WORKING PAPER

#675 | 2025

WP

Trading choices

LUCAS B. DYSKANT, ANDRÉ C. SILVA, BRUNO SULTANUM

NOVA SCHOOL OF BUSINESS & ECONOMICS

Trading choices*

Lucas B. Dyskant
Barclays Corporate & Investment Bank

André C. Silva
Nova SBE

Bruno Sultanum
University of Essex

May 2025

Abstract

We propose a model of over-the-counter markets based on three trading methods: principal inventory, agency risk-free, and all-to-all (A2A) trading. Principal and agency trading occur through dealers. A2A trading occurs directly through customer-customer trading. The model predicts that A2A size can remain stable while principal and agency trading change. Higher inventory costs shifts trading from principal to agency and decrease dealers' net positions. Bid-ask spreads can decrease even though transaction costs increase. High transaction costs can lead to multiple equilibria. The model shows how regulatory and technological changes affect trading choices, stability, and market indicators.

JEL classification: D53, G12, G18, G28.

Keywords: over-the-counter markets, intermediation costs, liquidity, corporate bond markets, financial market regulations, post-2008 regulations, Volcker rule.

May 19, 2025. Updates in this link.

*Silva: Nova School of Business and Economics, Universidade NOVA de Lisboa, Campus de Carcavelos, 2775-405 Carcavelos, Portugal, andre.silva@novasbe.pt. The views in this paper are those of the authors and not necessarily the views of Barclays Corporate & Investment Bank. Dyskant participated in this work in his personal capacity and not as part of his employment with Barclays. We benefited from discussions with Paco Buera, Briana Chang, Veronica Guerrieri, Zhiguo He, Yesol Huh, Ben Lester, Dmitry Orlov, Sebastien Plante, Erwan Quintin, Andreas Rapp, Philip Reny, Roberto Robatto, Sang Seo, Cathy Zhang and participants at seminars and conferences at the Wisconsin School of Business, Miami Business School, Purdue University, the IMF, the Board, the Philadelphia Fed, and the St. Louis Fed. Silva thanks the hospitality of the Kenneth C. Griffin Department of Economics at the University of Chicago, where part of this article was written. This work was partially funded by Fundação para a Ciência e a Tecnologia (UIDB/00124/2020, UIDP/00124/2020, UID/00124, Nova School of Business and Economics and Social Sciences DataLab - PINFRA/22209/2016), POR Lisboa and POR Norte (Social Sciences DataLab, PINFRA/22209/2016).

1 Introduction

We introduce a model of over-the-counter (OTC) markets that takes into account the choice among three trading methods: principal inventory trading, agency risk-free trading, and all-to-all (A2A) customer-customer trading. Prices and market size as well as search and inventory costs determine trading choices by customers and dealers. The model helps analyze how regulatory and technological shifts influence market structure, particularly OTC markets such as the corporate bond market, where A2A, principal, and agency trading frequencies have changed in the recent years.

Our model builds on search models of OTC markets following Duffie, Gârleanu, and Pedersen (2005) (DGP). Customers either search for dealers, who intermediate principal and agency trading, or search directly for other customers in A2A trading. Dealers face two different intermediation costs: inventory costs when providing on-the-spot liquidity in principal trades, and brokerage costs when facilitating trading by matching customers in agency trades. Customers can avoid these costs by trading with other customers in A2A markets. We assume that absent brokerage and inventory costs, principal trading offers the most effective search technology, followed by agency trading and then A2A trading.

The intermediation costs create a segmentation of the market into customer-dealer principal and agency trading, and A2A trading, based on the trade-off between costs and immediacy. Customers pay the highest intermediation costs in principal inventory trades to obtain immediacy, and pay no fees for A2A trading. Agency trading represents an intermediary case where customers pay intermediation costs lower than those of principal trade, and in exchange trades faster than in A2A markets, but slower than in principal trade. Agency trading represents customer liquidity provision as described by Choi, Huh, and Shin (2024).

Customers in the model have different valuations for holding an asset. Proposition 1 determines the equilibrium price and asset holdings with no frictions. Our strategy is to consider the cases with agency and principal trading separately and then to combine the cases together. We first consider the case with A2A and agency trading. We characterize which customers choose agency or A2A, the distribution of asset owners, and the equilibrium price (Lemmas 1–6 and Proposition 2). Next, we provide the analogous characterization for A2A and principal trading (Lemmas 7–12 and Proposition 3). We then determine the thresholds for each market, and characterize the equilibrium with the three trading methods together (Propositions 4 and 5).

We can have different market configurations (Section 6). We can have a large principal market together with small agency and A2A trading, which was a common configuration before 2010; or a smaller principal market but larger agency market, as we have today. An increase in agency trading does not necessarily imply an increase in A2A trading. Therefore, the model allows for a stable share of A2A trading alongside an increase in agency trading, as observed since the introduction of A2A platforms. In addition to intermediation costs, the markets can move according to search-intensity, bargaining, and other parameters. We interpret search and bargaining parameters changes as being the result of technological changes whereas the intermediation costs are the result of regulatory changes.

We find that high intermediation costs can lead to multiple equilibria (Proposition 6). We guarantee uniqueness for low intermediation costs (Figure 4 shows unique equilibrium and Figure 5 shows multiple equilibria). Equilibrium multiplicity occurs because customers predict the probability of matching with a dealer or another customer based on their trading choice. They cannot coordinate their search. For the same parameters, one equilibrium may feature a small agency market while another may feature a large agency market. The same occurs for A2A markets. The benefits of coordination increase as intermediation costs rise, and small intermediation costs rule out equilibrium multiplicity.

Having established how the different markets interact in equilibrium, we turn to the effects of market composition on liquidity and on the bid-ask spread (Section 7). Dealers engaged in principal trading hold inventory to provide liquidity on spot. We show that an increase in inventory costs imply a decrease in the dealers' net position (Proposition 7). The model predicts a decrease in the net position after an increase in costs caused by regulations, which we observe in the data (Figure 6).

Surprisingly, we show that technological improvements in A2A markets, such as more efficient search, do not affect dealers' net positions. As a result, we can have an active A2A market and undisturbed provision of on-spot liquidity with principal trading. This prediction agrees with the observation of an approximately stable share of A2A trading in parallel with the existence of dealer-intermediated trading.

The model predicts that customers benefit from the provision of liquidity in agency trading. They pay smaller ask prices and sell for higher bid prices in comparison with principal trading (Proposition 8). The predictions on the rewards of liquidity provision agree with the findings in Choi et al. (2024).

We show that the bid-ask spread can decrease even though inventory costs increase (Proposition 9). An increase in dealer costs induces customer to search for other customers. For a sufficient increase in dealer costs, this change in composition decreases aggregate bid-ask spread (Lester et al. (2023) also find a non-monotone bid-ask spread in a model with information frictions). Standard indicators of illiquidity rely on observed transactions. A decrease in the aggregate bid-ask spread implies an improvement in standard indicators of liquidity. The model then generates a change in the structure of the corporate bond market together with improvement of indicators of market illiquidity.[1]

Another unexpected result obtained from the interaction of the A2A, agency and principal trading is on the bid-ask spread. We show that technological improvements in the A2A market increase the bid-ask spread rather than decrease it (Proposition 9). According to the model, customers with moderate utility types move to an enhanced A2A market. The more extreme utility types then face higher bid-ask spreads when they use the services of dealers.

After the 2008 financial crisis, several regulations were enacted with the objective of avoiding future financial crises. In particular, the Dodd-Frank Wall Street Reform and Consumer Protection Act in 2010. To some extent, the Dodd-Frank Act and similar regulations in other countries accomplished their goal, as there was not a comparable financial crisis after 2008. However, the focus of academics, practitioners and government officials turned to how such regulations affect financial markets in normal times, when the economy is not under distress. There are indications that the form in which trades take place in over-the-counter markets has changed after the regulations were put in place.[2]

The model is able to explain changes on the corporate bond market after the Dodd-Frank

[1]The model implies a new measure of illiquidity, based on equilibrium prices, search frictions, and the fraction of the market that engages in inventory trading or customer liquidity provision. We explore this aspect further in Dyskant et al. (2025).

[2]During a 2015 congressional hearing, for example, Rep. French Hill questioned the Federal Reserve Chair at the time, Janet Yellen, on whether regulations were to blame for the deterioration of liquidity on different bond markets. Yellen replied: "I am not ruling out the possibility that regulations could play a role here, it is simply we have not been able to understand through a lot of different factors and we need to look at it more to sort out just what is going on and what the different influences are, but I am not ruling that out."

Act. The Act includes the Volcker rule, which prohibits institutions from trading that uses the inventory of assets purchased earlier with the intention of profiting from a higher sale price (proprietary trade). In the context of the model, we interpret the introduction of the Volcker rule as an increase in inventory costs. According to the model, the increase in inventory costs implies (1) the increase in agency trading or A2A trading; (2) longer trade executions; and (3) a potential decrease in measured illiquidity, even though market participants indicate more difficulty in trading. These predictions are in line with the observations on the corporate bond market after the post-2008 regulations.

2 Related literature

Our main contribution is to provide an equilibrium framework to understand how trading in OTC markets can be organized into principal, agency and A2A trading. Other papers that discuss the trade-off between principal and agency trading are Cimon and Garriott (2019), Kargar et al. (2021), Saar et al. (2023), and An and Zheng (2023). Cimon and Garriott (2019) and Kargar et al. (2021) have static models and do not consider how different trading mechanism result in different trading speeds. Saar et al. (2023) take this trade-off into consideration when studying bank-dealer competition. They find that the increase of regulatory costs might increase market efficiency if matchmaking is a superior form of trade. An and Zheng (2023) is closer to our model as a search and matching is used to model the OTC market. They concentrate on the incentives to the dealer to provide immediacy or matchmaking. They find that the bid-ask spread decreases with higher regulatory costs, whereas we find a non-monotone relation. Moreover, we find the possibility of multiple equilibria. We are the first to consider together the three trading mechanisms in OTC markets (principal, agency, and A2A) in equilibrium.

Our framework is a search and matching model in the spirit of DGP as investors have to search for a counterparty and trades are motivated by sudden changes in their characteristics. In the case of DGP, sudden changes in utility types; in our case, changes in issuance or asset maturity as in Bethune et al. (2022). We add a dealer with access to an interdealer market as Lagos and Rocheteau (2009). Our dealer is also called in the literature middleman or marketmaker. Hugonnier et al. (2020) also studies dealer intermediation in a model where dealers hold inventory, with the focus in what we refer to as principal trade.

We have an arbitrary distribution of utility types, as Hugonnier, Lester, and Weill (2022) (HLW), which generates dispersion in the execution trading times and on investor decisions. The utility types are known by the trading parties, as in HLW, and not private information as in BST. Heterogeneity in utility types is especially important in our case because we want to investigate changes in the masses of agents that engage in different trading modes.[3]

Customers choose to participate in customer-dealer or customer-customer matches. In the spirit of Guerrieri et al. (2010), agents direct their search in financial markets. Agents in our model hold either 0 or 1 units of the asset, as DGP, HLW, and other papers in the search and matching literature. Lagos and Rocheteau (2009) allow for arbitrary asset positions.

Our model discusses how direct matches between customers in OTC markets coexist with a marketmaker. Early search models that analyze the existence of marketmakers in OTC markets are Gehrig (1993) and Yavas (1996), with static models and equilibria defined as the outcome of a strategic game between marketmakers and customers (without principal trade). We show that prices, bid-ask spreads and other factors depend on factors that affect the dynamics such as intertemporal discounting and expectations of future matches. Dynamic factors have important effects on equilibrium variables, a point also made by DGP. We also show how a structure of principal, agency, and A2A market interact with interdealer prices and bid and ask prices in general equilibrium.

An early model that incorporates the dynamics in a search setting, but without the consideration of different OTC trade arrangements, is Spulber (1996). In Spulber, all trades require a middleman. Rust and Hall (2003) extend this model with the introduction of a monopolistic marketmaker that publicizes bid and ask prices. Both show how dynamic considerations, such as the time discount, affect equilibrium variables. We extend these models in different ways. We show, for example, how the technology of each market can affect the number of customers involved in each trade. We also show that each trade arrangement can coexist with other trade arrangements independently. Agency or another form of trade might disappear, for example, and different combinations of trading modes can be formed.

The improvement in the traditional measures of market liquidity after the 2008 financial

[3]It is common in the OTC to limit the utility types to a small number such as two types (DGP 2005, 2007; Vayanos and Wang 2007; and others). We allow for continuum of utility types.

crisis has been documented by Bessembinder et al. (2018). We confirm this finding with recent data for the BPW and Amihud liquidity measures (after Bao, Pan, and Wang (2011) and Amihud (2002)). The improvement in measured market liquidity could suggest that regulations had a minor impact on financial market liquidity. However, Bessembinder et al. (2018) and Choi et al. (2024) report a decrease in dealer trade frequency. Especially, Choi et al. (2024) indicate a change in the composition of the provision of liquidity. The provision of liquidity has increasingly been made by customers rather than dealers. In practice, there is a perceived movement of customers from principal trade to A2A trade or agency trade.[4]

We endogenize the market structure of OTC markets into different trading mechanisms. Hugonnier et al. (2020) and Afonso and Lagos (2015) endogenize the market structure by focusing on the intermediation chains, Farboodi et al. (2023) by the concentration of the frequency of trades around certain customers, interpreted as marketmakers, and Bethune et al. (2022) by screening ability. Intermediation chains arise in Hugonnier et al. (2020) as customers in different sides of the market are matched with a sequence of dealers. Intermediation arises endogenously in Afonso and Lagos (2015) as banks borrow and lend federal reserves to other counterparties according to their asset positions. In Farboodi et al. (2023), participants can increase their contact rate by making a payment. We maintain the contact rate fixed across trading methods, but we allow investors to change their method of trading. We have heterogeneity in the contact rate across methods of trading.[5] Intermediation arises endogenously in Bethune et al. (2022) as dealers with screening technology avoid negotiation breakdowns during bargaining.

In Lagos and Rocheteau (2009) and Duffie et al. (2005) there is an exogenous marketmaker or middleman that facilitates trade. We endogenize the choice to contact a middleman and the intensity of trade (agency or principal). Dealers and customers have exogenously given masses in Hugonnier et al. (2020). In our case, the masses of customers engaged in different trading methods are endogenously determined by preference and technological parameters

[4]We obtain similar findings on the importance of illiquidity for the yield spreads as Li and Yu (2023) and Wu (2024), which worked in independent papers. Our empirical results are different in some aspects (such as our focus on the BPW and Amihud measures) and they complement their findings. However, our focus is on the model to explain the movements in the corporate bond market. We discuss the liquidity premium in Dyskant et al. (2025).

[5]Üslü (2019) studies the endogenous formation of intermediaries with a richer model with heterogeneous meeting rates of intermediaries. He finds that intermediaries that trade more frequently earn higher markups. In our case, intermediaries that engage in principal trade, who trade faster, also earn higher markups.

such as costs and trading speed. Markets can change and the marketmaker can disappear. Analogously, if the technology of the the marketmaker is superior, the A2A market can disappear. The trading methods are connected, but they can exist independently. We describe different components of the OTC market in a relatively parsimonious way: with heterogeneity in utility, trade stimulated by asset issuance and maturity, and different matching technologies.

We have inventories in our model as dealers can buy an asset or sell short on spot to fast fulfill an order in case of principal trade. However, inventory is not maintained over time as the position is reversed in the end of the period. In contrast, Cohen et al. (2024) rule out short sales so that dealers have to carry on inventory over time. Dealers in Weill (2007) and Üslü (2019) also have positive inventory over time. These papers analyze the consequences of having positive inventory over time. We define inventory here to distinguish between principal and agency trading. Our modeling choice is motivated by our focus on the determinants for the choice on the method of trading.

Our paper is also related to the empirical literature that studies the effects of the post-2008 crisis regulations and the new composition of the market. See, among others, Dick-Nielsen et al. (2012), Anderson and Stulz (2017), Bessembinder et al. (2018), Bao et al. (2018), and Dick-Nielsen and Rossi (2018). Duffie (2012) discusses expected effects of the Volcker rule and Duffie (2018) discusses the regulatory effects more comprehensively. For long-run effects, related to the composition of the market, see, for example, Goldstein et al. (2017), Musto et al. (2018), Falato et al. (2021), Jiang et al. (2022), Choi et al. (2024), Li and Yu (2023), and Wu (2024). Our model can be applied to analyze the effects of regulatory changes.

We provide a comprehensive picture of the OTC market, with its components in terms of trading methods together. As it is an equilibrium model, we show how prices interact with the size of each market. As a result, we can study how changes in technology, preferences, or regulations changes the size of each component, bid-ask spreads, and other equilibrium values.

3 Model

3.1 Environment

We model over-the-counter markets as markets in which agents take decisions under search frictions. As discussed in Section 2, our model builds more directly on Duffie et al. (2005), Hugonnier et al. (2022), and Lagos and Rocheteau (2009).

Agents, time, goods and assets There are two types of agents in a continuum: a measure one of customers and an infinite measure of dealers. Time is continuous and infinite. There is a unique good and an endogenous supply $s \geq 0$ of assets. A unit of the asset pays a unit flow of dividends in the form of goods, which cannot be traded. The agent holding an asset consumes its dividends. All agents are infinitely-lived, discount the future at rate $r > 0$, and have transferable utility. We interpret transferable utility as linear utility over a numéraire good used to make payments. Customers hold either zero or one unit of the asset. We refer to customers who do not hold the asset as non-owners, and those who do as owners. Dealers can hold discrete amounts i of the asset, $i \in \{-1, 0, 1\}$. We interpret a negative dealer position $i < 0$ as either the dealer borrowing the asset to sell, a common practice for short selling, or holding a position below the target for its private account.[6]

Preferences Customers are heterogeneous in the utility ν that they derive from consuming the dividend flow. We refer to ν as the customer utility type. Types are fixed over time, common knowledge, independent across customers, and initially drawn from a given cumulative distribution F. The distribution F has support $\mathbb{R}$, and a continuous density $f > 0$. We assume that $\int \nu^2 f(\nu) d\nu < \infty$ and that there is no free disposal of assets. The assumption that the distribution of types has unbounded support is convenient because we do not have to consider corner solutions. However, we can obtain our main results with a bounded support $[\underline{\nu}, \bar{\nu}]$ if we assume that the density is sufficiently low at the extremes. Similarly, the assumption of no free disposal can be replaced with the assumption that the measure of agents with $\nu < 0$ is

[6] We could alternatively model dealers holding inventory in order to sell the bond; a form of "inventory-in-advance" constraint proposed by Cohen et al. (2024). Both versions imply similar results because we can pick the cost of shorting the bond to match the cost of holding inventory. We keep the current version where dealers can take a negative position because it simplifies the analysis.

sufficiently low. Dealers do not derive utility from holding assets. They hold inventory only to profit from intermediation. They pay a flow cost $c^l \geq 0$ when they take a long position, $i = 1$, and pay a flow cost $c^s \geq 0$ when they take a short position, $i = -1$.[7]

Decentralized market There is an OTC asset market in which customers choose to search for another customer or for a dealer. Customers who find a dealer then choose between agency and principal trade. Figure 1 depicts the sequence of actions of customers. We describe the sequence of trade in detail below. Customers cannot search for both simultaneously.

When searching for another customer, the customer finds one with Poisson arrival rate $\lambda_C/2 > 0$. This implies a meeting rate of λ_C (the two customers search for each other at $\lambda_C/2$). Customers searching for customers do not find customers searching for dealers. We interpret customer-to-customer trading as performed in all-to-all trade platforms (A2A), which allow customers to search for a counterparty without the participation of dealers.

When searching for a dealer, the customer finds one with Poisson arrival rate $\lambda_D^0 > 0$. After the customer and the dealer meet, they can trade in two ways: risky principal trade or agency trade. In a risky principal trade, the dealer trades with the customer on its own account on spot, and then joins an interdealer market to rebalance its portfolio. In an agency trade, the dealer joins the interdealer market on behalf of the customer and completes the transaction afterwards. Agency trading take longer to complete because the dealer has to search in the interdealer market. In risky principal trading, the dealer provides immediacy to the customer by trading on its own account. To simplify the model and notation, we assume that dealers exit the economy once they intermediate a transaction. In the case of principal trade, the dealer exits the economy after rebalancing the portfolio. In the case of agency trades, the dealer exits the economy after completing the trade on behalf of the customer.

Interdealer market Following Lagos and Rocheteau (2009), dealers access a competitive market with Poisson arrival rate $\lambda_D^1 > 0$, where the endogenous asset price is p. Dealers pay an intermediation cost $\tau \geq 0$ when trading on behalf of customers in an agency trade. That is, the net revenue of selling an asset in an agency trade is $p - \tau$, and the cost of buying an asset in an agency trade is $p + \tau$.

[7]This formulation captures the use of repurchase agreements to either go short or long on bonds, as introduced in Cimon and Garriott (2019). A similar formulation is used in Saar et al. (2023).

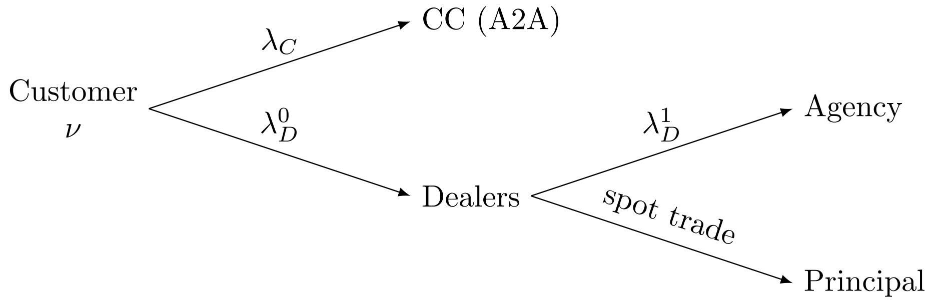


Figure 1: Sequence of actions in the OTC market. A customer of type ν chooses to search for another customer in the customer-customer market (all-to-all trade platform), or to search for a dealer. When meeting a dealer, customers choose between agency and principal trade.

Asset supply Asset issuance and maturity determine the supply of assets. Customers issue new assets at no cost with Poisson arrival rate $\eta > 0$, and an asset matures with Poisson arrival rate $\mu > 0$. The asset disappears from a portfolio and from the economy at maturity. In an agency trade, if the customer state changes because of asset maturity or issuance, the match ends. Dealers do not issue assets, and assets in their possession do not mature. This assumption simplifies the model, but is not important for the qualitative results. Asset issuance and maturity in our model follows Bethune et al. (2022) and implies that a steady state with positive trade emerges without introducing time-varying types.

Bargaining We use Nash bargaining to model trade. In a customer-dealer trade, the customer bargaining power is $\theta_D \in [0, 1]$. In a customer-customer trade, the buyer bargaining power is $\theta_C^n \in [0, 1]$ and the seller bargaining power is $\theta_C^o = 1 - \theta_C^n$. We assume the following relation between the search and bargaining parameters.

Assumption 1. *The search technology of the dealers is sufficiently better than the search technology of the customers. That is,* $\frac{\lambda_D^0 \lambda_D^1 \theta_D}{r+\mu+\eta+\lambda_D^0+\lambda_D^1} > \lambda_C \max\{\theta_C^o, \theta_C^n\}$.

Assumption 1 implies that customers are better off searching for dealers when dealers face no principal trade cost ($c^s = c^l = 0$) or no agency trade cost ($\tau = 0$).

3.2 Value functions and reservation values

We first specify the value function of customers according to their search choices. When searching for dealers, the value function is $V_D^o(\nu)$ for owners and $V_D^n(\nu)$ for non-owners. When searching for customers, the value function is $V_C^o(\nu)$ for owners, and $V_C^n(\nu)$ for non-owners. The value function yields the maximum between the two trading choices,

$$V^o(\nu) = \max\{V_D^o(\nu), V_C^o(\nu)\}, \text{ for owners,} \tag{1}$$

$$V^n(\nu) = \max\{V_D^n(\nu), V_C^n(\nu)\}, \text{ for non-owners.} \tag{2}$$

The value function of customers waiting in an agency trade with a dealer is $\tilde{V}_D^o(\nu)$ for owners, and $\tilde{V}_D^n(\nu)$ for non-owners. The reservation value of customers is the compensation that makes them indifferent between owning an asset or not, $\Delta(\nu) = V^o(\nu) - V^n(\nu)$.

The value function of a dealer with asset position i is W^i. The reservation values of a dealer when buying an asset, selling it, and in agency trade are $W^l = W^1 - W^0$, $W^s = W^0 - W^{-1}$, and W^a. As there is an infinite measure of dealers, the probability of a specific dealer to serve a customer is zero. This implies $W^0 = 0$. Moreover, as dealers exit the economy after an agency trade, they will have to be matched again to future customers, which implies that the dealers' reservation value after meeting a customer for agency trade is zero, $W^a = W^0 = 0$.

Searching for dealers The value function of a type-ν owner searching for a dealer satisfies

$$rV_D^o(\nu) = \nu - \mu\Delta(\nu) + \lambda_D^0\theta_D \max\Big\{\underbrace{W^l - \Delta(\nu)}_{\substack{\text{Principal}\\ \text{trade}}},\ \underbrace{\tilde{V}_D^o(\nu) - V_D^o(\nu)}_{\substack{\text{Agency}\\ \text{trade}}},\ 0\Big\}. \tag{3}$$

The first and second terms on the right-hand side are the utility flow of holding the asset and the expected loss of reservation value in case of asset maturity. The third term, $\lambda_D^0\theta_D \max\{W^l - \Delta(\nu), \tilde{V}_D^o(\nu) - V_D^o(\nu), 0\}$, is the profit of an owner when meeting a dealer. The pair owner-dealer has three options: they can trade in a principal trade so that the dealer purchases the asset immediately, they can wait to access the interdealer market in an agency trade, or they can decide not to trade. The gains from trade are respectively $W^l - \Delta(\nu)$, $\tilde{V}_D^o(\nu) - V_D^o(\nu)$, and zero. The owner keeps a share θ_D of the gains from trade.

The value function of a type-ν owner matched with a dealer in an agency trade satisfies

$$r\tilde{V}_D^o(\nu) = \nu - \mu[\tilde{V}_D^o(\nu) - V^n(\nu)] + \lambda_D^1 \max\{p - \tau - [\tilde{V}_D^o(\nu) - V^n(\nu)],\ 0\}. \tag{4}$$

The expected loss in case of maturity, in the second term, now takes into account that the customer has already been matched. The third term is the profit of an owner when selling the asset in the interdealer market.

The value function of a type-ν non-owner searching for dealers satisfies

$$rV_D^n(\nu) = \eta\Delta(\nu) + \lambda_D^0\theta_D \max\Big\{\underbrace{\Delta(\nu) - W^s}_{\substack{\text{Principal}\\ \text{trade}}},\ \underbrace{\tilde{V}_D^n(\nu) - V_D^n(\nu)}_{\substack{\text{Agency}\\ \text{trade}}},\ 0\Big\}. \tag{5}$$

The first term is the expected gain in reservation value in case of asset issuance. The second term is the profit of a non-owner when meeting a dealer. The pair owner-dealer has three options: the non-owner can buy the asset from the dealer in a principal trade, the non-owner can wait the dealer to access the interdealer market for an agency trade, or they can decide not to trade. The gains from trade are respectively $\Delta(\nu) - W^s$, $\tilde{V}_D^n(\nu) - V_D^n(\nu)$, and zero. The non-owner keeps a share θ_D of the gains from trade.

The value function of a type-ν non-owner matched with a dealer in an agency trade satisfies

$$r\tilde{V}_D^n(\nu) = \eta[V^o(\nu) - \tilde{V}_D^n(\nu)] + \lambda_D^1 \max\{V^o(\nu) - \tilde{V}_D^n(\nu) - (p+\tau),\ 0\}. \tag{6}$$

The expected gain in an asset issuance takes into account that the non-owner is matched. The second term is the profit of a non-owner when buying an asset in the interdealer market.

Value functions for the dealers The value functions of a dealer holding long and short positions satisfy respectively

$$rW^1 = rW^l = -c^l + \lambda_D^1(p - W^l), \tag{7}$$

$$rW^{-1} = -rW^s = -c^s + \lambda_D^1(W^s - p). \tag{8}$$

The dealer long on an asset pays the flow cost c^l and sells the asset at the price p in the interdealer market, which can be accessed at the rate λ_D. Similarly, the dealer short on an asset pays the flow cost c^s, and buys the asset at the price p in the interdealer market.[8]

Let the distributions of customer types for owners and non-owners be denoted by Φ^o and Φ^n, respectively. As each owner holds exactly one unit of the asset, the measure of assets is $s = \int d\Phi^o(\nu)$. Owners try to sell an asset and non-owners try to buy an asset.

