq \bar{E}(n + a),$$
$$E' - (\tilde{p} - P)a \geq \bar{E}(n' - a).$$

The first constraint is the surplus sharing Kalai (1977) constraint, which stipulates that the post-trade surplus of each dealer must be equal when there is equal bargaining power. The second equation reflects feasibility: dealers can trade only assets they have. Since dealers cannot consume simultaneously with trading, the last two equations ensure that post-trade both agents remain in compliance with the regulatory constraint given by (4). Note that agents do not necessarily always trade, as $a = 0$ is in the set of possible outcomes.

Investor trade. Investor trade occurs in a similar manner. Let $a > 0$ denote a dealer buy and an investor sell. The dealer maximizes their post-trade surplus with respect to (a, P),

$$\max_{P \geq 0, a} \left[W(E + (\tilde{p} - P)a, n + a) - W(E, n) \right], \tag{12}$$

subject to:

$$W(E + (\tilde{p} - P)a, n + a) - W(E, n) = V(E' - (\tilde{p} - P)a, n' - a, i) - V(E', n', i),$$
$$-n \leq a \leq n',$$
$$E + (\tilde{p} - P)a \geq \bar{E}(n + a),$$
$$E' - (\tilde{p} - P)a \geq 0.$$

Again, the first constraint is the Kalai (1977) surplus-sharing constraint, the second is feasibility, and the third is the dealer's regulatory compliance constraint. The last line reflects the investor's non-negative equity requirement.

3.3 Stationary equilibrium

I consider a steady state in which the reference price $\tilde{p}$ is constant ($d\tilde{p}_t = 0$).

3.3.1 Dealer value function and optimal consumption policy

Assuming appropriate differentiability and using Ito's Lemma, the dealer's Hamilton-Jacobi-Bellman (HJB) equation is:

$$\begin{aligned} rW(E, n) = \max_c \Big\{ & u(c) + \partial_E W(E, n) \left(\mu_\eta + rE - c + n \left(\mu_D - r\tilde{p} \right) \right) \\ & + \frac{1}{2} \partial_E^2 W(E, n) \left(\sigma_\eta^2 + \sigma_D^2 n^2 + 2n\rho\sigma_\eta\sigma_D \right) \\ & + \lambda_c \int_{E', n', i} \left[W(E + (\tilde{p} - P^*) a^*, n + a^*) - W(E, n) \right]^+ dH(E', n', i) \\ & + \lambda_d \int_{E', n'} \left[W(E + (\tilde{p} - P^*) a^*, n + a^*) - W(E, n) \right]^+ dF(E', n') \Big\}, \end{aligned} \tag{13}$$

The first term is the flow utility that dealers earn from consuming the numeraire good c; the second and third terms capture risk aversion due to the stochastic nature of the endowment and the OTC asset dividend (drift and volatility, respectively); the final two terms capture the expected change in value due to inter-dealer and investor trade, where (a^*, P^*) are determined by the bargaining process outlined in the previous section, and so are functions of the trade counterparty type - either (E', n') or (E', n', i).

Now consider the dealer's optimal consumption policy. The dealer maximizes their consumption c given their current level of equity E and a fixed stock of OTC risk assets n. As stated previously, the dealer must satisfy the regulatory VaR constraint as defined by (6). As a consequence a dealer's consumption policy $c^*(E, n)$ must satisfy at all times:

$$\begin{aligned} c^*(E,n) \leq c^{VaR}(E,n) = \left(\mu_\eta + n\left(\mu_D - r\tilde{p}\right)\right) + \frac{r}{e^{r\Delta t}-1}\left(e^{r\Delta t}E - \bar{E}(n)\right) \\ + \frac{r}{e^{r\Delta t}-1}\Phi^{-1}(\epsilon)\sqrt{\frac{1}{2}\frac{1}{r}\left(\sigma_\eta^2 + n^2\sigma_D^2 + 2\rho n \sigma_\eta \sigma_D\right)\left(e^{2r\Delta t}-1\right)}. \end{aligned} \tag{14}$$

Where $\Phi^{-1}(\epsilon)$ is the inverse cumulative distribution function (CDF) of a standard normal, otherwise known as the quantile function. This maximum level of consumption, $c^{VaR}(E, n)$, follows directly from the regulatory VaR constraint as defined in (6) as shown in Appendix A.1. To satisfy the regulatory VaR constraint, dealers must consume in such a way that the *drift* in their equity, given the nature of the stochastic processes of the endowment and dividends, is large enough to recover the regulatory minimum over the time horizon $t + \Delta t$ with probability $(1 - \epsilon)\%$.

A state-based boundary condition. I implement this regulatory VaR implied maximum consumption through a state-based boundary condition following the logic of Achdou et al. (2022). Achdou et al. (2022) handle constraints in continuous time by implementing what is essentially a constraint on the shape of the value function: as agents realize negative income shocks and approach their constraints, the boundary condition ensures that agents consume in such a way that they drift away from the constraint. The state-based boundary condition in my framework is $\forall n$:

$$\partial_E W(E,n) \geq u'\left(c^{VaR}(E,n)\right). \tag{15}$$

Under this condition, optimal consumption is characterized by the standard Euler equation

$$u'(c^*(E,n)) = \partial_E W(E,n), \tag{16}$$

which, given that (15) holds, always satisfies the regulatory VaR constraint as in (6).

There are two key differences between the implementation of the state-based boundary condition in this framework and that of Achdou et al. (2022). First, in theory, dealers can potentially realize a negative income shock such that the regulatory constraint is violated despite their efforts to remain in compliance (with probability $\epsilon\%$). In practice, when solving the model numerically, however, the discretization of the value function required by the numerical solution renders the effective probability of realizing a negative income shock that pushes dealers into violation is zero. As a consequence, in the computed equilibrium solution, the regulatory constraint is violated due to income shocks with effective probability zero. Second, the state-based boundary constraint in this framework is *endogenous*: dealers can periodically adjust their stock of risky OTC assets n, and thereby adjust the regulatory minimum level of equity they must hold $\bar{E}(n)$.

3.3.2 Investor value function and optimal consumption policy

The investor's HJB equation takes a similar form to that of dealers, with two key changes that reflect the two fundamental differences between dealers and investors. The first difference is that an investor's liquidity preference, the correlation parameter ρ^i, changes periodically. The second is that, while dealers have access to both the investor and inter-dealer markets, investors can trade only with dealers in the investor market. The investor's Hamilton-Jacobi-Bellman (HJB) equation is:

$$
\begin{aligned}
rV(E,n,i) = \max_c \{ & u(c) + \partial_E V(E,n,i)\left(\mu_\eta + rE - c + n(\mu_D - r\tilde{p})\right) \\
& + \frac{1}{2}\partial_E^2 V(E,n,i)\left(\sigma_\eta^2 + \sigma_D^2 n^2 + 2n\rho^i\sigma_\eta\sigma_D\right) \\
& + \gamma\pi(i')\left[V(E,n,i') - V(E,n,i)\right] \\
& + \lambda_c \int_{E',n'} \left[V(E - (\tilde{p} - P^*)\, a^*, n - a^*, i) - V(E,n,i)\right]^+ dF(E',n') \Big\}.
\end{aligned} \tag{17}
$$

The first term is the flow utility that investors earn from consuming the numeraire good c; the second and third terms capture risk-aversion due to the stochastic nature of income that depends on the investor's type i; the fourth term denotes the investor's change in value due to the liquidity preference type switching process; the final term captures the expected change in value due to investor trade where (a^*, P^*) are determined by bargaining with a dealer of type (E', n').

The investor's consumption policy, like dealers, is also implemented via a boundary condition, as discussed in the previous section. Optimal investor consumption always satisfies the standard Euler equation

$$
u'(c^*(E,n,i)) = \partial_E V(E,n,i), \tag{18}
$$

given the following boundary condition is satisfied:

$$
\partial_E V(E,n,i) \geq u'\left(c^{VaR}(E,n,i)\right). \tag{19}
$$

Where $c^{VaR}(E,n,i)$ is defined as:

$$
\begin{aligned}
c^{VaR}(E,n,i) = & \left(\mu_\eta + n\left(\mu_D - r\tilde{p}\right)\right) + \frac{r}{e^{r\Delta t} - 1}\left(e^{r\Delta t}E\right) \\
& + \frac{r}{e^{r\Delta t} - 1}\Phi^{-1}(\epsilon)\sqrt{\frac{1}{2}\frac{1}{r}\left(\sigma_\eta^2 + n^2\sigma_D^2 + 2\rho^i n\sigma_\eta\sigma_D\right)\left(e^{2r\Delta t} - 1\right)}.
\end{aligned} \tag{20}
$$

Therefore, investors always satisfy their non-negative equity VaR constraint as defined in equation (6).

The steady state distributions of agent types, the reference price of the risky OTC asset $\tilde{p}$, the market clearing condition, and the stationary equilibrium definition are all relegated for brevity to Appendix B.

3.4 The regulatory cost of OTC inventories

The regulatory costs of OTC inventories in this model arise endogenously from dealers' risk-based capital requirements. The risk-based requirement induces regulatory costs through two mechanisms. The first is by directly increasing dealer capital requirements. By definition of the risk-weighted capital ratio, a larger OTC position raises the minimum equity requirement: $\partial_n \bar{E}(n) > 0$. As inventories increase, dealers must commit more capital to satisfy the requirement, reducing their flexibility to handle adverse income shocks.

The second is through an indirect precautionary savings mechanism, summarized by Proposition 1. In line with standard incomplete markets models, for any given inventory level $n \geq 0$ there are precautionary savings. Specifically, the regulatory constraint raises the shadow value of equity as a dealer's regulatory slack shrinks, and as a consequence, for a fixed stock of inventories, dealers cut their consumption and boost their savings in the risk-free asset to move away from the constraint. However, what is novel is that since OTC assets are risky, a dealer's portfolio volatility increases with their OTC inventories. As a consequence, the precautionary savings effect becomes larger the more OTC assets a dealer has (Proposition 1.1).

Although dealer OTC inventories are sticky, they are not fixed: dealers can adjust them as OTC trading opportunities arrive. A novel result of the model is that, since the marginal value of equity varies with OTC inventories, so does the marginal value of an additional unit of OTC assets. Specifically, the reservation value (willingness to pay) for a trade in which a dealer buys one unit of the OTC asset and their regulatory slack remains fixed is declining as a dealer becomes more constrained (Proposition 1.2). In other words, a dealer approaching their regulatory limit wishes to reduce their inventories in order to create additional regulatory slack through (1) holding more cash and (2) reducing the volatility of their portfolio. Again, because of the OTC asset's risk profile, this effect is larger the more OTC assets a dealer has.

Figures 1a and 1b plot the marginal values of equity and OTC inventories, respectively, as a function of a dealer's distance-to-constraint for different asset levels: $n'' > n' > n > 0$. As shown, as a dealer becomes more constrained (distance-to-constraint falls), the former increases and the latter decreases. In between infrequent trading opportunities, dealers cut their consumption and boost their savings in the risk-free asset to move away from their bound. When a trading opportunity arises, they wish to offload OTC inventories to hold more cash and reduce the volatility of their portfolio. This effect is larger the more assets a dealer has.

Proposition 1. *Define for regulatory slack $s \geq 0$ the corresponding level of equity given a stock of OTC inventories $n \geq 0$: $E_s(n) := \bar{E}(n) + s$. Assume that the reference price is such that $\mu_D - r\tilde{p}$, if positive, is sufficiently small. In the neighborhood of the regulatory constraint:*

1. *The marginal value of equity increases as a dealer becomes more constrained (regulatory slack s declines), and this increase is steeper the larger the dealer's OTC inventories. That is, for $n' > n$:*

$$\partial_E W(E_s(n'), n') > \partial_E W(E_s(n), n) > 0.$$

2. *For trade size $a = 1$ and trade price P such that regulatory slack s is constant, the reservation value is:*

$$\Delta W(E_s, n) = W(E_s(n+1), n+1) - W(E_s(n), n).$$

The reservation value of an additional unit of OTC assets in a fixed-slack trade declines as a dealer becomes more constrained (regulatory slack s declines), and this decline is steeper the larger the dealer's OTC inventories:

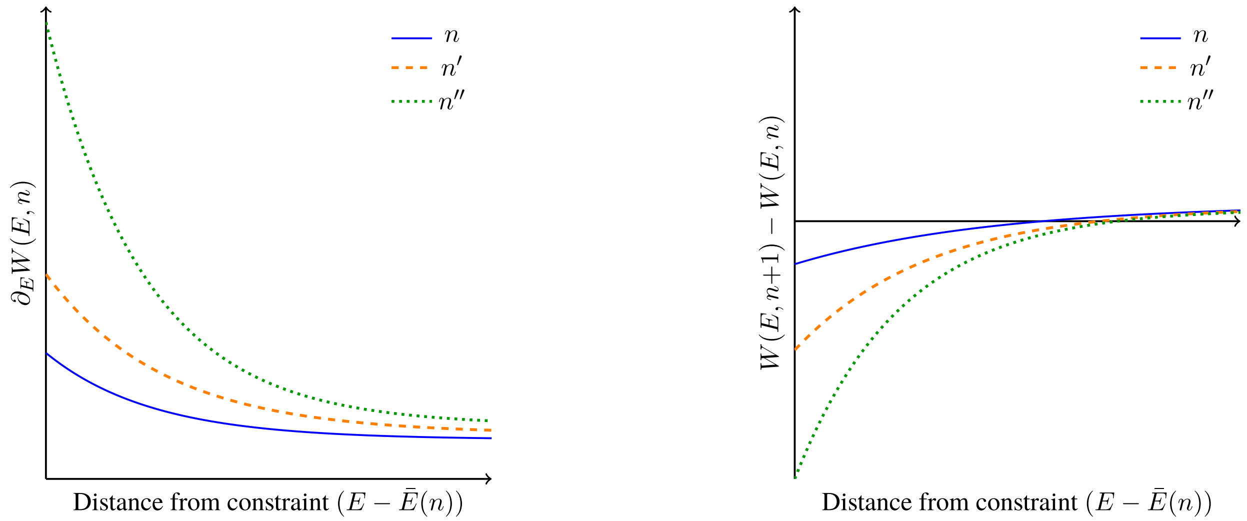
$$\partial_E \Delta W(E_s, n) > 0.$$

Proof: see Appendix A.2.

