Simultaneous Multilateral Search*

Sergei Glebkin[†] Bart Zhou Yueshen[‡] Ji Shen[§]

This version: March 11, 2021

* This paper benefits tremendously from discussions with Markus Baldauf, Jean-Edouard Colliard, Pierre Collin-Dufresne, Jérôme Dugast, Uday Rajan, Ioanid Roşu, Dimitri Vayanos, Kumar Venkataraman, Sebastian Vogel, Zhuo Zhong, and Haoxiang Zhu. In addition, comments and feedback are greatly appreciated from participants at conferences and seminars at EFA 2020, Collegio Carlo Alberto, EPFL, LSE, and INSEAD. There are no competing financial interests that might be perceived to influence the analysis, the discussion, and/or the results of this article.

† INSEAD; glebkin@insead.edu; Boulevard de Constance, Fontainebleau 77300, France.

‡ INSEAD; b@yueshen.me; 1 Ayer Rajah Avenue, Singapore 138676.

§ Peking University; jishen@gsm.pku.edu.cn; No. 38 Xueyuan Road, Haidian District, Beijing 100871, China.

Simultaneous Multilateral Search

Abstract

This paper studies simultaneous multilateral search (SMS) in over-the-counter markets: When searching, a customer simultaneously contacts several dealers and trades with the one offering the best quote. Higher search intensity (how often one can search) improves welfare, but higher search capacity (how many dealers one can contact) or higher transparency about dealer inventories might be harmful. When the market is in distress, customers might inefficiently favor bilateral bargaining (BB) over SMS. Such preference for BB speaks to the sluggish adoption of SMS trading, like request-for-quote protocols, in over-the-counter markets. Furthermore, a market-wide shift to SMS may not be socially optimal.

Keywords: request-for-quote, over-the-counter market, search, bargaining

(There are no competing financial interests that might be perceived to influence the analysis, the discussion, and/or the results of this article.)

1 Introduction

Search is a key feature in over-the-counter (OTC) markets. Duffie, Gârleanu, and Pedersen (2005, hereafter DGP) pioneered the theoretical study of OTC markets in a framework of random matching and *bilateral bargaining* (BB): Investors search for counterparties and are randomly matched over time. Upon successful matching, a buyer and a seller engage in Nash bargaining and split the trading gain according to their endowed bargaining power.

However, investors' interaction is not always bilateral. For example, in recent years, there is a rise of electronic trading in OTC markets, mainly in the form of Request-for-Quote (RFQ). In such marketplaces, where many corporate bonds and derivatives are traded, a customer contacts multiple dealers for quotes and then trades with the one offering the best price. Hendershott and Madhavan (2015) report that more than 10% of trades in the \$8tn corporate bond market is completed via RFQ. O'Hara and Zhou (2020) document a continued growth of RFQ-based trading of corporate bonds, but only sluggishly, with the highest trading volume share below 14% in their sample.

This paper develops a theoretical model, tailored to the above one-to-many searching. Specifically, a customer is allowed to query *multiple* dealers *at the same time*, hence the name "Simultaneous Multilateral Search" (SMS). The objective is twofold. First, we examine how the SMS technology affects assets allocation and welfare. Second, we study how customers choose to search: Do they favor SMS over BB? Are their choices efficient? How do we understand the sluggish growth of SMS-type of electronic trading (O'Hara and Zhou, 2020)?

Section 2 sets up the model following Hugonnier, Lester, and Weill (2020, hereafter HLW), where a continuum of customers trade an asset through a continuum of homogeneous dealers.[1] All agents can hold either zero or one unit of the asset. The customers are subject to stochastic valuation shocks. Those who hold the asset but have low valuation want to sell, while those without the asset but with high valuation want to buy. They actively search for dealers according to independent

[1] While HLW allow heterogeneous dealers to study dealer intermediation chains, we only introduce homogeneous dealers to model how customers contact them via SMS, like via an RFQ protocol in the real-world.

Poisson processes with intensity ρ. We modify the search process in three ways to model SMS: (i) each searching customer can reach n dealers; (ii) with probability q, the searching customer can commit to a reserve price, i.e., indicating a worst price she is willing to accept from the dealers; and (iii) the n dealers independently quote to the searching customer. Effectively, a searching customer runs a first-price auction, probabilistically with a reserve price, among the n dealers.[2]

Section 3 characterizes the equilibrium and discusses novel findings. One notable result is that the two search parameters, the intensity ρ (how frequently one can search) and the capacity n (how many potential counterparties one can reach), have contrasting implications for various equilibrium objects. In particular, a higher ρ always pushes the equilibrium asset allocation toward the Walrasian level, reducing the sizes of both the buyer- and the seller-customers, thus improving welfare. In contrast, a larger n can drive up the size of the short-side customers, away from the efficient Walrasian level, possibly *reducing welfare*.

The key mechanism is a dealer "bottleneck," arising from the asymmetric effects of n on the matching of the two sides of the market. To see this, suppose the asset is in excess supply and 90% of the dealers have inventory while the other 10% do not. Let us examine what happens when the capacity increases from $n = 2$ to $n = 3$: For a customer-seller, the matching rate with a no-inventory dealer increases from $1 - 0.9^2 = 19\%$ to $1 - 0.9^3 = 27.1\%$. Such an improvement in matching significantly adds to the asset inflow to dealers from customer-sellers. However, the outflow rate—the matching between customer-buyers and dealer-sellers—only increases by 0.9%, from $1 - 0.1^2 = 99\%$ to $1 - 0.1^3 = 99.9\%$. The negligible increase of the outflow rate is not at all enough to balance the significant rise in the inflow rate. That is, the asset is "clogged" at the dealers, creating a bottleneck that leaves more customer-buyers unmatched.[3] This leads to a surge

[2] Notably, the framework nests BB as a special case when $n = 1$: The searching customer randomly contacts one dealer and sets the price with probability q. With probability $1 - q$, the dealer sets the price. The parameter q thus serves as the customer's Nash bargaining power parameter, as in DGP and HLW.

[3] It is the increase of the unmatched customer-buyers that eventually balances the asset inflow to and the outflow from dealers in equilibrium. Whereas the inflow increases with n via the higher matching *rate*, the outflow increases via the increment in the larger customer-buyer population *size*.

in unrealized trading gains and may reduce welfare. To emphasize, this bottleneck effect is unique to the search capacity n. It always manifests except in the knife-edge case, in which exactly half of the dealers have inventory. In contrast, the search intensity ρ does not create any asymmetry in matching and always improves welfare.