Let $\Pi = \{\pi_{-1}, \pi_0^o, \pi_0^n, \pi_1\}$ denote the inventory distribution among dealers. π_{-1} is the measure of dealers in principal trade who sold an asset, and π_1 is the measure of dealers in principal trade who bought an asset. $\pi_0^o(\nu)$ is the measure of dealers in agency trade matched to owners of type below ν, and $\pi_0^n(\nu)$ is the measure of dealers in agency trade matched to non-owners of type below ν. The distributions of dealers in principal trade do not depend on ν as dealers trade on spot. Dealers in agency trade trade on behalf of a customer with type ν. Let $\pi_0 = \pi_0^o + \pi_0^n$ be the measure of dealers in agency trade waiting to access the interdealer market to trade on behalf of the customer.[9]

Let the sets Ω_D^o and Ω_C^o denote the owners that search for dealers and that search for other customers. Analogously, Ω_D^n and Ω_C^n denote the sets of non-owners that search for dealers and that search for other customers. $\{\Omega_D^o, \Omega_C^o\}$ and $\{\Omega_D^n, \Omega_C^n\}$ form two partitions of $\mathbb{R}$. We denote the search partitions of customers by $\mathcal{P} = \{\Omega_D^o, \Omega_C^o, \Omega_D^n, \Omega_C^n\}$. We can further partition the sets Ω_D^o and Ω_D^n into partitions of principal trade and agency trade, $\{\Omega_D^{o,p}, \Omega_D^{o,a}\}$ and $\{\Omega_D^{n,p}, \Omega_D^{n,a}\}$, where the superscripts p and a denote principal trade or agency trade. Denote the trade-mode partition by $\mathcal{P}_D = \{\Omega_D^{o,p}, \Omega_D^{o,a}, \Omega_D^{n,p}, \Omega_D^{n,a}\}$.

Customers searching for customers The value function of an owner of type ν searching for a non-owner satisfies

$$rV_C^o(\nu) = \nu - \mu\Delta(\nu) + \lambda_C \theta_C^o \int_{\Omega_C^n} [\Delta(\tilde{\nu}) - \Delta(\nu)] \mathbb{1}_{\{\Delta(\tilde{\nu}) > \Delta(\nu)\}} d\Phi^n(\tilde{\nu}). \tag{9}$$

The first term of the value function is the utility flow of holding the asset. The second term is the expected loss of the reservation value in case of asset maturity. The third term is the expected profits of an owner when meeting a non-owner. When trading with a non-owner of type $\tilde{\nu}$, an owner of type ν sells the asset if the reservation value of the counterparty, $\Delta(\tilde{\nu})$, is

[8] These expressions account for the fact that $W^0 = 0$, that is, the reservation value of a dealer holding a position of zero is equal to zero, as the likelihood of a particular dealer meeting a customer is zero.

[9] The distribution functions could change over time in general. We focus on steady-state equilibria and omit time subscripts.

higher than the reservation value of the owner, $\Delta(\nu)$. The gains from trade are $\Delta(\tilde{\nu}) - \Delta(\nu)$ and the owner keeps a share θ_C^o of it. We obtain the expected value of the gains from trade by integrating it in $\tilde{\nu}$ over Ω_C^n using the distribution of non-owners $\Phi^n(\tilde{\nu})$.

The value function of a non-owner of type ν searching for an owner satisfies

$$rV_C^n(\nu) = \eta\Delta(\nu) + \lambda_C\theta_C^n \int_{\Omega_C^o} [\Delta(\nu) - \Delta(\tilde{\nu})]\mathbb{1}_{\{\Delta(\nu)>\Delta(\tilde{\nu})\}} d\Phi^o(\tilde{\nu}). \tag{10}$$

The first term of the value function is the expected gain of reservation value in case of asset issuance. The second term is the expected profit of a non-owner when searching for an owner. A non-owner of type ν buys the asset from an owner of type $\tilde{\nu}$ if the reservation value of the non-owner is higher than the reservation value of the owner. The non-owner keeps a share θ_C^n of the gains from trade, $\Delta(\nu) - \Delta(\tilde{\nu})$. The expected gains from trade are obtained by integrating it in $\tilde{\nu}$ over Ω_C^o using the distribution of owners $\Phi^o(\tilde{\nu})$.

Value functions, reservation value and the optimal searching choice The value functions V^o and V^n of a customer of type ν satisfy

$$V^o(\nu) = \max\{V_D^o(\nu), V_C^o(\nu)\} \quad \text{and} \quad V^n(\nu) = \max\{V_D^n(\nu), V_C^n(\nu)\}, \tag{11}$$

and the reservation value function satisfies

$$\Delta(\nu) = V^o(\nu) - V^n(\nu). \tag{12}$$

We characterize the search partition $\mathcal{P} = \{\Omega_D^o, \Omega_C^o, \Omega_D^n, \Omega_C^n\}$ in the following way. For Ω_D^o, an owner searches for a dealer if it yields higher value than searching for a non-owner. In the same way, for Ω_D^n, a non-owner searches for a dealer if it yields higher value then searching for an owner. We have

$$\Omega_D^o = \{\nu \in \mathbb{R};\ V_D^o(\nu) \geq V_C^o(\nu)\} \quad \text{and} \quad \Omega_D^n = \{\nu \in \mathbb{R};\ V_D^n(\nu) \geq V_C^n(\nu)\}. \tag{13}$$

Analogously,

$$\Omega_C^o = \{\nu \in \mathbb{R};\ V_C^o(\nu) > V_D^o(\nu)\} \quad \text{and} \quad \Omega_C^n = \{\nu \in \mathbb{R};\ V_C^n(\nu) > V_D^n(\nu)\}. \tag{14}$$

Similarly, we characterize the search partition $\mathcal{P}_D = \{\Omega_D^{o,p}, \Omega_D^{o,a}, \Omega_D^{n,p}, \Omega_D^{n,a}\}$ in the following way. An owner-dealer pair or a non-owner-dealer use a principal trade if it yields higher value than waiting in an agency trade. Then for principal trade we have

$$\begin{aligned}
\Omega_D^{o,p} &= \left\{\nu \in \Omega_D^o;\ W^l - \Delta(\nu) \geq \tilde{V}_D^o(\nu) - V_D^o(\nu) \text{ and } W^l - \Delta(\nu) \geq 0\right\}, \\
\Omega_D^{n,p} &= \left\{\nu \in \Omega_D^n;\ \Delta(\nu) - W^s \geq \tilde{V}_D^n(\nu) - V_D^n(\nu) \text{ and } \Delta(\nu) - W^s \geq 0\right\}.
\end{aligned} \tag{15}$$

Analogously, for agency trade we have

$$\begin{aligned}
\Omega_D^{o,a} &= \left\{\nu \in \Omega_D^o;\ \tilde{V}_D^o(\nu) - V_D^o(\nu) > \max\{W^l - \Delta(\nu), 0\}\right\} \\
\Omega_D^{n,a} &= \left\{\nu \in \Omega_D^n;\ \tilde{V}_D^n(\nu) - V_D^n(\nu) > \max\{\Delta(\nu) - W^s, 0\}\right\}.
\end{aligned} \tag{16}$$

Customers search for dealers when indifferent between searching for a dealer or other customers (eq. 13). Customer-dealer pairs use principal trade when indifferent between agency or principal trade (eq. 15). Customers always search (there are no inactive customers). For the equilibrium considered, these assumptions are without loss of generality because there is a measure zero of customers that are indifferent in equilibrium.

3.3 Interdealer market clearing

The interdealer market clears when the measure of dealers selling assets is equal to the measure of dealers buying assets. Dealers can sell or buy assets either to rebalance their portfolios or to execute agency trades. The measure π_1 of dealers holding an asset satisfies

$$\dot{\pi}_1 = -\lambda_D^1 \pi_1 + \lambda_D^0 \int_{-\infty}^{\infty} \mathbb{1}_{\{\nu \in \Omega_D^{o,p}\}} d\Phi^o(\nu) = 0. \tag{17}$$

The measure π_{-1} of dealers short on an asset satisfies

$$\dot{\pi}_{-1} = -\lambda_D^1 \pi_{-1} + \lambda_D^0 \int_{-\infty}^{\infty} \mathbb{1}_{\{\nu \in \Omega_D^{n,p}\}} d\Phi^n(\nu) = 0. \tag{18}$$

The measure $\pi_0^o(\nu)$ of dealers with an owner type $\tilde{\nu} \leq \nu$ in an agency trade satisfies

$$\dot{\pi}_0^o(\nu) = -\left(\mu + \lambda_D^1\right) \pi_0^o(\nu) + \lambda_D^0 \int_{-\infty}^{\nu} \mathbb{1}_{\{\tilde{\nu} \in \Omega_D^{o,a}\}} d[\Phi^o(\tilde{\nu}) - \pi_0^o(\tilde{\nu})] = 0, \tag{19}$$

and the measure π_0^n of dealers with a non-owner type $\tilde{\nu} \leq \nu$ in an agency trade satisfies

$$\dot{\pi}_0^n(\nu) = -\left(\eta + \lambda_D^1\right) \pi_0^n(\nu) + \lambda_D^0 \int_{-\infty}^{\nu} \mathbb{1}_{\{\tilde{\nu} \in \Omega_D^{n,a}\}} d[\Phi^n(\tilde{\nu}) - \pi_0^n(\tilde{\nu})] = 0. \tag{20}$$

The interdealer market clears if

$$\lambda_D^1 \left(\pi_1 + \bar{\pi}_0^o\right) = \lambda_D^1 \left(\pi_{-1} + \bar{\pi}_0^n\right), \tag{21}$$

where $\bar{\pi}_0^o = \lim_{\nu \to \infty} \pi_0^o(\nu)$ and $\bar{\pi}_0^n = \lim_{\nu \to \infty} \pi_0^n(\nu)$. The interdealer market clears if the measure of dealers selling an asset on principal or agency trade equals the measure of dealers buying an asset on principal or agency trade.

3.4 The distribution of assets across customers

The cumulative distribution of owners and non-owners are given respectively by Φ^o and Φ^n. The change over time of the distribution of owners Φ^o satisfies

$$\begin{aligned}
\dot{\Phi}^o(\nu) &= \eta \Phi^n(\nu) - \mu \Phi^o(\nu) - \int_{-\infty}^{\nu} \left[\lambda_D^0 \mathbb{1}_{\{\tilde{\nu} \in \Omega_D^{o,p}\}} + \frac{\lambda_D^0 \lambda_D^1 \mathbb{1}_{\{\tilde{\nu} \in \Omega_D^{o,a}\}}}{\mu + \lambda_D^0 + \lambda_D^1}\right] d\Phi^o(\tilde{\nu}) \\
&+ \int_{-\infty}^{\nu} \left[\lambda_D^0 \mathbb{1}_{\{\tilde{\nu} \in \Omega_D^{n,p}\}} + \frac{\lambda_D^0 \lambda_D^1 \mathbb{1}_{\{\tilde{\nu} \in \Omega_D^{n,a}\}}}{\eta + \lambda_D^0 + \lambda_D^1}\right] d\Phi^n(\tilde{\nu}) \\
&- \lambda_C \int_{-\infty}^{\nu} \int_{\nu}^{\infty} \mathbb{1}_A d\Phi^n(\hat{\nu}) d\Phi^o(\tilde{\nu}) + \lambda_C \int_{-\infty}^{\nu} \int_{\nu}^{\infty} \mathbb{1}_A d\Phi^o(\hat{\nu}) d\Phi^n(\tilde{\nu}),
\end{aligned} \tag{22}$$

where $A = \{\tilde{\nu} \in \Omega_C^o, \hat{\nu} \in \Omega_C^n, \Delta(\hat{\nu}) > \Delta(\tilde{\nu})\}$. The first term on the right-hand side accounts for the inflow of new owners that have just issued an asset and the second term accounts for the outflow of owners because of asset maturity. The third and fourth terms account for owners searching for dealers. The third term for the outflow of owners with type below ν searching for dealers and that sell their asset, and the fourth for the inflow of non-owners with type below ν searching for dealers and that buy an asset. The fifth and sixth terms account for customers searching for other customers. The fifth term for the outflow of owners with type below ν searching for other customers, which sell their asset to non-owners of type above ν, and the sixth term for the inflow of non-owners with type below ν searching for other customers, which buy an asset from owners of type above ν. In an steady-state equilibrium, $\dot{\Phi}^o(\nu) = 0$ for all ν.

As the measure of customers F is exogenous, the measures of owners and non-owners satisfy the equilibrium condition

$$\Phi^o(\nu) + \Phi^n(\nu) = F(\nu). \tag{23}$$

All assets in the economy are held by owners or dealers. The stock of assets held by owners is

$$s = \int_{-\infty}^{\infty} d\Phi^o(\nu) = \Phi^o(\infty). \tag{24}$$

3.5 Equilibrium

We define a stationary equilibrium in the following way.

Definition 1. *An equilibrium is a family of value functions, reservation values, price, distributions and partitions,* $\{V^o, V^n, \Delta, W^l, W^s, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *satisfying equations* (3)–(24).

An equilibrium, even in steady state, can be a complicated object. To simplify it further, let $\Omega_C = \Omega_C^o = \Omega_C^n = (\nu_l, \nu_h)$ and $\Omega_D = \Omega_D^o = \Omega_D^n = (-\infty, \nu_l] \cup [\nu_h, \infty)$. Notice that we can have $\Omega_C^o = \Omega_C^n$ as we have some customers of type ν holding the asset and other customers of the same type that do not hold the asset. Moreover, let $\Omega_D^{o,a} = (\nu_l^a, \nu_l]$, $\Omega_D^{n,a} = [\nu_h, \nu_h^a)$, $\Omega_D^{o,p} = (-\infty, \nu_l^a]$, and $\Omega_D^{n,p} = [\nu_h^a, \infty)$. Define the following class of equilibrium.

Definition 2. *An equilibrium* $\{V^o, V^n, \Delta, W^l, W^s, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *is regular if*

i. $\Omega_C = \Omega_C^o = \Omega_C^n = (\nu_l, \nu_h)$,

ii. $\Omega_D = \Omega_D^o = \Omega_D^n = (-\infty, \nu_l] \cup [\nu_h, \infty)$, *iii.* $\Omega_D^{o,a} = (\nu_l^a, \nu_l]$, $\Omega_D^{n,a} = [\nu_h, \nu_h^a)$, $\Omega_D^{o,p} = (-\infty, \nu_l^a]$, *and* $\Omega_D^{n,p} = [\nu_h^a, \infty)$

for some ν_l^a, ν_l, ν_h, $\nu_h^a \in \mathbb{R}$ *satisfying* $\nu_l^a \leq \nu_l \leq \nu_h \leq \nu_h^a$, *with* $\nu_l < \nu_h$ *if* $c^l, c^s, \tau > 0$, *and the reservation value* $\Delta(\nu)$ *is continuous and strictly increasing.*

Figure 2 illustrates the partitions of a regular equilibrium. The motivation to look for an equilibrium with the characteristics of a regular equilibrium is the following. Customers with type close to each other, in $\Omega_C = (\nu_l, \nu_h)$, as they do not gain much from trading, choose to trade among themselves to avoid the costs associated with trading with a dealer, τ and c. Customers with moderate types, outside $\Omega_C = (\nu_l, \nu_h)$ but with $\nu > \nu_l^a$ and $\nu < \nu_h^a$, are willing to pay the higher dealer costs of agency trade, but not the cost associated with principal trade. Only customers with extreme types, in $\Omega_D^{o,p}$ or $\Omega_D^{n,p}$, are in a hurry to trade and are willing to pay the dealer costs for principal trade.

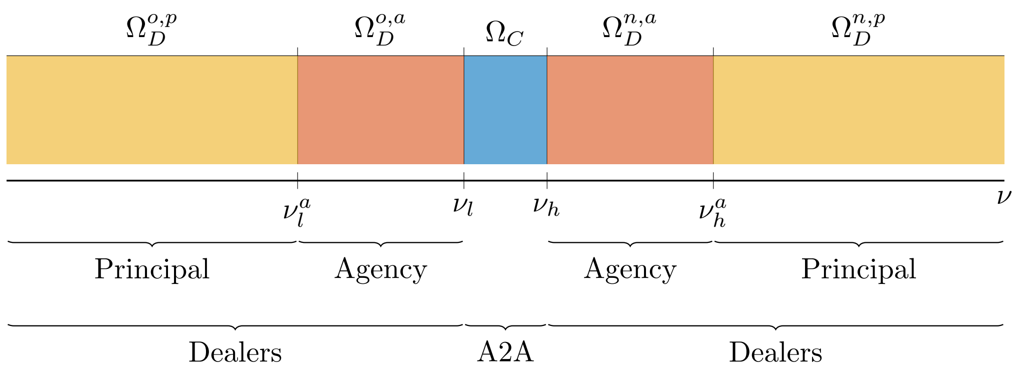


Figure 2: Partitions $\mathcal{P}$ and $\mathcal{P}_D$ in a regular equilibrium.

We impose $\nu_l < \nu_h$ when the intermediation cost is strictly positive. The reason is that we can always build an equilibrium where customers do not search for customers because they expect other customers to do the same. In this case, the probability of finding a customer is zero so customers may as well search for dealers. The assumption that $\nu_l < \nu_h$ if $c^l, c^s, \tau > 0$ rules out equilibria built on this sort of weak inequality. It is useful to first characterize the competitive limit and, in the next section, we characterize a regular equilibrium.

Proposition 1 (Competitive limit)**.** *If there are no search frictions, that is, if* $\lambda_D^0 = \lambda_D^1 = \lambda_C = +\infty$, *then the equilibrium satisfies the following:*

i. The equilibrium price is $p^* = \frac{\nu^*}{r+\mu+\eta}$, *where* $F(\nu^*) = \frac{\mu}{\mu+\eta}$.

ii. Customers type $\nu < \nu^*$ hold no asset.

iii. Customers type $\nu \geq \nu^*$ hold all assets.

With no search frictions, the price of the asset is equal to the present value of the valuation of the customer that is indifferent between owning and not owning the asset, ν^*. The valuation is discounted by the intertemporal discount r and the rates of maturity and issuance of the asset. Only customers with high valuation, $\nu > \nu^*$, own the asset. The value of ν^* is found such that the measure of high valuation customers is equal to the supply of assets. Customers with $\nu < \nu^*$ sell the asset instantaneously after issuance. Customers with $\nu \geq \nu^*$ buy the asset instantaneously after maturity.

4 Model solution

Consider first two limiting cases: one with high costs of principal trade, where all dealer intermediation occurs through agency trading, and another with high costs of agency, where all dealer intermediation occurs through principal trading.

4.1 Agency trading only

Let principal trade costs, c^l and c^s, be sufficiently high so that dealers engage in agency trading only. From equations (7) and (8), this implies $W^l = -\infty$, $W^s = \infty$, and $\Omega_D^{o,p} = \Omega_D^{n,p} = \emptyset$. The threshold types for the agency trade are then $\nu_l^a = -\infty$ and $\nu_h^a = \infty$. Dealers engage exclusively in agency risk-free trading.

A regular equilibrium has then two blocks. Given ν_l and ν_h, customers $\nu \in \Omega_D = (-\infty, \nu_l] \cup [\nu_h, \infty)$ engage in dealer-agency trades whereas customers $\nu \in \Omega_C = (\nu_l, \nu_h)$ trade bilaterally with other customers in all-to-all trades. We can solve these two blocks separately using the tools developed in Lagos and Rocheteau (2009) and HLW. The challenge is to characterize ν_l and ν_h that are consistent with the equilibrium equations (13) and (14). We need to find ν_l and ν_h such that customers searching for dealers, with $\nu \leq \nu_l$ or $\nu \geq \nu_h$, are not better off searching for other customers. Analogously, customers searching for customers with intermediary types $\nu_l < \nu < \nu_h$ are not better off searching for dealers.

4.1.1 Dealer-agency trading

The reservation value of a type-ν customer searching for a dealer is $\Delta(\nu) = V_D^o(\nu) - V_D^n(\nu)$. The value functions, $V_D^o(\nu)$ and $V_D^n(\nu)$, of a type-ν customer searching for a dealer when holding and not holding an asset are stated in equations (3)–(6). It is useful to define

$$\lambda_D^{al} = \frac{\lambda_D^0 \lambda_D^1}{r + \mu + \lambda_D^0 \theta_D + \lambda_D^1} \quad \text{and} \quad \lambda_D^{as} = \frac{\lambda_D^0 \lambda_D^1}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}. \tag{25}$$

Solving the four equations to isolate $\Delta(\nu)$ yields the following lemma.

Lemma 1. *Consider a regular equilibrium and the set of utility types $\Omega_D = (-\infty, \nu_l] \cup [\nu_h, +\infty)$. Then, the reservation value $\Delta(\nu)$ satisfies*

$$\Delta(\nu) = \begin{cases} \sigma_D^{al}[\nu + \lambda_D^{al}\theta_D(p - \tau)], & \nu \leq \nu_l, \\ \sigma_D^{as}[\nu + \lambda_D^{as}\theta_D(p + \tau)], & \nu \geq \nu_h, \end{cases} \tag{26}$$

where

$$\sigma_D^{al} = \frac{1}{r + \mu + \eta + \lambda_D^{al}\theta_D} \quad \textit{and} \quad \sigma_D^{as} = \frac{1}{r + \mu + \eta + \lambda_D^{as}\theta_D}. \tag{27}$$

Moreover, $\Delta(\nu_l) \leq p - \tau$ and $\Delta(\nu_h) \geq p + \tau$.

The derivative of the reservation value with respect to ν is given by σ_D^{al}, when $\nu \leq \nu_l$, and σ_D^{as}, when $\nu \geq \nu_h$. As HLW, we can interpret σ_D^{al} and σ_D^{as} as the local surplus at ν. It captures the trade surplus generated if the asset of an agent type ν is transferred to an agent type $\nu + d\nu$. For all $\nu \notin (\nu_l, \nu_h)$, the local surplus depends only on whether $\nu \leq \nu_l$ or $\nu \geq \nu_h$. That is because all customers type $\nu \leq \nu_l$ want to sell the asset and face the same price after bargaining and intermediation costs. As a result, the trade surplus is constant in this region. Similarly, all customers type $\nu \geq \nu_h$ want to buy the asset and face the same price after bargaining and intermediation costs.

The reason the trade surplus is different in the two regions is that customers type $\nu \leq \nu_l$ holding the asset to sell face the risk of the asset depreciating before it is sold in an agency trade, so they discount at the rate λ_D^{al}. Customers type $\nu \geq \nu_h$ waiting to buy an asset may have an issuance opportunity before they purchase the asset in an agency trade, so they discount at the rate λ_D^{as}.

σ_D^{al} and σ_D^{as} are indicators of market frictions. To see this, suppose that there are no search frictions for dealer intermediation. That is, $\lambda_D^0, \lambda_D^1 \to \infty$. In this case, $\sigma_D^{al} = \sigma_D^{as} = 0$ and so $\Delta(\nu)$ is constant in ν at $p - \tau$ or $p + \tau$. The reservation value is given by the competitive price plus or minus the intermediation cost. Higher values of σ_D^{al} and σ_D^{as} are then associated with higher search frictions. We later discuss an analogous measure of search frictions for the customer-customer market.

To establish the distributions of owners and non-owners Φ^o and Φ^n, we need to determine the distribution of dealers in agency trade, π_0^o and π_0^n. From Lemma 1, owners $\nu \leq \nu_l$ always sell in the interdealer market in agency trades and non-owners $\nu \geq \nu_h$ always buy. Then, (19) and (20) can be written as

$$-\left(\mu + \lambda_D^1\right) \pi_0^o(\nu) + \lambda_D^0 \left[\Phi^o(\nu) - \pi_0^o(\nu)\right] = 0, \ \nu \leq \nu_l,$$

$$-\left(\eta + \lambda_D^1\right) \pi_0^n(\nu) + \lambda_D^0 \left[\Phi^n(\nu) - \Phi^n(\nu_h) - \pi_0^n(\nu)\right] = 0, \ \nu \geq \nu_h.$$

Moreover, non-owners $\nu \leq \nu_l$ do not trade. It does not compensate for them searching for other customers because owners with reservation value below $\Delta(\nu)$ do not trade in the customer-customer market. It does not compensate for them buying an asset in agency because $\nu \leq \nu_l$ implies $\Delta(\nu) \leq p - \tau$, as established in lemma 1. The argument is analogous for owners of type $\nu \geq \nu_h$. Therefore, non-owners of type $\nu \leq \nu_l$ and owners of type $\nu \geq \nu_h$ do not trade. This leads to the following result.

Lemma 2 (Agency, dealer market)**.** *A regular equilibrium is such that the measures of owners and non-owners in agency trade satisfy*

$$\pi_0^o(\nu) = \begin{cases} \dfrac{\lambda_D^0 \Phi^o(\nu)}{\mu + \lambda_D^0 + \lambda_D^1}, & \nu \leq \nu_l, \\ \dfrac{\lambda_D^0 \Phi^o(\nu_l)}{\mu + \lambda_D^0 + \lambda_D^1}, & \nu > \nu_l, \end{cases} \tag{28}$$

$$\pi_0^n(\nu) = \begin{cases} 0, & \nu \leq \nu_l \\ \dfrac{\lambda_D^0 \left[\Phi^n(\nu) - \Phi^n(\nu_h)\right]}{\eta + \lambda_D^0 + \lambda_D^1}, & \nu \geq \nu_h. \end{cases} \tag{29}$$

Moreover, since there is no principal trade, $\pi_1 = \pi_{-1} = 0$.

We can now establish the distribution of owners and non-owners, Φ^o and Φ^n.

Lemma 3 (Distributions, dealer market)**.** *A regular equilibrium is such that the cumulative distribution of owners satisfies*

$$\Phi^o(\nu) = \begin{cases} \dfrac{\eta F(\nu)}{\mu+\eta+\tilde{\lambda}_D^{al}}, & \nu \leq \nu_l, \\ \dfrac{\eta}{\mu+\eta} - \dfrac{(\eta+\tilde{\lambda}_D^{as})[1-F(\nu)]}{\mu+\eta+\tilde{\lambda}_D^{as}}, & \nu \geq \nu_h, \end{cases} \tag{30}$$

where $\tilde{\lambda}_D^{al} = \frac{\lambda_D^0\lambda_D^1}{\mu+\lambda_D^0+\lambda_D^1}$, $\tilde{\lambda}_D^{as} = \frac{\lambda_D^0\lambda_D^1}{\eta+\lambda_D^0+\lambda_D^1}$, *and* $\Phi^o(\nu)+\Phi^n(\nu) = F(\nu)$. *Moreover,*

$$\frac{\eta\tilde{\lambda}_D^{al}F(\nu_l)}{\mu+\eta+\tilde{\lambda}_D^{al}} = \frac{\mu\tilde{\lambda}_D^{as}[1-F(\nu_h)]}{\mu+\eta+\tilde{\lambda}_D^{as}} \quad \textit{and} \quad \tilde{\lambda}_D^{al}\Phi^o(\nu_l)+\tilde{\lambda}_D^{as}\Phi^n(\nu_h) = \frac{\tilde{\lambda}_D^{as}\mu}{\mu+\eta}. \tag{31}$$

Define the separating utility type ν_s as the value of ν such that

$$\frac{\eta\tilde{\lambda}_D^{al}F(\nu_s)}{\mu+\eta+\tilde{\lambda}_D^{al}} = \frac{\mu\tilde{\lambda}_D^{as}[1-F(\nu_s)]}{\mu+\eta+\tilde{\lambda}_D^{as}}. \tag{32}$$

If $\tau = 0$, then $\nu_l = \nu_h = \nu_s$ and all customers trade with dealers. In this case, the marginal type ν_s separates the market into owners $\nu < \nu_s$ who want to sell the asset and non-owners $\nu \geq \nu_s$ who want to buy the asset.