To illustrate, Figure 2 plots an example of two dealer value functions for different stocks of OTC inventories, $n' > n$ for orange and blue, respectively. The vertical dashed lines at $\bar{E}(n)$ denote the minimum regulatory level of equity. $E_s(n)$ denotes the level of equity for a given level of regulatory *slack*, or distance-to-constraint. Say that the dealer with n assets considers buying OTC assets, increasing their inventory level to n'. The minimum level of regulatory slack the dealer needs to maintain their same pre-trade value, $W(E_s(n), n)$ is greater at n' assets than at their current level. Recall that when buying an asset, a dealer considers their total slippage, the difference between the reference price and the realized trade price: $\tilde{p} - P$. Under the first direct capital requirement mechanism, the dealer needs a sufficiently low buy price for the purchase to be regulatory-compliant. That is, the transaction must raise equity to at least the new minimum: $\bar{E}(n')$. Due to the second indirect precautionary savings mechanism, the dealer requires an additional discount on the buy price for the purchase to be incentive compatible. To execute the trade the dealer requires more regulatory slack at the new inventory level: $E^* > E_s(n')$. As dealer inventories increase, the amount of post-trade regulatory slack required to make a purchase incentive compatible also increases.

Thus, the true "balance sheet cost" of holding OTC inventories is measured by the additional equity a dealer demands due to this balance sheet level regulation when increasing their OTC inventories. As dealer inventories grow or capitalization falls, the marginal regulatory cost of holding additional OTC inventories also increases. This is in stark contrast to models in which balance sheet costs are modeled as per-unit linear costs or even quadratic costs. This work emphasizes how regulatory costs can vary over time as income shocks change dealer capitalization.

The fact that the risky OTC asset is traded in a market characterized by search frictions is critical to this dynamic. In a model with centralized trading, agents could continuously re-balance their asset levels freely. Agents with CARA utility would simply hold a fixed level of assets either at their preferred level or up to the regulatory limit. Unless agents face frictions in buying and selling the OTC asset, the threat of violating the constraint holds no power.



(a) Marginal Value of Equity vs. Distance-to-Constraint

(b) Marginal Value of OTC Assets vs. Distance-to-Constraint

Figure 1: Marginal Values vs. Distance-to-Constraint

In each graph, on the x-axis is the distance-to-constraint, $E - \bar{E}(n)$ of a dealer of state (E, n) with equity E and inventories n; 0 indicates that the dealer is just at the regulatory minimum level of equity, $E = \bar{E}(n)$.

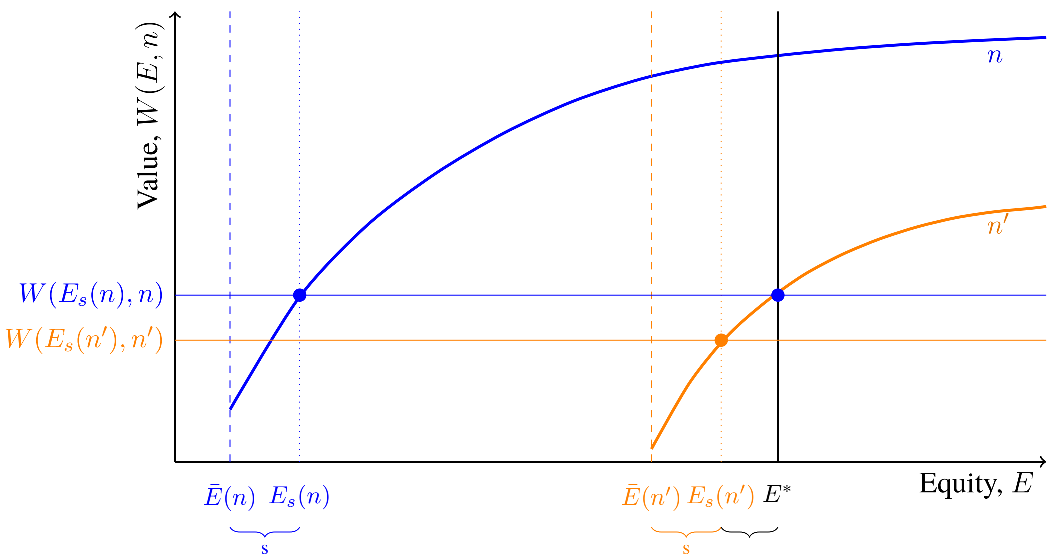


Figure 2: The regulatory cost of OTC inventories

This figure plots two example value functions $W(E, n)$ as a function of Equity (x-axis) for different levels of OTC inventories $n' > n$. The dashed vertical lines indicate the minimum level of equity a dealer with given OTC assets n or n' must hold, $\bar{E}(n)$ and $\bar{E}(n')$ respectively. $s \geq 0$ denotes a fixed level of regulatory slack such that $E_s(n) = \bar{E}(n) + s$.

4 Quantitative Application

Solving the steady state numerically amounts to a high-dimensional fixed point problem in the reference price $\tilde{p}$, agent value functions $\{W(E,n), V(E,n,i)\}$, agent trading rules $\{a^*, P^*\}$ for all possible trades, and agent distributions $\{F(E,n), H(E,n,i)\}$. Because the regulatory constraint introduces wealth effects into the bargaining process, standard tools in the search OTC literature cannot be applied to obtain analytical solutions to these objects.

In a similar environment without any regulatory capital requirement on the part of dealers, as in Duffie et al. (2007), Gârleanu (2009), Praz (2014), and Üslü (2019), CARA utility implies that an agent's valuation of assets is always independent of her equity position, given the assumption that agents are always well-capitalized. In these frameworks, one can replicate the risk-return trade-off faced by risk-averse agents using risk-neutral agents with a reduced-form quadratic utility over OTC assets. However, when a strict balance sheet constraint is introduced, the assumption that agents are always well-capitalized becomes unrealistic. As a consequence, this equity-independence feature of CARA utility no longer holds: agents must now also track their equity position and proximity to the regulatory constraint, and take their position into account when trading in the OTC asset. This introduces wealth effects into the bargaining problem, preventing the use of linearization techniques that allow these authors to obtain closed-form solutions.

The algorithm I have developed to solve the model numerically is available in Appendix C. At the heart of the algorithm is the method I use to compute agent value functions, given agent distributions and trading patterns. I extend the finite difference methods proposed by Achdou et al. (2022) to solve continuous-time heterogeneous-agent models with exogenous borrowing constraints. I adapt their method to account for what, in this model, is essentially a state-based and endogenous borrowing constraint.[8]

I discretize the state space for liquid risk-free assets to be $E_k \in [0, 10]$ with $K = 70$ points and

[8]One can re-arrange the risk-weighted asset ratio defined in (4) as $\omega \geq -(1-\Lambda\nu)\tilde{p}n$, which, since $1-\Lambda\nu > 0$, is a constraint on borrowing as it provides a lower bound for net risk-free assets, ω. Holding more OTC assets n allows dealers to back more borrowing; stricter regulation (higher Λ or ν) reduces the collateral value of OTC assets. The borrowing constraint is endogenous and state-based since it depends on dealer OTC asset holdings n, which in equilibrium are the outcome of dealer trading decisions.

risky OTC assets to be $n_j \in \{0, 1, 2, 3, 4, 5\}$. As the solution method requires obtaining the terms of trade (a, P) for each pair of agents, to reduce computation time, I assume unit trade $a \in \{0, -1, 1\}$. I restrict investors to only two types, high ($i = H$) and low ($i = L$), where $\rho^H < 0 < \rho^L$. High type investors have a stronger hedging motivation to hold OTC assets than low type investors. Finally, to better match the data, I incorporate two additional quantitative parameters.[9] First, I add a stock of non-trading risk-weighted assets $\psi \geq 0$, which is fixed and constant across dealers. Second, to ensure that holding $n_J = 5$ units of the OTC asset constitutes a *very* large position relative to equity, I introduce the regulatory parameter $\phi \geq 0$ to scale asset holdings within the regulatory constraint. The minimum level of equity with these two new quantitative parameters is:

$$E \geq \bar{E}(n) = \Lambda \left(\psi + \nu \tilde{p} n \times \phi\right)$$

Externally set parameters. Table 1 presents the externally set and policy-determined parameters. The following parameters are set externally. The discount rate is set to an annual rate of $r = 5\%$, and the supply of assets is $S = 2$. The prevalence of dealers in the economy relative to investors is set to $p_D = 1\%$. Conditional on receiving a preference shock, the probability an investor becomes a high type is $\pi_H = 25\%$ and a low type is $\pi_L = 1 - \pi_H = 75\%$. I assume that dealers have no hedging need for holding the OTC asset: $\rho = 0$. Dealers have no correlation between the dividend paid out by the OTC asset and their endowment income. I assume that low and high type investors have symmetric but opposite correlations: $\rho^L = |\rho^H|$ where $\rho^H < 0$ will be calibrated. I also choose the VaR settings such that the probability of violating the regulation constraint (for risk management purposes) is $\epsilon = 0.01\%$ over a time horizon of one quarter $\Delta t = 1$.

Policy parameters are chosen in accordance with US regulations prior to the GFC. I set the regulatory capital requirement to $\Lambda = 4\%$. As a baseline risk-weight, I use $\nu = 50\%$. Risk weights are most often set in 50-percentage-point increments, with a 50% risk weight used for relatively liquid investment-grade corporate bonds.[10] To study stricter regulation, I specifically examine a

[9]Proposition 1 is unchanged with the addition of these parameters.
[10]Risk weights of up to 200%, 250%, or even 500% can be used, depending on the asset's risk profile.

change in the capital requirement Λ from 4% to 8.5%, mirroring the changes to the common equity tier 1 risk-weighted asset ratio introduced by Basel II.5 and III for large, systemically important institutions. To compare model outcomes to a world without a risk-based capital requirement, I consider $\Lambda = 0\%$. In this case, dealers are simply subject to the same non-negative equity condition as investors.

Externally set parameters		
Variable	**Parameter**	**Value**
Discount rate (annual)	r	5%
Asset supply	S	2
Prevalence of dealers	p_D	1%
Investor type-switching, high	π_H	25%
Investor type-switching, low	π_L	$1 - \pi_H$
Dealer corr.	ρ	0
low type investor corr.	ρ^L	$\lvert\rho^H\rvert$
VaR probability	ϵ	0.01%
VaR time horizon (qtr)	Δt	1
Policy parameters		
Variable	**Parameter**	**Value**
OTC asset risk weight	ν	50%
Capital requirement	Λ	4%

Table 1: Externally set & policy parameters

Estimated parameters. The parameters of the stochastic process μ_η, σ_η, μ_D, and σ_D are directly estimated using quarterly balance sheet level data from the FR-Y9C filings of bank holding companies (BHCs). The data and the maximum likelihood procedure used are outlined in Appendix D. Table 2 presents the parameter estimates. As shown, the endowment is a low-return, low-volatility income stream, whereas the OTC asset has a higher return with higher volatility.

Calibrated parameters & moments. The remaining parameters to be calibrated are: λ_d, λ_c, ρ^H, κ, γ, ψ, and ϕ. I calibrate the baseline model to the pre-GFC period using investment-grade, risky-principal corporate bond transaction-level data from Academic TRACE (with the total sample

	(1) Endowment $d\eta_t$	(2) Dividend dD_t
$\hat{\mu}_\eta$	0.0101 (0.00076)	–
$\hat{\sigma}_\eta$	0.0717 (0.00078)	–
$\hat{\mu}_D$	–	0.0615 (0.0024)
$\hat{\sigma}_D$	–	0.1875 (0.0026)
N	50,704	10,081
LL	-103,086	-6,052.24
BHC controls	✓	✓
Sample controls	✓	✓
Crisis controls	✓	✓
Seasonality controls	✓	✓

Table 2: Stochastic process MLE estimates

Column (1) presents the maximum likelihood estimation (MLE) estimates of the parameters (μ_η, σ_η) for the process defined by equation (2). Column (2) presents the MLE estimates of the parameters (μ_D, σ_D) for the process defined by equation (1). Estimation details can be found in Appendix D.