Such bottlenecks arise not only with the search capacity n. In a specialized application, we allow customers to direct their searches to subsets of dealers of their choosing, based on noisy signals of dealer types. For example, a customer-buyer might have a rough idea of which dealers have inventory, based on recently reported trades. She then optimally directs her searches only to those dealers for higher matching probabilities. A key parameter is the signal quality ψ, which can be interpreted as the transparency of dealer inventories. We show that a similar bottleneck might arise when ψ increases: As customers can direct their searches more accurately, the matching on the short and on the long side is improved asymmetrically, hindering the efficient passing of the asset through dealers. Our model thus highlights a potential channel for how better inventory transparency might hurt welfare.

Another insight from the model is the dual role of dealer demographics in equilibrium: As is standard, dealer demographics (e.g., the proportion of dealers with inventories, i.e., dealer-sellers) affect matching (e.g., how likely a customer-buyer can find a counterparty to trade). New in this model, dealer demographics also affect the split of trading gains between customers and dealers. For example, if there are many dealers with inventories, when contacted by a customer-buyer, they will quote more competitively, as they know that the customer has also contacted $n-1$ other dealers, who very likely might also have inventories to sell. Such fiercer competition cuts more trading gains to the searching customer and less to the dealers. In this sense, SMS endogenizes the bargaining powers, which are by and large exogenous in existing search models. Further, in equilibrium, the dealers quote according to a mixed strategy (à la Burdett and Judd, 1983), creating price dispersion despite dealer homogeneity. Notably, the distribution of such price dispersion is also endogenously determined via dealer demographics.

Section 4 studies the customers' choices between BB and SMS. We show that the choice ultimately boils down to the comparison between the two technologies' expected trading gain intensities, which are the respective products of (i) the search intensity—how frequent one can search, (ii) the matching rate—how likely it is to find at least one counterparty, and (iii) the expected trading gain share above—how much trading gain one can get given a match.

At first glance, one might conclude that SMS has advantages for customers over BB in all three aspects above: (i) it offers faster connection (via electronic platforms), (ii) it allows customers to contact multiple dealers, and (iii) it encourages the competition among the contacted dealers, hence giving larger trading gain shares to customers. The analysis, however, reveals a potential downside of SMS, especially when customers have a low chance of committing to their reserve prices, i.e., when q is low in SMS. On MarketAxess for example, a customer always receives take-it-or-leave-it offers from dealers, effectively $q = 0$. In this case, the customer's expected trading gain is only determined by the endogenous competition among the contacted dealers. When such competition is insufficient, little trading gain is left for the customer, because any matched counterparty dealer will charge a monopoly price, knowing that she is likely the only counterparty that the customer can find (out of the n). In contrast, in BB, a customer always has some chances to secure the full trading gain, as long as q in BB is positive.

In equilibrium, the customers do not always use SMS, especially when the asset is in very imbalanced supply. Consider the case of excess supply. The large number of customer-sellers flood the dealer sector with the asset, making most of the dealers full in inventory. Consequently, the remaining customer-sellers find it very difficult to find dealer-buyers. Even if they do, using SMS, any matched dealer-buyer will knowingly charge a very low monopoly price. Instead, resorting to BB, customer-sellers can still negotiate prices with dealers. This prediction echoes the empirical finding in O'Hara and Zhou (2020) that when corporate bonds are under fire sell (i.e., in excess supply), the electronic trading volume share drops. Such an intrinsic tradeoff between SMS and BB could have hindered the adoption of electronic OTC trading in corporate bond markets.

The customers' endogenous choices between BB and SMS also have welfare and market design implications. The analysis shows that when the asset trades very fast, i.e., for sufficiently high search intensity ρ, a social planner strictly prefers SMS over BB, simply because SMS offers better matching, which creates large trading gains. Unlike the planner, who does not care about the split of trading gains, the customers might shy away from SMS because the trading gain split there is inferior compared to BB. Such inefficiency in technology adoption can be reduced by policies and market designs that incentivizes customers to use SMS. In the model, this can be achieved by setting a large enough q in SMS, e.g., by allowing customers to set reserve prices in RFQ platforms.

However, such patches might not always work, depending on the characteristics of the asset traded. For example, when the asset is intrinsically slow, i.e., for sufficiently low search intensity ρ, having all investors using SMS is not efficient. The intuition goes back to the bottleneck: In the case of excess supply, for example, the planner would like customer-sellers to use BB and buyers SMS to reduce the asset inflow into the dealers, so as to mitigate the bottleneck. Such a distinction between fast- and slow-moving assets is realistic and important. While corporate bonds on SMS trade in a few minutes (Hendershott and Madhavan, 2015), auctions of collateralized loan obligations (CLOs) can take a day or two (Hendershott et al., 2020). Asset-specific design and policies should be considered, as opposed to market-wide, blanket recommendations.

Contribution and related literature

The paper contributes to four strands of the literature. First, adding to the search models of OTC markets, this paper introduces the possibility for investors to contact *multiple* potential counterparties *at the same time*. Previous search models largely focus on BB as in DGP, Weill (2007), Vayanos and Weill (2008), Lagos and Rocheteau (2009), Lagos, Rocheteau, and Weill (2011), Üslü (2019), and HLW. A noteworthy difference is that in SMS, the split of trading gain between customers and dealers (their respective bargaining powers) is endogenous of the equilibrium dealer demographics. This feature distinguishes our model from other works that also have multiple dealers competing simultaneously for a given transaction. For example, Hendershott et al. (2017) consider a stylized model of how customers choose to form dealer networks. There, a buyer simultaneously contacts all dealers in her network, who then compete to find the asset for the customer. Similarly, in Wang (2017), any agent may query quotes within her network simultaneously. In these models, the split of trading gains between dealers and customers is exogenous. In Zhu (2012) and An (2020), customers *sequentially* contact possibly multiple dealers and the resulting endogenous trading gain splits arise due to other frictions like information asymmetry.