4.1.2 Customer-customer trading

For customers searching for other customers, the value functions and reservation values are obtained in the following way. For the value functions, from equations (9) and (10), the value functions of a type $\nu \in \Omega_C$ customer holding or not holding an asset satisfy

$$rV_C^o(\nu) = \nu - \mu\Delta(\nu) + \lambda_C\int_{\nu}^{\nu_h}\theta_C^o[\Delta(\tilde{\nu})-\Delta(\nu)]d\Phi^n(\tilde{\nu}), \text{ and} \tag{33}$$

$$rV_C^n(\nu) = \eta\Delta(\nu) + \lambda_C\int_{\nu_l}^{\nu}\theta_C^n[\Delta(\nu)-\Delta(\tilde{\nu})]d\Phi^o(\tilde{\nu}). \tag{34}$$

Given the definition of reservation value, $\Delta(\nu) = V_C^o(\nu) - V_C^n(\nu)$, in (12), the equations above imply the following lemma.

Lemma 4 (Reservation value, customer-customer market)**.** *A regular equilibrium satisfies*

$$\Delta(\nu) = \Delta(\nu_l) + \int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu}) d\tilde{\nu} \tag{35}$$

for all $\nu \in (\nu_l, \nu_h)$*, and*

$$\sigma_C(\nu) = \frac{1}{r + \mu + \eta + \lambda_C \Big\{ \theta_C^o \big[\Phi^n(\nu_h) - \Phi^n(\nu)\big] + \theta_C^n \big[\Phi^o(\nu) - \Phi^o(\nu_l)\big] \Big\}}. \tag{36}$$

for almost all $\nu \in (\nu_l, \nu_h)$*.*

The trade surplus between a seller ν and buyer $\nu + d\nu$ is approximately equal to $\sigma_C(\nu)d\nu$, which agrees with the interpretation discussed in HLW of $\sigma_C(\nu)d\nu$ as the local surplus. The function $\sigma_C(\nu)$ discounts the additional utility $d\nu$ by the discount rate r, the likelihood that the asset will mature μ, the loss in the likelihood of issuing an asset η, and the loss in option value from either meeting another buyer with higher valuation $\lambda_C \theta_C^o \big[\Phi^n(\nu_h) - \Phi^n(\nu)\big]$, or finding another seller with lower valuation, $\lambda \theta_C^n \big[\Phi^o(\nu) - \Phi^o(\nu_l)\big]$.

We now turn to the distributions Φ^o, Φ^n among customers searching other customers.

Lemma 5 (Distributions, customer-customer market)**.** *A regular equilibrium is such that the cumulative distribution of owners satisfies*

$$\begin{aligned} \tilde{\Phi}^o(\nu) &= F(\nu) - F(\nu_l) - \tilde{\Phi}^n(\nu) \\ &= -\tfrac{\mu+\eta+\lambda_C[F(\nu_h)-F(\nu)-s_C]}{2\lambda_C} + \tfrac{\sqrt{\{\mu+\eta+\lambda_C[F(\nu_h)-F(\nu)-s_C]\}^2 + 4\lambda_C \eta \big[F(\nu)-F(\nu_l)\big]}}{2\lambda_C}, \end{aligned} \tag{37}$$

$\nu \in (\nu_l, \nu_h)$*, where* $\tilde{\Phi}^o(\nu) \equiv \Phi^o(\nu) - \Phi^o(\nu_l)$ *and* $\tilde{\Phi}^n(\nu) \equiv \Phi^n(\nu) - \Phi^n(\nu_l)$*, and*

$$s_C \equiv \Phi^o(\nu_h) - \Phi^o(\nu_l) = \frac{\eta}{\mu + \eta} \big[F(\nu_h) - F(\nu_l)\big]. \tag{38}$$

Figure 3 shows a representation for the reservation value as a function of ν. Lemma 1 implies that $\Delta(\nu)$ is linear for $\nu \leq \nu_l$ and $\nu \geq \nu_h$. Moreover, $\Delta(\nu_l) \leq p - \tau$ and $\Delta(\nu_h) \geq p + \tau$. That is, owners with $\nu \leq \nu_l$ choose sell to dealers and non-owners with $\nu \geq \nu_h$ choose to buy from dealers. Lemmas 4 and 5 imply the nonlinear shape of Δ in (ν_l, ν_h). Customers that trade with dealers have $\nu \leq \nu_l$. Customers that trade with other customers have $\nu \in (\nu_l, \nu_h)$.

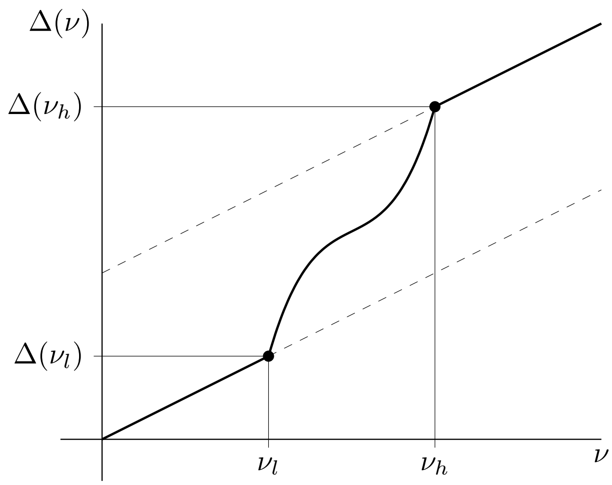


Figure 3: Reservation value as function of the customer type, $\Delta(\nu)$. Customers that trade with dealers have $\nu \leq \nu_l$. Customers that trade with other customers have $\nu \in (\nu_l, \nu_h)$.

4.1.3 Characterization

The results in Sections 4.1.1 and 4.1.2 establish necessary conditions for the equilibrium objects V^o, V^n, Δ, p, Φ^o, Φ^n, s, ν_l and ν_h. The equilibrium objects can all be written as functions of ν_l and ν_h. We now provide necessary conditions on ν_l and ν_h and show that, together with equations (5)–(38), these conditions are also sufficient for a regular equilibrium. These results provide a full characterization of the equilibrium.

Lemma 6. *A regular equilibrium satisfies*

$$2\tau\theta_D = \int_{\nu_l}^{\nu_h} w(\nu)\frac{\sigma_C(\nu) - \sigma_D^{as}}{\lambda_D^{as}\sigma_D^{as}} + [1 - w(\nu)]\frac{\sigma_C(\nu) - \sigma_D^{al}}{\lambda_D^{al}\sigma_D^{al}} d\nu, \tag{39}$$

where $w(\nu) = \frac{\theta_C^o\left[\Phi^n(\nu_h) - \Phi^n(\nu)\right]}{\theta_C^o\left[\Phi^n(\nu_h) - \Phi^n(\nu)\right] + \theta_C^n\left[\Phi^o(\nu) - \Phi^o(\nu_l)\right]}$. *Moreover,*

$$\begin{aligned} p &= \Delta(\nu_l) + \tau + \frac{\lambda_C\theta_C^o}{\lambda_D^{al}\theta_D}\int_{\nu_l}^{\nu_h}\int_{\nu_l}^{\nu}\sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^n(\nu) \\ &= \Delta(\nu_h) - \tau - \frac{\lambda_C\theta_C^n}{\lambda_D^{as}\theta_D}\int_{\nu_l}^{\nu_h}\int_{\nu}^{\nu_h}\sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^o(\nu). \end{aligned} \tag{40}$$

Lemmas 3 to 6 establish necessary conditions that are satisfied in all regular equilibria. In the proposition below, we show that these conditions are not only necessary but sufficient.

Therefore they fully characterize a regular equilibrium.

Proposition 2. *If a family* $\{V^o, V^n, \Delta, W^l, W^s, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *is a regular equilibrium, it satisfies equations* (26)–(40)*. Reversely, if* $\{\Delta, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *satisfies equations* (26)–(40)*, then* $\{V^o, V^n, \Delta, W^l, W^s, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *is a regular equilibrium where the value functions* W^l*,* W^s*,* V^o *and* V^n *are constructed using equations* (3)–(11)*.*

4.2 Principal trading only

Let now the agency trade cost τ be sufficiently high so that dealers undertake principal trades only. Equations (7) and (8) then imply that the threshold types for the agency trade are $\nu_l^a = \nu_l$ and $\nu_h^a = \nu_h$. Dealers and customers engage exclusively in principal trades and the sets of agency trades disappears, $\Omega_D^{o,a} = \Omega_D^{n,a} = \emptyset$.

4.2.1 Dealer-principal trading

The dealer reservation values can be obtained from equations (7) and (8). They yield

$$W^l = \frac{\lambda_D^1(p - \tilde{c}^l)}{r + \lambda_D^1} \quad \text{and} \quad W^s = \frac{\lambda_D^1(p + \tilde{c}^s)}{r + \lambda_D^1}, \tag{41}$$

where $\tilde{c}^l = \frac{c^l}{\lambda_D^1}$ and $\tilde{c}^s = \frac{c^s}{\lambda_D^1}$. The reservation value of a type-ν customer searching for a dealer is $\Delta(\nu) = V_D^o(\nu) - V_D^n(\nu)$. The value functions $V_D^o(\nu)$ and $V_D^n(\nu)$ of a type-ν customer searching for a dealer when holding or not holding an asset are stated in (3)–(6). Define

$$\lambda_D^p = \frac{\lambda_D^0 \lambda_D^1}{r + \lambda_D^1}. \tag{42}$$

Solving the four equations to isolate $\Delta(\nu)$ yields the following lemma.

Lemma 7. *Consider a regular equilibrium and the set of utility types* $\Omega_D = (-\infty, \nu_l] \cup [\nu_h, \infty)$*. Then, the reservation value* $\Delta(\nu)$ *satisfies*

$$\Delta(\nu) = \begin{cases} \sigma_D^p[\nu + \lambda_D^p \theta_D (p - \tilde{c}^l)], & \nu \le \nu_l, \\ \sigma_D^p[\nu + \lambda_D^p \theta_D (p + \tilde{c}^s)], & \nu \ge \nu_h, \end{cases} \tag{43}$$

where

$$\sigma_D^p = \frac{1}{r + \mu + \eta + \lambda_D^0 \theta_D}. \tag{44}$$

Moreover, $\Delta(\nu_l) \leq p - \tau$ *and* $\Delta(\nu_h) \geq p + \tau$*.*

The derivative of the reservation value with respect to ν is given by σ_D^p, when $\nu \leq \nu_l$, and σ_D^p, when $\nu \geq \nu_h$. The interpretation of σ_D^p and σ_D^p are analogous to the interpretation of σ_D^{al} and σ_D^{as}, as discussed in Section 4.1.

Before establishing the distributions of owners and non-owners Φ^o and Φ^n, we establish the distribution of the holdings of dealers, π_1 and π_{-1}. From Lemma 1, owners of type $\nu \leq \nu_l$ always sell in the interdealer market in principal trade, whereas non-owners of type $\nu \geq \nu_h$ always buy. Equations (17) and (18) can be rewritten as

$$-\lambda_D^1 \pi_1 + \lambda_D^0 \Phi^o(\nu_l) = 0, \quad \nu \leq \nu_l, \tag{45}$$

$$-\lambda_D^1 \pi_{-1} + \lambda_D^0 [1 - s - \Phi^n(\nu_h)] = 0, \quad \nu \geq \nu_h, \tag{46}$$

where $s = \Phi^o(\infty)$ are the asset holdings of owners. We then obtain the following result.

Lemma 8 (Principal, dealer market)**.** *A regular equilibrium is such that the distribution of dealers' holdings satisfies*

$$\pi_1 = \frac{\lambda_D^0 \Phi^o(\nu_l)}{\lambda_D^1} \quad \text{and} \quad \pi_{-1} = \frac{\lambda_D^0 [1 - s - \Phi^n(\nu_h)]}{\lambda_D^1}. \tag{47}$$

Moreover, since there is no agency trade, $\pi_0^o(\nu) = \pi_0^n(\nu) = 0$ *for all* ν*.*

We can now establish the distribution of owners and non-owners, Φ^o and Φ^n.

Lemma 9 (Distributions, dealer market)**.** *A regular equilibrium is such that the cumulative distributions of owners and non-owners satisfy*

$$\Phi^o(\nu) = \begin{cases} \dfrac{\eta F(\nu)}{\mu + \eta + \lambda_D^0}, & \nu \leq \nu_l, \\ \dfrac{\eta}{\mu + \eta} - \frac{(\eta + \lambda_D^0)[1 - F(\nu)]}{\mu + \eta + \lambda_D^0}, & \nu \geq \nu_h, \end{cases} \tag{48}$$

and $\Phi^o(\nu) + \Phi^n(\nu) = F(\nu)$*. Moreover,* ν_l *and* ν_h *satisfy*

$$\eta F(\nu_l) = \mu[1 - F(\nu_h)] \quad \text{and} \quad \Phi^o(\nu_l) + \Phi^n(\nu_h) = \frac{\mu}{\mu + \eta}. \tag{49}$$

Analogously to the previous section, define the separating utility type ν_s as

$$\eta F(\nu_s) = \mu[1 - F(\nu_s)] \Longrightarrow \nu_s = F^{-1}\left(\frac{\mu}{\eta + \mu}\right). \tag{50}$$

If $c^l = c^s = 0$, then $\nu_l = \nu_h = \nu_s$ and all customers trade with dealers. In this case, ν_s is the marginal type and separates the dealer-customer market into owners type $\nu < \nu_s$, who want to sell the asset, and non-owners type $\nu \geq \nu_s$, who want to buy the asset.

4.2.2 Customer-customer trading

The customer-customer market remains the same regardless of whether dealers engage in principal or agency trades. Therefore, we obtain results analogous to those for the case with agency only in Section 4.1.2.

Lemma 10 (Reservation value, customer-customer market)**.** *A regular equilibrium satisfies*

$$\Delta(\nu) = \Delta(\nu_l) + \int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu}) d\tilde{\nu} \tag{51}$$

for all $\nu \in (\nu_l, \nu_h)$*, and*

$$\sigma_C(\nu) = \frac{1}{r + \mu + \eta + \lambda_C \Big\{\theta_C^o \big[\Phi^n(\nu_h) - \Phi^n(\nu)\big] + \theta_C^n \big[\Phi^o(\nu) - \Phi^o(\nu_l)\big]\Big\}}. \tag{52}$$

for almost all $\nu \in (\nu_l, \nu_h)$*.*

Lemma 11 (Distributions, customer-customer market)**.** *A regular equilibrium is such that the cumulative distribution of owners satisfies*

$$\begin{aligned} \tilde{\Phi}^o(\nu) &= F(\nu) - F(\nu_l) - \tilde{\Phi}^n(\nu) \\ &= -\frac{\mu+\eta+\lambda_C[F(\nu_h)-F(\nu)-s_C]}{2\lambda_C} + \frac{\sqrt{\{\mu+\eta+\lambda_C[F(\nu_h)-F(\nu)-s_C]\}^2+4\lambda_C\eta\big[F(\nu)-F(\nu_l)\big]}}{2\lambda_C}, \end{aligned} \tag{53}$$

$\nu \in (\nu_l, \nu_h)$*, where* $\tilde{\Phi}^o(\nu) \equiv \Phi^o(\nu) - \Phi^o(\nu_l)$ *and* $\tilde{\Phi}^n(\nu) \equiv \Phi^n(\nu) - \Phi^n(\nu_l)$*, and*

$$s_C \equiv \Phi^o(\nu_h) - \Phi^o(\nu_l) = \frac{\eta}{\mu + \eta}\big[F(\nu_h) - F(\nu_l)\big]. \tag{54}$$

4.2.3 Characterization

The results in sections 4.2.1 and 4.2.2 establish necessary conditions for the equilibrium objects V^o, V^n, Δ, p, Φ^o, Φ^n, s, ν_l and ν_h. The equilibrium objects can all be written as functions of ν_l and ν_h. We now provide necessary conditions on ν_l and ν_h and show that, together with equations (53)–(54), these conditions are also sufficient for a regular equilibria. These results provide a full characterization of the equilibrium.

Lemma 12. *A regular equilibrium satisfies*

$$\lambda_D^p \theta_D (\tilde{c}^l + \tilde{c}^s) = \int_{\nu_l}^{\nu_h} \frac{\sigma_C(\nu) - \sigma_D^p}{\sigma_D^p} d\nu. \tag{55}$$

Moreover,

$$\begin{aligned} p &= \frac{\lambda_D^0}{\lambda_D^p}\Delta(\nu_l) + \tilde{c}^l + \frac{\lambda_C \theta_C^o}{\lambda_D^p \theta_D} \int_{\nu_l}^{\nu_h} \int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu}) d\tilde{\nu} d\Phi^n(\nu) \\ &= \frac{\lambda_D^0}{\lambda_D^p}\Delta(\nu_h) - \tilde{c}^s - \frac{\lambda_C \theta_C^o}{\lambda_D^p \theta_D} \int_{\nu_l}^{\nu_h} \int_{\nu}^{\nu_h} \sigma_C(\tilde{\nu}) d\tilde{\nu} d\Phi^o(\nu). \end{aligned} \tag{56}$$

Lemmas 9 to 12 establish necessary conditions that are satisfied in all regular equilibria. Analogously to the case with agency only, we show in the proposition below that these conditions are also sufficient. Therefore, they fully characterize a regular equilibrium.

Proposition 3. *If a family* $\{V^o, V^n, \Delta, W^l, W^s, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *is a regular equilibrium, it satisfies equations* (41)–(56)*. Reversely, if* $\{\Delta, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *satisfies equations* (41)–(56)*, then* $\{V^o, V^n, \Delta, W^l, W^s, p, \Phi^o, \Phi^n, \Pi, \mathcal{P}, \mathcal{P}_D\}$ *is a regular equilibrium where the value functions* V^o *and* V^n *are constructed using equations* (3)–(11)*.*

5 Principal vs agency trading

The decision to engage in agency or principal trading is based on the costs associated with each form of intermediation, the speed of trade, bargaining power, and customer valuation. Proposition 4 characterizes the regions where customers and dealers engage in principal or agency trade.

Proposition 4. *Consider a regular equilibrium. Owners type $\nu \leq \nu_l$ sell in principal trade when meeting a dealer if and only if*

$$\underbrace{p\theta_D(\sigma_D^p\lambda_D^p - \sigma_D^{al}\lambda_D^{al})}_{\substack{\text{Gain from payment} \\ \text{in fast execution}}} - \underbrace{\theta_D\sigma_D^p\lambda_D^p\tilde{c}^l}_{\substack{\text{Principal} \\ \text{trade cost}}} \geq \underbrace{(\sigma_D^{al} - \sigma_D^p)\nu}_{\substack{\text{Gain from value} \\ \text{in slow execution}}} - \underbrace{\theta_D\sigma_D^{al}\lambda_D^{al}\tau}_{\substack{\text{Agency} \\ \text{trade cost}}}. \tag{57}$$

Similarly, non-owners type $\nu \geq \nu_h$ buy in principal trade when meeting a dealer if and only if

$$\underbrace{(\sigma_D^{as} - \sigma_D^p)\nu}_{\substack{\text{Gain from value} \\ \text{in fast execution}}} - \underbrace{\theta_D\sigma_D^p\lambda_D^p\tilde{c}^s}_{\substack{\text{Cost of} \\ \text{principal trade}}} \geq \underbrace{p\theta_D(\sigma_D^p\lambda_D^p - \sigma_D^{as}\lambda_D^{as})}_{\substack{\text{Gain from payment} \\ \text{in slow execution}}} - \underbrace{\theta_D\sigma_D^{as}\lambda_D^{as}\tau}_{\substack{\text{Cost of} \\ \text{agency trade}}}. \tag{58}$$

Moreover, equation (57) *holds with equality at $\nu = \nu_l^a$ if $\nu_l^a < \nu_l$, and equation* (58) *holds with equality at $\nu = \nu_h^a$ if $\nu_h^a > \nu_h$.*

Equations (57) and (58) determine the choice between principal or agency trade. In equation (57), owners gain by selling assets fast with principal trade because they receive the payment earlier with its faster execution. As they own the asset, they also gain by holding the asset longer, which occurs with agency trade with its slower execution. This effect is more relevant as ν increases. Owners compare these gains net of the costs of principal and agency trade. As $\sigma_D^{al} > \sigma_D^p$, smaller values of ν make this equation more likely to be satisfied, which makes the choice of principal trade more likely. In (58), similarly, non-owners gain with principal trade because they will acquire the asset faster, but gain with agency trade because they delay payment. Higher values of ν makes principal trade more likely. We have agency trade for intermediary values of ν and principal trade for extreme values of ν.

Proposition 5. *Consider a regular equilibrium. The cumulative distribution of owners of* *type* $\nu \leq \nu_l$ *satisfies*

$$\Phi^o(\nu) = \begin{cases} \dfrac{\eta F(\nu)}{\mu + \eta + \lambda_D^0}, & \nu \leq \nu_l^a, \\ \dfrac{\eta F(\nu_l^a)}{\mu + \eta + \lambda_D^0} + \dfrac{\eta [F(\nu) - F(\nu_l^a)]}{\mu + \eta + \tilde{\lambda}_D^{al}}, & \nu_l^a < \nu \leq \nu_l, \end{cases} \tag{59}$$

and the cumulative distribution of non-owners of type $\nu \geq \nu_h$ *satisfies*

$$\Phi^n(\infty) - \Phi^n(\nu) = \begin{cases} \dfrac{\mu[1 - F(\nu_h^a)]}{\mu + \eta + \lambda_D^0} + \dfrac{\mu [F(\nu) - F(\nu_h)]}{\mu + \eta + \tilde{\lambda}_D^{as}}, & \nu_h \leq \nu < \nu_h^a, \\ \dfrac{\mu[1 - F(\nu)]}{\mu + \eta + \lambda_D^0}, & \nu \geq \nu_h^a, \end{cases} \tag{60}$$

where $\Phi^n(\infty) = 1 - s = \frac{\eta}{\mu+\eta}$*. Moreover, the market clearing condition implies*

$$\underbrace{\frac{\lambda_D^0 \eta F(\nu_l^a)}{\mu + \eta + \lambda_D^0}}_{\substack{\textit{Supply from} \\ \textit{principal}}} + \underbrace{\frac{\tilde{\lambda}_D^{al} \eta [F(\nu_l) - F(\nu_l^a)]}{\mu + \eta + \tilde{\lambda}_D^{al}}}_{\substack{\textit{Supply from} \\ \textit{agency}}} = \underbrace{\frac{\lambda_D^0 \mu [1 - F(\nu_h^a)]}{\mu + \eta + \lambda_D^0}}_{\substack{\textit{Demand from} \\ \textit{principal}}} + \underbrace{\frac{\tilde{\lambda}_D^{as} \mu [F(\nu_h^a) - F(\nu_h)]}{\mu + \eta + \tilde{\lambda}_D^{as}}}_{\substack{\textit{Demand from} \\ \textit{agency}}}. \tag{61}$$

Proposition 5 describes the distribution of owners and non-owners that trade with dealers, that is, with types $\nu \leq \nu_l$ or $\nu \geq \nu_h$. Owners $\nu \leq \nu_l^a$ choose principal trade, with measure determined by the arrival of dealer meetings, λ_D^0. Owners $\nu \in (\nu_l^a, \nu_l]$ choose agency trade, with measure determined by the arrival of dealer meetings combined with the arrival rate of access to the interdealer market $\tilde{\lambda}_D^{al}$. Similarly, non-owners $\nu \geq \nu_h^a$ choose principal trade, with measure determined by λ_D^0, and non-owners $\nu \in [\nu_h, \nu_h^a)$ choose agency trade, with measure determined by $\tilde{\lambda}_D^{as}$. The market clearing condition in (61) equates supply and demand of assets in the interdealer market, given by dealers carrying a position after principal trade and by those in agency trade.

6 Market composition and equilibrium multiplicity

The results above imply a procedure for solving for an equilibrium, which involves the determination of the equilibrium values of the thresholds ν_l^a, ν_l, ν_h, and ν_h^a, and the equilibrium price p of the asset. To facilitate the exposition, consider a symmetric economy.

Assumption 2. *An economy is symmetric if $\mu = \eta$ and the distribution of types ν is symmetric around its mean.*

In a symmetric economy, $\lambda_D^{as} = \lambda_D^{al} \equiv \lambda_D^a$ and so $\sigma_D^{al} = \sigma_D^{as} \equiv \sigma_D^a$. For the remainder of the paper we focus on this case. We also assume, for the current section, that $\tilde{c}^l = \tilde{c}^s \equiv \tilde{c}$.

According to lemmas 3 and 6, ν_h can be expressed as a function of $\nu_l \in (-\infty, \nu_s]$. Therefore, define the functions $G,\ H : (-\infty, \nu_s] \to \mathbb{R}$ as

$$G(\nu_l) = \frac{1}{2}\int_{\nu_l}^{g(\nu_l)} \frac{\sigma_C(\nu) - \sigma_D^a}{\lambda_D^a \theta_D \sigma_D^a} d\nu, \tag{62}$$

$$H(\nu_l) = \frac{1}{2}\int_{\nu_l}^{h(\nu_l)} \frac{\sigma_C(\nu) - \sigma_D^p}{\lambda_D^p \theta_D \sigma_D^p} d\nu. \tag{63}$$

$g(\nu_l)$ yields the value of ν_h such that the measure of dealers in agency trade that want to sell is equal to the measure of those that want to buy, and analogously for $h(\nu_l)$ for principal trade. In a symmetric economy, we have

$$g(\nu_l) = h(\nu_l) = F^{-1}[1 - F(\nu_l)], \quad \nu_l \in (-\infty, \nu_s]. \tag{64}$$

Note that G and H are strictly positive for $\nu_l < \nu_s$, and both are zero at $\nu_l = \nu_s$.

Suppose that the parameters are such all trade with dealers is made with agency. According to lemma 6, an equilibrium ν_l solves

$$\tau = G(\nu_l). \tag{65}$$

After obtaining ν_l, we obtain the other equilibrium variables through proposition 2. Analogously, if all trade with dealers is principal trade, then lemma 12 implies

$$\tilde{c} = H(\nu_l^a), \tag{66}$$

where $\nu_l^a = \nu_l$, and we obtain the other equilibrium variables through proposition 3.

Given the equilibrium price p, define the principal-agency condition line $q(\nu)$ by

$$q(\nu) = \frac{\sigma_D^p \lambda_D^p - \sigma_D^a \lambda_D^a}{\sigma_D^p} p + \frac{\sigma_D^a \lambda_D^a}{\sigma_D^p} \tau - \frac{\sigma_D^a - \sigma_D^p}{\theta_D \sigma_D^p} \nu. \tag{67}$$

This equation will help us implement proposition 4. The value of $q(\nu_l)$ is equal to the relative gain from principal trade relative to agency trade, not taking into account the cost of principal trade. It is a linear function of ν because customers evaluate the asset linearly and because the interdealer price p is obtained competitively. As $\sigma_D^a > \sigma_D^p$, the slope of q is negative. The gain of principal trade declines as the valuation of customers increases. Owners gain less by selling the asset fast as ν increases.

According to proposition 4, owners $\nu \leq \nu_l$ sell in principal trade if and only if

$$\lambda_D^p \tilde{c} \leq q(\nu). \tag{68}$$

If there is a value ν_l^a that satisfies this equation with equality such that $\nu_l^a < \nu_l$ then we have found an equilibrium ν_l^a for which principal and agency trades coexist. Owners $\nu \leq \nu_l^a$ engage in principal trade, agents $\nu_l^a < \nu \leq \nu_l$ engage in agency trade, and agents $\nu_l < \nu \leq \nu_s$ engage in A2A trade with other customers. There are analogous thresholds ν_h, ν_h^a for non-owners with $\nu \geq \nu_s$ who want to buy the asset. Agents select themselves into different trading methods.

If ν_l^a that satisfies (68) is such that $\nu_l^a > \nu_l$ then there is no agency trade. This situation is more likely if the cost of principal trade $\tilde{c}$ is low relative to the cost of agency trade τ. There is principal trade only. We use equation (66) to determine the equilibrium value of ν_l^a.