Variable	**Param.**	**Value**	**Moment**	**Model**	**Data**
Risk aversion	κ	0.4	RWA ratio, mean (%)	21.61	11.14
Rate of investor pref. shock	γ	0.083	RWA ratio, std. dev.	11.48	8.66
high type investor corr.	ρ^H	-0.5	Investor Spread, mean (bps)	6.05	15.05
Meeting intensity, Investor (qtr)	λ_c	12	Investor Spread, std. (bps)	5.46	18.40
Meeting intensity, Inter-dealer (qtr)	λ_d	25	Investor / Inter-dealer Turnover	3.88	1.69
Non-trading RWA	ψ	80	RWA ratio, mean (95pct. TA) (%)	17.21	8.78
Regulatory scaling parameter	ϕ	3	TA / Total Assets, mean (95pct. TA) (%)	17.5	1.53

Table 3: Calibrated parameters & moments

Note that *RWA*: risk-weighted assets and *TA*: trading assets. *95pct. TA* indicates that variable is calculated using the BHCs that have trading assets $\geq$ the 95th percentile of trading assets.

varying from 2006 to 2024). Table 3 presents the calibrated parameters, the target moments, and the model's moments relative to the data. Since λ_d controls how fast dealers trade in the inter-dealer market, I target the ratio of investor turnover to inter-dealer turnover. Since ρ^H determines the hedging need of the high type investor and λ_c controls how often investors receive trading opportunities, I target the average spread in the investor market and standard deviation, respectively. I use the spread definition of Choi et al., 2024, defined as the difference between the trade price relative to a dealer reference price. For a given trade t in bond b, the spread in basis points is:

$$Spread_{b,t} = 2 \times Q \times \left(\log(P_{b,t}) - \log(\tilde{p}^{b}_{dealer})\right) \times 10,000. \tag{21}$$

From the dealer's perspective, $Q = 1$ if the trade is a sale and $Q = -1$ if the trade is a buy. A positive spread indicates a profit for the dealer buying or selling to investors or other dealers. Following Choi et al., 2024, in the data I use as the dealer reference price the volume-weighted average (using trades with volumes greater than \$100,000) of the inter-dealer price for each bond in each quarter. In the model, the dealer reference price is the volume-weighted average price using inter-dealer trades.

Spreads are calculated in the data and in the model as follows. To find the average investor spread across all investment-grade corporates in the data, I first compute the volume-weighted average spread for each bond over all investor trades within a quarter. I then average of these volume-weighted quarterly investor spreads across all bonds traded in the investor market pre-GFC. In the model, for each potential trade I calculate the spread using the dealer reference price. Using the agents' steady state distributions, I obtain the average spread for realized investor trades.

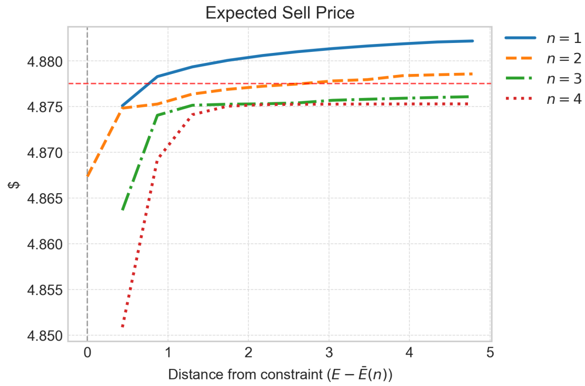
The remaining parameters are calibrated to moments from the balance sheets of the top 25% largest BHCs pre-GFC. Agent risk aversion κ and the rate at which investors receive preference shocks γ are calibrated using the first and second moments of dealer risk-weighted asset ratios. The stock of non-trading risk-weighted assets ψ is calibrated to the average risk-weighted ratio of BHCs that hold a high amount of trading assets (specifically, more than the 95th percentile of trading assets). For the regulatory scaling parameter ϕ, I target the average ratio of trading assets to total assets for these same BHCs with high trading asset holdings.

4.1 Baseline model outcomes

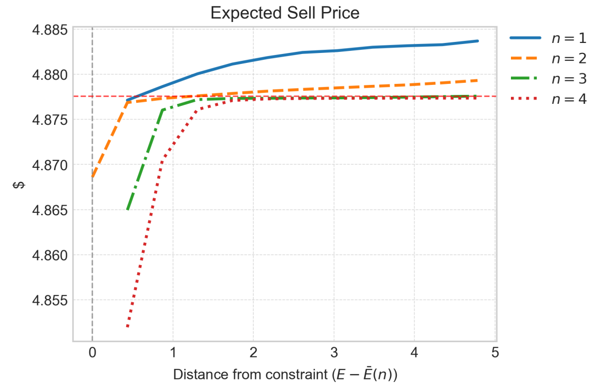
Summary statistics for each of the three capital requirement cases $\Lambda = \{0\%, 4\%, 8.5\%\}$ are presented in Table 4 columns (1) to (3) respectively. Consider the baseline, $\Lambda = 4\%$ as in column

Statistic	Policy (Λ, ν) (1) **(0%, 0%)**	(2) **(4%, 50%)**	(3) **(8.5%, 50%)**	$\Delta_{2,3}\%$
Equilibrium reference price ($\tilde{p}$)	4.878	4.878	4.875	-0.08
Average inter-dealer price	4.877	4.877	4.876	-0.01
std. dev. inter-dealer price	0.00080	0.00084	0.000897	+6.79
Average investor price	4.878	4.878	4.874	-0.08
std. dev. investor price	0.00159	0.00167	0.00181	+8.38
Average inter-dealer spread (bps)	-2.86e-17	-0.013	-0.017	-32.54
std. dev. inter-dealer spread	3.314	3.475	3.678	+5.83
Average investor spread (bps)	5.760	6.048	6.578	+8.78
std. dev. investor spread	5.015	5.462	6.134	+12.30
Inter-dealer trade (%)	9.467	10.164	11.175	+9.95
Investor trade (%)	70.952	70.940	70.155	-1.11
Investor / inter-dealer trades	4.163	3.877	3.488	-10.03
Inter-dealer turnover, qtr (%)	42.600	45.740	50.286	+9.90
Investor turnover, qtr (%)	177.380	177.350	175.389	-1.10
Inter-dealer Intermediary Index, mean	0.0381	0.0417	0.0435	+4.31
Investor Intermediary Index, mean	0.6371	0.6126	0.5935	-3.11
Equity				
Mean	15.006	16.550	18.070	+9.18
Std. dev.	9.572	8.570	7.540	-12.03
25th pct.	5.652	8.260	10.434	+26.27
50th pct.	15.217	16.520	18.261	+10.55
75th pct.	24.347	24.780	25.652	+3.52
Risk-weighted asset ratio (%)				
Mean	–	21.610	23.638	+9.36
Std. dev.	–	11.480	10.120	-11.84
25th pct.	–	10.430	13.734	+31.66
50th pct.	–	20.870	23.070	+10.54
75th pct.	–	31.860	31.998	+0.43
Dealer OTC assets				
Mean	2.289	2.239	2.189	-2.23
Std. dev.	0.936	0.950	0.950	-0.05
low type Investor OTC assets				
Mean	1.974	1.975	1.973	-0.10
Std. dev.	1.002	1.011	1.012	+0.05
high type Investor OTC assets				
Mean	2.066	2.065	2.076	+0.53
Std. dev.	0.979	0.986	0.986	-0.02

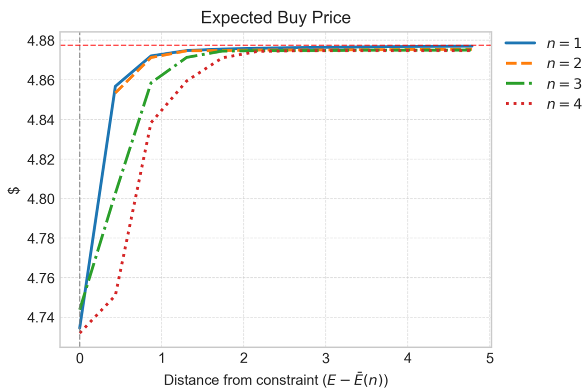
Table 4: Model statistics for different capital requirements Λ



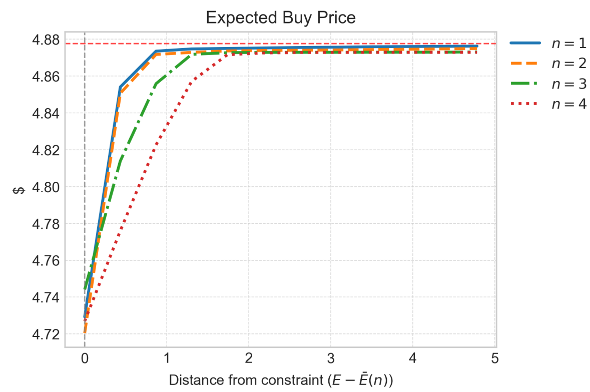
(a) Expected sell price, Inter-dealer market



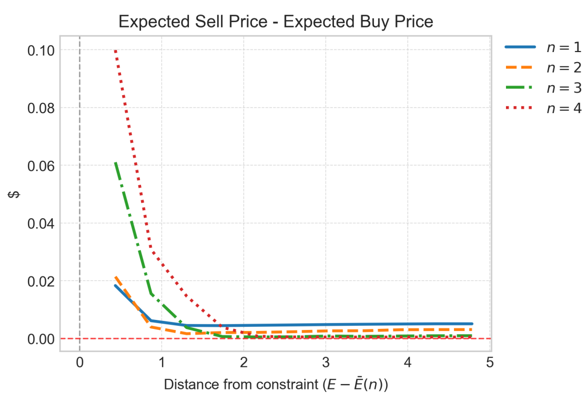
(b) Expected sell price, Investor market



(c) Expected buy price, Inter-dealer market



(d) Expected buy price, Investor market

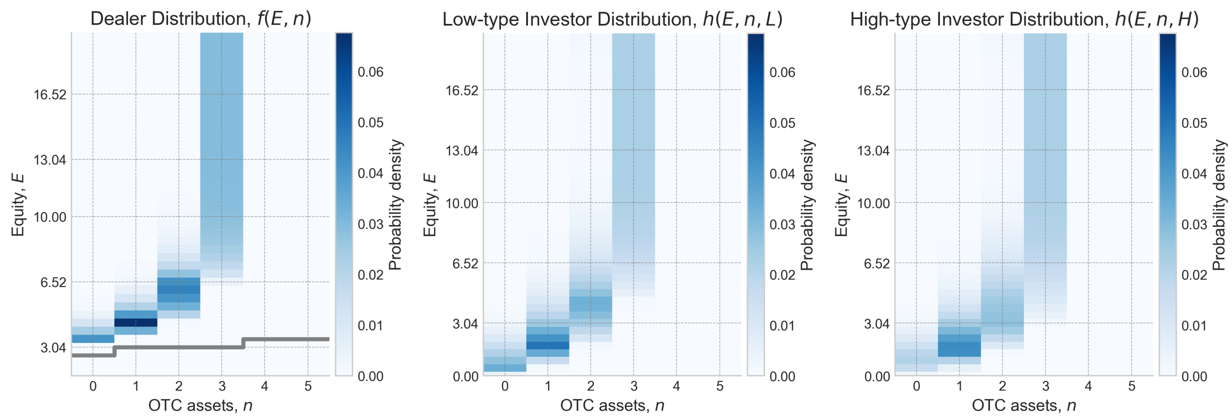


(e) Expected sell price less expected buy price, Inter-dealer market

(f) Expected sell price less expected buy price, Investor market

Figure 3: Prices & Spreads by market

In each graph, on the x-axis is the distance-to-constraint, $E - \bar{E}(n)$, of a dealer of state (E, n) with equity E and inventories n; 0 indicates that the dealer is just at the regulatory minimum level of equity, $E = \bar{E}(n)$. On the y-axis is the dollar value of the variable plotted. *Expected Sell Price* (*Expected Buy Price*) is the average price at which a (E, n) dealer expects to sell (buy) one unit of the OTC asset for, based on the equilibrium distribution of agents.



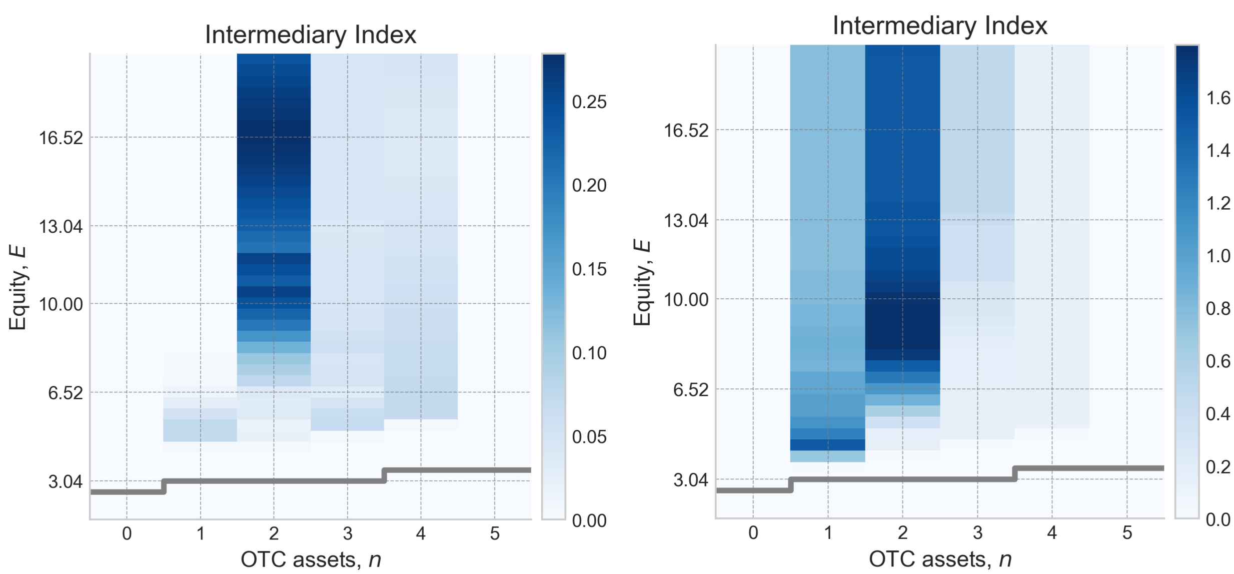
(a) Distribution of dealer types $f(E,n)$

(b) Distribution of investor types, low $h(E,n,L)$

(c) Distribution of investor types, high $h(E,n,H)$

Figure 4: Agent Distributions

Each heatmap shows the steady state equilibrium distributions of dealers, low type investors, and high type investors. On the x-axis is an agent's OTC asset holdings n, and on the y-axis is the agent's equity E. For dealers, the solid gray step function denotes the regulatory minimum level of equity they must hold $\bar{E}(n)$ for each stock of OTC inventories n. The darker the color, the more often agents spend time in that state.