Second, this paper contributes to the theory of electronic OTC markets. Vogel (2019) studies a hybrid OTC market where investors can trade in both the traditional voice market and the electronic RFQ platform. Liu, Vogel, and Zhang (2017) compare the the electronic RFQ protocol in an OTC market with a centralized exchange market. Both papers model the RFQ trading similarly to the current paper, with the key difference being that their RFQ matching rates are exogenous, whereas they are endogenous of dealer demographics in this paper. Riggs et al. (2019) study the RFQ trading in Swap Exchange Facilitites. Their analysis highlights order size as an important determinant of customers' choice of trading mechanism. Our analysis complements theirs and highlights another factor, dealer demographics. In a different line, Saar et al. (2019) compare dealers' market making (direct liquidity provision) and matchmaking (searching on the customers' behalf for counterparties) and study the effects of bank dealers' balance sheet costs.

Third, there is a growing body of literature comparing centralized versus decentralized trading (Pagano, 1989; Chowdhry and Nanda, 1991) in various aspects. Babus and Parlatore (2017) study the endogenous formation of fragmented markets due to investors' strategic behavior. Glode and Opp (2019) compare the efficiency of OTC and limit-order markets in a setting where investors endogenously acquire expertise. Lee and Wang (2019) study uninformed and informed investors' venue choice through an adverse selection channel. Dugast, Üslü, and Weill (2019) examine banks' choice among centralized trading, OTC trading, or both, in a setting where the banks differ in their risky asset endowment and in their capacity of OTC trading. This paper instead compares the conventional voice trading versus the relatively new electronic trading within the OTC setting.

Finally, this paper contributes to the auctions literature with uncertain number of bidders (see, e.g., the survey by Klemperer, 1999) and to the literature on pricing with heterogeneously informed consumers (e.g., Butters, 1977; Varian, 1980; and Burdett and Judd, 1983). Apart from the above literature on OTC markets, applications of such "random pricing" mechanisms are also seen recently in exchange trading, such as Jovanovic and Menkveld (2015) and Yueshen (2017). The main insight from this paper is that such uncertainty about the number of quoters (bidders) can arise endogenously from the search process.

2 Model setup

Time is continuous. We consider the trading of an asset in fixed supply s.

Customers and dealers. There is a continuum of *c*ustomers with mass m_c and a continuum of *d*ealers with mass m_d. Both groups of agents are risk-neutral, discount future utility at the same rate r, and can each hold either zero or one unit of the asset. An owner of the asset will be denoted by o and a non-owner n.

The two groups of agents differ in the flow utility from holding the asset. A customer owner derives $y(t) \in \{y_h, y_l\}$ (*h*igh or *l*ow), which evolves stochastically according to a continuous time Markov chain: $\mathbb{P}(y(t + \mathrm{d}t) = y_h | \, y(t) = y_l) = \lambda_u \mathrm{d}t$ and $\mathbb{P}(y(t + \mathrm{d}t) = y_l | \, y(t) = y_h) = \lambda_d \mathrm{d}t$, where λ_d and λ_u are the respective switching intensities. A dealer-owner instead derives constant instantaneous utility y_d—they are not subject to preference shocks.

In summary, there are four types of customers, $\{ho, hn, lo, ln\}$, and two types of dealers, $\{do, dn\}$. Their population measures at any time t are denoted by $m_\sigma(t)$ for $\sigma \in \{ho, hn, lo, ln, do, dn\}$, with $m_{ho}(t) + m_{hn}(t) + m_{lo}(t) + m_{ln}(t) = m_c$ and $m_{do}(t) + m_{dn}(t) = m_d$.

Search. The setup above follows Hugonnier, Lester, and Weill (2020) but with all dealers having the same preference. In particular, customers cannot contact each other and have to search for dealers to trade with. We generalize how customers interact with dealers by introducing a trading technology characterized by $\{\rho, n, q\}$. Using the technology, at a Poisson process with intensity ρ, each customer can contact up to n dealers.[4] The Poisson processes are independent. Each contact by a customer-buyer (-seller) is a "match" if the contacted dealer is of type-do (-dn). The probability for *one* such match is written as π_{do} (π_{dn}).[5] Such probabilities can be functions of the shares of the searching customer's target dealer type; that is, $\pi_{do} = \pi\left(\frac{m_{do}}{m_d}\right)$ and $\pi_{dn} = \pi\left(\frac{m_{dn}}{m_d}\right)$. We assume $\pi(x)$ has support $x \in [0, 1]$ and is monotone increasing with $\pi(0) = 0$, $\pi(1) = 1$, $\pi'(0) < \infty$, and $\pi(x) \geq x$. Consider the following two examples:

- *Pure random matching:* Each dealer is selected from the whole dealer population at random. In this case, $\pi(x) = x \in [0, 1]$, which is standard, as in DGP and HLW.
- *Random matching with signals:* Right before contacting the dealers, each customer can observe signals $\{\delta_i\}$, $i \in [0, m_d]$. Each signal $\delta_i \in \{1, 0\}$ reveals correctly the inventory of dealer i with probability $\psi \in (\frac{1}{2}, 1]$: $\psi = \mathbb{P}[\delta_i = 1 | \sigma_i = do] = \mathbb{P}[\delta_i = 0 | \sigma_i = dn]$. One can think of ψ as the signal quality and interpret it as the transparency of dealer inventories. (The signals are conditionally independent of each other.) Customers can *direct their search* to the subset of dealers with a particular realization of a signal. Within the chosen subset, the search is random. A customer-buyer (-seller) would then like to direct her search only to the subset of dealers whose signals equal one (zero). Define

$$\pi(x) := \frac{\psi x}{\psi x + (1 - \psi)(1 - x)}. \tag{1}$$

  Then by Bayes' rule, each contact by a customer-buyer (-seller) has success rate $\pi_{do} = \pi\left(\frac{m_{do}}{m_d}\right)$ ($\pi_{dn} = \pi\left(\frac{m_{dn}}{m_d}\right)$). Note that $\pi(x)$ degenerates to $\pi(x) = x$ if $\psi \to \frac{1}{2}$ (uninformative signals).

It is worth mentioning that the assumption $\pi(x) \geq x$ is natural: *The least* a customer can do is to

[4] Customers can choose to contact fewer than n dealers. Since there is no cost of contacting more, in equilibrium, they will always choose to contact n dealers. With contact costs, investors in Riggs et al. (2019) choose an interior number of contacts. Such a cost does not bring novel insights in the current model and, hence, is set to zero.