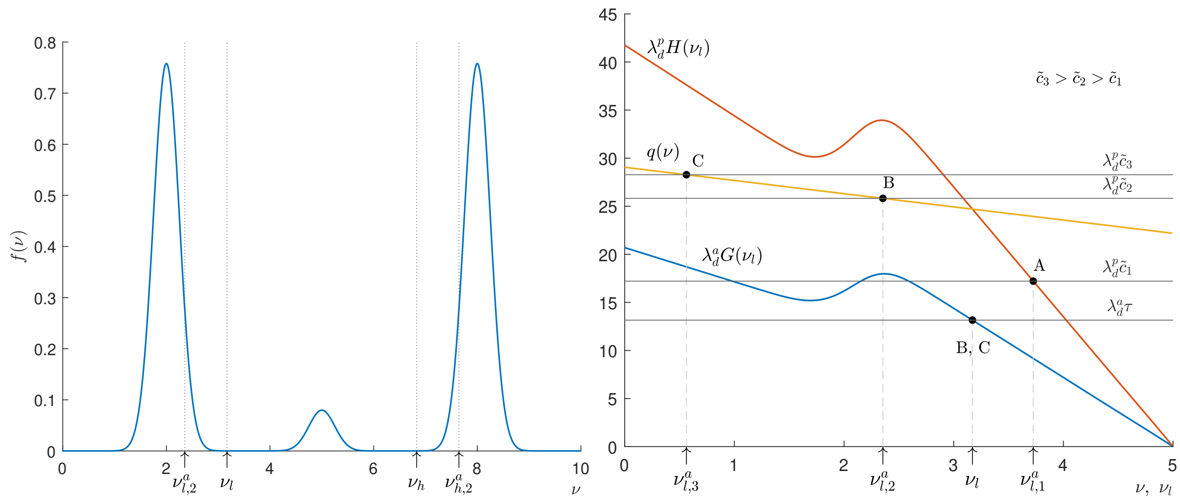
The equilibrium might not be unique. For small τ, equation (65) implies ν_l close to its highest possible value, ν_s, and unique. However, depending on the distribution of customer types, we can have multiple equilibria as agents choose the trading method depending on their expectations for the number of agents as counterparties.

Figures 4 and 5 show how to obtain the equilibrium. They show cases of unique and multiple equilibria. The meeting rates are such that it is faster to find a dealer than finding another customer in A2A markets, $\lambda_C = 1$ and $\lambda_0 = 5$, and that it is easier to find a customer with the help of the dealer, $\lambda_1 = 3 > \lambda_C$. Panel 4a shows the distribution of customers for $\nu \in [0, 10]$. There is a concentration of customers with low and high types.[10]

Panels 4b-5b show different equilibrium patterns for low, median, and high τ, and different values of $\tilde{c}$. We have unique equilibrium for low τ. We have multiple equilibrium in the agency

[10] The distribution is a combination of three normal distributions $N(\mu, \sigma)$, $f(\nu) = 0.475N(2, 0.25) + 0.05N(5, 0.25) + 0.475N(8, 0.25)$. Other parameters are $r = 0.05$, $\mu = \eta = 0.15$, and $\theta_D = \theta_C^n = \theta_C^o = 0.5$.

or principal markets for median and high τ depending on the value of $\tilde{c}$. The panels show the functions G and H multiplied by λ_D^a λ_D^p so that they can be shown in the same graph and compared with the agency and principal costs. We explain each case below.[11]



(a) Density $f(\nu)$ and thresholds of equilibrium B

(b) Unique equilibrium for low τ

Figure 4: Low agency cost τ and an increase in principal flow cost $\tilde{c}$ (right). Equilibrium A, with low $\tilde{c}$, has no agency trade as $\nu_l < \nu_l^a$. Equilibria B and C have coexistence of principal and agency trade. Agency trade increases as the cost of principal trade $\tilde{c}$ increases. The market composition changes.

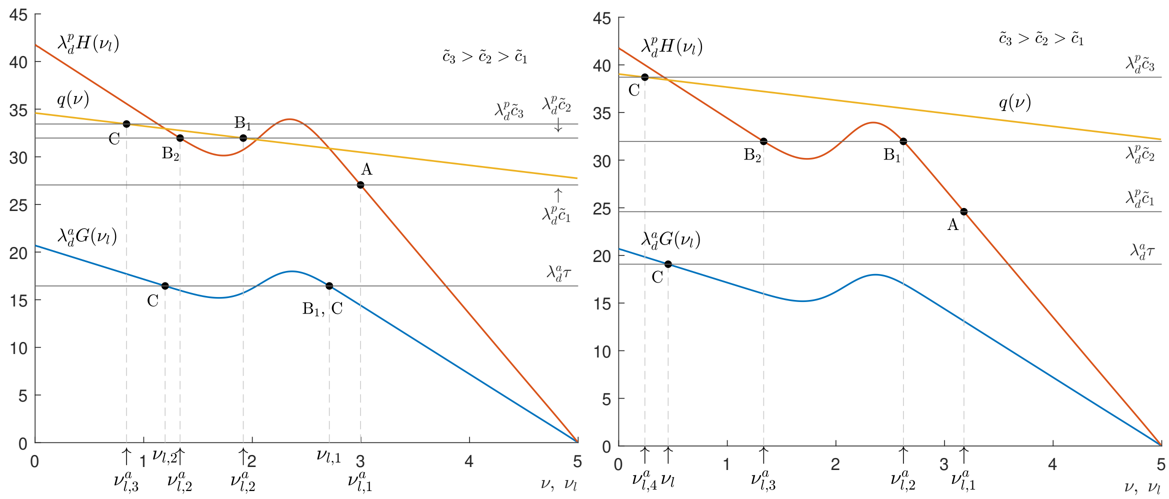
Panel 4b shows the case with low τ. Given τ and equation (65), an equilibrium ν_l is given by the intersection of $\lambda_D^a G$ and $\lambda_D^a \tau$. For $\tilde{c} = \tilde{c}_1$, the value of ν_l^a implied by the principal-agency condition line q is such that $\nu_l^a > \nu_l$. Therefore, the pair ν_l with this ν_l^a cannot be an equilibrium. The equilibrium is given by $\nu_{l,1}^a$, for which $\lambda_D^p H$ and $\lambda_D^p \tilde{c}_1$ intersect. This is equilibrium A. We have unique equilibrium with principal trade only.

As $\tilde{c}$ increases in panel 4b, we move to equilibria B and C. For these equilibria, $\nu_l^a < \nu_l$, where ν_l^a is such that $\lambda_D^p \tilde{c} = q(\nu_l^a)$. Agency trade coexists with principal trade in each case. Panel 4a shows the values of ν_l^a, ν_l, ν_h and ν_h^a of equilibrium B. We have principal trade for $\nu \leq \nu_l^a$, agency trade for $(\nu_l^a, \nu_l]$, A2A trade for (ν_l, ν_h), and again agency and principal trade for $[\nu_h, \nu_h^a)$ and $\nu \geq \nu_h^a$. As the cost of principal trade increases, the market for agency trade increases.

Panel 5a shows the case with median τ. We have multiple equilibrium in this case for agency trade, as $\lambda_D^a \tau$ crosses $\lambda_D^a G$ in multiple points. The thresholds $\nu_{l,1}$, $\nu_{l,2}$ determine

[11] The panels show $\nu \in (-\infty, \nu_s]$ and they determine ν_l and ν_l^a. There are symmetrical panels for $\nu \in [\nu_s, \infty)$ to determine ν_h and ν_h^a.

stable equilibria and we concentrate on these values (we discuss stability later in this section). Equilibrium A occurs with low principal cost $\tilde{c}_1$ and has unique equilibrium with principal trade only, similar to equilibrium A in panel 5b.



(a) Multiple equilibrium for median τ

(b) Multiple equilibrium for high τ

Figure 5: Multiple equilibrium with median and for high τ. Left: Cost $\tilde{c}_2$ implies two equilibria, one with principal and agency (B_1) and one with principal only (B_2). Cost $\tilde{c}_3$ has principal and two possible equilibria with agency trade. Right: Equilibrium B has multiple equilibria with principal only (no agency trade). The interaction between parameters imply different market structures. Principal trade can shrink abruptly in the case of an increase in $\tilde{c}$.

With $\tilde{c}_2$ in panel 5a, we have multiple equilibrium with agency or principal only. If the agency market is large, with $\nu_{l,1}$ at equilibrium B_1, then the principal-agency condition line q together with $\lambda_D^p \tilde{c}_2$ determine $\nu_{l,2}^a < \nu_{l,1}$, which is the equilibrium threshold for the principal market. The market has agency and principal trades in coexistence (in addition to the A2A trade for $\nu \in (\nu_{l,1}, \nu_{h,1})$). If the agency market is small, with $\nu_{l,2} < \nu_{l,1}$ as threshold, then q and $\lambda_D^p \tilde{c}_2$ would determine $\nu_{l,2}^a > \nu_{l,2}$, which rules out $\nu_{l,2}$ as equilibrium. The equilibrium $\nu_{l,2}^a$ is then determined by $\lambda_D^p H$ and q, at B_2, and has principal only. The market has two possible equilibria: B_1 with agency and principal, and B_2 with principal only. A change from B_1 to B_2 implies less revenues for dealers either from principal trade or from facilitating agency trades. It implies an abrupt loss of business for dealers.

As $\tilde{c}$ increases to $\tilde{c}_3$ in panel 5a, we have multiple equilibria for the agency market and unique equilibrium for the principal market, C. q and $\lambda_D^p \tilde{c}_3$ determine $\nu_{l,3}^a$, which is smaller than $\nu_{l,2}$ and $\nu_{l,1}$. These two values are then equilibrium thresholds for the agency market and $\nu^a_{l,3}$ is an equilibrium threshold for the principal market. If the equilibrium changes from $\nu_{l,1}$ and $\nu_{l,2}$ then the set of principal trades would not change, but the set of agency trades facilitated by dealers would shrink from $(\nu^a_{l,3}, \nu_{l,1}]$ to $(\nu^a_{l,3}, \nu_{l,2}]$. As the principal market remains constant but the agency market decreases, more customers would coordinate in the A2A market. Dealers would lose brokerage fees.

Panel 5b shows multiple equilibria in the principal market. τ is high and consequently the agency market is small. We have a unique ν_l. On the other hand, as $\tilde{c}$ increases, we can have unique equilibrium with principal only (A), multiple equilibria with principal only (B_1, B_2), and unique equilibrium with principal and agency (C). For equilibria B_1 and B_2, the principal-agency condition line q with $\lambda^p_D \tilde{c}_2$ implies $\nu^a_l > \nu_l$ which rules out the agency market. The equilibrium ν^a_l is then determined by $\lambda^p_D H$ and $\lambda^p_D \tilde{c}_2$, which implies two stable values for ν^a_l. We then have principal only but the volume of principal trades can decrease abruptly for A2A if the equilibrium changes from $\nu^a_{l,2}$ to $\nu^a_{l,3}$.

Multiple equilibria arise because G or H may not be monotone. The non-monotonicity occurs because more customers search for customers in A2A if they are convinced that others will follow this strategy. When they do so, the probability of matching is higher and the gain of searching in A2A increases. As the figures above show, multiple equilibrium occurs when intermediation costs are sufficiently high.

Proposition 6. *There exists a $\bar{\tau} > 0$ and $\bar{\tilde{c}} > 0$ such that a regular equilibrium is unique for any pair $(\tilde{c}, \tau) \in [0, \bar{\tilde{c}}) \times [0, \bar{\tau})$.*

The reinforcement of searching when others search (strategic complementarity) is not strong enough to generate multiplicity when intermediation costs are small. Assumption 1 implies that the technology of dealers is superior to A2A net of trading costs. If τ is small, no matter how many customers search for customers, it is still preferable to search for dealers. Multiplicity happens only if τ is large enough so that the measure of customers searching for customers affects the decision on the trading mode.

About stability, we argue that an equilibrium threshold ν is stable when it is determined in a region where G or H are decreasing. We focus on the argument for G, as it is the same for H.

An interpretation of G is that it is a proxy for the expected difference in valuation of an owner with ν_l and a non-owner with ν_h, both using A2A trade. A large G implies that a non-owner might need to pay a substantial amount to buy the asset. If $G(\nu_l) > \tau$, it is better to switch from A2A to dealer agency trade. A buyer might pay τ in an agency trade, but the total payment would still be smaller than the expected price to pay in A2A trade. A switch from A2A to dealer-agency implies an increase in ν_l and a smaller interval (ν_l, ν_h).

$G(\nu_l)$ decreases with ν_l if the valuation of agents that engage in A2A trades gets closer to each other as ν_l increases. This is the case in panel 5a for equilibria B_1 and C. An increase in τ decreases the gain of agency trades. The equilibrium ν_l would decrease and the set of A2A trades would increase. Similarly, an increase in λ_C makes A2A trades more effective. It implies a downward shift in G. For a constant τ, it would decrease ν_l and increase the set of dealer agency trades.

There is a point in between equilibria B_1 and C, however, that determines an additional equilibrium. G intersects with τ, but G is increasing in this region. This region includes the higher density of utility types shown in panel 5a. An increase in τ would increase the equilibrium ν_l. Dealer agency trades would increase with τ. Similarly, an increase in λ_C, would shift G downward and decrease the set of A2A trades.

These counterfactual effects are related with the instability of equilibrium for intermediary value of ν_l. For this equilibrium, suppose that a small set of agents to the left of ν_l switch their decisions from dealer agency trade to A2A trade. The set of agents in A2A would increase to $(\nu_l - \epsilon, \nu_h')$, where $\nu_h' = g(\nu_l - \epsilon)$. We would then have $G(\nu_l - \epsilon) < G(\nu_l) < \tau_2$, which implies that it is beneficial for an agent to the left of $\nu_l - \epsilon$ also to switch from dealer-agency to A2A. All agents to the left would behave in the same way, which would increase further the set of agents in CC trades, until the equilibrium with $\nu_{l,2}$ in C reached.

The equilibrium is stable for B_1 and C. A switch of a small set of agents to the left of $\nu_{l,2}$ from dealers to A2A would increase G. It would be better to return to trade with dealers. The same reasoning can be applied to a switch from A2A to dealers to the right of $\nu_{l,2}$ and also to the other stable equilibrium at $\nu_{l,1}$.

A stable equilibrium is therefore associated with a region where G is decreasing; and an unstable equilibrium with a region where G is increasing. In regions where G is decreasing, small perturbations in A2A or in dealer agency trades would make agents return to their previous decisions on the counterparty. In regions where G is increasing, such perturbations would make agents switch the trading counterparty permanently toward a new equilibrium. We then focus on regions where G is decreasing and, therefore, have stable equilibria.

7 The effect of market composition on liquidity

The US corporate bond market is undergoing a shift in how trades are executed. The framework in which dealers hold large inventories and engage in principal trade is moving towards agency and A2A trading. Different sources have documented this change. Choi et al. (2024) provide evidence for the shift toward agency trades. Using TRACE data, they document a significant increase in the fraction of agency trades in corporate bonds since 2011. Kargar et al. (2023) provide evidence of the shift toward all-to-all trades. Using MarketAxess data, they document that the platform represented about 21% of trade volume in corporate bonds by the third quarter of 2022.

Market participants in the US corporate bond market attribute this shift largely to two factors. The 2010 Dodd-Frank Act, which increased capital requirements and the costs of inventory, and the establishment of electronic bond markets, which made it easier for dealers to engage in agency trading and fostered the growth of A2A trading platforms.[12]

Financial regulations and the rise of electronic trading exert competing effects on market efficiency and liquidity. Financial regulations increase dealers' holding costs, which leads to longer execution time, wider bid-ask spreads, and lower trade volume. The rise of electronic trading increases access to counterparties, which leads to shorter execution times, tighter bid-ask spreads, and higher trade volume.

We use our model to understand the impact of these changes. In particular, we study the impact of changes in the parameters of the model on A2A trade speed λ_C, and on the cost of holding long positions c^l on dealers' net position and the bid-ask spread.

7.1 Dealers' net positions

A key indicator of dealer activity and market structure is the dealers' net positions. In the model, dealers can hold long (+1) or short (-1) positions in the asset after engaging in principal

[12] See `https://www.greenwich.com/press-release/primary-dealers-role-bond-market-evolving`.

trades. Dealers involved in agency trading are pure intermediaries and do not contribute to asset holdings. Therefore, the dealers' net positions is

$$\mathcal{H} = \pi_1 - \pi_{-1}, \tag{69}$$

where π_1 is the measure of dealers with a long position after buying the asset from a customer in a principal trade, and π_{-1} is the measure of dealers with a short position after selling the asset to a customer in a principal trade. A positive value of $\mathcal{H}$ implies that the dealers hold more assets in inventory than they have shorted.[13]

Proposition 7. *Consider a region with a unique regular equilibrium in which agency trade is active for both buying and selling bonds. Then:*

(a) The dealers' net position is decreasing in c^l.

(b) If $c^l = c^s$, the dealers' net position does not depend on λ_C.

When the cost for dealers to hold long positions c^l increases, then principal trades in which customers sell to dealers become less attractive. To see this, note that the set of customers selling to dealers in principal trades are those of type $\nu \leq \nu_l^a$. From Proposition 4,

$$\underbrace{p\theta_D(\sigma_D^p\lambda_D^p - \sigma_D^{al}\lambda_D^{al})}_{\substack{\text{Gain from payment} \\ \text{in fast execution}}} - \underbrace{\theta_D\sigma_D^p\lambda_D^p\frac{c^l}{\lambda_D^1}}_{\substack{\text{Principal} \\ \text{trade cost}}} = \underbrace{(\sigma_D^{al} - \sigma_D^p)\nu_l^a}_{\substack{\text{Gain from value} \\ \text{in slow execution}}} - \underbrace{\theta_D\sigma_D^{al}\lambda_D^{al}\tau}_{\substack{\text{Agency} \\ \text{trade cost}}}. \tag{70}$$

An increase in c^l reduces the left-hand side of this equation. As a result, the equilibrium ν_l^a decreases, which increases the size of the agency market. This shift leads to a decrease in π_1 and a decrease in the net position $\mathcal{H}$. Notice that λ_C does not affect (70). Therefore, if the agency market is active, then an increase in trading speed in A2A (increase in λ_C) would have a small impact on the dealers' net position. It would only impact indirectly through σ_C and G.

In a symmetric economy, an improvement in trading speed A2A does not affect at all the dealers' net positions (part b of the proposition). An increase in λ_C enhances the efficiency

[13]Net positions capture the overall directional exposure of dealers, but similar results can be derived for the gross positions of dealers in our model.

of A2A trades, which reduces the volume of customer-dealer interactions as more customers choose direct trading. In a symmetric economy, this reduction affects the magnitude of dealers' intermediation equally on both the buying and selling sides, which leaves the net directional exposure, $\mathcal{H} = \pi_1 - \pi_{-1}$, unchanged. Therefore, although λ_C influences the scale of dealer activity, it does not change the balance between long and short positions.

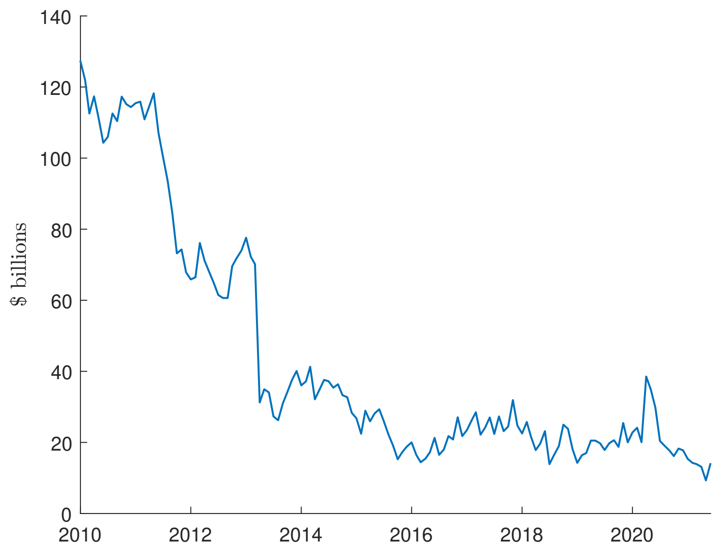


Figure 6: Primary dealers median net monthly position in corporate debt. Data: Federal Reserve Bank of New York, Primary Dealer Statistics.

The model predicts a decline in the dealers' net positions as principal trade costs increase. This prediction agrees with observed trends in the US corporate bond market. Figure 6 shows the median monthly net position of primary dealers in corporate debt instruments, which exceeded \$100 billion prior to 2011 and reflects a market structure heavily reliant on dealer inventories for principal trades. Following the implementation of post-2008 regulations, the net position decreased significantly to less than \$40 billion between 2013 and 2020, and has recently fallen below \$20 billion. This decline is consistent with an increase in c^l as higher capital and regulatory costs discourage dealers from holding large inventories, shifting intermediation toward agency and all-to-all trading as predicted by proposition 7.

7.2 Bid-ask spread

Bid-ask spread in principal trade Define $p_{D,P}^{buy}$ as the average price paid by a customer in principal trade when buying an asset from a dealer (the ask price) and p_D^{sell} as the average price received by a customer in principal trade when selling an asset to the dealer (the bid price). We define the bid-ask spread in principal trade BA_P as the average difference between how much customers pay when buying an asset and how much customers receive when selling an asset. The bid-ask spread in principal trade is then

$$\text{BA}_P = p_{D,P}^{buy} - p_{D,P}^{sell}. \tag{71}$$

In equilibrium, we have

$$\begin{aligned}
p_{D,P}^{buy} &= \frac{\int_{\nu_h^a}^{\infty}[\theta_D W^s + (1-\theta_D)\Delta(\nu)]d\Phi^n(\nu)}{\Phi^n(\infty) - \Phi^n(\nu_h^a)} = \frac{\int_{\nu_h^a}^{\infty}\left[\theta_D \frac{\lambda_D^1(p+\tilde{c}^s)}{r+\lambda_D^1} + (1-\theta_D)\Delta(\nu)\right]d\Phi^n(\nu)}{\Phi^n(\infty) - \Phi^n(\nu_h^a)} \\
&= (1-\theta_D)\mathbb{E}_{\Phi^n}\left[\Delta(\nu)|\nu \geq \nu_h^a\right] + \theta_D \frac{\lambda_D^1(p+\tilde{c}^s)}{r+\lambda_D^1}. \qquad (72) \\
p_{D,P}^{sell} &= \frac{\int_{-\infty}^{\nu_l^a}[\theta_D W^l + (1-\theta_D)\Delta(\nu)]d\Phi^o(\nu)}{\Phi^o(\nu_l^a)} = \frac{\int_{-\infty}^{\nu_l^a}\left[\theta_D \frac{\lambda_D^1[p-\tilde{c}^l]}{r+\lambda_D^1} + (1-\theta_D)\Delta(\nu)\right]d\Phi^o(\nu)}{\Phi^o(\nu_l^a)} \\
&= (1-\theta_D)\mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu \leq \nu_l^a\right] + \theta_D \frac{\lambda_D^1[p-\tilde{c}^l]}{r+\lambda_D^1}, \qquad (73)
\end{aligned}$$

Substituting (72) and (73) into (71), we get

$$\text{BA}_P = (1-\theta_D)\left\{\mathbb{E}_{\Phi^n}\left[\Delta(\nu)|\nu \geq \nu_h^a\right] - \mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu \leq \nu_l^a\right]\right\} + \theta_D \frac{c^s + c^l}{r+\lambda_D^1}, \tag{74}$$

where we used that $\tilde{c}^s = c^s/\lambda_D^1$ and $\tilde{c}^l = c^l/\lambda_D^1$.

Bid-ask spread in agency trade Analogously, the bid-ask spread in agency trade BA_A is defined as

$$\text{BA}_A = p_{D,A}^{buy} - p_{D,A}^{sell}. \tag{75}$$

In an agency trade, the dealer connects an owner and a buyer in the interdealer market. When we write the value function of owners searching for dealers, Equation (3), we assume that the owner pays the dealer a fraction $1-\theta^D$ of the gains from this service when they meet. In practice, the payment should be made when the trade is executed, not when the customer meets the dealer. To disentangle this payment, suppose that the dealer charges a fee $f^o(\nu)$ to the owner of type ν so the final price in an agency trade is $p - \tau - f^o(\nu)$. The same applies to non-owners when buying an asset. In this case, the final price in an agency trade is $p + \tau + f^n(\nu)$. The fees $f^o(\nu)$ and $f^n(\nu)$ are formally defined implicitly by

$$\theta_D[\tilde{V}_D^o(\nu) - V_D^o(\nu)] = \tilde{\tilde{V}}_D^o(\nu) - V_D^o(\nu) \quad \text{and} \quad \theta_D[\tilde{V}_D^n(\nu) - V_D^n(\nu)] = \tilde{\tilde{V}}_D^n(\nu) - V_D^n(\nu),$$

where

$$r\tilde{\tilde{V}}_D^o(\nu) = \nu - \mu[\tilde{\tilde{V}}_D^o(\nu) - V^n(\nu)] + \lambda_D^1 \max\{p - \tau - f^o(\nu) - [\tilde{\tilde{V}}_D^o(\nu) - V^n(\nu)], 0\},$$
$$r\tilde{\tilde{V}}_D^n(\nu) = \eta[V^o(\nu) - \tilde{\tilde{V}}_D^n(\nu)] + \lambda_D^1 \max\{V^o(\nu) - \tilde{\tilde{V}}_D^n(\nu) - (p + \tau + f^n(\nu)), 0\}.$$

The above equalities guarantee that the fees $f^o(\nu)$ and $f^n(\nu)$ are consistent with the bargaining between the dealer and the customer, so the surplus share of customers is θ_D.

In equilibrium, the fees paid to the dealer upon successful execution of an agency trade are

$$f^o(\nu^o) = (1 - \theta_D^o)[p - \tau - \Delta(\nu^o)] \quad \text{and} \quad f^n(\nu^n) = (1 - \theta_D^n)[\Delta(\nu^n) - p - \tau] \tag{76}$$

by an owner of type $\nu^o \in \Omega_D^{o,a}$ and non-owner of type $\nu^n \in \Omega_D^{n,a}$, where

$$\theta_D^o = \frac{\theta_D(\lambda_D^0 + r + \mu + \lambda_D^1)}{r + \mu + \lambda_D^0\theta_D + \lambda_D^1} \quad \text{and} \quad \theta_D^n = \frac{\theta_D(\lambda_D^0 + r + \eta + \lambda_D^1)}{r + \eta + \lambda_D^0\theta_D + \lambda_D^1}.$$

We are now in a position to compute $p_{D,A}^{sell}$ and $p_{D,A}^{buy}$. In a regular equilibrium, these are

$$p_{D,A}^{sell} = \frac{\int_{\nu_l^a}^{\nu_l}[p - \tau - f^o(\nu)]d\Phi^o(\nu)}{\Phi^o(\nu_l) - \Phi^o(\nu_l^a)} = \theta_D^o(p - \tau) + (1 - \theta_D^o)\mathbb{E}_{\Phi^o}[\Delta(\nu)|\nu_l^a \le \nu \le \nu_l], \tag{77}$$
$$p_{D,A}^{buy} = \frac{\int_{\nu_h}^{\nu_h^a}[p + \tau + f^n(\nu)]d\Phi^n(\nu)}{\Phi^n(\nu_h^a) - \Phi^n(\nu_h)} = \theta_D^n(p + \tau) + (1 - \theta_D^n)\mathbb{E}_{\Phi^n}[\Delta(\nu)|\nu_h \le \nu \le \nu_h^a]. \tag{78}$$

Substituting (77) and (78) into (75), we get

$$\begin{aligned} \text{BA}_A = {} & (\theta_D^n - \theta_D^o)p + (\theta_D^n + \theta_D^o)\tau \\ & + (1 - \theta_D^n)\mathbb{E}_{\Phi^n}[\Delta(\nu)|\nu_h \le \nu \le \nu_h^a] - (1 - \theta_D^o)\mathbb{E}_{\Phi^o}[\Delta(\nu)|\nu_l^a \le \nu \le \nu_l]. \end{aligned} \tag{79}$$

Proposition 8 (Agency vs principal trading prices)**.** *Consider a region with a unique regular* *equilibrium in which agency trade is active for both buying and selling bonds. Then, the following relations holds.*

(a) The average price customers pay to buy an asset in an agency trade, $p_{D,A}^{buy}$*, is lower than the average price they pay in a principal trade,* $p_{D,P}^{buy}$*.*

(b) The average price customers receive to sell an asset in an agency trade, $p_{D,A}^{sell}$*, is higher than the average price they receive in a principal trade,* $p_{D,P}^{sell}$*.*

(c) The bid-ask spread charged by dealers in agency trades, BA_A*, is smaller than the bid-ask spread charged by dealers in principal trades,* BA_P*.*

Proposition 8 highlights the trade-off between trade speed and trade cost of agency and principal trading. Parts (a) and (b) show that agency trading offers better prices: buyers pay a lower average price in agency trades ($p_{D,A}^{buy} < p_{D,P}^{buy}$), while sellers receive a higher average price ($p_{D,A}^{sell} > p_{D,P}^{sell}$). Principal trading offers immediacy as customers can trade on spot with the dealer, but this convenience comes at a price premium for buyers or a discount for sellers. Agency trading, although slower, offers better prices because the dealer acts purely as an intermediary, bearing no inventory cost. Part (c) of the proposition formalizes this price difference as a smaller bid-ask spread with agency trading ($\mathrm{BA}_A < \mathrm{BA}_P$). These predictions on the favorable prices of agency trading agree with the evidence in Choi et al. (2024).