(a) Intermediary index, Inter-dealer market

(b) Intermediary index, Investor market

Figure 5: Intermediation quality by market

Each heatmap represents the intermediation index for the inter-dealer and investor markets as defined in equation (22). On the x-axis is a dealer's OTC asset holdings n and on the y-axis is the dealer's equity E. The solid gray step function denotes the regulatory minimum level of equity they must hold $\bar{E}(n)$ for each stock of OTC inventories n. The darker the color, the higher the index, and the better an intermediary an (E,n) dealer is.

(2). As Proposition 1 shows, the regulatory constraint distorts dealers' marginal valuations of both equity and the risky OTC asset near the regulatory bound. These distortions affect dealer market-making behavior and, in turn, equilibrium outcomes. In this section, I discuss three key implications of balance sheet constraints for dealer trading behavior across the distribution of dealer states.

First, the regulatory constraint affects dealer pricing strategies. Figures 3 (a)-(d) show how a dealer's expected sell and buy prices in both the inter-dealer and investor markets vary with their distance-to-constraint. On the x-axis is the dealer's distance from their constraint, their regulatory slack. Each line color and style denotes a different OTC inventory level n, from 1 to 4. The *Expected Sell Price* (*Expected Buy Price*) is the average price at which a (E, n) dealer expects to sell (buy) one unit of the OTC asset for, based on the equilibrium distribution of agents. As a dealer's distance-to-constraint falls, their sell price declines; they are willing to sell their OTC inventory at fire sale prices. Additionally, this effect is stronger the more assets a dealer has, due to the OTC asset's risk profile. Similarly, as a dealer becomes more constrained, they are less willing to buy the asset and will do so only at a steeply discounted price. Both price declines directly reflect the changing marginal values of dealers in the region near the constraint.

While both expected buy and sell prices fall as a dealer becomes more constrained, the wedge between them rises. Figures 3 (e) and (f) plot the difference between the expected sell and buy prices for a dealer of state (E, n) for both the inter-dealer and investor markets. The difference, or spread, reflects how costly it is for investors to trade with that type of dealer. As each graph shows, as a dealer's slack declines, the wedge between their expected sell and buy prices increases. This effect is more pronounced the more OTC assets a dealer has. Unlike other models with ad-hoc regulatory costs, this framework can match the cross-sectional empirical evidence from Adrian et al. (2017) and Breckenfelder and Ivashina (2021a), among others, who show that bonds traded by more constrained dealers have higher trading costs.

Second, OTC inventories exhibit mean-reversion, and, as agents become less well-capitalized, their target inventories fall. In other words, agents have a preferred asset level they would like to hold; when they are close to their regulatory constraint, this desired inventory level is lower.

Figure 4 presents a heatmap of the distribution of agent types, where the gray step function for dealers represents the minimum level of equity $\bar{E}(n)$ they must have for each level of risky OTC asset n. Figure 4 shows that agents maintain a target inventory, generally holding $n = 3$ units of the risky OTC asset when they are well-capitalized. When well-capitalized agents receive trading opportunities, they balance the future trading value of assets against any deviation from their target inventory level. This result is consistent with Üslü (2019), in which the author shows that risk-neutral traders with quadratic utility over OTC assets have a target inventory.[11] It is also consistent with the empirical evidence, for example from Dick-Nielsen and Rossi (2019), which shows that dealer inventories are mean-reverting.

However, Figure 4 also shows that there are significant masses of agents with low OTC inventories and low-capitalization. As agents receive negative income shocks, they would rather exchange some of their risky OTC inventories not only to hold more cash to relax their regulatory constraint (for dealers) or to satisfy their non-negative equity constraint (for investors), but also to reduce the overall portfolio volatility. As a result, their desired inventories of risky OTC assets decline.

The third and final observation about the cross-section of dealers is that more constrained dealers tend to be worse intermediaries. Figure 5 presents the intermediation index for each dealer type (E, n) in each market, inter-dealer and investor. The intermediary index is a simple measure of how much trading volume a dealer of state (E, n) expects to facilitate on each side of the market:

$$\text{Intermediary Index}(E, n) = \sqrt{E[\text{Trade Volume}_{sell}|E, n] \times E[\text{Trade Volume}_{buy}|E, n]}. \tag{22}$$

A dealer of state (E, n) is a "good" intermediary if they are willing to both buy and sell at high volumes. As shown, dealers intermediate much more in the investor market than they do for themselves. What is common across the two markets is that the best intermediaries tend to be dealers who are reasonably well-capitalized (i.e., have sufficient regulatory slack) and have an intermediate amount of assets, $n = 2$. As a dealer's regulatory slack declines, they become less

[11] As shown in the Appendix of Üslü (2019)'s paper, behind this risk-neutral-quadratic-utility trader is the assumption that they are always well-capitalized.

willing to buy the asset, and so are a worse intermediary by this definition. This effect is consistent with the empirical evidence from Adrian et al. (2017) and Bao et al. (2018) that shows that bonds traded by more constrained dealers are generally less liquid, with lower turnover and trade volumes.

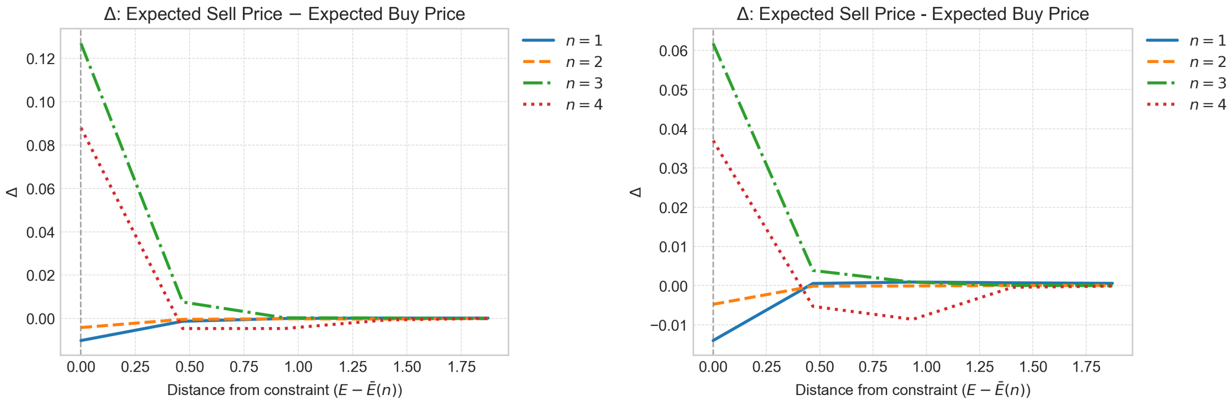
Consider now dealers with fewer assets, $n = 1$. They are weaker intermediaries than those with $n = 2$ assets because they are generally less willing to sell. However, as these dealers approach their regulatory limits, they become "better" intermediaries because they are more willing to sell to move away from the constraint.

4.1.1 No-constraint benchmark

To understand how the presence of the regulatory constraint affects market outcomes on average, I compare the baseline economy in column (2) of Table 4 with the economy without a regulatory constraint, column (1). That is, all agents are only subject to a simple non-negative equity condition.

Using the equilibrium distribution of agents, I obtain summary measures of trading costs for each market, captured by spreads. First, the introduction of the regulatory constraint increases average spreads and price dispersion. On average, investor spreads, as calculated using equation (21), rise from 5.760 to 6.048. Because the minimum regulatory level of equity $\bar{E}(n)$ has shifted upward for positive OTC inventories $n > 0$, dealers now spend more time in the region near the constraint where trading is most costly for investors. There is also more price dispersion: the standard deviation of investor trade prices rises from 0.00159 to 0.00167.

There is also higher turnover in the inter-dealer market when the regulatory constraint is in place, compared to a simple non-negative equity requirement. Since the regulatory requirement is stricter, dealers trade more actively to re-balance their portfolios to maintain enough regulatory slack. As a result, assets change hands more frequently, leading to higher turnover. At the same time, dealers hold smaller inventories overall, as maintaining their previous OTC inventory positions has become more costly. Furthermore, the OTC asset positions of high- and low-liquidity investors are less extreme, since it is now more expensive for them to reach their preferred allocations.

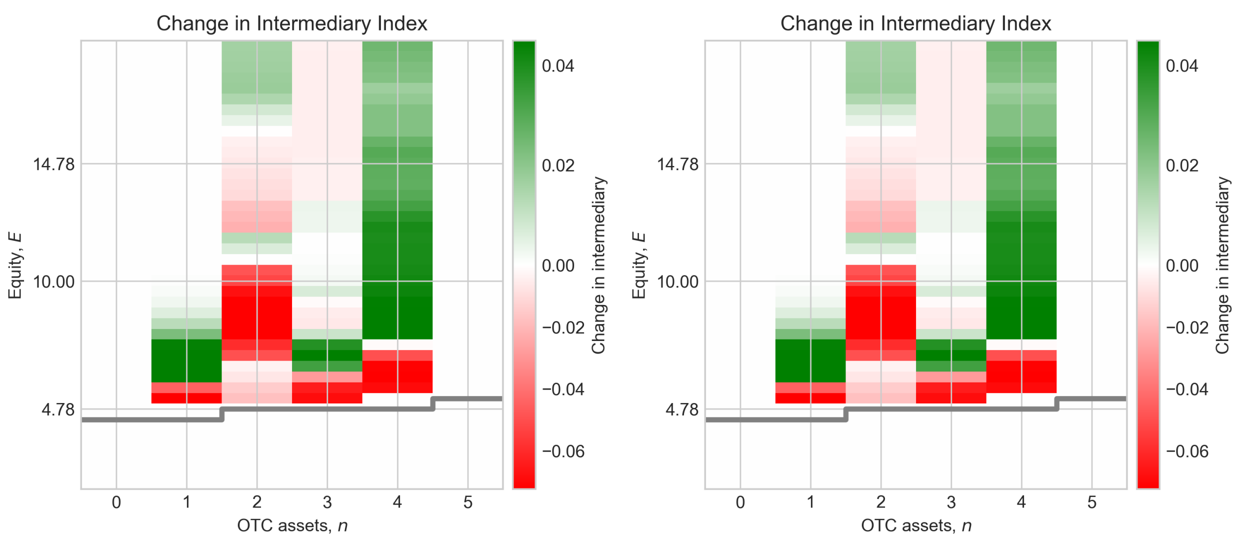


(a) Change in the average return on purchases of the OTC assets, Inter-dealer market

(b) Change in the average return on purchases of the OTC assets, Investor market

Figure 6: Change in the average return on purchases of the OTC assets following an increase in the risk weight $\nu\%$

In each graph, on the x-axis is the distance-to-constraint, $E - \bar{E}(n)$ of a dealer of state (E, n) with equity E and inventories n; 0 indicates that the dealer is just at the regulatory minimum level of equity, $E = \bar{E}(n)$. On the y-axis is the dollar value of the variable plotted. Plotted is the difference in *Expected Sell Price-Expected Buy Price* from Figure 3 for the same distance-to-constraint following stricter regulation.



(a) Inter-dealer market

(b) Investor market

Figure 7: Change in Intermediation Index by market

Each heatmap represents the change in the intermediation index for the inter-dealer and investor markets from Figure 5. On the x-axis are the dealer's OTC asset holdings n, and on the y-axis is equity E. The solid gray step function denotes the regulatory minimum $\bar{E}(n)$. Green indicates a positive change; red a negative change; darker colors indicate larger magnitudes.

5 Stricter Regulation and OTC trade

Now consider the counterfactual of stricter regulation in which the capital requirement, Λ, increases from 4% to 8.5%. Column (3) of Table 4 reports the corresponding model statistics for this new steady state, while column (4) shows the percentage change relative to the low capital requirement economy, column (2). Compared with the benchmark case, dealers must hold a higher minimum level of equity, and they spend more time, on average, in the region near their constraint. This section highlights three important consequences for OTC market outcomes.