[5] The probability of finding *at least one* matching dealer for a customer-buyer (-seller) is then $1 - (1 - \pi_{do})^n$ ($1 - (1 - \pi_{dn})^n$).

randomly search among all dealers, in which case each contact by a customer-buyer (-seller) has probability $\pi_{do} = \frac{m_{do}}{m_d}$ ($\pi_{dn} = \frac{m_{dn}}{m_d}$) to be a match as in "pure random matching."

Price determination. Upon contacting the dealers:

- With probability q, the customer makes a take-it-or-leave-it offer (TIOLIO) to all the contacted dealers, who then choose to accept the offer or walk away. If more than one dealer accepts, the customer randomly chooses one to trade with.
- With probability $1 - q$, the n dealers simultaneously make independent TIOLIOs to the customer, who then chooses the best quote or walks away.

A contacted dealer may be unable to accommodate the contacting customer due to the inventory constraint (i.e., not a match). Importantly, each dealer makes his decision independently, not knowing the types of the other $(n - 1)$ contacted dealers. To note, this specific price determination mechanism does not affect the results about demographics and welfare (Sections 3.1–3.4).

Parameter values and supports. We normalize the customer mass to $m_c = 1$ and require the dealer mass $m_d > 0$. We also require $s \in (0, 1 + m_d)$ so as to study asset allocation meaningfully. The customers' preference-switching intensities are strictly positive, i.e., $\lambda_u > 0$ and $\lambda_d > 0$. We set $y_h > y_l$ so that some customers are a "high" type and some "low." An additional constraint on y_d will be introduced in Proposition 1 to ensure that there is positive trading gains. The technology parameters have supports $\rho \in (0, \infty)$, $n \in \mathbb{N}$ (the natural numbers), and $q \in [0, 1]$.

Remarks

Remark 1. The trading technology is general enough to encompass some of the most common protocols in OTC trading. For example, the case of $n = 1$ can be thought of as customers reaching dealers by phone and negotiating the terms of trade via bargaining (BB, as in DGP and in HLW). The case of $n > 1$ captures technologies that allow a customer to reach multiple dealers in one click, hence the name "simultaneous multilateral search" (SMS). For example, this is the case for the RFQ protocol on electronic platforms (like MarketAxess and Swap Execution Facilities, SEFs); for auctions like bid/offer-wanted-in-competition (B/OWIC); and in housing markets where a seller can be in touch with possibly many buyers at the same time.

Remark 2. In practice, customers can choose how to get in touch with dealers. They can always call dealers (BB) but they can also click buttons on electronic platforms like RFQs (SMS). After exploring the equilibrium properties of one general technology in Section 3, we study how customers choose between "call" and "click" in Section 4.

Remark 3. The general trading technology is governed by three parameters:

- The search intensity ρ, inherited from DGP and HLW, implies that the technology connects a customer with dealers in exponential waiting time with mean $1/\rho$. For example, auctions on MarketAxess vary in length, from 5 to 20 minutes (Hendershott and Madhavan, 2015). Trading of collateralized loan obligations (CLOs) is typically organized through B/OWIC by email (Hendershott et al., 2020) and can take a considerably longer time.
- The search capacity n, new in this paper, flexibly nests BB ($n = 1$) with SMS protocols that allow contacting multiple dealers. For example, on the MD2C platform operated by Bloomberg Fixed Income Trading, clients can select up to $n = 6$ quotes (Fermanian, Guéant, and Pu, 2017). On the Bloomberg Swap Execution Facility (SEF), this upper bound is set to $n = 5$ (Riggs et al., 2019).
- The probability q reflects the customer's "baseline bargaining power" when using the trading technology. When $n = 1$, q essentially serves as the customers' Nash bargaining power as in DGP and HLW. When $n > 1$, q can be thought of as the customers' ability to communicate, and commit to, their reserve prices. Such ability is modeled as the customers' probability to make a TIOLIO to the n dealers.[6] On a typical RFQ platform such as MarketAxess, q is effectively zero, as customers can only solicit quotes from dealers but cannot set reserve prices (O'Hara and Zhou, 2020). Instead, when trading is less formally organized, q can be larger. The BWIC to

[6] In our setup with homogeneous dealers, both setting a reserve price and making a TIOLIO allow customers to extract full trading surplus. They are therefore equivalent.

sell CLOs is conducted by email, and it is possible that customers communicate their indicative reserve price through such emails. In housing markets, sellers often post indicative asks that are negotiable.

Remark 4. Dealers sometimes broadcast indicative bids and offers to customers on electronic platforms (called "dealer runs," Section III.B of Bessembinder, Spatt, and Venkataraman, 2020). This feature likens the RFQ interpretation of the current model to models of "directed search," where dealers first post quotes and then customers direct their queries to selected dealers (see, e.g., Wright et al., 2020). In Appendix A, we derive a directed search model and show that it is the limiting case of $\psi \to 1$, i.e., when the signals of dealer inventories become perfect in our special case of "random matching with signals." Intuitively, under directed search, dealers post prices that reveal their types (with or without inventories), and such prices effectively become fully-revealing signals of their inventories. In the other extreme of $\psi \to \frac{1}{2}$ we get "pure random matching," as in DGP and HLW. More generally, we show in Appendix A that the directed search equilibrium obtains when the search capacity $n \to \infty$ in our main model. Thus, our setup bridges "random matching" and "directed search". See also Shi (2019) for an alternative setup connecting random matching with directed search.

Remark 5. This paper focuses on SMS technology like RFQ platforms, which, to our knowledge, only lets customers search for dealers, not the other way, in most real-world applications. The model thus shuts down dealer-to-customer searches. However, in reality, dealers probably do take initiatives to reach customers (though not via SMS) to, e.g., arrange riskless principal trades. Studying dealers' search for customers will be an interesting and fruitful future research that goes beyond the scope of this paper. See, e.g., Saar et al. (2019) for an analysis of dealers' matchmaking versus market making.

3 Stationary equilibrium

There are three sets of equilibrium objects: 1) the demographics $\{m_\sigma\}$ (Section 3.1); 2) the value functions $\{V_\sigma\}$ (Section 3.2); and 3) the split of trading gains (Section 3.5, together with the dealer's pricing strategies). These objects in general can depend on time. We look for a stationary Markov perfect equilibrium, under which the objects are time-invariant constants. We also focus on symmetric equilibrium; that is, the agents of the same type have the same value functions and receive the same fraction of the trading gains. Along the way, Section 3.3 and 3.4 also discuss how welfare is affected by search parameters.