Bid-ask spread Having established the differences between prices in agency and principal trading, we now turn to the overall bid-ask spread in the market, which combines the effects of both trading methods. We define the bid-ask spread BA as the weighted average of the bid-ask spreads in principal and agency trading,

$$\mathrm{BA} = \omega_P \, \mathrm{BA}_P + \omega_A \, \mathrm{BA}_A, \tag{80}$$

where

$$\omega_P = \frac{\lambda_D^0[\Phi^o(\nu_l^a) + 1 - s - \Phi^n(\nu_h^a)]}{\lambda_D^0[\Phi^o(\nu_l^a) + 1 - s - \Phi^n(\nu_h^a)] + \lambda_D^1[\pi_0^o(\nu) - \pi_0^o(\nu_l^a) + \pi_0^n(\nu_h^a) - \pi_0^n(\nu)]}, \tag{81}$$

$$\omega_A = 1 - \omega_P = \frac{\lambda_D^1[\pi_0^o(\nu_l) - \pi_0^o(\nu_l^a) + \pi_0^n(\nu_h^a) - \pi_0^n(\nu_h)]p_{D,A}^{buy}}{\lambda_D^0[\Phi^o(\nu_l^a) + 1 - s - \Phi^n(\nu_h^a)] + \lambda_D^1[\pi_0^o(\nu_l) - \pi_0^o(\nu_l^a) + \pi_0^n(\nu_h^a) - \pi_0^n(\nu_h)]}. \tag{82}$$

Proposition 9. *Consider a region with a unique regular equilibrium. Then the bid-ask spread* BA *satisfies the following:*

(a) For sufficiently small c^l and c^s, BA *is increasing in c^l.*

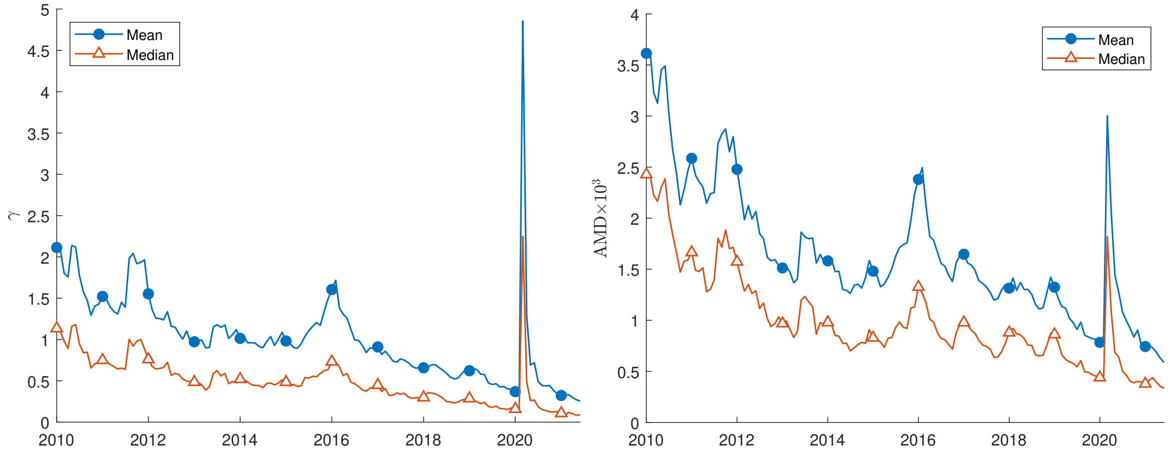
(b) There exists a region of τ, c^l and c^s where BA *is decreasing in c^l.*

(c) BA *is increasing in λ_C, the search efficiency in the customer-customer market, if $c^l = c^s$.*

The bid-ask spread of the dealers BA is a weighted average of the principal-trade spread BA_P and the agency-trade spread BA_A. An increase in the inventory cost c^l affects BA directly and indirectly. The direct effect is to increase principal trade costs BA_P. The indirect effect is to alter trade composition by increasing the weight of the agency bid-ask spread BA_A.

When inventory costs are low, principal trades dominate the market because its immediacy compensates the higher costs with comparison with agency. The direct effect prevails as it impacts the majority of the market. In contrast, when inventory costs are high, agency trades dominate the market. The indirect effect becomes more significant as the lower-cost agency trades exert a greater influence on the overall spread. The BA spread might decrease even though a trading method with higher search frictions prevails.

A surprising result is that the search efficiency in the customer-customer market, λ_C, always increases the bid-ask spread. An increase in λ_C enhances the efficiency of the A2A market, which reduces the incentive for customers to seek dealers. As a result, the A2A interval (ν_l, ν_h) increases, as ν_l decreases and ν_h increases. This adjustment amplifies the difference in the reservation values for both the principal and the agency trading. As dealer transactions increasingly involve high-valuation buyers and low-valuation sellers, then BA_P and BA_A increase. This effect implies an increase in the overall bid-ask spread BA.



(a) γ (b) AMD

Figure 7: Illiquidity measures: γ and AMD.

Figure 7 shows two widely used transaction-cost measures: the γ measure (Bao et al. 2011), and the Amihud (AMD) measure (Amihud 2002), both detailed in appendix B. Both measures have consistently declined since 2010, which suggests an increase in dealers' inventory costs c^l through the framework of the model.

The rise in c^l, which we associate with post-2008 regulations such as the Volcker Rule, implies a shift from principal to agency trades. This shift aligns with the observation that liquidity provision has been made more frequently by customers rather than dealers, as shown in Choi et al. (2024). The increase in customer liquidity provision coincides with regulatory cost increases and agrees with the predictions of the model.

As c^l increases, inventory expenses increase and the principal bid-ask spread BA_P increases (eq. 74), which shifts customers towards agency trading and increases the weight of ω_A on the overall BA spread (eq. 80). A substantial shift amplifies the influence of BA_A, which reduces the observed BA as agency trades, with lower spreads, dominate. Yet, the bid-ask spread BA_P, for the immediacy of principal trading, increases sharply with c^l. This increase is consistent with the 40%–60% cost increase for unmatched dealer-to-customer trades post-2012 documented by Choi et al. (2024). The contrast between lower spreads in agency trades versus sharply higher costs for dealer-provided immediacy reflects market restructuring induced by regulatory changes. Customers assume greater liquidity provision and immediacy is supplied as a premium service.

Although we posit that the search efficiency in the customer-customer market, λ_C, also increased post-2008, we do not observe the rise in dealers' bid-ask spreads (BA) predicted by Proposition 9. We conjecture three potential explanations for this discrepancy. First, the compositional shift toward agency trades, driven by rising c^l as discussed earlier, may have outweighed the effect of a higher λ_C. Second, concurrent improvements in trading technology for both customers and dealers could have mitigated a shift from dealer intermediation to A2A platforms, stabilizing spreads. Finally, Proposition 9 assumes a unique equilibrium, whereas multiple equilibria may exist; the economy might be locked in one with a higher proportion of dealer-intermediated trades, as explored in Section 6.

8 Conclusions

We propose a model of over-the-counter markets which recognizes that the execution of trades takes different forms. In practice, customers can pay a premium for immediacy with principal-inventory trading, wait longer to complete a trade with agency trading, or use electronic platforms in A2A trading. Dealers participate in trading with inventory or having agency to trade on behalf of customers. Each form of trading has its own characteristics in terms of efficiency and costs. We incorporate these features in a model where customers and dealers choose in equilibrium the optimal trading method.

We show that we can have coexistence of different trading methods in equilibrium. Higher trading costs are associated with multiple equilibria. An increase in inventory costs are associated with a decrease in dealers' net positions. An improvement of the search technology of A2A markets is associated with an increase in the bid-ask spread. An increase in inventory costs increases the bid-ask spread for low initial costs, but might eventually decrease the bid-ask spread as the composition of the market shifts from principal to agency trading.

Our results connect the technological and regulatory framework with over-the-counter markets through the composition of trading. The model can be used to make predictions about different changes in over-the-counter markets. It has been shown that the liquidity premium has changed after the Dodd-Frank regulations (Li and Yu 2023, Wu 2024). We use a version of the model to study the liquidity premium in Dyskant et al. (2025). The model can also be used to obtain the consequences of regulatory changes on welfare. Another result to explore is the measure of illiquidity based on the consumer surplus. This measure takes into account search frictions and the prevalence of principal, agency, and A2A trading.

References

Afonso, Gara, and Ricardo Lagos, 2015, Trade dynamics in the market for federal funds, *Econometrica* 83, 263–313.

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

An, Yu, and Zeyu Zheng, 2023, Immediacy provision and matchmaking, *Management Science* 69, 1245–1263.

Anderson, Mike, and René M. Stulz, 2017, Is post-crisis bond liquidity lower?, Working Paper 23317, National Bureau of Economic Research.

Bao, Jack, Maureen O'Hara, and Xing (Alex) Zhou, 2018, The Volcker rule and corporate bond market making in times of stress, *Journal of Financial Economics* 130, 95–113.

Bao, Jack, Jun Pan, and Jiang Wang, 2011, The illiquidity of corporate bonds, *The Journal of Finance* 66, 911–946.

Bessembinder, Hendrik, Stacey Jacobsen, William Maxwell, and Kumar Venkataraman, 2018, Capital commitment and illiquidity in corporate bonds, *The Journal of Finance* 73, 1615–1661.

Bethune, Zachary, Bruno Sultanum, and Nicholas Trachter, 2022, An information-based theory of financial intermediation, *The Review of Economic Studies* 89, 2381–2444.

Choi, Jaewon, Yesol Huh, and Sean Seunghun Shin, 2024, Customer liquidity provision: Implications for corporate bond transaction costs, *Management Science* 70, 187–206.

Cimon, David, and Corey Garriott, 2019, Banking regulation and market making, *Journal of Banking & Finance* 109, 105653.

Cohen, Assa, Mahyar Kargar, Benjamin Lester, and Pierre-Olivier Weill, 2024, Inventory, market making, and liquidity in OTC markets, *Journal of Economic Theory* 222, 105917.

Dick-Nielsen, Jens, 2009, Liquidity biases in TRACE, *The Journal of Fixed Income* 19, 43–55.

Dick-Nielsen, Jens, 2014, How to clean Enhanced TRACE data, Working Paper.

Dick-Nielsen, Jens, Peter Feldhütter, and David Lando, 2012, Corporate bond liquidity before and after the onset of the subprime crisis, *Journal of Financial Economics* 103, 471–492.

Dick-Nielsen, Jens, and Marco Rossi, 2018, The cost of immediacy for corporate bonds, *The Review of Financial Studies* 32, 1–41.

Duffie, Darrell, 2012, Market making under the proposed Volcker rule, Report to the Securities Industry and Financial Markets Association, Stanford University.

Duffie, Darrell, 2018, Financial regulatory reform after the crisis: An assessment, *Management Science* 64, 4835–4857.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen, 2005, Over-the-counter markets, *Econometrica* 73, 1815–1847.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen, 2007, Valuation in over-the-counter markets, *The Review of Financial Studies* 20, 1865–1900.

Dyskant, Lucas B., Andre C. Silva, and Bruno Sultanum, 2025, The Volcker rule and the liquidity premium in corporate bond markets, Working Paper.

Falato, Antonio, Ali Hortaçsu, Dan Li, and Chaehee Shin, 2021, Fire-sale spillovers in debt markets, *The Journal of Finance* 76, 3055–3102.

Farboodi, Maryam, Gregor Jarosch, and Robert Shimer, 2023, The emergence of market structure, *The Review of Economic Studies* 90, 261–292.

Gehrig, Thomas, 1993, Intermediation in search markets, *Journal of Economics & Management Strategy* 2, 97–120.

Goldstein, Itay, Hao Jiang, and David T. Ng, 2017, Investor flows and fragility in corporate bond funds, *Journal of Financial Economics* 126, 592–613.

Guerrieri, Veronica, Robert Shimer, and Randall Wright, 2010, Adverse selection in competitive search equilibrium, *Econometrica* 78, 1823–1862.

Hugonnier, Julien, Benjamin Lester, and Pierre-Olivier Weill, 2020, Frictional intermediation in over-the-counter markets, *The Review of Economic Studies* 87, 1432–1469.

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

Jiang, Hao, Yi Li, Zheng Sun, and Ashley Wang, 2022, Does mutual fund illiquidity introduce fragility into asset prices? evidence from the corporate bond market, *Journal of Financial Economics* 143, 277–302.

Kargar, Mahyar, Benjamin Lester, David Lindsay, Shuo Liu, Pierre-Olivier Weill, and Diego Zúñiga, 2021, Corporate bond liquidity during the covid-19 crisis, *The Review of Financial Studies* 34, 5352–5401.

Kargar, Mahyar, Benjamin Lester, Sébastien Plante, and Pierre-Olivier Weill, 2023, Sequential search for corporate bonds, Working Paper.

Lagos, Ricardo, and Guillaume Rocheteau, 2009, Liquidity in asset markets with search frictions, *Econometrica* 77, 403–426.

Lester, Benjamin, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones, 2023, Market-making with search and information frictions, *Journal of Economic Theory* 212, 105714.

Li, Jian, and Haiyue Yu, 2023, Investor composition and the liquidity component in the U.S. corporate bond market, Working Paper.

Mahanti, Sriketan, Amrut Nashikkar, Marti Subrahmanyam, George Chacko, and Gaurav Mallik, 2008, Latent liquidity: A new measure of liquidity, with an application to corporate bonds, *Journal of Financial Economics* 88, 272–298.

Musto, David, Greg Nini, and Krista Schwarz, 2018, Notes on bonds: Illiquidity feedback during the financial crisis, *The Review of Financial Studies* 31, 2983–3018.

Rust, John, and George Hall, 2003, Middlemen versus market makers: A theory of competitive exchange, *Journal of Political Economy* 111, 353–403.

Saar, Gideon, Jian Sun, Ron Yang, and Haoxiang Zhu, 2023, From market making to matchmaking: Does bank regulation harm market liquidity?, *The Review of Financial Studies* 36, 678–732.

Spulber, Daniel F., 1996, Market making by price-setting firms, *The Review of Economic Studies* 63, 559–580.

Üslü, Semih, 2019, Pricing and liquidity in decentralized asset markets, *Econometrica* 87, 2079–2140.

Vayanos, Dimitri, and Tan Wang, 2007, Search and endogenous concentration of liquidity in asset markets, *Journal of Economic Theory* 136, 66–104.

Weill, Pierre-Olivier, 2007, Leaning against the wind, *The Review of Economic Studies* 74, 1329–1354.

Wu, Botao, 2024, Post-crisis regulations, trading delays, and increasing corporate bond liquidity premium, Working Paper.

Yavas, Abdullah, 1996, Search and trading in intermediated markets, *Journal of Economics & Management Strategy* 5, 195–216.

INTERNET APPENDIX

A Proofs

Proof of Proposition 1

Proof. The supply of assets in the market satisfies $\dot{s} = (1-s)\eta - s\mu$. Setting $\dot{s} = 0$ implies $s = \frac{\eta}{\mu+\eta}$. The owners value function satisfy

$$rV^o = \nu - \mu p^* \tag{83}$$

and the non-owners value function satisfy

$$rV^n = \eta p^*. \tag{84}$$

Customers choose to be owners if and only if $V^o - p^* \geq V^n$. That is,

$$\nu - \mu p^* - rp^* \geq \eta p^* \Longleftrightarrow \nu \geq (r+\mu+\eta)p^*. \tag{85}$$

Define $\nu^* = (r+\mu+\eta)p^*$. The measure of owners has to equal the supply of assets. Therefore,

$$1 - F(\nu^*) = s \Longleftrightarrow F(\nu^*) = \frac{\mu}{\mu+\eta}. \tag{86}$$

■

Proof of Lemma 1

Proof. First note that we can rewrite $\tilde{V}_D^o(\nu)$ in equation (4) as

$$\begin{aligned} r\tilde{V}_D^o(\nu) &= \nu - \mu[\tilde{V}_D^o(\nu) - V^n(\nu)] + \lambda_D^1 \max\{p - \tau - [\tilde{V}_D^o(\nu) - V^n(\nu)], 0\} \\ &= \nu - \mu[\tilde{V}_D^o(\nu) - V_D^o(\nu) + \Delta(\nu)] + \lambda_D^1 \max\{p - \tau - [\tilde{V}_D^o(\nu) - V_D^o(\nu) + \Delta(\nu)], 0\}, \end{aligned}$$

where we used that $\Delta(\nu) = V_D^o(\nu) - V^n(\nu) \implies -V^n(\nu) = -V_D^o(\nu) + \Delta(\nu)$. Take the difference between the equation above and (3) to obtain

$$\tilde{V}_D^o(\nu) - V_D^o(\nu) = \frac{\lambda_D^1 \max\{p-\tau-[\tilde{V}_D^o(\nu) - V^n(\nu)], 0\} - \lambda_D^0\theta_D \max\left\{\tilde{V}_D^o(\nu) - V_D^o(\nu), 0\right\}}{r+\mu}.$$

The above equation implies that $\tilde{V}_D^o(\nu) - V_D^o(\nu) \geq 0$. Otherwise, we would have

$$\tilde{V}_D^o(\nu) - V_D^o(\nu) = \frac{\lambda_D^1 \max\{p-\tau-[\tilde{V}_D^o(\nu) - V^n(\nu)], 0\}}{r+\mu} < 0,$$

which is a contradiction. Since $\tilde{V}_D^o(\nu) - V_D^o(\nu) \geq 0$, we can rewrite the difference as

$$\tilde{V}_D^o(\nu) - V_D^o(\nu) = \frac{\lambda_D^1 \max\{p-\tau-\Delta(\nu), \tilde{V}_D^o(\nu) - V_D^o(\nu)\}}{r+\mu+\lambda_D^0\theta_D+\lambda_D^1}.$$

Finally, if $\max\{p-\tau-\Delta(\nu), \tilde{V}_D^o(\nu) - V_D^o(\nu)\} = \tilde{V}_D^o(\nu) - V_D^o(\nu)$, then the above equation implies that $\tilde{V}_D^o(\nu) - V_D^o(\nu) = 0$. As a result, we can conclude that

$$\tilde{V}_D^o(\nu) - V_D^o(\nu) = \frac{\lambda_D^1 \max\{p-\tau-\Delta(\nu), 0\}}{r+\mu+\lambda_D^0\theta_D+\lambda_D^1}.$$

After replacing $\tilde{V}_D^o(\nu) - V_D^o(\nu)$ above in the equation (3) we obtain

$$rV_D^o(\nu) = \nu - \mu\Delta(\nu) + \lambda_D^l\theta_D \max\{p-\tau-\Delta(\nu), 0\}, \tag{87}$$

where $\lambda_D^l = \frac{\lambda_D^0\lambda_D^1}{r+\mu+\lambda_D^0\theta_D+\lambda_D^1}$. Analogously,

$$rV_D^n(\nu) = \eta\Delta(\nu) + \lambda_D^s\theta_D \max\{\Delta(\nu) - (p+\tau), 0\}, \tag{88}$$

where $\lambda_D^s = \frac{\lambda_D^0\lambda_D^1}{r+\eta+\lambda_D^0\theta_D+\lambda_D^1}$. Take the difference between equations (87) and (88) to obtain

$$r\Delta(\nu) = \nu - \mu\Delta(\nu) - \eta\Delta(\nu) + \lambda_D^{al}\theta_D \max\{p-\tau-\Delta(\nu), 0\} - \lambda_D^{as}\theta_D \max\{\Delta(\nu) - p - \tau, 0\},$$

which implies

$$\Delta(\nu) = \frac{\nu + \lambda_D^{al}\theta_D \max\{p-\tau, \Delta(\nu)\} + \lambda_D^{as}\theta_D \min\{p+\tau, \Delta(\nu)\}}{r+\mu+\eta+[\lambda_D^{al}+\lambda_D^{as}]\theta_D} \tag{89}$$

for all types $\nu \in (-\infty, \nu_l] \cup [\nu_h, \infty)$. Equation (89) is associated with a functional operator satisfying all Blackwell's conditions for a contraction on the space of continuous functions with bounded variation $\frac{1}{r+\mu+\eta}$. Then, by the contraction mapping theorem, there is a unique function Δ satisfying the equation (89). Also note that if $\tau = 0$, then the results follow directly from equation (89). So we focus on the case with $\tau > 0$.

As Δ is strictly increasing and continuous, we must have that $\Delta(\nu_l) \leq p - \tau$. To see this, notice first that, if $p - \tau < \Delta(\nu_l) < p + \tau$, then the customer would not trade with a dealer because of transaction costs, as the reservation value of a potential seller is higher than the highest bid price of a dealer, $p - \tau$, and the reservation value of a potential buyer is smaller than the lowest ask price of a dealer, $p + \tau$. The last terms in equations (3) and (5) would be zero. Therefore, searching for a dealer is equivalent to be inactive. In this case, the customer would be better off searching for customers type $\nu \in (\nu_l, \nu_h)$ to obtain a share of the gains from trade. This implies that $\nu_l \notin \Omega_D$, which is a contradiction. Implicit in this argument is the fact that the densities of Φ^o and Φ^n are bounded away from zero in the set (ν_l, ν_h) because of issuance and maturity (see proof of Lemma 3), and $\nu_l \neq \nu_h$ (which holds by assumption on a regular equilibrium with $\tau > 0$).

Moreover, if $\Delta(\nu_l) \geq p + \tau$, then either $p - \tau < \Delta(\nu) < p + \tau$ for some customer type $\nu \in \Omega_D$ or $\Delta(\nu) \geq p + \tau$ for all customer type $\nu \in \Omega_D$. The first cannot hold because again it would imply $\nu \notin \Omega_D$. The second would be inconsistent with interdealer market clearing because all customers searching for a dealer would want to buy assets as their reservation value would be greater than or equal to the highest ask price.

Therefore, we must have $\Delta(\nu_l) \leq p - \tau$. An analogous argument applies for ν_h in the opposite direction. That is, $\Delta(\nu_h) \geq p + \tau$. With $\Delta(\nu_l) \leq p - \tau$ and $\Delta(\nu_h) \geq p + \tau$, we can solve for the max relations in equation (89), which then implies equation (26). ■

Proof of Lemma 2

Proof. The proof of this result can be found in the text. Specifically, equations (19) and (20) can be written as

$$-\left(\mu + \lambda_D^1\right) \pi_0^o(\nu) + \lambda_D^0 \left[\Phi^o(\nu) - \pi_0^o(\nu)\right] = 0$$

for $\nu \leq \nu_l$, and

$$-\left(\eta + \lambda_D^1\right)\pi_0^n(\nu) + \lambda_D^0\left[\Phi^n(\nu) - \Phi^n(\nu_h) - \pi_0^n(\nu)\right] = 0.$$

for $\nu \geq \nu_h$. Moreover, non-owners of type $\nu \leq \nu_l$ and owners of type $\nu \geq \nu_h$ do not trade, which leads to the result stated in Lemma 2. ■

Proof of Lemma 3

Proof. Equation (22) implies

$$\dot{\Phi}^o(\infty) = \eta\Phi^n(\infty) - \mu\Phi^o(\infty) \tag{90}$$

as, when ν goes to infinity, both the inflow and outflow from trading goes to zero. Then, from $\dot{\Phi}^o(\infty) = 0$ and equation (23), we have that

$$\eta[F(\infty) - \Phi^o(\infty)] - \mu\Phi^o(\infty) = 0 \iff \Phi^o(\infty) = \frac{\eta}{\mu + \eta}, \tag{91}$$

which characterizes the total supply of assets $s = \Phi^o(\infty)$, equal, by definition, to the measure of owners. This also establishes that the measure of non-owners is given by

$$\Phi^n(\infty) = F(\infty) - \Phi^o(\infty) = 1 - \Phi^o(\infty) \implies \Phi^n(\infty) = \frac{\mu}{\mu + \eta}. \tag{92}$$

And by definition we have $1 - s = \Phi^n(\infty)$. Consider now the case $\nu \leq \nu_l$. According the law of motion for Φ^o, given by equation (22), we have

$$\dot{\Phi}^o(\nu) = \eta\Phi^n(\nu) - \mu\Phi^o(\nu) - \lambda_D^1\pi_0^o(\nu) = \eta\Phi^n(\nu) - \mu\Phi^o(\nu) - \tilde{\lambda}_D^{al}\Phi^o(\nu), \tag{93}$$

as no-owners with $\tilde{\nu} \leq \nu \leq \nu_l$ will neither purchase the asset in agency or from other customers. Substituting $\Phi^n(\nu) = F(\nu) - \Phi^o(\nu)$ and setting $\dot{\Phi}^o(\nu) = 0$ implies

$$\Phi^o(\nu) = \frac{\eta F(\nu)}{\mu + \eta + \tilde{\lambda}_D^{al}}, \quad \nu \leq \nu_l. \tag{94}$$

For the case $\nu \geq \nu_h$, it is useful to work with the measure of non-owners of type above ν,

$\Phi^n(\infty) - \Phi^n(\nu)$. Using equations (23) and (22), we have

$$
\begin{aligned}
0 &= \dot{\Phi}^n(\infty) - \dot{\Phi}^n(\nu) && (95)\\
&= -\eta[\Phi^n(\infty) - \Phi^n(\nu)] - \lambda_D^1[\pi_0^n(\infty) - \pi_0^n(\nu)] + \mu[\Phi^o(\infty) - \Phi^o(\nu)] && (96)\\
&= -\eta[\Phi^n(\infty) - \Phi^n(\nu)] - \tilde{\lambda}_D^{as}[\Phi^n(\infty) - \Phi^n(\nu)] + \mu[\Phi^o(\infty) - \Phi^o(\nu)] && (97)\\
&= -\eta[1 - s - F(\nu) + \Phi^o(\nu)] - \tilde{\lambda}_D^{as}[1 - s - F(\nu) + \Phi^o(\nu)] + \mu[s - \Phi^o(\nu)] && (98)\\
&= -(\eta + \lambda_D)[1 - F(\nu)] + (\mu + \eta + \tilde{\lambda}_D^{as})[s - \Phi^o(\nu)] && (99)
\end{aligned}
$$

$$
\implies s - \Phi^o(\nu) = \frac{(\eta + \tilde{\lambda}_D^{as})[1 - F(\nu)]}{\mu + \eta + \tilde{\lambda}_D^{as}}. \tag{100}
$$

As $s = \frac{\eta}{\mu+\eta}$, we have

$$
\Phi^o(\nu) = \frac{\eta}{\mu + \eta} - \frac{(\eta + \tilde{\lambda}_D^{as})[1 - F(\nu)]}{\mu + \eta + \tilde{\lambda}_D^{as}}, \quad \nu \geq \nu_h. \tag{101}
$$

Now let us show that $\frac{\eta \tilde{\lambda}_D^{al} F(\nu)}{\mu+\eta+\tilde{\lambda}_D^{al}} = \frac{\mu \tilde{\lambda}_D^{as}[1-F(\nu)]}{\mu+\eta+\tilde{\lambda}_D^{as}}$. According to the market clearing condition (21) and Lemmas 1 and 2,

$$
\lambda_D^1 \pi_0^o(\infty) = \lambda_D^1 \pi_0^n(\infty) \implies \tilde{\lambda}_D^{al} \Phi^o(\nu_l) = \tilde{\lambda}_D^{as} \left[\Phi^n(\infty) - \Phi^n(\nu_h)\right]. \tag{102}
$$

We know that $\Phi^o(\nu_l) = \frac{\eta F(\nu_l)}{\mu+\eta+\tilde{\lambda}_D^{al}}$. Moreover,

$$
\begin{aligned}
\Phi^n(\infty) - \Phi^n(\nu_h) &= F(\infty) - \Phi^o(\infty) - [F(\nu_h) - \Phi^o(\nu_h)]\\
&= 1 - s - \left[F(\nu_h) - \frac{\eta}{\mu + \eta} + \frac{(\eta + \tilde{\lambda}_D^{as})[1 - F(\nu_h)]}{\mu + \eta + \tilde{\lambda}_D^{as}}\right] = \frac{\mu[1 - F(\nu_h)]}{\mu + \eta + \tilde{\lambda}_D^{as}}. && (103)
\end{aligned}
$$