First, stricter regulation leads to higher investor spreads, both in the cross-section and on average, given the new equilibrium distribution of agents. Figure 6 presents the change in dealer expected spreads as a function of their distance-to-constraint for both the inter-dealer and investor markets. Specifically, I compare expected spreads (the difference between the dealer's expected sell and buy prices) in both regulatory cases ($\Lambda = 4\%, \Lambda = 8.5\%$) for a fixed amount of regulatory slack. As shown, expected spreads rise as dealers become more constrained in both markets. Consequently, average investor spreads rise, as does spread variation. These patterns are consistent with the higher observed average costs for trades facilitated by bank-affiliated dealers post-GFC, as documented by Adrian et al. (2017), Breckenfelder and Ivashina (2021b), and Choi et al. (2024).

Second, stricter regulation results in a decline in average dealer inventories, mirroring what has been observed post-GFC (Bessembinder et al., 2018; Dick-Nielsen & Rossi, 2019; Goldstein & Hotchkiss, 2020). This occurs for two reasons. Near the constraint, the marginal value of the OTC asset falls while the marginal value of equity rises, which reduces dealers' willingness to hold risky OTC assets. Moreover, since the minimum equity requirement is now higher, dealers spend more time near their regulatory constraint where their target inventories are lower. Consequently, dealers have, on average, fewer OTC inventories. The model also adds nuance regarding investor behavior: since trade is more costly, investors take less extreme positions. Because of lower demand, the reference price $\tilde{p}$ falls.

Third, stricter regulation reduces dealer intermediation in the investor market, thereby lowering investor turnover. There is a decline in the average intermediation index and turnover in the investor market. Again, this is consistent with empirical evidence showing that bank-affiliated dealers deploy less capital and provide less liquidity due to tighter post-crisis regulations (Adrian et al., 2017; Bao et al., 2018; Bessembinder et al., 2018; Dick-Nielsen & Rossi, 2019). Figure 7 shows how the intermediation index changes following an increase in the capital requirement Λ across dealer states (E, n) in both markets: green areas indicate an increase in the intermediation index, while red areas indicate a decline. Compared to Figure 5, Figure 7 shows that dealer states that were previously good intermediaries before the regulatory change, facilitating both high buy and sell volumes, are now closer to their regulatory constraints since equity minimums have increased. As a consequence, they are less willing to buy and therefore act as worse intermediaries in both markets.

At the same time, however, dealers rely more on the inter-dealer market to reallocate inventories among themselves: stricter regulation results in a smaller ratio of inter-dealer to investor trades, higher inter-dealer turnover, and higher inter-dealer intermediation on average. Dealers tap the inter-dealer market more often to facilitate fewer investor trades. Because dealers now have, on average, less regulatory slack, their own trading needs increase for risk-management purposes. In equilibrium, this shift generates a reallocation of intermediation activity: a rise in inter-dealer market trade, intermediation, and turnover, but a decline in these same values in the investor market. This dynamic, absent in frameworks that impose ad-hoc trading costs or assume centralized inter-dealer markets, provides a novel explanation for the post-GFC decline in investor liquidity attributable to regulatory changes.

In summary, the model captures how tighter regulation reshapes OTC market dynamics in three key ways that not only reflect the empirical evidence but also highlight important nuances that ad-hoc modeling approaches fail to capture. With stricter regulation via higher capital requirements, investor spreads rise both on average and in the cross-section when dealers are constrained and have large OTC inventories; dealers hold smaller inventories and investors take less extreme positions; and intermediation quality in the investor market declines as dealers increasingly turn to the inter-dealer market to manage their regulatory compliance.

6 Conclusion

In this paper, I ask how post-GFC regulatory changes in Basel risk-weighted capital requirements have affected dealer market-making behavior in OTC markets. To answer this question, I develop a model of a two-tiered dealer-intermediated OTC market in which risk-averse dealers choose portfolios of risk-free liquid assets and risky OTC traded assets subject to a risk-based capital requirement. This regulatory constraint alters the marginal valuation of the risky OTC traded asset for dealers near their regulatory bound on equity. I solve the model numerically using an original iterative algorithm based on the finite difference methods proposed by Achdou et al. (2022). The model shows that dealers alter their trading behavior near the regulatory bound, charging higher transaction costs when they have less regulatory slack. Following an increase in the capital requirement, the distortions introduced by this constraint raise investor trading costs and reduce intermediation quality on average. Dealers lean away from facilitating trades for investors, trading more often in the more liquid inter-dealer market to manage their distance from their regulatory constraints.

To understand how these changes in dealer market-making behavior affect market resilience, I am currently working on obtaining the transition dynamics (impulse response functions) in response to shocks to (1) dealer capitalization and (2) investor liquidity preferences in the low and high regulatory environments. Will better capitalized dealers provide more or less liquidity in times of stress? Because inventory stickiness shapes the cost of holding assets, I am also extending this steady state analysis to examine how outcomes vary with search frictions. Are regulatory costs larger in less liquid markets?

This paper raises questions about the broader implications of post-GFC regulatory changes for both financial markets and the real economy. Stricter capital requirements (via risk-weighted asset ratios) can affect liquidity in financial markets and, in turn, growth, investment, and other real outcomes. This paper also raises the question of what an optimal policy balancing reduced risk in the banking sector against potentially reduced liquidity and resilience in decentralized financial markets might look like.

References

Achdou, Y., Han, J., Lasry, J. M., Lions, P. L., & Moll, B. (2022). Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. *Review of Economic Studies*, *89*(1), 45–86. https://doi.org/10.1093/restud/rdab002

Adrian, T., Boyarchenko, N., & Shachar, O. (2017). Dealer balance sheets and bond liquidity provision [Publisher: Elsevier B.V.]. *Journal of Monetary Economics*, *89*, 92–109. https://doi.org/10.1016/j.jmoneco.2017.03.011

Adrian, T., Capponi, A., Fleming, M., Vogt, E., & Zhang, H. (2020). Intraday market making with overnight inventory costs [Publisher: Elsevier B.V.]. *Journal of Financial Markets*, *50*(40), 100564. https://doi.org/10.1016/j.finmar.2020.100564

Adrian, T., Kiff, J., & Shin, H. S. (2018). Liquidity, leverage, and regulation 10 years after the global financial crisis. *Annual Review of Financial Economics*, *10*, 1–24. https://doi.org/10.1146/annurev-financial-110217-023113

Amihud, Y., & Mendelson, H. (1980). Dealership Market: Market-Making with Inventory. *Journal of Financial Economics*, *8*, 31–53.

An, Y., & Zheng, Z. (2023). Immediacy Provision and Matchmaking [Publisher: INFORMS Inst.for Operations Res.and the Management Sciences]. *Management Science*, *69*(2), 1245–1263. https://doi.org/10.1287/mnsc.2022.4355

Anderson, C. S., McArthur, D. C., & Wang, K. (2023). Internal risk limits of dealers and corporate bond market making [Publisher: Elsevier B.V.]. *Journal of Banking and Finance*, *147*, 106653. https://doi.org/10.1016/j.jbankfin.2022.106653

Bao, J., O'Hara, M., & Zhou, X. (2018). The Volcker Rule and corporate bond market making in times of stress [Publisher: Elsevier B.V.]. *Journal of Financial Economics*, *130*(1), 95–113. https://doi.org/10.1016/j.jfineco.2018.06.001

Bessembinder, H., Jacobsen, S., Maxwell, W., & Venkataraman, K. (2018). Capital Commitment and Illiquidity in Corporate Bonds. *Journal of Finance*, *73*(4), 1615–1661. https://doi.org/10.1111/jofi.12694

Bethune, Z., Sultanum, B., & Trachter, N. (2022). An Information-based Theory of Financial Intermediation. *The Review of Economic Studies*, *89*(5), 2381–2444. https://doi.org/10.1093/restud/rdab092

Breckenfelder, J., & Ivashina, V. (2021a). *Bank Balance Sheet Constraints and Bond Liquidity* (tech. rep.) (Publication Title: ECB Working Paper Issue: 2589). European Central Bank. https://doi.org/10.2139/ssrn.3929823

Breckenfelder, J., & Ivashina, V. (2021b). *Bank leverage constraints and bond market illiquidity during the COVID-19 crisis* (tech. rep.) (Publication Title: Research Bulletin Volume: 89). ECB. https://www.ecb.europa.eu/pub/economic-research/resbull/2021/html/ecb.rb211124~d9e3f578d2.en.html

Carrasco, J. A., & Harrison, R. (2023). Costly multi-unit search [Publisher: Elsevier B.V.]. *European Economic Review*, *154*(March), 104432. https://doi.org/10.1016/j.euroecorev.2023.104432

Carrasco, J. A., & Smith, L. (2017). Search at the margin. *American Economic Review*, *107*(10), 3146–3181. https://doi.org/10.1257/aer.20151579

Cetorelli, N., & Stern, S. (2015). *Same Name, New Businesses : Evolution in the Bank Holding Company* (tech. rep.).

Choi, J., Huh, Y., & Seunghun Shin, S. (2024). Customer Liquidity Provision: Implications for Corporate Bond Transaction Costs. *Management Science*, *70*(1), 187–206. https://doi.org/10.1287/mnsc.2022.4646

Cimon, D., & Garriott, C. (2019). Banking regulation and market making [Publisher: Elsevier B.V.]. *Journal of Banking and Finance*, *109*, 105653. https://doi.org/10.1016/j.jbankfin.2019.105653

Cohen, A., Kargar, M., Lester, B., & Weill, P.-O. (2024). Inventory, market making, and liquidity in OTC markets. *Journal of Economic Theory*, *222*, 105917. https://doi.org/10.1016/j.jet.2024.105917

Dick-Nielsen, J., & Rossi, M. (2019). The cost of immediacy for corporate bonds. *Review of Financial Studies*, *32*(1), 1–41. https://doi.org/10.1093/rfs/hhy080

Donaldson, J. R., Piacentino, G., & Thakor, A. (2018). Warehouse banking. *Journal of Financial Economics*, *129*(2), 250–267. https://doi.org/10.1016/j.jfineco.2018.04.011

Duffie, D. (2023). *Resilience redux in the US Treasury market* (tech. rep.) (Issue: 2022). Jackson Hole.

Duffie, D., Fleming, M., Keane, F., Nelson, C., Shachar, O., & Van Tassel, P. (2023). *Dealer capacity and US Treasury market functionality* (tech. rep.) (Publication Title: Staff Report). Federal Reserve Bank of New York.

Duffie, D., Gârleanu, N., & Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica*, *73*(6), 1815–1847. https://doi.org/10.1111/j.1468-0262.2005.00639.x

Duffie, D., Gârleanu, N., & Pedersen, L. H. (2007). Valuation in Over-the-Counter Markets. *Review of Financial Studies*, *20*(6), 1865–1900.

Farboodi, M., Jarosch, G., & Shimer, R. (2023). The emergence of market structure. *The Review of Economic Studies*, *90*(1), 261–292.

Fleskes, N. (2025). Liquidity Crises and Endogenous Dealer Capacity in Over-the-counter Markets.

Gârleanu, N. (2009). Portfolio choice and pricing in illiquid markets [Publisher: Elsevier Inc.]. *Journal of Economic Theory*, *144*(2), 532–564. https://doi.org/10.1016/j.jet.2008.07.006

Garman, M. B. (1976). Market microstructure. *Journal of Financial Economics*, *3*(3), 257–275. https://doi.org/10.1016/0304-405X(76)90006-4

Goldstein, M. A., & Hotchkiss, E. S. (2020). Providing liquidity in an illiquid market: Dealer behavior in US corporate bonds [Publisher: Elsevier B.V.]. *Journal of Financial Economics*, *135*(1), 16–40. https://doi.org/10.1016/j.jfineco.2019.05.014

Gong, G. X., & Wright, R. (2024). Middlemen in Search Equilibrium with Intensive and Extensive Margins. *International Economic Review*, *65*(4), 1657–1679. https://doi.org/10.1111/iere.12721

Gu, C., Wang, L., & Wright, R. (2024). Intermediaries, Inventories And Endogenous Dynamics In Frictional Markets. https://doi.org/10.2139/ssrn.4716108

Ho, T., & Stoll, H. R. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial Economics*, *9*(1), 47–73. https://doi.org/10.1016/0304-405X(81)90020-9

Ho, T. S. Y., & Stoll, H. R. (1983). The Dynamics of Dealer Markets Under Competition. *The Journal of Finance*, *38*(4), 1053–1074. https://doi.org/10.1111/j.1540-6261.1983.tb02282.x

Hugonnier, J., Lester, B., & Weill, P.-O. (2020). Frictional Intermediation in Over-The-Counter Markets. *Review of Economic Studies*, *87*(3), 1432–1469. https://doi.org/10.1093/restud/rdz037

Hugonnier, J., Lester, B. R., & Weill, P.-O. (2021). Heterogeneity in Decentralized Asset Markets. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2646096

Jarosch, G., Menzio, G., & Farboodi, M. (2016). Tough Middlemen.

Johri, A., & Leach, J. (2002). Middlemen and the allocation of heterogeneous goods. *International Economic Review*, *43*(2), 347–361. https://doi.org/10.1111/1468-2354.t01-1-00018

Kalai, E. (1977). Proportional solutions to bargaining situations: Interpersonal utility comparisons. *Econometrica: Journal of the Econometric Society*, 1623–1630.