3.1 Demographics

There are in total six demographic variables, $\{m_{ho}, m_{ln}, m_{hn}, m_{lo}, m_{do}, m_{dn}\}$, one for each agent type. The following three conditions must hold in equilibrium by definition:

(2) market clearing: $m_{ho} + m_{lo} + m_{do} = s;$

(3) total customer mass: $m_{ho} + m_{ln} + m_{hn} + m_{lo} = 1;$ and

(4) total dealer mass: $m_{do} + m_{dn} = m_d.$

In a stationary equilibrium, the total measure of $h$igh type customers must be time-invariant; i.e., the net flow during any instance $\mathrm{d}t$ must be zero:

(5) net flow of high type customers: $(m_{lo} + m_{ln})\lambda_u \cancel{\mathrm{d}t} - (m_{ho} + m_{hn})\lambda_d \cancel{\mathrm{d}t} = 0,$

which also ensures that the net flow of low type customers is zero.

Two more equations are needed in order to pin down the six demographic variables. These two last conditions arise from trading. In equilibrium, only two types of customers want to trade with dealers: The lo-type wants to sell to dn-buyer, and the hn-type wants to buy from do-seller. The other two types, ho and ln, stand by and do not trade (which is a conjecture for now, and we will later verify it in Lemma 2, after Proposition 1).

Consider the inflows to and the outflows from the the lo-sellers. In a short period of $\mathrm{d}t$, a measure of $m_{lo}\rho\mathrm{d}t$ of sellers will be searching, each having probability

$$v_{lo} = v(\pi_{dn}) := 1 - (1 - \pi_{dn})^n$$

to find at least one dn-buyer (out of n) to trade with.[7] Hence, there is an outflow of $\rho m_{lo}v_{lo}\mathrm{d}t$ due to the searching lo-sellers. In addition, due to preference shocks, there is an inflow of $\lambda_d m_{ho}\mathrm{d}t$ and an outflow of $\lambda_u m_{lo}\mathrm{d}t$. In a stationary equilibrium, the sum of the in/outflows above must be zero:

$$\text{(6)} \qquad \text{net flow of } lo\text{-sellers:} \qquad -\rho m_{lo}v_{lo} - \lambda_u m_{lo} + \lambda_d m_{ho} = 0.$$

Analogously, define $v_{hn} = v(\pi_{do})$ as the probability for a searching hn-buyer to find at least one do-seller. Then the zero net flow condition for hn-buyers becomes

$$\text{(7)} \qquad \text{net flow of } hn\text{-buyers:} \qquad -\rho m_{hn}v_{hn} - \lambda_d m_{hn} + \lambda_u m_{ln} = 0,$$

which is the last equation needed to pin down the stationary demographics:

Lemma 1 (Stationary demographics). The demographics Equations (2)-(7) uniquely pin down the population sizes $\{m_{ho}, m_{ln}, m_{hn}, m_{lo}\} \in (0, 1)^4$ and $\{m_{do}, m_{dn}\} \in (0, m_d)^2$.

There are other stationarity conditions. For example, the hn-buyer-initiated trading volume amounts to $\rho m_{hn}v_{hn}$, while the lo-seller-initiated volume is $\rho m_{lo}v_{lo}$. They are also, respectively, the asset flow out of and into the dealer sector. Therefore, the trading volume intensity t must satisfy

$$\text{(8)} \qquad t := \rho m_{hn}v_{hn} = \rho m_{lo}v_{lo},$$

for otherwise the dealer-owner mass, m_{do}, will not be stable. Indeed, Equation (8) is guaranteed by (6) − (7) + (5). Appendix C.1 shows that Equations (2)-(7) are indeed sufficient for the stationarity of all other types of agents.

[7] The exact law of large numbers in Duffie, Qiao, and Sun (2019) is applied so that the fractions of the populations of each type are their expected values. See also Sun (2006) and Duffie and Sun (2007, 2012).

3.2 Value functions

Denote by V_σ a type-σ agent's value function. Then the reservation values for the asset are

$$R_l := V_{lo} - V_{ln},\ R_h := V_{ho} - V_{hn},\ \text{and}\ R_d := V_{do} - V_{dn}$$

for the low-type customers, the high-type customers, and the dealers, respectively. For an lo-seller and a dn-dealer to trade, the price p must fall between $R_l \leq p \leq R_d$; and likewise, for an hn-buyer and a do-dealer to trade, the price must fall between $R_d \leq p \leq R_h$. Such prices split the trading gains, which are written as, respectively,

$$\Delta_{dl} := R_d - R_l \text{ and } \Delta_{hd} := R_h - R_d \tag{9}$$

for the two kinds of trades. For now, we make the conjecture that there are positive trading gains: $R_l \leq R_d \leq R_h$, which will be guaranteed by a condition on y_d (see Proposition 1 below).

3.2.1 The split of the trading gain

The split of the trading gain between an hn-buyer and a do-dealer can always be written as $\gamma_{hn}\Delta_{hd}$ and $(1-\gamma_{hn})\Delta_{hd}$, where $\gamma_{hn} \in [0,1]$ represents the hn-buyer's "expected trading gain share." Likewise, the split of trading gain between a dn-dealer and an lo-seller can be written as $(1-\gamma_{lo})\Delta_{dl}$ and $\gamma_{lo}\Delta_{dl}$ for some lo-seller's expected trading gain share $\gamma_{lo} \in [0,1]$.

The shares $\{\gamma_{hn}, \gamma_{lo}\}$ reflect how in expectation prices are set between trading pairs. For now we allow general $\{\gamma_{hn}, \gamma_{lo}\} \in [0,1]^2$, so that our results up to Section 3.5 hold for arbitrary price-setting mechanisms, up to two minimal assumptions: (i) a trade occurs whenever at least one of the n contacts is a match; and (ii) the shares $\{\gamma_{hn}, \gamma_{lo}\}$ do not depend on the value functions $\{V_\sigma\}$ (but they can depend on any other endogenous variables). We complete the equilibrium characterization by deriving $\{\gamma_{hn}, \gamma_{lo}\}$ endogenously in Section 3.5, verifying also the two minimal assumptions.