Thus, $\frac{\eta \tilde{\lambda}_D^{al} F(\nu)}{\mu+\eta+\tilde{\lambda}_D^{al}} = \frac{\mu \tilde{\lambda}_D^{as}[1-F(\nu)]}{\mu+\eta+\tilde{\lambda}_D^{as}}$. Finally, the result that $\tilde{\lambda}_D^{al} \Phi^o(\nu_l) = \frac{\mu \tilde{\lambda}_D^{as}}{\mu+\eta} - \tilde{\lambda}_D^{as} \Phi^n(\nu_h)$ comes from equation (102) and the fact that $\Phi^n(\infty) = F(\infty) - \Phi^o(\infty) = 1 - \frac{\eta}{\mu+\eta} = \frac{\mu}{\mu+\eta}$. ■

Proof of Lemma 4

Proof. By taking the difference between equations (33) and (34), we know that the reservation value satisfies

$$\Delta(\nu) = \frac{\nu + \lambda_C \int_{\nu}^{\nu_h} \theta_C^o[\Delta(\tilde{\nu}) - \Delta(\nu)]d\Phi^n(\tilde{\nu}) - \lambda_C \int_{\nu_l}^{\nu} \theta_C^n[\Delta(\nu) - \Delta(\tilde{\nu})]d\Phi^o(\tilde{\nu})}{r + \mu + \eta}. \tag{104}$$

Moreover, because Δ is continuous and monotone, equation (104) implies that Δ is Lipschitz continuous in the interval (ν_l, ν_h). To see this, note that we can rearrange equation (104) to show that

$$\left|\frac{\Delta(\nu + t) - \Delta(\nu)}{t}\right| \leq \left|\frac{1 + 2\lambda_C \sup_x f(x)[\Delta(\nu_h) - \Delta(\nu_l)]}{r + \mu + \eta}\right|$$

for all ν and $\nu + t$ in the interval (ν_l, ν_h), where f is the density of the distribution F. Given that Δ is Lipschitz continuous in the interval (ν_l, ν_h), Δ is differentiable almost everywhere in the interval (ν_l, ν_h) and satisfies $\Delta(\nu) = \Delta(\nu_l) + \int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu})d\tilde{\nu}$, where $\sigma_C(\nu)$ denote the derivative of Δ. Using this result, take the derivative on both sides of equation (104) to obtain

$$\sigma_C(\nu) = \frac{1 - \lambda_C \theta_C^o[\Phi^n(\nu_h) - \Phi^n(\nu)]\sigma_C(\nu) - \lambda_C \theta_C^n[\Phi^o(\nu) - \Phi^o(\nu_l)]\sigma_C(\nu)}{r + \mu + \eta}. \tag{105}$$

We then obtain $\sigma_C(\nu)$ by rearranging the equation above. ∎

Proof of Lemma 5

Proof. We have $\dot{\tilde{\Phi}}^o(\nu) = \dot{\Phi}^o(\nu) - \dot{\Phi}^o(\nu_l)$. From equation (22), we have

$$\begin{aligned}
\dot{\tilde{\Phi}}^o(\nu) &= \eta\tilde{\Phi}^n(\nu) - \mu\tilde{\Phi}^o(\nu) - \lambda_C \int_{\nu_l}^{\nu} \int_{\nu}^{\nu_h} d\Phi^n(\hat{\nu})d\Phi^o(\tilde{\nu}) \\
&= \eta\tilde{\Phi}^n(\nu) - \mu\tilde{\Phi}^o(\nu) - \lambda_C\tilde{\Phi}^o(\nu)\big[\Phi^n(\nu_h) - \Phi^n(\nu)\big] \\
&= \eta\tilde{\Phi}^n(\nu) - \mu\tilde{\Phi}^o(\nu) - \lambda_C\tilde{\Phi}^o(\nu)\big[F(\nu_h) - F(\nu)\big] + \lambda_C\tilde{\Phi}^o(\nu)\big[\tilde{\Phi}^o(\nu_h) - \tilde{\Phi}^o(\nu)\big] \\
&= \eta\big[F(\nu) - F(\nu_l)\big] - \tilde{\Phi}^o(\nu)\Big\{\mu + \eta + \lambda_C\big[F(\nu_h) - F(\nu) - \tilde{\Phi}^o(\nu_h)\big]\Big\} - \lambda_C\tilde{\Phi}^o(\nu)^2. \quad (106)
\end{aligned}$$

We can then solve the quadratic equation above with $\dot{\tilde{\Phi}}^o(\nu) = 0$ to obtain equation (37). Equation (38) is obtained by the solution of the quadratic equation for $\nu = \nu_h$. ∎

Proof of Lemma 6

Proof. In a regular equilibrium, customers of type ν_l are indifferent between searching for dealers or customers. The reason is that equations (3)–(10) and the continuity of Δ imply the continuity of V_C^o, V_C^n, V_D^o and V_D^n. As a result,

$$\begin{aligned} rV_C^o(\nu_l) &= \nu_l - \mu\Delta(\nu_l) + \lambda_C \int_{\nu_l}^{\nu_h} \theta_C^o[\Delta(\nu) - \Delta(\nu_l)]d\Phi^n(\nu) \\ &= \nu_l - \mu\Delta(\nu_l) + \lambda_C\theta_C^o \int_{\nu_l}^{\nu_h}\int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^n(\nu) \\ &= \nu_l - \mu\Delta(\nu_l) + \lambda_D^l\theta_D[p - \tau - \Delta(\nu_l)] = rV_D^o(\nu_l). \end{aligned} \tag{107}$$

This implies that

$$p = \Delta(\nu_l) + \tau + \frac{\lambda_C\theta_C^o}{\lambda_D^l\theta_D}\int_{\nu_l}^{\nu_h}\int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^n(\nu). \tag{108}$$

Similarly,

$$\begin{aligned} rV_C^n(\nu_h) &= \eta\Delta(\nu_h) + \lambda_C \int_{\nu_l}^{\nu_h} \theta_C^n[\Delta(\nu_h) - \Delta(\nu)]d\Phi^o(\nu) \\ &= \eta\Delta(\nu_h) + \lambda_C\theta_C^n \int_{\nu_l}^{\nu_h}\int_{\nu}^{\nu_h} \sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^o(\nu) \\ &= \eta\Delta(\nu_h) + \lambda_D^s\theta_D[\Delta(\nu_h) - p - \tau] = rV_D^n(\nu_h), \end{aligned} \tag{109}$$

which implies

$$p = \Delta(\nu_h) - \tau - \frac{\lambda_C\theta_C^n}{\lambda_D^s\theta_D}\int_{\nu_l}^{\nu_h}\int_{\nu}^{\nu_h} \sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^o(\nu). \tag{110}$$

Equating the two expressions for the price, and using lemma 4, we obtain

$$\begin{aligned} 2\tau = {} & \lambda_D\theta_D\int_{\nu_l}^{\nu_h}\sigma_C(\nu)d\nu \\ & - \frac{\lambda_C\theta_C^n}{\lambda_D^s\theta_D}\int_{\nu_l}^{\nu_h}\int_{\nu}^{\nu_h} \sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^o(\nu) - \frac{\lambda_C\theta_C^o}{\lambda_D^l\theta_D}\int_{\nu_l}^{\nu_h}\int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu})d\tilde{\nu}d\Phi^n(\nu). \end{aligned} \tag{111}$$

Applying integration by parts in the last two terms we obtain

$$\begin{aligned} 2\tau &= \int_{\nu_l}^{\nu_h}\sigma_C(\nu)d\nu \\ &\quad - \int_{\nu_l}^{\nu_h}\left\{\frac{\lambda_C\theta_C^n}{\lambda_D^s\theta_D}[\Phi^o(\nu) - \Phi^o(\nu_l)] + \frac{\lambda_C\theta_C^o}{\lambda_D^l\theta_D}[\Phi^n(\nu_h) - \Phi^n(\nu)]\right\}\sigma_C(\nu)d\nu \\ &= \int_{\nu_l}^{\nu_h}\left\{1 - \frac{\lambda_C\theta_C^n}{\lambda_D^s\theta_D}[\Phi^o(\nu) - \Phi^o(\nu_l)] - \frac{\lambda_C\theta_C^o}{\lambda_D^l\theta_D}[\Phi^n(\nu_h) - \Phi^n(\nu)]\right\}\sigma_C(\nu)d\nu. \end{aligned} \tag{112}$$

Define $w(\nu) = \frac{\theta_C^o\left[\Phi^n(\nu_h)-\Phi^n(\nu)\right]}{\theta_C^o\left[\Phi^n(\nu_h)-\Phi^n(\nu)\right]+\theta_C^n\left[\Phi^o(\nu)-\Phi^o(\nu_l)\right]}$. Then we can rewrite the above equation as

$$2\tau = \int_{\nu_l}^{\nu_h} \frac{w(\nu)\left\{\lambda_D^s\theta_D - \lambda_C\theta_C^n\left[\Phi^o(\nu)-\Phi^o(\nu_l)\right] - \lambda_C\theta_C^o\left[\Phi^n(\nu_h)-\Phi^n(\nu)\right]\right\}}{\lambda_D^s\theta_D}\sigma_C(\nu)d\nu$$
$$+\int_{\nu_l}^{\nu_h} \frac{\left[1-w(\nu)\right]\left\{\lambda_D^l\theta_D - \lambda_C\theta_C^n\left[\Phi^o(\nu)-\Phi^o(\nu_l)\right] - \lambda_C\theta_C^o\left[\Phi^n(\nu_h)-\Phi^n(\nu)\right]\right\}}{\lambda_D^l\theta_D}\sigma_C(\nu)d\nu. \tag{113}$$

From the definition of $\sigma_C(\nu)$, we have $\lambda_C\theta_C^o\left[\Phi^n(\nu_h)-\Phi^n(\nu)\right] + \lambda_C\theta_C^n\left[\Phi^o(\nu)-\Phi^o(\nu_l)\right] = \frac{1}{\sigma_C(\nu)} - (r+\mu+\eta)$. Substituting above and rearranging implies

$$2\tau\theta_D = \int_{\nu_l}^{\nu_h} w(\nu)\frac{\sigma_C(\nu)-\sigma_D^{as}}{\lambda_D^{as}\sigma_D^{as}} + \left[1-w(\nu)\right]\frac{\sigma_C(\nu)-\sigma_D^{al}}{\lambda_D^{al}\sigma_D^{al}}d\nu, \tag{114}$$

where we used that $r+\mu+\eta+\lambda_D^{al}\theta_D = 1/\sigma_D^{al}$ and $r+\mu+\eta+\lambda_D^{as}\theta_D = 1/\sigma_D^{as}$. ■

Proof of Proposition 2

Proof. The necessity of equations (26)–(40) are established in Lemmas 3–6. So let us focus on the sufficiency. Consider a family $\{\Delta, p, \Phi^o, \Phi^n, \nu_l, \nu_h\}$ satisfying equations (26)–(40) and value functions V^o and V^n constructed using equations (3)–(11) given the family $\{\Delta, p, \Phi^o, \Phi^n, \nu_l, \nu_h\}$. Let us show that the family $\{V^o, V^n, \Delta, p, \Phi^o, \Phi^n, \nu_l, \nu_h\}$ is a regular equilibrium—that is, it satisfies equations (3)–(23) and definition 2.

Equations (3)–(11)**:** These equations are satisfied by the construction of V^o and V^n.

Equations (13)–(14)**:** First let us show that $V_D^o(\nu) \geq V_C^o(\nu)$ for all $\nu \leq \nu_l$.

$$V_D^o(\nu) \geq V_C^o(\nu) \Longleftrightarrow \lambda_D\theta_D[(p-\tau)-\Delta(\nu)] \geq \lambda_C\theta_C^o\int_{\nu_l}^{\nu^h}[\Delta(\tilde{\nu})-\Delta(\nu)]d\Phi^n(\tilde{\nu})$$
$$\Longleftrightarrow (p-\tau)-\Delta(\nu) \geq \tfrac{\lambda_C\theta_C^o}{\lambda_D\theta_D}\int_{\nu_l}^{\nu^h}[\Delta(\tilde{\nu})-\Delta(\nu_l)]d\Phi^n(\tilde{\nu})$$
$$+\tfrac{\lambda_C\theta_C^o}{\lambda_D\theta_D}[\Phi^n(\nu_h)-\Phi^n(\nu_l)][\Delta(\nu_l)-\Delta(\nu)].$$

From equation (40) we know that $\frac{\lambda_C \theta_C^o}{\lambda_D \theta_D} \int_{\nu_l}^{\nu^h} [\Delta(\tilde{\nu}) - \Delta(\nu_l)] d\Phi^n(\tilde{\nu}) = (p - \tau) - \Delta(\nu_l)$, therefore

$$V_D^o(\nu) \geq V_C^o(\nu) \iff \Delta(\nu_l) - \Delta(\nu) \geq \frac{\lambda_C \theta_C^o}{\lambda_D \theta_D} [\Phi^n(\nu_h) - \Phi^n(\nu_l)][\Delta(\nu_l) - \Delta(\nu)].$$

Assumption 1 implies that $\frac{\lambda_C \theta_C^o}{\lambda_D \theta_D} \in (0, 1)$. From (48), $\Phi^n(\nu_h) - \Phi^n(\nu_l) = \frac{\mu[F(\nu_h) - F(\nu_l)]}{\mu + \eta} \in [0, 1)$. Using (26) in Lemma 1, we have

$$V_D^o(\nu) \geq V_C^o(\nu) \iff \nu_l - \nu \geq \frac{\lambda_C \theta_C^o}{\lambda_D \theta_D} [\Phi^n(\nu_h) - \Phi^n(\nu_l)][\nu_l - \nu].$$

We can then see that $V_D^o(\nu) \geq V_C^o(\nu)$ holds. Moreover, it holds with strictly inequality for all $\nu < \nu_l$. The proofs that $V_D^n(\nu) \geq V_C^n(\nu)$ for all $\nu \leq \nu_l$; $V_D^o(\nu) \leq V_C^o(\nu)$ for all $\nu \in (\nu_l, \nu_h)$; and $V_D^n(\nu) \leq V_C^n(\nu)$ for all $\nu \in (\nu_l, \nu_h)$ are analogous.

Equation (12)**:** Let us start with $\nu \leq \nu_l$. In this case we have $V^o(\nu) = V_D^o(\nu)$ and $V^n(\nu) = V_D^n(\nu)$ based on equation (13). Then, from (3) and (5) we have

$$\begin{aligned} V^o(\nu) - V^n(\nu) &= \frac{\nu + \lambda_D \theta_D (p - \tau) - (\mu + \eta + \lambda_D \theta_D) \Delta(\nu)}{r} \\ &= \frac{(r + \mu + \eta + \lambda_D \theta_D) \Delta(\nu) - (\mu + \eta + \lambda_D \theta_D) \Delta(\nu)}{r} = \Delta(\nu). \end{aligned}$$

The result for $\nu \geq \nu_h$ is analogous. For $\nu \in (\nu_l, \nu_h)$ we have

$$\begin{aligned} r[V^o(\nu) - V^n(\nu)] &= \nu - (\mu + \eta) \Delta(\nu) \\ &\quad + \lambda_C \int_{\nu}^{\nu_h} \theta_C^o [\Delta(\tilde{\nu}) - \Delta(\nu)] d\Phi^n(\tilde{\nu}) - \lambda_C \int_{\nu_l}^{\nu} \theta_C^n [\Delta(\nu) - \Delta(\tilde{\nu})] d\Phi^o(\tilde{\nu}) \end{aligned}$$

Replacing equation (35) and applying integration by parts we get

$$\begin{aligned} r[V^o(\nu) - V^n(\nu)] &= \nu - (\mu + \eta) \Delta(\nu_l) - (\mu + \eta) \int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu}) d\tilde{\nu} \\ &\quad + \lambda_C \int_{\nu}^{\nu_h} \theta_C^o [\Phi^n(\nu_h) - \Phi^n(\tilde{\nu})] \sigma_C(\tilde{\nu}) d\tilde{\nu} - \lambda_C \int_{\nu_l}^{\nu} \theta_C^n [\Phi^o(\tilde{\nu}) - \Phi^o(\nu_l)] \sigma_C(\tilde{\nu}) d\tilde{\nu} \end{aligned}$$

$$
\begin{aligned}
&= \nu - (\mu+\eta)\Delta(\nu_l) + \lambda_C \int_{\nu}^{\nu_h} \theta_C^o[\Phi^n(\nu_h) - \Phi^n(\tilde{\nu})]\sigma_C(\tilde{\nu})d\tilde{\nu} \\
&- \int_{\nu_l}^{\nu} \left\{\mu+\eta+\lambda_C\theta_C^n[\Phi^o(\tilde{\nu}) - \Phi^o(\nu_l)]\right\}\sigma_C(\tilde{\nu})d\tilde{\nu} \\
&= \nu - (\mu+\eta)\Delta(\nu_l) + \lambda_C \int_{\nu}^{\nu_h} \theta_C^o[\Phi^n(\nu_h) - \Phi^n(\tilde{\nu})]\sigma_C(\tilde{\nu})d\tilde{\nu} \\
&- \nu + \nu_l - \int_{\nu_l}^{\nu} \left\{r+\lambda_C\theta_C^o[\Phi^n(\nu_h) - \Phi^n(\tilde{\nu})]\right\}\sigma_C(\tilde{\nu})d\tilde{\nu} \\
&= \nu_l - (r+\mu+\eta)\Delta(\nu_l) + \lambda_C \int_{\nu_l}^{\nu_h} \theta_C^o[\Phi^n(\nu_h) - \Phi^n(\tilde{\nu})]\sigma_C(\tilde{\nu})d\tilde{\nu} + r\Delta(\nu).
\end{aligned}
$$

Now we can replace equation (40) to obtain

$$
\begin{aligned}
r[V^o(\nu) - V^n(\nu)] &= \nu_l - (r+\mu+\eta)\Delta(\nu_l) + \lambda_C\theta_C^o[p-\tau-\Delta(\nu_l)] + r\Delta(\nu) \\
&= \nu_l + \lambda_C\theta_C^o(p-\tau) - (r+\mu+\eta+\lambda_C\theta_C^o)\Delta(\nu_l) + r\Delta(\nu) = r\Delta(\nu),
\end{aligned}
$$

where the last equality we obtained using equation (26) applied to $\Delta(\nu_l)$.

Equation (21)**:** The left-hand side of Equation (21) is given by

$$
\lambda_D \int_{\Omega_D^o} \mathbb{1}_{\{\Delta(\nu)<p-\tau\}} d\Phi^o(\nu) = \lambda_D \int_{-\infty}^{\nu_l} d\Phi^o(\nu) = \lambda_D \Phi^o(\nu_l).
$$

The right-hand side is

$$
\lambda_D \int_{\Omega_D^n} \mathbb{1}_{\{\Delta(\nu)>p+\tau\}} d\Phi^n(\nu) = \lambda_D \int_{\nu_h}^{\infty} d\Phi^n(\nu) = \lambda_D \left[\Phi^n(\infty) - \Phi^n(\nu_h)\right].
$$

Therefore, we have market clearing if, and only if, $\Phi^o(\nu_l) = \Phi^n(\infty) - \Phi^n(\nu_h)$. This equation holds because, from the second equation of (30), $\Phi^o(\infty) = \frac{\eta}{\mu+\eta} \implies \Phi^n(\infty) = 1 - \Phi^o(\infty) = \frac{\mu}{\mu+\eta}$, and, from equation (31), $\frac{\mu}{\mu+\eta} - \Phi^n(\nu_h) = \Phi^o(\nu_l)$.

Equation (22)**:** First, consider $\nu \leq \nu_l$. Then, equation (22) is given by

$$
\dot{\Phi}^o(\nu) = \eta\Phi^n(\nu) - \mu\Phi^o(\nu) - \lambda_D\Phi^o(\nu) = \eta F(\nu) - (\eta - \mu - \lambda_D)\Phi^o(\nu).
$$

Equation (30) states that $\Phi^o(\nu) = \frac{\eta F(\nu)}{\eta-\mu-\lambda_D}$. Thus, $\dot{\Phi}^o(\nu) = \eta F(\nu) - \eta F(\nu) = 0$. Consider now $\nu \geq \nu_h$. Then, equation (22) is given by

$$
\begin{aligned}
\dot{\Phi}^o(\nu) &= \eta\Phi^n(\nu) - \mu\Phi^o(\nu) - \lambda_D\Phi^o(\nu_l) + \lambda_D[\Phi^n(\nu) - \Phi^n(\nu_h)] \\
&= (\eta+\lambda_D)F(\nu) - (\mu+\eta+\lambda_D)\Phi^o(\nu) - \lambda_D[\Phi^o(\nu_l) + \Phi^n(\nu_h)].
\end{aligned}
$$

Using the second equation of (30) and (31), we have

$$
\begin{aligned}
\dot{\Phi}^o(\nu) &= (\eta+\lambda_D)F(\nu) + (\eta+\lambda_D)[1-F(\nu)] - \frac{\eta(\mu+\eta+\lambda_D)}{\mu+\eta} - \frac{\lambda_D\mu}{\mu+\eta} \\
&= \eta+\lambda_D - \frac{\eta(\mu+\eta)+\lambda_D(\mu+\eta)}{\mu+\eta} = \eta+\lambda_D - (\eta+\lambda_D) = 0.
\end{aligned}
$$

Finally, let us consider $\nu \in (\nu_l, \nu_h)$. In this case we have

$$
\begin{aligned}
\dot{\Phi}^o(\nu) &= \eta\Phi^n(\nu) - \mu\Phi^o(\nu) - \lambda_D\Phi^o(\nu_l) - \lambda_C[\Phi^o(\nu) - \Phi^o(\nu_l)][\Phi^n(\nu_h) - \Phi^n(\nu)] \\
&= \eta[\Phi^n(\nu) - \Phi^n(\nu_l)] - \mu[\Phi^o(\nu) - \Phi^0(\nu_l)] + \eta\Phi^n(\nu_l) - \mu\Phi^o(\nu_l) - \lambda_D\Phi^o(\nu_l) \\
&- \lambda_C[\Phi^o(\nu) - \Phi^o(\nu_l)][F(\nu_h) - F(\nu)] + \lambda_C[\Phi^o(\nu) - \Phi^o(\nu_l)][\Phi^o(\nu_h) - \Phi^o(\nu)].
\end{aligned}
$$

We have shown that $\eta\Phi^n(\nu_l) - \mu\Phi^o(\nu_l) - \lambda_D\Phi^o(\nu_l) = 0$ when considering the case $\nu \leq \nu_l$. By using this result and the notation $\tilde{\Phi}^o(\nu) = \Phi^o(\nu) - \Phi^0(\nu_l)$ and $s_C = \tilde{\Phi}^o(\nu_h)$, we obtain

$$
\begin{aligned}
\dot{\Phi}^o(\nu) &= \eta[F(\nu) - F(\nu_l)] - (\mu+\eta)\tilde{\Phi}^o(\nu) \\
&- \lambda_C\tilde{\Phi}^o(\nu)[F(\nu_h) - F(\nu)] + \lambda_C\tilde{\Phi}^o(\nu)\tilde{\Phi}^o(\nu_h) - \lambda_C\tilde{\Phi}^o(\nu)^2 \\
&= \eta[F(\nu) - F(\nu_l)] - (\mu+\eta)\tilde{\Phi}^o(\nu) \\
&- \tilde{\Phi}^o(\nu)\left\{\mu+\eta+\lambda_C[F(\nu_h) - F(\nu) - s_C]\right\} - \lambda_C\tilde{\Phi}^o(\nu)^2.
\end{aligned}
$$

The distribution $\tilde{\Phi}^o(\nu)$, as defined in equation (37), is the positive root of the equation above. Therefore, $\dot{\Phi}^o(\nu) = 0$.

Equation (23): This is directly stated in equations (30) and (37).

We showed that the family $\{V^o, V^n, \Delta, p, \Phi^o, \Phi^n, \nu_l, \nu_h\}$ is an equilibrium. That is, that it satisfies equations (3)–(23). It is easy to see that it must also be a regular equilibrium because equation (40) implies that $\nu_l \leq \nu_h$ with strict inequality if $\tau > 0$, and equations (26) and (35) imply that Δ is continuous and strictly increasing. ■

Proof of Lemmas 7–12 and Proposition 3 The proofs of these results follow the same steps as those in Lemmas 1–6 and Proposition 2, and are therefore we omit them for brevity.

Proof of Proposition 4

Proof. First, consider the decision of an owner of type $\nu \leq \nu_l$ when meeting a dealer. Let us compute the reservation value of a customer type $\nu \leq \nu_l$ that always sells in agency, $\Delta^a(\nu)$, compared to always selling in principal, $\Delta^p(\nu)$. For selling in principal we have

$$
\begin{aligned}
r\Delta^p(\nu) = r[V_D^o(\nu) - V_D^n(\nu)] &= \nu - (\mu+\eta)\Delta^p(\nu) + \lambda_D^0\theta_D[W^l - \Delta^p(\nu)] \\
\implies \Delta^p(\nu) &= \frac{\nu + \lambda_D^0\theta_D W^l}{r+\mu+\eta+\lambda_D^0\theta_D} = \sigma_D^p[\nu + \lambda_D^p\theta_D(p-\tilde{c}^l)],
\end{aligned}
$$

where $\lambda_D^p = \frac{\lambda_D^0\lambda_D^1}{r+\lambda_D^1}$ and $\sigma_D^p = \frac{1}{r+\mu+\eta+\lambda_D^0\theta_D}$. For selling in agency, we have

$$
\begin{aligned}
r\Delta^a(\nu) = r[V_D^o(\nu) - V_D^n(\nu)] &= \nu - (\mu+\eta)\Delta^a(\nu) + \lambda_D^0\theta_D[\tilde{V}_D^o(\nu) - V_D^o(\nu)] \\
&= \nu - (\mu+\eta)\Delta^a(\nu) + \frac{\lambda_D^0\lambda_D^1\theta_D[p-\tau-\Delta^a(\nu)]}{r+\mu+\lambda_D^0\theta_D+\lambda_D^1} \\
&= \nu - (\mu+\eta)\Delta^a(\nu) + \lambda_D^{al}\theta_D[p-\tau-\Delta^a(\nu)] \\
\implies \Delta^a(\nu) &= \frac{\nu + \lambda_D^{al}\theta_D[p-\tau-\Delta^a(\nu)]}{r+\mu+\eta+\lambda_D^{al}\theta_D} = \sigma_D^{al}[\nu + \lambda_D^{al}\theta_D(p-\tau)],
\end{aligned}
$$

where $\lambda_D^{al} = \frac{\lambda_D^0\lambda_D^1}{r+\mu+\lambda_D^0\theta_D+\lambda_D^1}$ and $\sigma_D^{al} = \frac{1}{r+\mu+\eta+\lambda_D^{al}\theta_D}$.