Kargar, M., Passadore, J., & Silva, D. (2023). *Liquidity and Risk in OTC Markets: A Theory of Asset Pricing and Portfolio Flows* (tech. rep.). https://ssrn.com/abstract=3731019

Lagos, R., & Rocheteau, G. (2009). Liquidity in Asset Markets With Search Frictions. *Econometrica*, *77*(2), 403–426. https://doi.org/10.3982/ecta7250

Masters, A. (2007). Middlemen in search equilibrium. *International Economic Review*, *48*(1), 343–362. https://doi.org/10.1111/j.1468-2354.2007.00428.x

Nosal, E., Wong, Y. Y., & Wright, R. (2019). Intermediation in markets for goods and markets for assets [Publisher: Elsevier Inc.]. *Journal of Economic Theory*, *183*, 876–906. https://doi.org/10.1016/j.jet.2019.07.001

Nosal, E., Wong, Y.-Y., & Wright, R. (2015). More on Middlemen: Equilibrium Entry and Efficiency in Intermediated Markets. *Journal of Money, Credit and Banking*, *47*(S2), 7–37. https://doi.org/10.1111/jmcb.12213

Praz, R. (2014). Equilibrium Asset Pricing with Both Liquid and Illiquid Markets. *SSRN Electronic Journal*, (312417). https://doi.org/10.2139/ssrn.2464421

Rocheteau, G., & Weill, P.-O. (2011). Liquidity in frictional asset markets. *Journal of Money, Credit and Banking*, *43*(SUPPL. 2), 261–282. https://doi.org/10.1111/j.1538-4616.2011.00435.x

Rubinstein, A., & Wolinsky, A. (1987). Middlemen. *Quarterly Journal of Economics*, *102*(3), 581–593. https://doi.org/10.2307/1884218

Saar, G., Sun, J., Yang, R., & Zhu, H. (2019). From Market Making to Matchmaking: Does Bank Regulation Harm Market Liquidity? *SSRN Electronic Journal*, (May). https://doi.org/10.2139/ssrn.3399063

Shevchenko, A. (2004). Middlemen. *International Economic Review*, *45*(1), 695–722. http://onlinelibrary.wiley.com/doi/10.1111/j.1468-2354.2009.00532.x/full

Üslü, S. (2019). Pricing and Liquidity in Decentralized Asset Markets. *Econometrica*, *87*(6), 2079–2140. https://doi.org/10.3982/ecta14713

Vayanos, D., & Weill, P.-O. (2008). A Search-Based Theory of the On-the-Run. *The Journal of Finance*, *63*(3), 1361–1398. https://doi.org/10.1111/j.1540-6261.2008.01360.x

Wang, X., & Zhong, Z. ((2022). Post-crisis regulations, market making, and liquidity in over-the-counter markets [Publisher: Elsevier B.V.]. *Journal of Banking and Finance*, *134*. https://doi.org/10.1016/j.jbankfin.2021.106354

Watanabe, M. (2010). A model of merchants [Publisher: Elsevier Inc.]. *Journal of Economic Theory*, *145*(5), 1865–1889. https://doi.org/10.1016/j.jet.2010.02.011

Watanabe, M. (2020). Middlemen: A directed search equilibrium approach. *B.E. Journal of Macroeconomics*, *20*(2). https://doi.org/10.1515/bejm-2019-0258

Weill, P.-O. (2007). Leaning against the wind. *Review of Economic Studies*, *74*(4), 1329–1354. https://doi.org/10.1111/j.1467-937X.2007.00451.x

Wright, R., & Wong, Y.-Y. (2014). Buyers, Sellers, and Middlemen: Variations on Search-Theoretic Themes. *International Economic Review*, *55*(2), 375–397. https://doi.org/10.1111/iere.12053

Yang, M., & Zeng, Y. (2021). Coordination and Fragility in Liquidity Provision.

Appendix

A Derivations & proofs

A.1 Derivation of the regulatory VaR constraint & the implied maximum consumption $c^{VaR}(E,n)$

The regulatory VaR constraint is defined as in equation (6), repeated here for completeness:

$$VaR_\epsilon(R_{\Delta t}(n)) \geq 0.$$

The evolution of a dealer's equity given a fixed stock of OTC inventories n can be written as a linear stochastic differential equation:

$$dE_t = (rE_t - k)\,dt + \sigma^T dW_t.$$

where

$$k := -\left(\mu_\eta - c + n\left(\mu_D - r\tilde{p}\right)\right); \quad \sigma = \begin{bmatrix} \sigma_\eta \\ n\sigma_D \end{bmatrix}; \quad dW_t = \begin{bmatrix} dZ_{\eta,t} \\ dZ_{D,t} \end{bmatrix}.$$

Multiplying each side by e^{-rt} and integrating over $t + \Delta t$, I obtain:

$$E_{t+\Delta t} = e^{r\Delta t}E_t - k\frac{1}{r}\left(e^{r\Delta t} - 1\right) + e^{r\Delta t}\sigma^\top \int_0^{t+\Delta t} e^{-r(s-t)}dW_s.$$

Since W is a two dimensional Brownian motion, $E_{t+\Delta t}$ conditional on E_t is normally distributed:

$$E_{t+\Delta t}|E_t \sim N\left(\mu_{t+\Delta t}, \sigma^2_{t+\Delta t}\right).$$

Recall that the VaR is that value q such that

$$VaR_\epsilon(E_{t+\Delta t} - \bar{E}(n)) = \inf\{q \in \mathcal{R} : Pr\left(E_{t+\Delta t} - \bar{E}(n) \leq q\right) \geq \epsilon\}.$$

Since for a fixed stock of inventories n, $\bar{E}(n)$ is a constant,

$$E_{t+\Delta t} - \bar{E}(n)|E_t \sim N\left(\mu_{t+\Delta t} - \bar{E}(n), \sigma^2_{t+\Delta t}\right).$$

Therefore,

$$Pr\left[E_{t+\Delta t} - \bar{E}(n) < q\right] = \Phi\left(\frac{q - \left(\mu_{t+\Delta t} - \bar{E}(n)\right)}{\sigma_{t+\Delta t}}\right).$$

The VaR q is therefore

$$q = \mu_{t+\Delta t} - \bar{E}(n) + \sigma_{t+\Delta t}\Phi^{-1}(\epsilon).$$

The VaR requirement is this value, q, is non-negative:

$$VaR(E_{t+\Delta t} - \bar{E}(n)) = q \geq 0.$$

Setting the above value of $q \geq 0$ and solving for c provides the solution for $c^{VaR}(E, n)$ as stated in equation (16).

A.2 Proof of Proposition 1

Define for regulatory slack $s \geq 0$ the corresponding level of equity given a stock of OTC inventories $n \geq 0$: $E_s(n) := \bar{E}(n) + s$. Assume that the reference price is such that $\mu_D - r\tilde{p}$, if positive, is sufficiently small. In the neighborhood of the regulatory constraint:

1. *The marginal value of equity increases as a dealer becomes more constrained (regulatory slack s declines), and this increase is steeper the larger the dealer's OTC inventories (for $n' > n$):*

$$\partial_E W(E_s(n'), n') > \partial_E W(E_s(n), n) > 0.$$

2. *For trade size $a = 1$ and trade price P such that regulatory slack s is constant, the reservation value is:*

$$\Delta W(E_s, n) = W(E_s(n+1), n+1) - W(E_s(n), n).$$

   The reservation value of an additional unit of OTC assets in a fixed-slack trade declines as a dealer becomes more constrained (regulatory slack s declines), and this decline is steeper the larger the dealer's OTC inventories:

$$\partial_E \Delta W(E_s, n) > 0.$$

Sketch of proof: By optimal consumption and CARA utility, it follows that the value function is increasing and concave in equity. I first show that at the constraint point with zero slack, $E_s(n) = \bar{E}(n)$, the marginal value of equity is higher the more OTC assets a dealer has. For a given stock of inventories n, the boundary condition (15) is such that at $\bar{E}(n)$,

$$\partial_E W(\bar{E}(n), n) = u'\left(c^{VaR}(\bar{E}(n), n)\right).$$

By definition of $c^{VaR}(E, n)$, equation (14), the partial derivative with fixed slack is:

$$
\begin{aligned}
\partial_n c^{VaR}(E,n)|_{slack} =& \mu_D - r\tilde{p} \\
& + \frac{r}{e^{r\Delta t}-1}\Phi^{-1}(\epsilon)\frac{1}{2}\left[\frac{1}{2}\frac{1}{r}\left(\sigma_\eta^2 + n^2\sigma_D^2 + 2\rho n\sigma_\eta\sigma_D\right)\left(e^{2r\Delta t}-1\right)\right]^{-\frac{1}{2}} \\
& \times\left[\frac{1}{2}\frac{1}{r}\left(2n\sigma_D^2 + 2\rho\sigma_\eta\sigma_D\right)\left(e^{2r\Delta t}-1\right)\right].
\end{aligned}
$$

Since ϵ is a small positive, $\Phi^{-1}(\epsilon)$ is negative. By assumption, the reference price is such that $\mu_D - r\tilde{p}$, if positive, is sufficiently small. This is a reasonable assumption as for common values of ϵ, say 0.0001, the magnitude of $\Phi^{-1}(\epsilon)$ is quite large, -3.719. Therefore it is that:

$$
\partial_n c^{VaR}(E,n)|_{slack} < 0.
$$

Therefore for $n' > n$ at the constraint point:

$$
c^{VaR}(\bar{E}(n),n) > c^{VaR}(\bar{E}(n'),n')
$$

Because CARA utility is increasing and concave $u'(c) > 0, u''(c) < 0$, the above implies

$$
u'\left(c^{VaR}(\bar{E}(n),n)\right) < u'\left(c^{VaR}(\bar{E}(n'),n')\right).
$$

By the boundary condition (15):

$$
\partial_E W(\bar{E}(n'),n') > \partial_E W(\bar{E}(n),n)
$$

Combining the facts that (1) the marginal value of equity is larger the more assets a dealer has at the constraint point and (2) the value function is concave in equity (and is also continuous & twice differentiable), then for some small slack $s > 0$ (a small region near the constraint):

$$
\partial_E W(E_s(n'),n') > \partial_E W(E_s(n),n).
$$

For trade size $a = 1$ and trade price P such that regulatory slack is constant given $\tilde{p}$, the reservation value is:

$$
\Delta W(E_s,n) = W(E_s(n+1),n+1) - W(E_s(n),n).
$$

By the above, the partial derivative with respect to equity of a fixed-slack trade is therefore positive:

$$\partial_E \Delta W(E_s, n) = \partial_E W(E_s(n+a), n+a) - \partial_E W(E_s(n), n) > 0.$$

B Stationary Equilibrium

B.1 Equilibrium distributions of agent types

Define the probability that, post inter-dealer trade, a dealer of state (E', n') becomes state (E, n) to be $\alpha^c_{E,n,E',n'}$. Likewise for investor trade, $\alpha^c_{E,n,E',n'}$. The steady state distribution of dealers $F(E, n)$ with probability density function $f(E, n)$ satisfies the following Kolmogorov forward (Fokker-Planck) equation:

$$\begin{aligned}
0 = &- \partial_E \left\{ (\mu_\eta - c^*(E,n) + rE + n(\mu_D - r\tilde{p}))\, f(E,n) \right\} \\
&+ \frac{1}{2}\partial_E^2 \left\{ \left(\sigma_\eta^2 + \sigma_D^2 n^2 + 2n\rho\sigma_\eta\sigma_D\right) f(E,n) \right\} \\
&+ \lambda_d \left(\int \alpha^d_{E,n,E',n'} f(E',n') d(E',n') \right) \\
&- \lambda_d f(E,n) \left(\int \alpha^d_{E',n',E,n} d(E',n') \right) \\
&+ \lambda_c \left(\int \alpha^c_{E,n,E',n'} f(E',n') d(E',n') \right) \\
&- \lambda_c f(E,n) \left(\int \alpha^c_{E',n',E,n} d(E',n') \right).
\end{aligned} \tag{23}$$

The first two terms capture the changes in the distribution due to the stochastic diffusion processes on the endowment and the OTC asset dividend. The last four lines show the inflows and outflows of agents due to trade in the OTC asset in both the inter-dealer and investor markets.

Similarly, the steady state conditional distribution of investors $H(E, n, i)$ satisfies the following:

$$\begin{aligned}
0 = &- \partial_E \left\{ (\mu_\eta - c^*(E,n,i) + rE + n(\mu_D - r\tilde{p}))\, h(E,n,i) \right\} \\
&+ \frac{1}{2}\partial_E^2 \left\{ \left(\sigma_\eta^2 + \sigma_D^2 n^2 + 2n\rho^i\sigma_\eta\sigma_D\right) h(E,n,i) \right\} \\
&+ \gamma\pi(i)h(E,n,i') - \gamma\pi(i')h(E,n,i) \\
&+ \lambda_c \left(\int \alpha_{E,n,i,E',n',i} h(E',n',i) d(E',n',i) \right) \\
&- \lambda_c h(E,n,i) \left(\int \alpha_{E',n',i,E,n,i} d(E',n',i) \right).
\end{aligned} \tag{24}$$

Where $\alpha_{E,n,i,E',n',i}$ is the probability that post-trade an investor of state (E', n', i) becomes state

(E, n, i). As with dealers, the first two lines reflect changes in the distribution due to stochastic income. The third line reflects the change in the distribution due to the changes in the investor's liquidity preference type i. The last two lines are the inflows and outflows of investors due to trade.