3.2.2 Hamilton-Jacobi-Bellman equations

Consider first an ho-bystander. Over a short period $\mathrm{d}t$, the ho-bystander gets utility $y_h \mathrm{d}t$ from holding the asset; plus, with intensity $\lambda_d \mathrm{d}t$, she switches to lo-type and her value changes by $V_{lo} - V_{ho}$; minus $rV_{ho}\mathrm{d}t$ due to discounting. Hence, her HJB equation is

$$0 = y_h + \lambda_d \cdot (V_{lo} - V_{ho}) - rV_{ho}. \tag{10}$$

Similarly, an ln-bystander has HJB equation

$$0 = \lambda_u \cdot (V_{hn} - V_{ln}) - rV_{ln}. \tag{11}$$

Consider next an lo-seller. Over $\mathrm{d}t$ units of time, her value increases by $y_l \mathrm{d}t$ due to the asset holding. It may also change by $V_{ho} - V_{lo}$ with intensity $\lambda_u \mathrm{d}t$ due to a preference shock. The value also reduces by $rV_{lo}\mathrm{d}t$ due to discounting. Apart from these three, there is trading, from which she expects an instantaneous trading gain of $\rho \nu_{lo} \gamma_{lo} \Delta_{dl} \mathrm{d}t$—she searches for dealers at intensity ρ, finds at least one match out of the n contacts with probability ν_{lo}, and expects a trading gain share of γ_{lo}. For notation simplicity, we write $\zeta_{lo} := \rho \nu_{lo} \gamma_{lo}$ as an lo-seller's "expected trading gain intensity." Therefore, the HJB equation for an lo-seller is

$$0 = y_l + \lambda_u \cdot (V_{ho} - V_{lo}) - rV_{lo} + \zeta_{lo} \Delta_{dl}. \tag{12}$$

Similarly, an hn-buyer has

$$0 = \lambda_d \cdot (V_{ln} - V_{hn}) - rV_{hn} + \zeta_{hn} \Delta_{hd}, \tag{13}$$

where the expected trading gain intensity is $\zeta_{hn} := \rho \nu_{hn} \gamma_{hn}$.

Finally, consider the dealers. A do-seller's HJB equation has the similar structure as before

$$0 = y_d - rV_{do} + \zeta_{do} \Delta_{hd}, \tag{14}$$

just without the type-switching term because dealers do not receive preference shocks. To find a do-seller's expected trading gain intensity ζ_{do}, note that the total trading gain from all hn-buyer initiated trades amounts to $m_{hn}\rho\nu_{hn}\Delta_{hd}$. Since each hn-buyer expects $\zeta_{hn}\Delta_{hd}$, a do-seller gets the per capita remainder; that is,

$$\zeta_{do} := \frac{m_{hn}\rho\nu_{hn} - m_{hn}\zeta_{hn}}{m_{do}} = \frac{m_{hn}\rho\nu_{hn}}{m_{do}}(1-\gamma_{hn}).$$

Similarly, a dn-buyer has

$$0 = -rV_{dn} + \zeta_{dn}\Delta_{dl} \tag{15}$$

with expected trading gain intensity

$$\zeta_{dn} := \frac{m_{lo}\rho\nu_{lo} - m_{lo}\zeta_{lo}}{m_{dn}} = \frac{m_{lo}\rho\nu_{lo}}{m_{dn}}(1-\gamma_{lo}).$$

Recall from Equation (9) that both trading gains Δ_{hd} and Δ_{dl} are linear combinations of the value functions $\{V_\sigma\}$. Thus, Equations (10)-(15) constitute a linear system with six equations and six unknowns, solved by the proposition below.

Proposition 1 (Equilibrium value functions). Define the thresholds $\overline{y}_d$ and $\underline{y}_d$ as

$$\overline{y}_d := y_h - (y_h - y_l)\frac{\lambda_d}{\lambda_d + \lambda_u + r} \quad \text{and} \quad \underline{y}_d := y_l + (y_h - y_l)\frac{\lambda_u}{\lambda_d + \lambda_u + r}.$$

When $\underline{y}_d \le y_d \le \overline{y}_d$, the reservation values satisfy $0 < R_l < R_d < R_h$ and the value functions are the solution to the linear equation systems (10)-(15). (Note that $\overline{y}_d > \underline{y}_d$ always holds.)

The parameter constraint of $y_d \in (\overline{y}_d, \underline{y}_d)$ ensures positive trading gains, i.e., $\Delta_{hd} = R_h - R_d > 0$ and $\Delta_{dl} = R_d - R_l > 0$. (Similar conditions are also seen in the literature; e.g., Proposition 3 of HLW.) When $y_d \notin (\overline{y}_d, \underline{y}_d)$, intuitively, the dealers are no longer "intermediaries" between buyers and sellers and the economy might enter a steady state without trading.[8] In the rest of the analysis, therefore, we focus on the more interesting and empirically relevant case with trades by assuming that $y_d \in (\overline{y}_d, \underline{y}_d)$ always holds.

Finally, the lemma below verifies the conjecture that ho- and ln-customers do stand by:

[8] For example, suppose $y_d \notin (\overline{y}_d, \underline{y}_d)$ and $R_d > R_h$. Then do-dealers and hn-buyers do not trade, and by the stationarity condition (8), there must be no trade between dn-dealers and lo-sellers, either.

Lemma 2. When $\Delta_{hd} > 0$ and $\Delta_{dl} > 0$, both *ho*- and *ln*-customers stay out of trading.

3.3 Search technology and allocation efficiency

This subsection examines how allocation efficiency is affected by search technology. We are particularly interested in the contrast of the two search parameters, the intensity ρ and the capacity n—how fast customers can find dealers versus how many dealers can be reached in one "click."

We focus on the case in which the asset is in excess supply, formally defined below:

Lemma 3. The *hn*-buyers are on the short side of the market, i.e., $m_{hn} < m_{lo}$, if and only if

$$s > \eta + \frac{1}{2}m_d, \text{ where } \eta := \frac{\lambda_u}{\lambda_u + \lambda_d}. \tag{16}$$

Intuitively, the threshold on the right-hand side of (16) is the "intrinsic demand" for the asset: The fraction η is the population size of the steady-state high-type customers, who are natural holders of the asset. In addition, since the dealers are homogeneous, half of them are also natural holders of the asset. When the asset supply s exceeds such intrinsic demand, the *hn*-buyers are on the short side and the *lo*-sellers on the long side. (The case of excess demand, $s \leq \eta + \frac{1}{2}m_d$, is symmetric and is omitted for brevity.)