In a regular equilibrium, an owner type $\nu \leq \nu_l$ chooses principal trade if and only if the value from principal trade is greater than or equal to the value from agency trade, i.e.,

$$
\lambda_D^0\theta_D[W^l - \Delta^p(\nu)] \geq \lambda_D^0\theta_D[\tilde{V}_D^o(\nu) - V_D^o(\nu)] = \lambda_D^{al}\theta_D[p-\tau-\Delta^a(\nu)]
$$

$$
\iff \lambda_D^0\left[\frac{\lambda_D^1(p-\tilde{c}^l)}{r+\lambda_D^1} - \Delta^p(\nu)\right] = \lambda_D^p(p-\tilde{c}^l) - \lambda_D^0\Delta^p(\nu) \geq \lambda_D^{al}(p-\tau) - \lambda_D^{al}\Delta^a(\nu)
$$

$$
\iff \lambda_D^p(p-\tilde{c}^l) - \lambda_D^0\sigma_D^p[\nu + \lambda_D^p\theta_D(p-\tilde{c}^l)] \geq \lambda_D^{al}(p-\tau) - \lambda_D^{al}\sigma_D^{al}[\nu + \lambda_D^{al}\theta_D(p-\tau)]
$$

$$\iff \left[1-\lambda_D^0\sigma_D^p\theta_D\right]\lambda_D^p(p-\tilde{c}^l)-\left[1-\lambda_D^{al}\sigma_D^{al}\theta_D\right]\lambda_D^{al}(p-\tau)\geq\left[\lambda_D^0\sigma_D^p-\lambda_D^{al}\sigma_D^{al}\right]\nu.$$

Note the following relations,

$$\begin{aligned}
1-\lambda_D^0\sigma_D^p\theta_D &= 1-\frac{\lambda_D^0\theta_D}{r+\mu+\eta+\lambda_D^0\theta_D}=(r+\mu+\eta)\sigma_D^p,\\
1-\lambda_D^{al}\sigma_D^{al}\theta_D &= 1-\frac{\lambda_D^{al}\theta_D}{r+\mu+\eta+\lambda_D^{al}\theta_D}=(r+\mu+\eta)\sigma_D^{al},\\
\lambda_D^0\sigma_D^p-\lambda_D^{al}\sigma_D^{al} &= (r+\mu+\eta)(\lambda_D^0-\lambda_D^{al})\sigma_D^p\sigma_D^{al}=(r+\mu+\eta)\frac{\sigma_D^{al}-\sigma_D^p}{\theta_D}.
\end{aligned}$$

Replacing it in the previous equation we obtain that an owner type $\nu\leq\nu_l$ chooses principal trade if and only if

$$\begin{aligned}
\lambda_D^p\sigma_D^p(p-\tilde{c}^l)-\lambda_D^{al}\sigma_D^{al}(p-\tau) &\geq \frac{\sigma_D^{al}-\sigma_D^p}{\theta_D}\nu.\\
\iff p\theta_D(\lambda_D^p\sigma_D^p-\lambda_D^{al}\sigma_D^{al})-\lambda_D^p\sigma_D^p\theta_D\tilde{c}^l &\geq (\sigma_D^{al}-\sigma_D^p)\nu-\lambda_D^{al}\sigma_D^{al}\theta_D\tau,
\end{aligned}$$

which is equation (57). The proof for $\nu\geq\nu_h$ is analogous. ■

Proof of Proposition 5

Proof. We derive the distributions and market clearing by analyzing the steady-state laws of motion and trade flows.

Owners' Distribution ($\Phi^o(\nu)$, $\nu\leq\nu_l$): Owners ($\nu\leq\nu_l$) hold an asset and search for dealers to sell. The law of motion for $\Phi^o(\nu)$ is given by

$$\dot{\Phi}^o(\nu)=\eta\Phi^n(\nu)-\mu\Phi^o(\nu)-\lambda_D^1\pi_0^o(\nu)-\text{outflow from principal trades,}\tag{115}$$

where η is the asset issuance rate, μ is the maturity rate, $\pi_0^o(\nu)$ is the measure of owners in agency trades with dealers, and the outflow from principal trades depends on the trade choice. In steady state, $\dot{\Phi}^o(\nu)=0$.

- **Case $\nu\leq\nu_l^a$:** Owners choose principal trades (from Proposition 4). Upon meeting a dealer (rate λ_D^0), they sell immediately, and the dealer rebalances in the interdealer market (rate λ_D^1). The outflow is

$$\lambda_D^0 \Phi^o(\nu),$$

since all owners $\nu \leq \nu_l^a$ sell via principal trades, and $\pi_0^o(\nu) = 0$ (no agency trades). Therefore, $0 = \eta\Phi^n(\nu) - \mu\Phi^o(\nu) - \lambda_D^0\Phi^o(\nu)$. Substitute $\Phi^n(\nu) = F(\nu) - \Phi^o(\nu)$ (total customers have distribution F),

$$0 = \eta[F(\nu) - \Phi^o(\nu)] - \mu\Phi^o(\nu) - \lambda_D^0\Phi^o(\nu) \Longrightarrow \Phi^o(\nu) = \frac{\eta F(\nu)}{\mu + \eta + \lambda_D^0}.$$

- **Case $\nu_l^a < \nu \leq \nu_l$:** Owners type $\tilde{\nu} \in (\nu_l^a, \nu)$ choose agency trading. After meeting a dealer (rate λ_D^0), they enter an agency trade and sell in the interdealer market (rate λ_D^1). From equation (19),

$$0 = -(\mu + \lambda_D^1)\pi_0^o(\nu) + \lambda_D^0[\Phi^o(\nu) - \Phi^o(\nu_l^a) - \pi_0^o(\nu)] \Longrightarrow \pi_0^o(\nu) = \frac{\lambda_D^0[\Phi^o(\nu) - \Phi^o(\nu_l^a)]}{\mu + \lambda_D^1 + \lambda_D^0}.$$

The outflow from agency trades is $\lambda_D^1 \pi_0^o(\nu) = \tilde{\lambda}_D^{al}[\Phi^o(\nu) - \Phi^o(\nu_l^a)]$, where $\tilde{\lambda}_D^{al} = \frac{\lambda_D^0\lambda_D^1}{\mu + \lambda_D^1 + \lambda_D^0}$ (effective agency trade rate). Then,

$$0 = \eta[F(\nu) - \Phi^o(\nu)] - \mu\Phi^o(\nu) - \tilde{\lambda}_D^{al}[\Phi^o(\nu) - \Phi^o(\nu_l^a)] - \lambda_D^0\Phi^o(\nu_l^a)$$
$$\Longrightarrow \Phi^o(\nu) = \frac{\eta F(\nu) + (\lambda_D^{al} - \lambda_D^0)\Phi^o(\nu_l^a)}{\mu + \eta + \tilde{\lambda}_D^{al}} = \frac{\eta[F(\nu) - F(\nu_l^a)] + \eta F(\nu_l^a) + (\lambda_D^{al} - \lambda_D^0)\Phi^o(\nu_l^a)}{\mu + \eta + \tilde{\lambda}_D^{al}}.$$

From the previous case, we have $\Phi^o(\nu_l^a) = \frac{\eta F(\nu_l^a)}{\mu + \eta + \lambda_D^0}$ and so

$$\eta F(\nu_l^a) + (\lambda_D^{al} - \lambda_D^0)\Phi^o(\nu_l^a) = \eta F(\nu_l^a) + (\lambda_D^{al} - \lambda_D^0)\frac{\eta F(\nu_l^a)}{\mu + \eta + \lambda_D^0}$$
$$= \left[1 + \frac{\lambda_D^{al} - \lambda_D^0}{\mu + \eta + \lambda_D^0}\right]\eta F(\nu_l^a) = \frac{\mu + \eta + \lambda_D^{al}}{\mu + \eta + \lambda_D^0}\eta F(\nu_l^a).$$

replacing it in the previous equation yields

$$\Phi^o(\nu) = \frac{\eta[F(\nu) - F(\nu_l^a)] + \eta F(\nu_l^a) + (\lambda_D^{al} - \lambda_D^0)\Phi^o(\nu_l^a)}{\mu + \eta + \tilde{\lambda}_D^{al}} = \frac{\eta F(\nu_l^a)}{\mu + \eta + \lambda_D^0} + \frac{\eta[F(\nu) - F(\nu_l^a)]}{\mu + \eta + \tilde{\lambda}_D^{al}},$$

which matches equation (59).

Non-owners' Distribution ($\Phi^n(\infty) - \Phi^n(\nu)$, $\nu \geq \nu_h$): This proof is analogous to the proof for $\nu \leq \nu_l$ and we omit it here.

Market Clearing Condition: The interdealer market clears when supply equals demand (equation (21)):

$$\lambda_D^1(\pi_1 + \bar{\pi}_0^o) = \lambda_D^1(\pi_{-1} + \bar{\pi}_0^n).$$

- Supply side:
  - Principal ($\nu \leq \nu_l^a$): $\pi_1 = \frac{\lambda_D^0 \Phi^o(\nu_l^a)}{\lambda_D^1} = \frac{\lambda_D^0 \eta F(\nu_l^a)}{(\mu+\eta+\lambda_D^0)\lambda_D^1}$,
  - Agency ($\nu_l^a < \nu \leq \nu_l$): $\bar{\pi}_0^o = \frac{\tilde{\lambda}_D^{al}[\Phi^o(\nu_l) - \Phi^o(\nu_l^a)]}{\lambda_D^1} = \frac{\tilde{\lambda}_D^{al}\eta[F(\nu_l) - F(\nu_l^a)]}{\lambda_D^1(\mu+\eta+\tilde{\lambda}_D^{al})}$.
- Demand side:
  - Principal ($\nu \geq \nu_h^a$): $\pi_{-1} = \frac{\lambda_D^0[\Phi^n(\infty) - \Phi^n(\nu_h^a)]}{\lambda_D^1} = \frac{\lambda_D^0 \mu[1 - F(\nu_h^a)]}{(\mu+\eta+\lambda_D^0)\lambda_D^1}$,
  - Agency ($\nu_h < \nu < \nu_h^a$): $\bar{\pi}_0^n = \frac{\tilde{\lambda}_D^{as}[\Phi^n(\nu_h^a) - \Phi^n(\nu_h)]}{\lambda_D^1} = \frac{\tilde{\lambda}_D^{as}\mu[F(\nu_h^a) - F(\nu_h)]}{\lambda_D^1(\mu+\eta+\tilde{\lambda}_D^{as})}$.

Substituting into equation (21) confirms equation (61). ■

Proof of Proposition 6

Proof. To prove that there exist $\bar{\tau} > 0$ and $\bar{\tilde{c}} > 0$ such that a regular equilibrium is unique for any $(\tilde{c}, \tau) \in [0, \bar{\tilde{c}}) \times [0, \bar{\tau})$, we need to show that the equilibrium characterized by a pair (ν_l, ν_h) is unique. Note that a regular equilibrium satisfies either

(a) $\tau = G(\nu_l)$ with $\lambda_D^p \tilde{c} = q(\nu_l^a)$ and $\nu_l^a \leq \nu_l$; or

(b) $\tilde{c} = H(\nu_l)$ with $\nu_l^a = \nu_l$.

Moreover, $G(\nu_s) = H(\nu_s) = 0$. Therefore, it suffices to show that both G and H are monotone in a neighborhood of $(\bar{\nu}_l, \nu_s]$ and the result follows it.

First let us establish that

$$G(\nu_l) = \frac{1}{2} \int_{\nu_l}^{g(\nu_l)} \frac{\sigma_C(\nu) - \sigma_D^a}{\lambda_D^a \theta_D \sigma_D^a} d\nu$$

is strictly decreasing in $(\bar{\nu}_l, \nu_s]$. Its derivative is given by

$$G'(\nu_l) = \frac{1}{2}\left\{g'(\nu_l)\frac{\sigma_C(g(\nu_l)) - \sigma_D^a}{\lambda_D^a \theta_D \sigma_D^a} - \frac{\sigma_C(\nu_l) - \sigma_D^a}{\lambda_D^a \theta_D \sigma_D^a}\right\} + \frac{1}{2\lambda_D^a \theta_D \sigma_D^a}\int_{\nu_l}^{g(\nu_l)} \frac{\partial \sigma_C(\nu)}{\partial \nu_l} d\nu,$$

where $\sigma_D^a = \frac{1}{r+\mu+\eta+\lambda_D^a\theta_D}$ and $\sigma_C(\nu) = \frac{1}{r+\mu+\eta+\lambda_C\left\{\theta_C^o\left[\Phi^n(\nu_h)-\Phi^n(\nu)\right]+\theta_C^n\left[\Phi^o(\nu)-\Phi^o(\nu_l)\right]\right\}}$. The first term in $G'(\nu_l)$ is negative since $g'(\nu_l) = -\frac{f(\nu_l)}{f(g(\nu_l))} < 0$ and $\frac{\sigma_C(\nu)-\sigma_D^a}{\sigma_D^a} > 0$ for all $\nu \in [\nu_l, g(\nu_l)]$. Bound the first term above by

$$-\frac{1}{2\lambda_D^a\theta_D}\left[\frac{r+\eta+\mu+\lambda_D^a\theta_D}{r+\eta+\mu+\lambda_C \max\{\theta_C^o, \theta_C^n\}} - 1\right],$$

which is strictly negative by Assumption 1.

For the second term, we need the partial derivative $\frac{\partial\sigma_C(\nu)}{\partial\nu_l}$ to be bounded as $\nu_l \to \nu_s$. We can see this is the case upon inspection of equation (37):

$$\Phi^o(\nu) - \Phi^o(\nu_l) = -\frac{\mu+\eta+\lambda_C[F(\nu_h)-F(\nu)-s_C]}{2\lambda_C} + \frac{\sqrt{\{\mu+\eta+\lambda_C[F(\nu_h)-F(\nu)-s_C]\}^2+4\lambda_C\eta\left[F(\nu)-F(\nu_l)\right]}}{2\lambda_C}.$$

$\frac{\partial\sigma_C(\nu)}{\partial\nu_l}$ is bounded as $\nu_l \to \nu_s$, $g(\nu_l) \to \nu_s$. Therefore, the integral's limits shrink, and the bounded integrand ensures that the second term converges to zero. As a result, $G'(\nu_l) < 0$ in $(\bar{\nu}_l, \nu_s]$. The proof for H is analogous. This establishes a unique regular equilibrium. ■

Proof of Proposition 7

Proof. The dealers' net position is defined as $\pi_1 - \pi_{-1}$, where π_1 and π_{-1} are the measures of dealers holding long and short positions in principal trades, respectively, given by

$$\pi_1 = \frac{\lambda_D^0 \Phi^o(\nu_l^a)}{\lambda_D^1}, \quad \pi_{-1} = \frac{\lambda_D^0[\Phi^n(\infty) - \Phi^n(\nu_h^a)]}{\lambda_D^1}.$$

Then, the net position is

$$\mathcal{H} = \pi_1 - \pi_{-1} = \frac{\lambda_D^0}{\lambda_D^1}\left[\Phi^o(\nu_l^a)) - (\Phi^n(\infty) - \Phi^n(\nu_h^a))\right] = \frac{\lambda_D^0\{\eta F(\nu_l^a) - \mu[1 - F(\nu_h^a)]\}}{\lambda_D^1(\mu+\eta+\lambda_D^0)}.$$

The independence of λ_C comes direct from the equation above and the observation that

$\Phi^o(\nu_l^a)) = \Phi^n(\infty) - \Phi^n(\nu_h^a)$ in a symmetric economy. So, let us focus on showing that $\pi_1 - \pi_{-1}$ is decreasing in c^l.

From the proof of Lemma 6, we must have

$$\begin{aligned} f^1(\nu_l, \nu_h) &= \int_{\nu_l}^{\nu_h} w(\nu) \frac{\sigma_C(\nu) - \sigma_D^{as}}{\lambda_D^{as}\sigma_D^{as}} + [1 - w(\nu)] \frac{\sigma_C(\nu) - \sigma_D^{al}}{\lambda_D^{al}\sigma_D^{al}} d\nu - 2\tau\theta_D = 0 \\ p(\nu_l, \nu_h) &= \Delta(\nu_l) + \tau + \frac{\lambda_C \theta_C^o}{\lambda_D^l \theta_D} \int_{\nu_l}^{\nu_h} \int_{\nu_l}^{\nu} \sigma_C(\tilde{\nu}) d\tilde{\nu} d\Phi^n(\nu) \\ &= \Delta(\nu_h) - \tau - \frac{\lambda_C \theta_C^n}{\lambda_D^s \theta_D} \int_{\nu_l}^{\nu_h} \int_{\nu}^{\nu_h} \sigma_C(\tilde{\nu}) d\tilde{\nu} d\Phi^o(\nu). \end{aligned}$$

From Proposition 4,

$$f^2(\nu_l, \nu_h, \nu_l^a) = p(\nu_l, \nu_h)\theta_D(\lambda_D^p\sigma_D^p - \lambda_D^{al}\sigma_D^{al}) - \lambda_D^p\sigma_D^p\theta_D\tilde{c}^l - \left[(\sigma_D^{al} - \sigma_D^p)\nu_l^a - \lambda_D^{al}\sigma_D^{al}\theta_D\tau\right] = 0,$$

$$f^3(\nu_l, \nu_h, \nu_h^a) = (\sigma_D^{as} - \sigma_D^p)\nu_h^a - \theta_D\sigma_D^p\lambda_D^p\tilde{c}^s - \left[p(\nu_l, \nu_h)\theta_D(\sigma_D^p\lambda_D^p - \sigma_D^{as}\lambda_D^{as}) - \theta_D\sigma_D^{as}\lambda_D^{as}\tau\right] = 0.$$

From Proposition 5,

$$f^4(\nu_l, \nu_h, \nu_l^a, \nu_h^a) = \frac{\lambda_D^0\eta F(\nu_l^a)}{\mu+\eta+\lambda_D^0} + \frac{\tilde{\lambda}_D^{al}\eta[F(\nu_l)-F(\nu_l^a)]}{\mu+\eta+\tilde{\lambda}_D^{al}} - \left\{\frac{\lambda_D^0\mu[1-F(\nu_h^a)]}{\mu+\eta+\lambda_D^0} + \frac{\tilde{\lambda}_D^{as}\mu[F(\nu_h^a)-F(\nu_h)]}{\mu+\eta+\tilde{\lambda}_D^{as}}\right\} = 0.$$

We then obtain the system

$$\begin{cases} f^1(\nu_l, \nu_h) = 0, \\ f^2(\nu_l, \nu_h, \nu_l^a) = 0, \\ f^3(\nu_l, \nu_h, \nu_h^a) = 0, \\ f^4(\nu_l, \nu_h, \nu_l^a, \nu_h^a) = 0. \end{cases}$$

Applying the implicit function theorem, we get

$$\underbrace{\begin{bmatrix} \frac{\partial f^1}{\partial \nu_l} & \frac{\partial f^1}{\partial \nu_h} & 0 & 0 \\ \frac{\partial f^2}{\partial \nu_l} & \frac{\partial f^2}{\partial \nu_h} & \frac{\partial f^2}{\partial \nu_l^a} & 0 \\ \frac{\partial f^3}{\partial \nu_l} & \frac{\partial f^3}{\partial \nu_h} & 0 & \frac{\partial f^3}{\partial \nu_h^a} \\ \frac{\partial f^4}{\partial \nu_l} & \frac{\partial f^4}{\partial \nu_h} & \frac{\partial f^4}{\partial \nu_l^a} & \frac{\partial f^4}{\partial \nu_h^a} \end{bmatrix}}_{J} \begin{bmatrix} \frac{\partial \nu_l}{\partial c_l} \\ \frac{\partial \nu_h}{\partial c_l} \\ \frac{\partial \nu_l^a}{\partial c_l} \\ \frac{\partial \nu_h^a}{\partial c_l} \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda_D^p\sigma_D^p\theta_D \\ 0 \\ 0 \end{bmatrix}.$$

We can sign each of the terms of the matrix J, which implies the sign structure

$$\begin{bmatrix} - & + & 0 & 0 \\ + & + & - & 0 \\ - & - & 0 & + \\ + & + & + & + \end{bmatrix}.$$

By Cramer's rule, $\frac{\partial \nu_l^a}{\partial c_l} = \frac{|J_l^a|}{|J|}$, where J_l^a is the matrix obtained by replacing the third column of J with the colum vector $[0,\ \lambda_D^p \sigma_D^p \theta_D,\ 0,\ 0]^T$.

We determine the sign of $|J|$ first. Let's assign symbolic values with appropriate signs $(a, b, c, d, e, f, g, h, i, j, k, l > 0)$,

$$J = \begin{bmatrix} -a & b & 0 & 0 \\ c & d & -e & 0 \\ -f & -g & 0 & h \\ i & j & k & l \end{bmatrix}.$$

Compute the determinant using cofactor expansion along the first row, where $a_{13} = 0$, $a_{14} = 0$,

$$\det(J) = a_{11} \cdot \det(M_{11}) \cdot (-1)^{1+1} + a_{12} \cdot \det(M_{12}) \cdot (-1)^{1+2} = -a \cdot \det(M_{11}) - b \cdot \det(M_{12}).$$

Minor M_{11} (remove first row, first column):

$$\det(M_{11}) = \det \begin{bmatrix} d & -e & 0 \\ -g & 0 & h \\ j & k & l \end{bmatrix} = -h(dk + ej) - l(eg) = -hdk - hej - leg.$$

Minor M_{12} (remove first row, second column):

$$\det(M_{12}) = \det \begin{bmatrix} c & -e & 0 \\ -f & 0 & h \\ i & k & l \end{bmatrix} = -h(ck + ei) - l(ef) = -hck - hei - lef.$$

Substituting back, we obtain $\det(J) = -a(-hdk - hej - leg) - b(-hck - hei - lef) = a(hdk +$

$hej + leg) + b(hck + hei + lef) > 0$. All terms are positive, as $a, h, d, k, e, j, l, g, b, c, i, f > 0$. Therefore, $\det(J) > 0$.

Now, we determine the sign of $|J_l^a|$. Again let's assign symbolic values with appropriate signs $(a, b, c, d, e, f, g, h, i, j, k, l > 0)$,

$$J_l^a = \begin{bmatrix} -a & b & 0 & 0 \\ c & d & L & 0 \\ -f & -g & 0 & h \\ i & j & 0 & l \end{bmatrix}.$$

The third column has three zeros ($a_{13} = 0$, $a_{33} = 0$, $a_{43} = 0$). Expand along it:

$$\det(J_l^a) = 0 + L \cdot \det(M_{23}) \cdot (-1) + 0 + 0 = -L \cdot \det(M_{23}).$$

Minor M_{23} (remove second row, third column):

$$\det(M_{23}) = \det \begin{bmatrix} -a & b & 0 \\ -f & -g & h \\ i & j & l \end{bmatrix} = -h(-aj - bi) + l(ag + bf) = haj + hbi + lag + lbf > 0.$$

Thus, $\det(J_l^a) = -L \cdot \det(M_{23}) < 0$, and we can conclude that $\frac{\partial \nu_l^a}{\partial c_l} = \frac{|J_l^a|}{|J|} < 0$.

We can follow the same approach to obtain $\frac{\partial \nu_h^a}{\partial c_l} = \frac{|J_h^a|}{|J|}$, where J_h^a is the matrix obtained by replacing the fourth column of J with the column vector $[0,\ \lambda_D^p \sigma_D^p \theta_D,\ 0,\ 0]^T$. Again let's assign symbolic values with appropriate signs $(a, b, c, d, e, f, g, h, i, j, k, l > 0)$,

$$J_h^a = \begin{bmatrix} -a & b & 0 & 0 \\ c & d & -e & L \\ -f & -g & 0 & 0 \\ i & j & k & 0 \end{bmatrix}.$$

The fourth column has three zeros ($a_{14} = 0$, $a_{34} = 0$, $a_{44} = 0$). Expand along it:

$$\det(J_h^a) = 0 + L \cdot \det(M_{24}) + 0 + 0 = h \cdot \det(M_{24}).$$

Minor M_{24} (remove second row, fourth column):

$$\det(M_{24}) = \det \begin{bmatrix} -a & b & 0 \\ -f & -g & 0 \\ i & j & k \end{bmatrix} = agk + fbk > 0.$$

Thus, $\det(J_h^a) = L \cdot \det(M_{24}) > 0$ and we can conclude that $\frac{\partial \nu_h^a}{\partial c_l} = \frac{|J_h^a|}{|J|} > 0$. Finally,

$$\frac{\partial \mathcal{H}}{\partial c_l} = \frac{\lambda_D^0 \left\{ \eta f(\nu_l^a) \frac{\partial \nu_l^a}{\partial c_l} + \mu f(\nu_h^a) \frac{\partial \nu_h^a}{\partial c_l} \right\}}{\lambda_D^1(\mu + \eta + \lambda_D^0)},$$

which we can show to be negative after replacing $\frac{\partial \nu_l^a}{\partial c_l}$ and $\frac{\partial \nu_h^a}{\partial c_l}$, and using that $\eta = \mu$. ■

Proof of Proposition 8

Proof. We aim to prove that in a region with a unique regular equilibrium where agency trade is active for both buying and selling bonds, $p_{D,A}^{buy} \leq p_{D,P}^{buy}$.

From (72), the average price paid by non-owners buying assets in principal trade is

$$p_{D,P}^{buy} = (1-\theta_D)\mathbb{E}_{\Phi^n}[\Delta(\nu)|\nu \geq \nu_h^a] + \theta_D \frac{\lambda_D^1(p + \tilde{c}^s)}{r + \lambda_D^1} \geq (1-\theta_D)\Delta(\nu_h^a) + \theta_D \frac{\lambda_D^1(p + \tilde{c}^s)}{r + \lambda_D^1} = \bar{p}_P^{buy},$$

where the inequality comes from the expectation being taken over $\nu \geq \nu_h^a$ and Δ being strictly increasing. $\bar{p}_P^{buy}$ is the price paid in principal trade of a customer type ν_h^a. From equation (78), the average price paid by non-owners buying assets in agency trade is

$$p_{D,A}^{buy} = (1-\theta_D^n)\mathbb{E}_{\Phi^n}[\Delta(\nu)|\nu_h \leq \nu \leq \nu_h^a] + \theta_D^n(p + \tau) \leq (1-\theta_D^n)\Delta(\nu_h^a) + \theta_D^n(p+\tau) = \bar{p}_A^{buy},$$

where $\theta_D^n = \frac{\theta_D(\lambda_D^0 + r + \eta + \lambda_D^1)}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}$, and the inequality comes from the expectation being taken over $\nu \in [\nu_h, \nu_h^a]$ and Δ being strictly increasing. $\bar{p}_A^{buy}$ is the price paid in agency trade of a customer type ν_h^a.