B.2 The reference price of the OTC asset $\tilde{p}$

The reference price of the OTC asset $\tilde{p}$ is the volume-weighted average of all the realized trade prices in the economy. Denote the volume-weight of a trade between two dealers, conditional on receiving a trade opportunity and trading a non-zero amount of the asset to be $w_{trade}(E, n, E', n')$. Likewise for dealers and investors, $w_{trade}(E, n, E', n', i)$. Let the weights sum to 1. The weighted reference price $\tilde{p}$ is therefore:

$$\begin{aligned}\tilde{p} = &\int\int \left(w_{trade}(E, n, E', n') \times P^*(E, n, E', n')\right) dF(E', n')dF(E, n)\\ &+ \int\int \left(w_{trade}(E, n, E', n', i) \times P^*(E, n, E', n', i)\right) dH(E', n', i)dF(E, n).\end{aligned} \tag{25}$$

B.3 Market clearing

Finally, market clearing requires that all asset shares $S > 0$ be held by agents. With two unit measures of agents, it must be that the steady state weighted average of asset holdings of all agents equals the supply in shares:

$$p_D \int ndF(E, n) + (1 - p_D) \int ndH(E, n, i) = S. \tag{26}$$

B.4 Stationary equilibrium definition

A stationary equilibrium is a set of value functions $W(E, n), V(E, n, i)$, consumption policies $c^*(E, n), c^*(E, n, i)$, distributions $f(E, n)$,$h(E, n, i)$, trading rules (a^*, P^*) for the OTC asset within each market (inter-dealer, investor) between each potential pair of traders, and a reference price $\tilde{p}$ such that, taking the model primitives $\{r, \lambda_d, \lambda_c, \mu_\eta, \sigma_\eta, \mu_D, \sigma_D, \kappa, \rho, \rho^i, \gamma, \pi(i), \Lambda, \nu, p_D, \epsilon, \Delta t\}$ as given, the following are satisfied:

1. Value functions $W(E, n), V(E, n, i)$ satisfy (13), (17), (15) and (19);
2. Consumption policies $c^*(E, n), c^*(E, n, i)$ satisfy (16) and (18);
3. Trading rules (a^*, P^*) satisfy the bargaining protocol defined in (11) and (12);
4. Distributions $f(E, n)$,$h(E, n, i)$ satisfy (23) and (24);

5. The reference price $\tilde{p}$ is defined as in (25) and market clearing holds as in (26).

C Stationary equilibrium solution method

Solving the stationary equilibrium involves two main sub-problems. The first involves solving for the intensities with which agents change their type in the E and n dimensions due to optimal trade outcomes as a function of value functions and distributions. This step is solved by simply finding the grid point that best satisfies the Kalai (1977) bargaining problems for each possible trade, if indeed one exists.[12]

The second involves solving for the agent value functions (taking as given switching intensities due to trade), and then solving for the implied agent distributions. I solve for value functions using finite-difference methods as outlined in Achdou et al. (2022) with some modifications. For investors, I use the standard finite difference methodology. For dealers, however, due to the regulatory constraint, the boundary of the equity grid is staggered: dealers with more OTC asset holdings must hold higher equity. The modifications to the finite difference method I make can be found in the next section, Appendix C.1. I solve the steady state using the algorithm defined in Algorithm 1.

C.1 State-based boundary conditions and finite-difference methods

In this section I outline the finite-difference algorithm based on Achdou et al. (2022) modified to handle the state-based boundary condition for dealers as it appears in this model. Let there be I liquid asset grid points E_i with step size ΔE in the domain $[E_L, E_H]$ and J OTC asset holding points n_j with step size $\Delta n = 1$ in the domain $[0, n_H]$. Let $T_{ij,i'j'}$ be the trade outcome probabilities due to trade in the OTC market: $T_{ij,i'j'}$ represents the probability of switching from type ij to type $i'j'$ in the defined grid. Note that $\sum_{i',j'} T_{ij,i'j'} = 1$. Define the following:

$$s_{ij} = \mu_\eta - c_{ij} + rE_i + n_j(\mu_D - r\tilde{p});$$
$$\sigma_j^2 = \sigma_\eta^2 + \sigma_D^2 n_j^2 + 2n_j\rho^m\sigma_\eta\sigma_D.$$

Note the relevant forward and backward differences are:

$$\partial_{E,F} w_{ij} = \frac{w_{i+1,j} - w_{ij}}{\Delta E};$$

$$\partial_{E,B} w_{ij} = \frac{w_{ij} - w_{i-1,j}}{\Delta E};$$

[12] This is a bottleneck; with 70 liquid asset grid points and up to 5 units of the OTC asset held, obtaining the optimal trade outcomes for all agents takes about 60 seconds on a Windows laptop with a 12th Gen Intel Core i7 processor using Python 3.11 when using parallelization.

Table 5: Algorithm for Steady State Equilibrium

Algorithm 1: Steady State Equilibrium

Initial guess $\tilde{p}$; tolerances $\mathrm{tol}_1, \mathrm{tol}_2$; weights w_1, w_2; max iterations X, Y; agent grids on (e_k, n_j), (e_k, n_j, i). Converged reference price $\tilde{p}$, distributions f, h, and value functions W, V.

Start:;
Define minimum equity levels $k^*(n) = \bar{E}(n)$ given $\tilde{p}$.;
Solve for $\mathbb{W}^{(0)} = \{W^{(0)}, V^{(0)}\}$ and $\mathbb{D}^{(0)} = \{f^{(0)}, h^{(0)}\}$ without OTC trade using $k^*(n)$ (use finite differences; see appx. C.1).;
Solve for OTC trade $\mathbb{T}^{(0)}$ given $\mathbb{W}^{(0)}$ and $\mathbb{D}^{(0)}$.;
for $x \leftarrow 1$ **to** X **do**
    `// Inner loop: Value functions, OTC trade, and distributions`
    $\mathbb{W}_0^{(x)} \leftarrow \mathbb{W}^{(x-1)}$;
    $\mathbb{D}_0^{(x)} \leftarrow \mathbb{D}^{(x-1)}$;
    $\mathbb{T}_0^{(x)} \leftarrow \mathbb{T}^{(x-1)}$;
    for $y \leftarrow 1$ **to** Y **do**
        Compute optimal trade intensities $\mathbb{T}_y^{(x)}$ using $\mathbb{W}_{y-1}^{(x)}$ and $\mathbb{D}_{y-1}^{(x)}$.;
        Obtain $\mathbb{W}_y^{(x)}$.;
        Solve for $\mathbb{D}_y^{(x)}$ given $\mathbb{W}_y^{(x)}$.;
        $c_{11} \leftarrow \max\{|\mathbb{W}_y^{(x)} - \mathbb{W}_{y-1}^{(x)}|\}$;
        $c_{12} \leftarrow \max\{|\mathbb{T}_y^{(x)} - \mathbb{T}_{y-1}^{(x)}|\}$;
        $c_{13} \leftarrow \max\{|\mathbb{D}_y^{(x)} - \mathbb{D}_{y-1}^{(x)}|\}$;
        if $\max\{c_{11}, c_{12}, c_{13}\} < tol_1$ **then**
            $\mathbb{W}^{(x)} \leftarrow \mathbb{W}_y^{(x)}$;
            $\mathbb{T}^{(x)} \leftarrow \mathbb{T}_y^{(x)}$;
            $\mathbb{D}^{(x)} \leftarrow \mathbb{D}_y^{(x)}$;
            break;
        else
            $\mathbb{W}_y^{(x)} \leftarrow w_1\, \mathbb{W}_y^{(x)} + (1 - w_1)\, \mathbb{W}_{y-1}^{(x)}$;
            $\mathbb{T}_y^{(x)} \leftarrow w_1\, \mathbb{T}_y^{(x)} + (1 - w_1)\, \mathbb{T}_{y-1}^{(x)}$;
            $\mathbb{D}_y^{(x)} \leftarrow w_1\, \mathbb{D}_y^{(x)} + (1 - w_1)\, \mathbb{D}_{y-1}^{(x)}$;
    `// Outer loop: Reference price`
    Compute $\tilde{p}_x$ from $\{\mathbb{W}^{(x)}, \mathbb{D}^{(x)}, \mathbb{T}^{(x)}\}$.;
    $c_2 \leftarrow |\tilde{p}_x - \tilde{p}|$.;
    if $c_2 < tol_2$ **then**
        stop
    else
        $\tilde{p} \leftarrow w_2 \tilde{p}_x + (1 - w_2)\tilde{p}$ and update minimum equity levels $k^*(n) = \bar{E}(n)$ given new $\tilde{p}$

$$\partial_E^2 w_{ij} = \frac{w_{i,j+1} - 2w_{ij} + w_{i,j-1}}{(\Delta E)^2}.$$

The discretized Dealer HJB is w_{ij}, given a guess w_{ij}^n the next guess w_{ij}^{n+1} satisfies:

$$\begin{aligned}\frac{w_{ij}^{n+1} - w_{ij}^n}{\Delta} + r w_{ij}^{n+1} =& u(c_{ij}^n) + \partial_E w_{ij}^{n+1} s_{ij}^n + \frac{1}{2}\partial_E^2 w_{ij}^{n+1}\sigma_j^2 \\ &+ \lambda \sum_{i',j'} T_{ij,i'j'}\left(w_{i'j'}^{n+1} - w_{ij}^{n+1}\right).\end{aligned}$$

The implicit guess using the upwind scheme is a system of $I \times J$ equations:

$$\begin{aligned}\frac{w_{ij}^{n+1} - w_{ij}^n}{\Delta} + r w_{ij}^{n+1} =& u(c_{ij}^n) + \partial_{E,F} w_{ij}^{n+1}\left[s_{ij,F}^n\right]^+ + \partial_{E,B} w_{ij}^{n+1}\left[s_{ij,B}^n\right]^- + \frac{1}{2}\partial_E^2 w_{ij}^{n+1}\sigma_j^2 \\ &+ \lambda \sum_{i',j'} T_{ij,i'j'}\left(w_{i'j'}^{n+1} - w_{ij}^{n+1}\right).\end{aligned} \tag{27}$$

Note that optimal agent consumption is always:

$$c_{ij}^n = (u')^{-1}(\partial_E w_{i,j}^n).$$

Where the marginal value of equity is:

$$\begin{aligned}\partial_E w_{i,j} =& \partial_E w_{ij,F} 1\{s_{ij,F} > 0\} + \partial_E w_{ij,B} 1\{s_{ij,B} < 0\} \\ &+ u'(\mu_\eta + rEi + n_j(\mu_D - r\tilde{p})) 1\{s_{ij,F} \le 0 \le s_{ij,B}\}.\end{aligned}$$

Writing equation (27) in matrix notation and using the backwards and forward difference equations, I obtain the following:

$$\begin{aligned}\frac{w_{ij}^{n+1} - w_{ij}^n}{\Delta} + r w_{ij}^{n+1} =& u(c_{ij}^n) + \frac{w_{i+1,j}^{n+1} - w_{i,j}^{n+1}}{\Delta E}\left[s_{ij,F}^n\right]^+ + \frac{w_{i,j}^{n+1} - w_{i-1,j}^{n+1}}{\Delta E}\left[s_{ij,B}^n\right]^- \\ &+ \frac{w_{i+1,j}^{n+1} - 2w_{ij}^{n+1} + w_{i-1,j}^{n+1}}{(\Delta E)^2}\frac{\sigma_j^2}{2} \\ &+ \lambda \sum_{i',j'} T_{ij,i'j'}\left(w_{i'j'}^{n+1} - w_{ij}^{n+1}\right).\end{aligned}$$

Which can be written for generic grid points ij as:

$$\begin{aligned}\frac{w_{ij}^{n+1} - w_{ij}^n}{\Delta} + r w_{ij}^{n+1} =& u(c_{ij}^n) + z_{ij}^n w_{i+1,j}^{n+1} + y_{ij}^n w_{i,j}^{n+1} + x_{ij}^n w_{i-1,j}^{n+1} \\ &+ \lambda \sum_{i',j'} T_{ij,i'j'} w_{i'j'}^{n+1}.\end{aligned} \tag{28}$$

Where:

$$z_{ij}^{n} = \left(\frac{1}{\Delta E}\left[s_{ij,F}^{n}\right]^{+} + \frac{1}{2}\sigma_j^2\frac{1}{(\Delta E)^2}\right);$$

$$y_{ij}^{n} = \left(\frac{1}{\Delta E}\left[s_{ij,B}^{n}\right]^{-} - \frac{1}{\Delta E}\left[s_{ij,F}^{n}\right]^{+} - \sigma_j^2\frac{1}{(\Delta E)^2} - \lambda\right);$$

$$x_{ij}^{n} = \left(\frac{1}{2}\sigma_j^2\frac{1}{(\Delta E)^2} - \frac{1}{\Delta E}\left[s_{ij,B}^{n}\right]^{-}\right).$$

Up until now, this method has been consistent with Achdou et al. (2022). The modification lies in how I handle the boundary condition. As in Achdou et al. (2022), implementing the boundary condition requires a modification to the backwards difference of the value function at the boundary point. The difference here is that now the boundary point for dealers is no longer the bottom of the grid only, and I have multiple points which need to be constrained for all levels of equity that violate the constraint: $E \leq \bar{E}(n_j)$ for each n_j. Therefore, for intermediate OTC asset holdings $n_j \leq n_H$, there is a region of the grid where consumption is constrained.