Trading volume and customer sizes. Figure 1(a) illustrates how the trading volume intensity t increases with both the search intensity ρ and the capacity n. This is intuitive: Both parameters improve the matching efficiency between customers and dealers. Instead, Figure 1(b) and 1(c) contrast between how ρ and n affect *differently* the customer population sizes. Specifically, while n monotonically reduces m_{lo} (the long side), it can increase m_{hn} (the short side) when ρ is sufficiently small. That is, a larger search capacity n might exacerbate inefficient allocation.

Proposition 2 (Search technology and customer sizes). The trading volume intensity t increases in both n and ρ. The search intensity ρ always reduces both m_{hn} and m_{lo}. The search capacity n always reduces the long-side customer mass but has ambiguous effect on the short-side customer mass. In particular, when ρ is sufficiently small, the short-side customer mass increases with n.

The bottleneck effect. The reason that a larger search capacity n can hurt welfare is because it creates a "bottleneck" at dealers, hindering the efficient passing through of the asset from *lo*-sellers to *hn*-buyers. To see how, first note that a larger search capacity n does help matching: Both the probabilities of v_{lo} and v_{hn} of finding at least one dealer counterparty increase with n (Lemma D.2). However, the magnitudes of the increases are far from equal. The increment in v_{lo} is much more substantial than that in v_{hn}, as illustrated in Figure 2. This is because the *hn*-buyers are on the short side of the market and there are many more *do*-dealers to be easily found (than *dn*-dealers for the long side *lo*-sellers). Correspondingly, v_{hn} is much closer to its upper bound of 100% and cannot increase by as much as v_{lo}. Thus, all else being equal, an increase in n matches many more *lo*-*dn* pairs than *hn*-*do* pairs.

The *lo*-*dn* trades let the asset flow into the dealer sector, while the *hn*-*do* trades let the asset flow out of the dealers. The above asymmetric effects of n—the substantially larger increment in the inflow than in the outflow—clog the asset at the dealers, hence the "bottleneck." Therefore, with a larger n, the dealers take in a lot of the asset from *lo*-sellers but only give out little to *hn*-buyers. Consequently, the *hn*-buyer size m_{hn} increases and the *lo*-seller size m_{lo} reduces.

Summing up, there are two pairs of asymmetric effects: In terms of matching probability, v_{lo} increases much faster than v_{hn}. In terms of population sizes, m_{lo} shrinks, whereas m_{hn} balloons. These effects ensure the stationarity of dealers in equilibrium, with $\rho v_{lo} m_{lo} = \rho v_{hn} m_{hn}$ (Equation 8).

It is worth emphasizing that the bottleneck arises only with the search capacity n but *not* with the intensity ρ. This is because ρ scales up both the inflow $\rho v_{lo} m_{lo}$ and the outflow $\rho v_{hn} m_{hn}$ and there is no asymmetry. This is a novel finding, thanks to the flexibility of n. For example, while the SMS setup is borrowed from HLW, the bottleneck does not manifest there, as their customers and dealers only meet bilaterally (i.e., $n = 1$).