A customer type ν_h^a is indifferent between buying in principal or agency trade. Therefore,

$$\Delta(\nu_h^a) - \bar{p}_P^{buy} = \tilde{\tilde{V}}_D^n(\nu_h^a) - V_D^n(\nu_h^a).$$

From the definition of $\tilde{\tilde{V}}_D^n(\nu_h^a)$ we have that

$$\tilde{\tilde{V}}_D^n(\nu_h^a) - V_D^n(\nu_h^a) = \lambda_D^1 \frac{\Delta(\nu_h^a) - \bar{p}_A^{buy}}{r + \eta + \lambda_D^0 + \lambda_D^1},$$

which implies

$$\frac{\Delta(\nu_h^a) - \bar{p}_P^{buy}}{\Delta(\nu_h^a) - \bar{p}_A^{buy}} = \frac{\lambda_D^1}{r + \eta + \lambda_D^0 + \lambda_D^1} > 1 \iff p_{D,A}^{buy} < p_{D,P}^{buy}.$$

The proof that $p_{D,A}^{sell} > p_{D,P}^{sell}$ is analogous and we omit it here. Finally, as $p_{D,A}^{buy} < p_{D,P}^{buy}$ and $p_{D,A}^{sell} > p_{D,P}^{sell}$, we must have

$$\mathrm{BA}_A = p_{D,A}^{buy} - p_{D,A}^{sell} < p_{D,P}^{buy} - p_{D,P}^{sell} = \mathrm{BA}_P\,.$$

■

Proof of Proposition 9

Part (a)

Proof. From Proposition 4, we have that for small enough c^l and c^s the agency market is inactive and therefore

$$\mathrm{BA} = \mathrm{BA}_P = (1-\theta_D)\left\{\mathbb{E}_{\Phi^n}\left[\Delta(\nu)\middle|\nu \geq \nu_h\right] - \mathbb{E}_{\Phi^o}\left[\Delta(\nu)\middle|\nu \leq \nu_l\right]\right\} + \theta_D \frac{c^s + c^l}{r + \lambda_D^1}.$$

Which implies that

$$\frac{\partial\,\mathrm{BA}}{\partial \tilde{c}^l} = (1-\theta_D)\left\{\frac{\partial \mathbb{E}_{\Phi^n}\left[\Delta(\nu)\middle|\nu \geq \nu_h\right]}{\partial \nu_h} \cdot \frac{\partial \nu_h}{\partial \tilde{c}^l} - \frac{\partial \mathbb{E}_{\Phi^o}\left[\Delta(\nu)\middle|\nu \leq \nu_l\right]}{\partial \nu_l} \cdot \frac{\partial \nu_l}{\partial \tilde{c}^l} + \frac{\theta_D}{r + \lambda_D^1}.\right\}$$

The terms $\frac{\partial \mathbb{E}_{\Phi^n}\left[\Delta(\nu)\middle|\nu \geq \nu_h\right]}{\partial \nu_h}$ and $\frac{\partial \mathbb{E}_{\Phi^o}\left[\Delta(\nu)\middle|\nu \leq \nu_l\right]}{\partial \nu_l}$ are positive because $\Delta(\nu)$ is strictly increasing. So it suffices to show that $\frac{\partial \nu_l}{\partial \tilde{c}^l} < 0$ and $\frac{\partial \nu_h}{\partial \tilde{c}^l} > 0$. From Proposition 3

$$\frac{\lambda_D^p \theta_D (c^l + c^s)}{\lambda_D^1} = \int_{\nu_l}^{\nu_h} \frac{\sigma_C(\nu) - \sigma_D^p}{\sigma_D^p} d\nu \quad \text{and} \quad \eta F(\nu_l) = \mu[1 - F(\nu_h)].$$

Applying total differentiation we get that

$$\frac{\lambda_D^p \theta_D \sigma_D^p}{\lambda_D^1} = \left\{ \sigma_C(\nu_h) - \sigma_D^p + \int_{\nu_l}^{\nu_h} \frac{\partial \sigma_C(\nu)}{\partial \nu_h} d\nu \right\} \frac{\partial \nu_h}{\partial \tilde{c}^l} - \left\{ \sigma_C(\nu_l) - \sigma_D^p + \int_{\nu_l}^{\nu_h} \frac{\partial \sigma_C(\nu)}{\partial \nu_l} d\nu \right\} \frac{\partial \nu_l}{\partial \tilde{c}^l}$$

$$0 = \mu f(\nu_h) \frac{\partial \nu_h}{\partial \tilde{c}^l} + \eta f(\nu_l) \frac{\partial \nu_l}{\partial \tilde{c}^l}.$$

Solving the system we obtain that

$$\frac{\partial \nu_l}{\partial \tilde{c}^l} = \frac{-\lambda_D^p \theta_D \sigma_D^p}{\lambda_D^1 \left\{ \sigma_C(\nu_h) - \sigma_D^p + \int_{\nu_l}^{\nu_h} \frac{\partial \sigma_C(\nu)}{\partial \nu_h} d\nu \right\} \frac{\eta f(\nu_l)}{\mu f(\nu_h)} + \lambda_D^1 \left\{ \sigma_C(\nu_l) - \sigma_D^p + \int_{\nu_l}^{\nu_h} \frac{\partial \sigma_C(\nu)}{\partial \nu_l} d\nu \right\}}.$$

Now note that for c^l and c^s small, we have $\nu_l \approx \nu_h$. Thus,

$$\frac{\partial \nu_l}{\partial \tilde{c}^l} \approx \frac{-\lambda_D^p \theta_D \sigma_D^p}{\lambda_D^1 \left\{ \sigma_C(\nu_h) - \sigma_D^p \right\} \frac{\eta f(\nu_l)}{\mu f(\nu_h)} + \lambda_D^1 \left\{ \sigma_C(\nu_l) - \sigma_D^p \right\}} < 0,$$

where the inequality comes from the fact that $\sigma_C(\nu_h) > \sigma_D^p$. Since

$$0 = \mu f(\nu_h) \frac{\partial \nu_h}{\partial \tilde{c}^l} + \eta f(\nu_l) \frac{\partial \nu_l}{\partial \tilde{c}^l} = 0,$$

we also obtain that $\frac{\partial \nu_h}{\partial \tilde{c}^l} > 0$, which concludes the proof. ■

Part (b)

Proof. To show that BA is decreasing in c^l in some region it suffices to show that the bid-ask spread when $\tau \approx 0$ and $c^l = c^s = \infty$ is strictly smaller than when $c^l = c^s \approx 0$. Let BA^∞ be the bid-ask spread when $c^l = c^s = \infty$ and $\tau \approx 0$. In this case, there is no principal trade and

$$\text{BA}^\infty = \text{BA}_A = (1 - \theta_D^n) \left\{ \mathbb{E}_{\Phi^n} \left[\Delta(\nu) | \nu \geq \nu_h \right] - \mathbb{E}_{\Phi^o} \left[\Delta(\nu) | \nu \leq \nu_l \right] \right\} + 2\theta_D^n \tau,$$

where

$$\theta_D^n = \frac{\theta_D(\lambda_D^0 + r + \eta + \lambda_D^1)}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}.$$

Using the definition of Δ given in Lemma 1, we have that

$$\Delta(\nu) = \begin{cases} \sigma_D^a[\nu + \lambda_D^a \theta_D (p - \tau)], & \nu \leq \nu_l, \\ \sigma_D^a[\nu + \lambda_D^a \theta_D (p + \tau)], & \nu \geq \nu_h, \end{cases}$$

where

$$\sigma_D^a = \frac{1}{r + \mu + \eta + \lambda_D^a \theta_D} \quad \text{and} \quad \lambda_D^a = \frac{\lambda_D^0 \lambda_D^1}{r + \mu + \lambda_D^0 \theta_D + \lambda_D^1} = \frac{\lambda_D^0 \lambda_D^1}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}.$$

Thus,

$$\begin{aligned} \text{BA}^\infty &= (1 - \theta_D^n) \left\{ \mathbb{E}_{\Phi^n} \left[\Delta(\nu) \middle| \nu \geq \nu_h \right] - \mathbb{E}_{\Phi^o} \left[\Delta(\nu) \middle| \nu \leq \nu_l \right] \right\} + 2\theta_D^n \tau \\ &= (1 - \theta_D^n) \left\{ \sigma_D^a \left[\mathbb{E}(\nu | \nu \geq \nu_h) - \mathbb{E}(\nu | \nu \leq \nu_l) + 2\lambda_D^a \theta_D \tau \right] \right\} + 2\theta_D^n \tau \\ &= (1 - \theta_D^n) \sigma_D^a \left\{ \mathbb{E}(\nu | \nu \geq \nu_h) - \mathbb{E}(\nu | \nu \leq \nu_l) \right\} + 2\tau \left\{ \theta_D^n + (1 - \theta_D^n) \sigma_D^a \lambda_D^a \theta_D \right\}. \end{aligned}$$

Now we lower the cost c to find a new equilibrium such that the threshold ν_l and ν_h stay the same, and where $\nu_l^a = \nu_l$ and $\nu_h^a = \nu_h$ — that is, the agency market is inactive. In this case,

$$\text{BA} = \text{BA}_P = (1 - \theta_D) \left\{ \mathbb{E}_{\Phi^n} \left[\Delta(\nu) \middle| \nu \geq \nu_h \right] - \mathbb{E}_{\Phi^o} \left[\Delta(\nu) \middle| \nu \leq \nu_l \right] \right\} + \theta_D \frac{c^s + c^l}{r + \lambda_D^1}.$$

Using the definition of Δ given in Lemma 7, we have that

$$\Delta(\nu) = \begin{cases} \sigma_D^p[\nu + \lambda_D^p \theta_D (p - \tilde{c}^l)], & \nu \leq \nu_l, \\ \sigma_D^p[\nu + \lambda_D^p \theta_D (p + \tilde{c}^s)], & \nu \geq \nu_h, \end{cases}$$

where

$$\sigma_D^p = \frac{1}{r + \mu + \eta + \lambda_D^0 \theta_D} \quad \text{and} \quad \lambda_D^p = \frac{\lambda_D^0 \lambda_D^1}{r + \lambda_D^1}.$$

Thus,

$$\begin{aligned}
\text{BA} &= (1-\theta_D)\left\{\mathbb{E}_{\Phi^n}\left[\Delta(\nu)|\nu \geq \nu_h\right] - \mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu \leq \nu_l\right]\right\} + \theta_D \frac{c^s + c^l}{r + \lambda_D^1} \\
&= (1-\theta_D)\left\{\sigma_D^p\left[\mathbb{E}(\nu|\nu \geq \nu_h) - \mathbb{E}(\nu|\nu \leq \nu_l) + \frac{\lambda_D^0 \lambda_D^1}{r + \lambda_D^1}\theta_D \frac{c^s + c^l}{\lambda_D^1}\right]\right\} + \theta_D \frac{c^s + c^l}{r + \lambda_D^1} \\
&= (1-\theta_D)\sigma_D^p\left\{\mathbb{E}(\nu|\nu \geq \nu_h) - \mathbb{E}(\nu|\nu \leq \nu_l)\right\} + \theta_D \frac{c^s + c^l}{r + \lambda_D^1}\left\{(1-\theta_D)\lambda_D^0 \sigma_D^p + 1\right\}.
\end{aligned}$$

When $c^l = c^s \approx 0$ and $\tau \approx 0$ the last terms drop in BA and BA^∞. So, to show that $\text{BA}^\infty < \text{BA}$, it suffices to show that $(1-\theta_D^n)\sigma_D^a < (1-\theta_D)\sigma_D^p$. Note that

$$1 - \theta_D^n = 1 - \frac{\theta_D(\lambda_D^0 + r + \eta + \lambda_D^1)}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1} = \frac{(1-\theta_D)(r + \eta + \lambda_D^1)}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}.$$

The left-hand side is

$$(1-\theta_D^n)\sigma_D^a = \left(\frac{(1-\theta_D)(r + \eta + \lambda_D^1)}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}\right) \cdot \frac{1}{r + \mu + \eta + \lambda_D^a \theta_D}.$$

As $\lambda_D^a = \frac{\lambda_D^0 \lambda_D^1}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}$, the denominator of σ_D^a is

$$r + \mu + \eta + \frac{\theta_D \lambda_D^0 \lambda_D^1}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1} = \frac{(r + \mu + \eta)(r + \eta + \lambda_D^0 \theta_D + \lambda_D^1) + \theta_D \lambda_D^0 \lambda_D^1}{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}.$$

Therefore,

$$\sigma_D^a = \frac{r + \eta + \lambda_D^0 \theta_D + \lambda_D^1}{(r + \mu + \eta)(r + \eta + \lambda_D^0 \theta_D + \lambda_D^1) + \theta_D \lambda_D^0 \lambda_D^1}.$$

$$\implies (1-\theta_D^n)\sigma_D^a = \frac{(1-\theta_D)(r + \eta + \lambda_D^1)}{(r + \mu + \eta)(r + \eta + \lambda_D^0 \theta_D + \lambda_D^1) + \theta_D \lambda_D^0 \lambda_D^1}.$$

The right-hand side is

$$(1-\theta_D)\sigma_D^p = (1-\theta_D) \cdot \frac{1}{r + \mu + \eta + \lambda_D^0 \theta_D} = \frac{1-\theta_D}{r + \mu + \eta + \lambda_D^0 \theta_D}.$$

The inequality becomes:

$$\frac{(1-\theta_D)(r + \eta + \lambda_D^1)}{(r + \mu + \eta)(r + \eta + \lambda_D^0 \theta_D + \lambda_D^1) + \theta_D \lambda_D^0 \lambda_D^1} < \frac{1-\theta_D}{r + \mu + \eta + \lambda_D^0 \theta_D}.$$

Since $1-\theta_D > 0$, divide both sides by $1-\theta_D$:

$$\frac{r+\eta+\lambda_D^1}{(r+\mu+\eta)(r+\eta+\lambda_D^0\theta_D+\lambda_D^1)+\theta_D\lambda_D^0\lambda_D^1} < \frac{1}{r+\mu+\eta+\lambda_D^0\theta_D}.$$

Simplifying this expression yields $0 < \mu\lambda_D^0\theta_D$. As $\mu, \lambda_D^0, \theta_D > 0$, the inequality holds. ■

Part (c)

Proof. Finally, let us show part (c). We know that $\text{BA} = \omega_P \,\text{BA}_P + \omega_A \,\text{BA}_A$, where

$$\begin{aligned}
\text{BA}_P &= (1-\theta_D)\left\{\mathbb{E}_{\Phi^n}\left[\Delta(\nu)|\nu \geq \nu_h^a\right] - \mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu \leq \nu_l^a\right]\right\} + \theta_D \frac{c^s+c^l}{r+\lambda_D^1},\\
\text{BA}_A &= (1-\theta_D^n)\Big\{\mathbb{E}_{\Phi^n}\left[\Delta(\nu)|\nu_h \leq \nu \leq \nu_h^a\right] - \mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu_l^a \leq \nu \leq \nu_l\right]\Big\} 2\theta_D^n \tau,\\
\omega_P &= \frac{\lambda_D^0[\Phi^o(\nu_l^a)+1-s-\Phi^n(\nu_h^a)]}{\lambda_D^0[\Phi^o(\nu_l^a)+1-s-\Phi^n(\nu_h^a)] + \lambda_D^1[\pi_0^o(\nu_l)-\pi_0^o(\nu_l^a)+\pi_0^n(\nu_h^a)-\pi_0^n(\nu_h)]},\\
\omega_A = 1-\omega_P &= \frac{\lambda_D^1[\pi_0^o(\nu_l)-\pi_0^o(\nu_l^a)+\pi_0^n(\nu_h^a)-\pi_0^n(\nu_h)]p_{D,A}^{buy}}{\lambda_D^0[\Phi^o(\nu_l^a)+1-s-\Phi^n(\nu_h^a)] + \lambda_D^1[\pi_0^o(\nu)-\pi_0^o(\nu_l^a)+\pi_0^n(\nu_h^a)-\pi_0^n(\nu)]}.
\end{aligned}$$

There are two possibilities for equilibrium: either agency is active or not.

Case 1 (agency not active) In this case, we have that ν_l and ν_h are determined by

$$f^1(\nu_l,\nu_h) = F(\nu_l)+F(\nu_h) = 1 \quad \text{and} \quad f^2(\nu_l,\nu_h,\lambda_C) = \int_{\nu_l}^{\nu_h} \sigma_C(\nu)-\sigma_D^p d\nu = \lambda_D^p\theta_D\sigma_D^p(\tilde{c}^l+\tilde{c}^s).$$

Applying the implicit function theorem we then get that

$$\underbrace{\begin{bmatrix} \frac{\partial f^1}{\partial \nu_l} & \frac{\partial f^1}{\partial \nu_h} \\ \frac{\partial f^2}{\partial \nu_l} & \frac{\partial f^2}{\partial \nu_h} \end{bmatrix}}_{J} \begin{bmatrix} \frac{\partial \nu_l}{\partial \lambda_C} \\ \frac{\partial \nu_h}{\partial \lambda_C} \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{\partial f^2}{\partial \lambda_C} \end{bmatrix}$$

Since there is a unique equilibrium, we must have that $\frac{\partial f^2}{\partial \nu_l} < 0$ and $\frac{\partial f^2}{\partial \nu_h} > 0$. Otherwise we can find multiple pairs (ν_l, ν_h) satisfying the equilibrium conditions—which contradicts the assumption that we are in a region with a unique equilibrium.

But then we have that

$$|J| = f(\nu_l)\frac{\partial f^2}{\partial \nu_h} - f(\nu_h)\frac{\partial f^2}{\partial \nu_l} > 0.$$

Applying Cramer's rule.

$$\frac{\partial \nu_l}{\partial \lambda_C} = \frac{|J_l|}{|J|} = \frac{1}{|J|}\left|\begin{bmatrix} 0 & \frac{\partial f^1}{\partial \nu_h} \\ -\frac{\partial f^2}{\partial \lambda_C} & \frac{\partial f^2}{\partial \nu_h} \end{bmatrix}\right| = \frac{f(\nu_h)\frac{\partial f^2}{\partial \lambda_C}}{|J|}$$

$$\frac{\partial \nu_h}{\partial \lambda_C} = \frac{|J_h|}{|J|} = \frac{1}{|J|}\left|\begin{bmatrix} \frac{\partial f^1}{\partial \nu_l} & 0 \\ \frac{\partial f^2}{\partial \nu_l} & -\frac{\partial f^2}{\partial \lambda_C} \end{bmatrix}\right| = -\frac{f(\nu_l)\frac{\partial f^2}{\partial \lambda_C}}{|J|}$$

Note that

$$\frac{\partial f^2}{\partial \lambda_C} = \int_{\nu_l}^{\nu_h} \frac{\partial \sigma_C(\nu)}{\partial \lambda_C} d\nu.$$

And since

$$\sigma_C(\nu) = \frac{1}{r + \mu + \eta + \lambda_C\Big\{\theta_C^o\big[\Phi^n(\nu_h) - \Phi^n(\nu)\big] + \theta_C^n\big[\Phi^o(\nu) - \Phi^o(\nu_l)\big]\Big\}}$$

we have

$$\frac{\partial \sigma_C(\nu)}{\partial \lambda_C} = \frac{-\Big\{\theta_C^o\big[\Phi^n(\nu_h) - \Phi^n(\nu)\big] + \theta_C^n\big[\Phi^o(\nu) - \Phi^o(\nu_l)\big]\Big\}}{\Big[r + \mu + \eta + \lambda_C\Big\{\theta_C^o\big[\Phi^n(\nu_h) - \Phi^n(\nu)\big] + \theta_C^n\big[\Phi^o(\nu) - \Phi^o(\nu_l)\big]\Big\}\Big]^2} < 0.$$

Therefore, $\frac{\partial \nu_l}{\partial \lambda_C} < 0$ and $\frac{\partial \nu_h}{\partial \lambda_C} > 0$. Since in this case

$$\text{BA} = \text{BA}_P = (1 - \theta_D)\left\{\mathbb{E}_{\Phi^n}\big[\Delta(\nu)|\nu \geq \nu_h\big] - \mathbb{E}_{\Phi^o}\big[\Delta(\nu)|\nu \leq \nu_l\big]\right\} + \theta_D\frac{c^s + c^l}{r + \lambda_D^1},$$

it is easy to see that BA increases with λ_C.

Case 2 (agency active) In this case, we have that ν_l and ν_h are determined by

$$f^1(\nu_l, \nu_h) = F(\nu_l) + F(\nu_h) = 1 \quad \text{and} \quad f^2(\nu_l, \nu_h, \lambda_C) = \int_{\nu_l}^{\nu_h} \sigma_C(\nu) - \sigma_D^a d\nu = 2\lambda_D^a\theta_D\sigma_D^a\tau.$$

Given the these equations, the proof that $\frac{\partial \nu_l}{\partial \lambda_C} < 0$ and $\frac{\partial \nu_h}{\partial \lambda_C} > 0$ follow the same steps for the case with inactive agency and we omit it here.

We know that $\text{BA} = \omega_P \text{BA}_P + \omega_A \text{BA}_A = \omega_P[\text{BA}_P - \text{BA}_A] + \text{BA}_A$, where

$$\text{BA}_P = (1-\theta_D)\left\{\mathbb{E}_{\Phi^n}\left[\Delta(\nu)|\nu \geq \nu_h^a\right] - \mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu \leq \nu_l^a\right]\right\} + \theta_D \frac{c^s + c^l}{r + \lambda_D^1},$$

$$\text{BA}_A = (1-\theta_D^n)\Big\{\mathbb{E}_{\Phi^n}\left[\Delta(\nu)|\nu_h \leq \nu \leq \nu_h^a\right] - \mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu_l^a \leq \nu \leq \nu_l\right]\Big\}2\theta_D^n \tau,$$

$$\omega_P = \frac{\lambda_D^0[\Phi^o(\nu_l^a) + 1 - s - \Phi^n(\nu_h^a)]}{\lambda_D^0[\Phi^o(\nu_l^a) + 1 - s - \Phi^n(\nu_h^a)] + \lambda_D^1[\pi_0^o(\nu_l) - \pi_0^o(\nu_l^a) + \pi_0^n(\nu_h^a) - \pi_0^n(\nu_h)]},$$

$$\omega_A = 1 - \omega_P = \frac{\lambda_D^1[\pi_0^o(\nu) - \pi_0^o(\nu_l^a) + \pi_0^n(\nu_h^a) - \pi_0^n(\nu)]p_{D,A}^{buy}}{\lambda_D^0[\Phi^o(\nu_l^a) + 1 - s - \Phi^n(\nu_h^a)] + \lambda_D^1[\pi_0^o(\nu_l) - \pi_0^o(\nu_l^a) + \pi_0^n(\nu_h^a) - \pi_0^n(\nu_h)]}.$$

Note that BA_P is independent of ν_l and ν_h, and, thus, of λ_C. BA_A decreases with ν_l, because of the term $-\mathbb{E}_{\Phi^o}\left[\Delta(\nu)|\nu_l^a \leq \nu \leq \nu_l\right]$, and increases with ν_h, because of the term $\mathbb{E}_{\Phi^n}[\Delta(\nu)|\nu_h \leq \nu \leq \nu_h^a]$. Therefore, BA_A increases with λ_C. Finally, ω_P decreases with ν_l and increases with ν_h, and thus, increases with λ_C. Therefore,

$$\frac{\partial \text{BA}}{\partial \lambda_C} = \frac{\partial \omega_P}{\partial \lambda_C}[\text{BA}_P - \text{BA}_A] + (1-\omega_P)\frac{\partial \text{BA}_A}{\partial \lambda_C} > 0.$$

■

B Data and illiquidity measures

B.1 Data

We use corporate bonds transactions data from the TRACE Enhanced (ETRACE) database from January 2005 to June 2021. This initial dataset provides us with a total of 171,140,493 trades as well as with 283,250 uniquely-identifiable bonds.[14]

[14]The Trade Reporting and Compliance Engine (TRACE) is the "FINRA-developed vehicle that facilitates the mandatory reporting of over-the-counter secondary market transactions in eligible fixed income securities." The bond transactions report was implemented in different phases. It started with Phase I, on July 2002, for investment grade bonds and with issue size greater than or equal to $1 bi, and it continued later with the requirements expanded in Phase II in 2003. The complete implementation occurred in 2005, with Phase III. The report of corporate bond transactions is mandatory for all broker-dealers FINRA members. Therefore, Phase III virtually contains complete coverage of all public transactions. For consistency of the selection into the dataset, our dataset focus on Phase III. The Enhanced TRACE differs from the Standard TRACE in that it discloses more detailed information in individual transactions, e.g., actual trade size.

We use the procedure in Dick-Nielsen (2009) and Dick-Nielsen (2014) to filter out errors, cancellations, reversals and double counting as well as transactions missing individual CUSIP identification. We subsequently drop trades missing yield information and trades that are either on a when-issued basis, in a non-secondary market, with a special condition, automatic give-ups, or in equity-linked notes.[15]

To avoid having many bonds in our sample that trade only momentarily, we add the following two conditions: (1) the bond must have existed in ETRACE for at least one complete year; and (2) the bond must have traded at least 75% of its relevant trading days (BPW and Anderson and Stulz 2017). Bonds must also have sufficient trades to satisfy the conditions necessary to calculate their individual illiquidity measures. Having applied all these trade-based criteria, we are left with 55,753,160 transactions in 5,410 unique issues.

We use Bloomberg to collect bond information on issuance and maturity dates, provisions, coupons, currency denomination, amount outstanding, and ratings. We use the amount outstanding of each issue at the last business day of each month. We exclude trades that took place outside the range of issuance and maturity dates of an issue, and bonds for which the outstanding amount at the last business day of that specific month was zero. Defaulting bonds are eliminated from the sample for as long as they are considered in default, and so are bonds with missing information. We only keep in our sample callable or non-provisional, fixed-rate bonds issued in the US.

We calculate the individual yield spread as the difference between the yield of the corporate bond and the yield of the government bond with the same maturity, as in BPW. The constant maturity yield curve is obtained from the Federal Reserve Bank of St. Louis FRED dataset. We use linear interpolation to calculate the yield of the government bond matching the exact maturity of the corporate bond. The monthly cross-sectional yield spread of a corporate bond is then calculated as the average daily spread in the month.

We use the Eikon dataset to collect each issuer's daily 5-year Credit Default Swap (CDS) quotes, which we use to proxy for the issuer's credit risk. Our measure of credit risk for each monthly cross-section is the average of the issuer's end-of-day CDS spreads. As this data is

[15]To remove any potentially erroneous trades still remaining in the database, we also add a price filter for trades with prices deviating more than 25% from the daily average. This procedure cleans only about 0.1% of the trades.

sufficiently large for the bonds in our database from December 2007, we redefine our sample period to begin in December 2007. We use stock prices to calculate the annualized equity return volatility of each issuer. Bonds missing CDS and equity volatility data are excluded from our dataset. We collect the daily stock prices of the issuers from CRSP.

Our final bond sample consists of 44,168,687 trades in 4,315 unique issues, which are distributed over a period of 163 months starting from December 2007. In total, we have 196,345 combinations of bond-month observations. The number of observations varies between monthly cross-sections depending on, among other things, newly-issued and matured bonds, trade frequency, and issues satisfying our selection criteria in the observed cross-section. Our final sample is predominantly composed by investment grade bonds. Table I shows summary statistics of our calculated illiquidity measures together with data on other variables for the period. The illiquidity measures are described in the next section.

B.2 Illiquidity measures

We discuss the behavior of two measures of illiquidity: the γ measure, proposed by BPW, and the Amihud measure, proposed by (Amihud, 2002). Figure 7 shows the evolution of the measures over time.

The γ measure (BPW) is given by the covariance of subsequent price changes. The γ_i measure for bond i is defined as

$$\gamma_i = -\text{Cov}(\Delta p_{it}, \Delta p_{it+1}), \tag{116}$$

where $\Delta p_{it} = p_{it} - p_{it-1}$ and p_{it} is the logarithm of the clean price P_{it} of bond i on trade t. The clean price is the bond price minus accrued interest since the last coupon payment. We require a bond to have at least ten pairs of consecutive annualized-returns to estimate γ_i. The objective of the measure is to extract a transitory component from observed prices. This transitory component is interpreted as the impact of illiquidity, as efficient markets with no trading frictions imply uncorrelated returns.

The Amihud measure for each bond is given by the average of absolute returns divided by the volume of trades,

$$\text{AMD}_{id} = \frac{1}{N_{id}} \sum_{j=1}^{N_{id}} \frac{|r_{ij}|}{V_{id}}, \tag{117}$$

where N_{id} is the number of available returns r_{ij} of bond i on day d, and V_{id} is the volume of trade of bond i on day d in millions of dollars. We require at least two trades on each day to estimate AMD_{id}.

High Amihud measure implies high price change per unit of volume, that is, high impact or order flow. Liquid markets should not show large changes in price relative to volume. Therefore, a high Amihud measure is interpreted as lack of market liquidity.

We define a measure of aggregate market level illiquidity over time by taking the median, mean or volume-weighted average of bond measures in each cross-section. Figure 7 shows the aggregate measures for the corporate bond market γ and AMD over time. γ and AMD increase when liquidity worsens. Both illiquidity measures strongly increased during the financial crisis. After the crisis, liquidity gradually improved. The covid shock was large but brief and did not affect the trend. Table II shows the correlations between γ, AMD and other variables.[16]

[16] Additional measures of illiquidity are given, among others, by Mahanti et al. (2008) and Dick-Nielsen et al. (2012). Mahanti et al. present a liquidity measure based on the accessibility of the issues. Dick-Nielsen et al. introduce a measure computed by an average of different illiquidity measures.

Table I
Summary statistics

Summary statistics (mean, median and standard deviation) for our sample. The observations are the bond-month combinations. The mean, median and standard deviation are averages of the respective cross-sectional measures for each month. Spread is the corporate bond yield spread detailed in section B.1. γ and AMD are the illiquidity measures detailed in section B.2.

Panel A: Illiquidity measures, spread, and CDS			
Variable	Mean	Median	SD
γ	1.76	0.82	3.09
AMD ($\times 10^3$)	2.40	1.43	3.23
Spread	1.90	1.35	1.93
CDS ($\times 10^{-2}$)	1.46	0.92	1.89
Panel B: sample information			
Observations	196,345		
N Bonds	4,315		
Investment Grade	88%		
Callable	38%		
N Firms	490		
N Trades	44,168,687		

Table II
Correlations between illiquidity measures and other variables

Correlations between our main illiquidity measures, γ and AMD, and the spread, CDS, and other commonly used liquidity measures. Spread is the corporate bond yield spread with respect to the US Treasury with the same maturity (appendix B.1). Maturity is the issue's time to maturity. Maturity and age are calculated in years at the last business day of each month. Turnover is the traded volume divided by the amount outstanding. ZTD is the percentage of zero-trading days.

	γ	AMD	Spread	CDS	Volume	Frequency	Maturity	Age	Turnover	ZTD
γ	1.00									
AMD	.466	1.00								
Spread	.385	.444	1.00							
CDS	.290	.347	.816	1.00						
Volume	−.002	−.055	.040	.056	1.00					
Frequency	.047	.196	.146	.140	.420	1.00				
Maturity	.163	.149	.092	−.026	.097	−.052	1.00			
Age	.017	.109	.079	.056	−.202	−.001	−.075	1.00		
Turnover	.013	−.008	.125	.127	.588	.303	.110	−.209	1.00	
ZTD	−.051	−.198	−.080	−.088	−.199	−.356	.084	.016	−.034	1.00

Nova School of Business & Economics
Campus de Carcavelos
Rua da Holanda 1
2775-405 Carcavelos | Portugal
novasbe.pt