I impose the boundary condition as follows. For all i, j such that $E_i < E(n_j)$, the boundary condition needs to be satisfied for the backward and forward difference. Identify the boundary point of equity, or the minimum level of equity for each j that satisfies the constraint, as $\bar{E}(n_{j^*})$ and the grid point as i^*j^*. The backward and forward differences for all $\hat{i} \in \{i < i^*\}$ must satisfy the boundary condition (15):

$$\partial_{E,B} w_{\hat{i}j^*} = u'(c_{\hat{i},j^*}^{VaR});$$

$$\partial_{E,F} w_{\hat{i}j^*} = u'(c_{\hat{i},j^*}^{VaR}).$$

Now, consider the point at which the regulatory constraint first binds, i^*j^*. The backward difference at point i^*j^* is:

$$\partial_{E,B} w_{i^*j^*} = \frac{w_{i^*j^*} - w_{i^*-1,j^*}}{\Delta E} = u'(c_{i^*,j^*}^{VaR}).$$

Therefore the value function at the point $i^* - 1, j^*$ in the equity grid is by definition:

$$w_{i^*-1,j^*} = w_{i^*j^*} - \Delta E u'(c_{i^*,j^*}^{VaR}).$$

Similarly, the value function for any point below the constraint point is:

$$w_{\hat{i}-1,j^*} = w_{\hat{i}j^*} - \Delta E u'(c_{\hat{i},j^*}^{VaR}).$$

Therefore equation (28) at the boundary point i^*, j^* is:

$$\begin{aligned}\frac{w_{i^*j^*}^{n+1} - w_{i^*j^*}^{n}}{\Delta} + rw_{i^*j^*}^{n+1} =& u(c_{i^*j^*}^{n}) - x_{i^*j^*}^{n}\Delta E u'(c_{i^*,j^*}^{VaR}) \\ &+ z_{i^*j^*}^{n} w_{i^*+1,j^*}^{n+1} + \left(x_{i^*j^*}^{n} + y_{i^*j^*}^{n}\right) w_{i^*j^*}^{n+1} \\ &+ \lambda \sum_{i',j'} T_{ij,i'j'} w_{i'j'}^{n+1}.\end{aligned}$$

And for any point below the constraint point, $\hat{i} \in \{i < i^*\}$, is:

$$\begin{aligned}\frac{w_{\hat{i}j^*}^{n+1} - w_{\hat{i}j^*}^{n}}{\Delta} + rw_{\hat{i}j^*}^{n+1} =& u(c_{\hat{i}j^*}^{n}) - x_{\hat{i}j^*}^{n}\Delta E u'(c_{\hat{i},j^*}^{VaR}) \\ &+ z_{\hat{i}j^*}^{n} w_{\hat{i}+1j^*}^{n+1} + \left(x_{\hat{i}j^*}^{n} + y_{\hat{i}j^*}^{n}\right) w_{\hat{i}j^*}^{n+1} \\ &+ \lambda \sum_{i',j'} T_{ij,i'j'} w_{i'j'}^{n+1}.\end{aligned}$$

The upper bound of the equity grid is handled in the standard manner as in Achdou et al. (2022) by requiring that the marginal value of wealth at the upper most equity grid point I for all OTC asset holdings n_j be zero:

$$\partial_E w(E_I, n_j) = 0 \implies w_{Ij} = w_{I+1,j}.$$

And so equation (28) at the upper boundary point Ij is:

$$\frac{1}{\Delta} w_{Ij}^{n+1} - \frac{1}{\Delta} w_{Ij}^{n} + rw_{Ij}^{n+1} = u(c_{Ij}^{n}) + x_{Ij}^{n} w_{I-1,j}^{n+1} + \left(y_{Ij}^{n} + z_{Ij}^{n}\right) w_{I,j}^{n+1}.$$

Note that such boundary constraints are not necessary on the OTC asset grid j as the constraints on the bargaining process prevent agents from trading in violation of feasibility or regulation.

Equation (28) can therefore be written for all valid points on the i, j grid as:

$$\begin{aligned}\frac{w_{ij}^{n+1} - w_{ij}^{n}}{\Delta} + rw_{ij}^{n+1} =& u(c_{ij}^{n}) + z_{ij}^{n} w_{i+1,j}^{n+1} + y_{ij}^{n} w_{i,j}^{n+1} + x_{ij}^{n} w_{i-1,j}^{n+1} \\ &+ \lambda \sum_{i',j'} T_{ij,i'j'} w_{i'j'}^{n+1}.\end{aligned} \tag{29}$$

Where:

$$z_{ij}^{n*} = \begin{cases} z_{ij}^{n} & \text{if } i < I, \\ 0 & \text{if } i = I; \end{cases}$$

$$y_{ij}^{n*} = \begin{cases} y_{ij}^n & \text{if } i < I \text{and } i > i_j^*, \\ y_{ij}^n + x_{ij}^n & \text{if } i < I \text{and } i \le i_j^*, \\ y_{ij}^n + z_{ij}^n & \text{if } i = I; \end{cases}$$

$$x_{ij}^{n*} = \begin{cases} x_{ij}^n & \text{if } i > i_j^*, \\ 0 & \text{if } i \le i_j^*. \end{cases}$$

Define A^n to be a sparse matrix with $I \times J$ columns and $I \times J$ rows with y_{ij}^{n*} on the center diagonal, x_{ij}^{n*} on the lower diagonal, and z_{ij}^{n*} on the upper diagonal. Define $u_{ij}^n = u(c_{ij}^n)$. Define w^n as a vector of length $I \times J$, with entries $(w_{11}, ..., w_{I1}, w_{12}..., w_{I2}, w_{1J}, ..., w_{IJ})$. Define b^n to be a vector of length $I \times J$ with entries:

$$b_{ij}^n = \begin{cases} u_{ij}^n - x_{ij}^n \Delta \omega u'(c_{i,j^*}^{VaR}) & \text{if } i \le i_j^*, \\ u_{ij}^n & \text{otherwise.} \end{cases}$$

Define T to be a matrix with $I \times J$ columns and $I \times J$ rows populated with the trading probabilities $T_{ij,i'j'}$. Therefore the difference equation can be written in matrix form as:

$$\underbrace{\frac{1}{\Delta}\left(w^{n+1} - w^n\right) + rw^{n+1}}_{\text{dim=}(I\times J)\times 1} = \underbrace{u^n}_{\text{dim=}(I\times J)\times 1} + \underbrace{A^n + \lambda T}_{\text{dim=}(I\times J)\times(I\times J)} \times \underbrace{w^{n+1}}_{\text{dim=}(I\times J)\times 1}.$$

Which can be solved following the standard iterative procedure outlined in Achdou et al. (2022).

D Estimation of stochastic processes

I use quarterly balance sheet level data from the FR-Y9C filings of bank holding companies (BHCs) to estimate the constants in the stochastic processes in equations (1) and (2), $\mu_\eta, \sigma_\eta, \mu_D$, and σ_D, via maximum likelihood. FR-Y9C filings are publicly available financial statements that financial institutions are required to prepare and submit for regulatory purposes. I filter the dataset to only include domestic US based bank holding companies with strictly positive levels of assets: the data spans 76 quarters from Q1 2006 to Q4 2024, and includes 1580 distinct BHC's. Figure (8) displays the count of BHC's in the sample over time. The sample composition changes in 2015 and 2018 due to increases in the size thresholds for BHC reporting requirements.

I define an agent's endowment income in the data, $d\hat{\eta}$, as a BHC's recorded consolidated net income less its income from trading assets, defined as the sum of total interest on its stock of trading assets and trading revenues. I define an agent's dividend income per unit of risky OTC asset in the

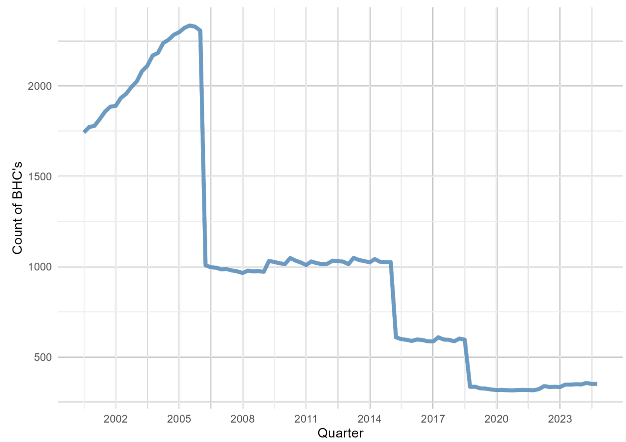


Figure 8: Number of bank holding companies (BHCs) in the sample over time

data, $d\hat{D}$, to be the sum of a BHC's total dividend income from trading assets divided by their stock of trading assets.

Since equity in this model spans $[0, 30]$, I need to scale $d\hat{\eta}$ appropriately before estimating the endowment process. I first normalize the line item for total equity, $\hat{E}$ between 0 and 30 after first winsorizing at the 1% level. Then I scale the endowment income $d\hat{\eta}$ proportionally using the ratio of normalized and original equity to obtain $d\hat{\eta}^*$. I scale dividend income from trading assets using this same scaling factor.

Since inventories in this model span $[0, 5]$, I repeat the scaling procedure above. I normalize total trading assets to between 0 and 5, after first winsorizing at the 1% level. I obtain the scaled per unit dividend $d\hat{D}^*$ by dividing scaled dividend income by scaled trading assets. Table 6 reports the summary statistics for the scaled variables of interest.

	All BHCs					**BHCs with Trading Assets**				
	Mean	*Std. Dev.*	*Min*	*Median*	*Max*	*Mean*	*Std. Dev.*	*Min*	*Median*	*Max*
Equity, $\hat{E}^*$	0.96	3.79	0.0037	0.0775	30.00	4.27	7.82	0.0037	0.7677	30.00
Endowment income, $d\hat{\eta}^*$	0.03	0.14	-0.3092	0.0033	1.07	0.11	0.29	-0.30917	0.02372	1.07
Dividend income, $d\hat{D}^*$	–	–	–	–	–	0.06435	0.16880	0.00000	0.006231	1.33
Total assets (log)	14.61	1.56	5.89	14.14	22.16	16.57	1.97	11.17	16.33	22.16
ROA (%)	5.78	723.15	-198.13	0.43	118,286.15	0.50	0.93	-21.76	0.46	19.37
No. Subsidiaries	1.44	1.77	1.00	1.00	57.00	1.79	2.48	1.00	1.00	40.00
Age (years)	2.45	3.00	0.000	2.00	33.00	2.41	2.61	0.000	2.00	26.00

Table 6: Summary Statistics

Variables with a * are normalized to match the equity and asset grids. Dividend income $d\hat{D}$ is reported only for BHCs with trading assets.

I now derive the log-likelihood function for the endowment process (the procedure is identical for the dividend process). First, I allow the drift and volatility of an agent's endowment and dividend to vary by key BHC characteristics: size (log of total assets), age (years), ROA (%), complexity (no. of subsidiaries), and whether the BHC ever holds trading assets. I also control for the GFC (years 2007-2009) and Covid-19 (years 2020 and 2021), in case there is exceptional volatility during these periods. Since the reporting requirements for the sample change over time, I include a control for the sample period (pre-2015, 2015-2017, and post-2017). Finally, since data is reported quarterly, I include an indicator variable for Q1 to help control for seasonality. Table 6 reports the summary statistics for these control variables. Let X_{it} denote a vector of all these controls, where all continuous variables are winsorized at the 3% level.

Since the unit of time in this model is one quarter and the data are observed quarterly, the time step is $\Delta t = 1$. The observed endowment income for BHC i at time t then follows:

$$d\eta_{it} = \mu_{it} + \varepsilon_{it}, \qquad \varepsilon_{it} \mid X_{it} \sim \mathcal{N}(0, \sigma_{it}^2 \Delta t),$$

where the drift μ_{it} is linear function of the controls:

$$\mu_{it} = m\Delta t + X_{it}^{\top}\beta.$$

I allow for heteroskedasticity in volatility, with the log-variance linear in the same controls:

$$\log \sigma_{it}^2 = a_0 + X_{it}^{\top}\gamma.$$

Therefore, conditional on X_{it}, the distribution of endowment income is:

$$d\eta_{it} \mid X_{it} \sim \mathcal{N}\left(m\Delta t + X_{it}^{\top}\beta,\ \exp(a_0 + X_{it}^{\top}\gamma)\Delta t\right).$$

The log-likelihood function for the full panel of observations is:

$$\ell(m, \beta, a_0, \gamma) = \sum_{t=1}^{T}\sum_{i=1}^{N_t}\left[-\frac{1}{2}\log(2\pi) - \frac{1}{2}\left(a_0 + X_{it}^{\top}\gamma + \log \Delta t\right) - \frac{\left(\eta_{it} - m\Delta t - X_{it}^{\top}\beta\right)^2}{2\exp\left(a_0 + X_{it}^{\top}\gamma\right)\Delta t}\right]. \tag{30}$$

I estimate the parameter vector (m, β, a_0, γ) by maximizing the log-likelihood in (30) using the observed scaled net income $d\hat{\eta}_{it}^*$ from the FR-Y9C filings for each BHC. The log-likelihood function for the dividend process is of a similar form, for which the observed scaled return on trading assets $d\hat{D}_{it}^*$ is used.

Using the delta method, I obtain estimates and standard errors of the drift and volatility of the endowment and dividend processes conditional on the following BHC characteristics and time periods: (1) BHCs that, at some point in the sample, hold trading assets; (2) the time period is pre-2015, non-crisis, non-Q1; and (3) BHCs that have an average age, size, complexity, and ROA. The MLE estimates at these values are presented in Table 2 in the body of the text.