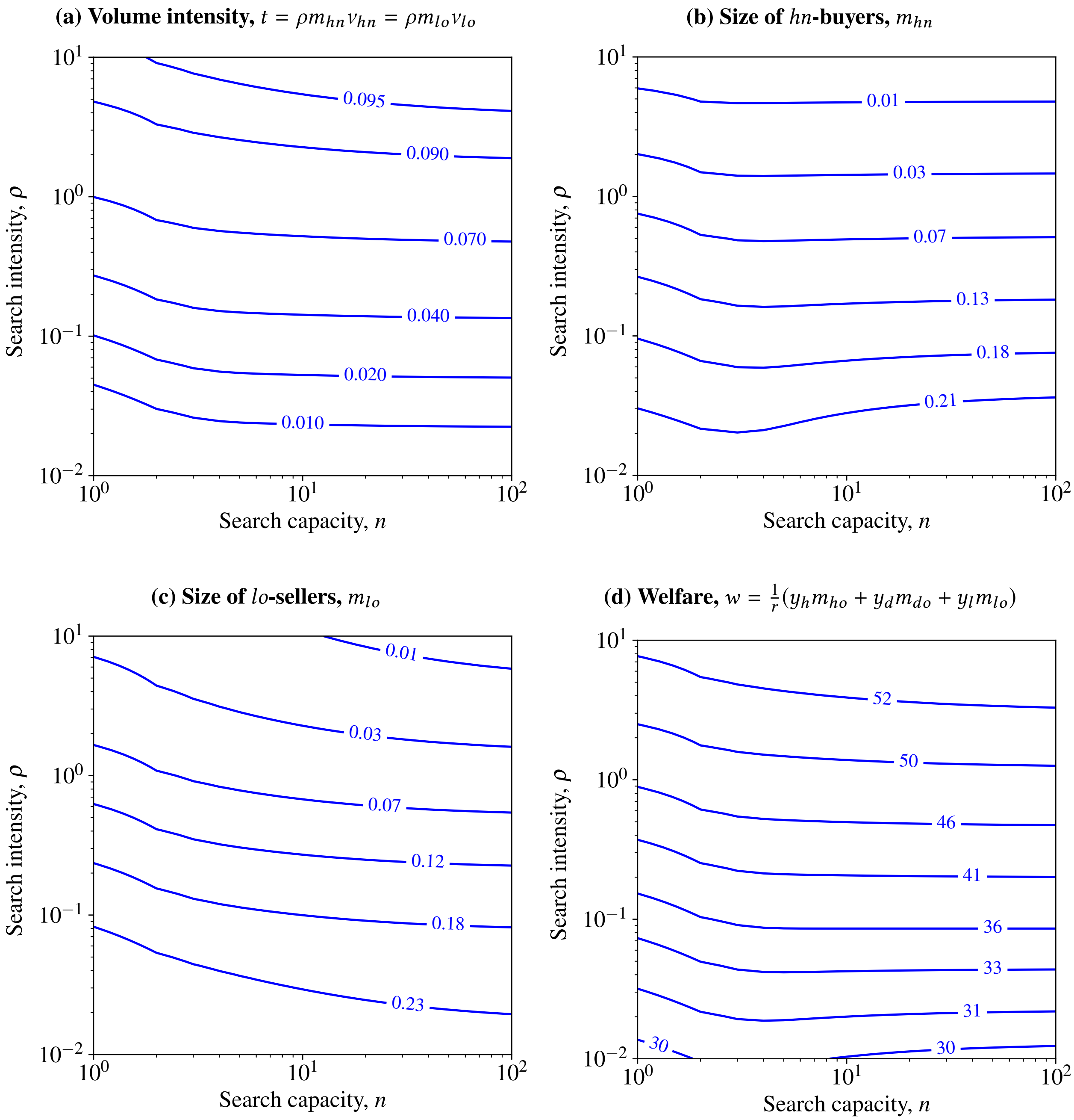

Figure 1: Customer sizes, trading volume, and welfare. This figure plots how the search intensity ρ and the search capacity n affect trading volume intensity in (a), customer sizes in (b) and (c), and welfare in (d). Apart from ρ and n, the other parameters are set at $s = 0.6$, $m_d = \lambda_d = \lambda_u = r = 0.1$, $y_h = 10.0$, $y_d = 3.5$, and $y_l = 0.0$. The customers' intrinsic bargaining power q is irrelevant here as the equilibrium demographics do not depend on it; see Equations (2)-(7) and Lemma 1. Additional, q is not relevant for welfare, as it only affects the split of the trading gain but not the total size of it.

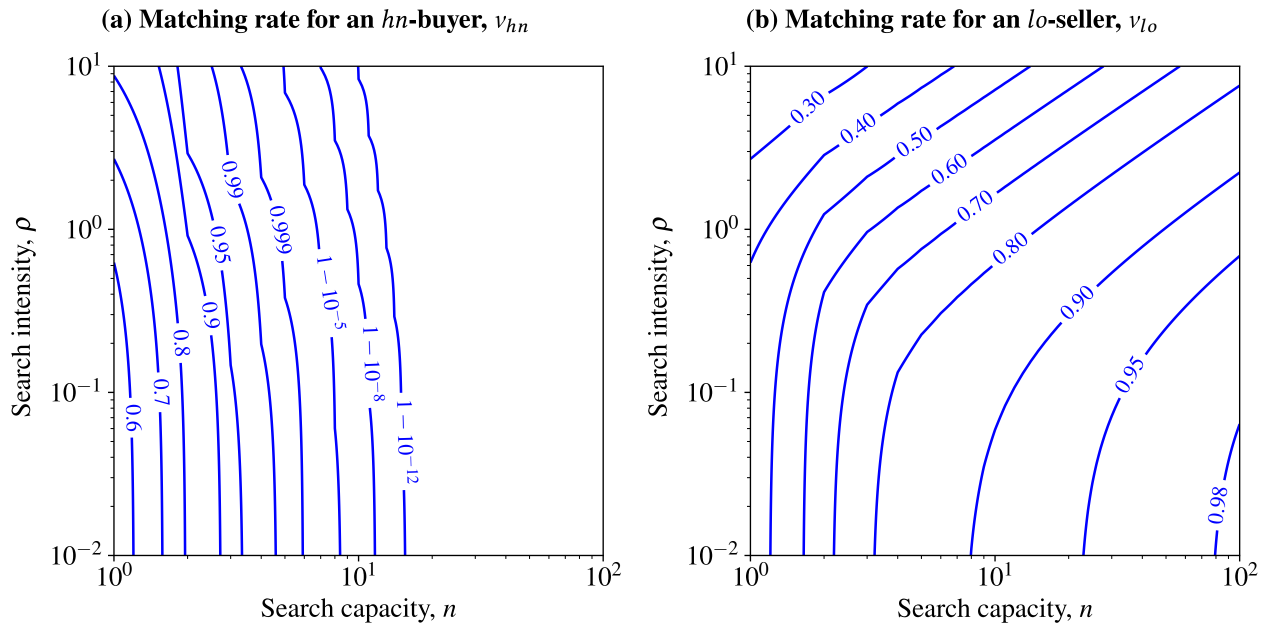

Figure 2: Matching rates. This figure plots how the search intensity ρ and the search capacity n affect the customers' matching rates with dealers. Other than ρ and n, the parameters are set at $s = 0.6$, $m_d = \lambda_d = \lambda_u = r = 0.1$, $y_h = 10.0$, $y_d = 3.5$, and $y_l = 0.0$. The customers' intrinsic bargaining power q is irrelevant here, as the equilibrium demographics do not depend on it; see Equations 2-7 and Lemma 1.

Welfare. Welfare in this economy is the present value of all asset-owners' utility flows:

$$w := \frac{1}{r}(y_h m_{ho} + y_d m_{do} + y_l m_{lo}).$$

Note that, unsurprisingly, only the demographics matter for welfare, because the pricing strategies (Section 3.5) only affect the split of trading gains, not the size of the "pie." We rewrite welfare as

$$w = \frac{y_d}{r} m_{do} + \frac{1}{r} \underbrace{\frac{m_{ho} y_h + m_{lo} y_l}{m_{ho} + m_{lo}}}_{=:\hat{y}} (m_{ho} + m_{lo}) = \frac{y_d}{r} m_{do} + \frac{\hat{y}}{r}(s - m_{do}), \tag{17}$$

where the second equality follows the market clearing condition (2). That is, the m_{do} mass of dealers get a flow utility y_d, while the $(s - m_{do})$ mass of customer-owners get average flow utility $\hat{y}$.

When the search intensity ρ is low, $\hat{y}$ is approximately

$$\hat{y} := \frac{m_{ho} y_h + m_{lo} y_l}{m_{ho} + m_{lo}} \approx \eta y_h + (1 - \eta) y_l \tag{18}$$

because the stationarity condition (6) implies that $m_{lo}\lambda_u \approx m_{ho}\lambda_d$ and hence $\frac{m_{ho}}{m_{lo}} \approx \frac{\lambda_u}{\lambda_d} = \frac{\eta}{1-\eta}$.

Expressions (17) and (18) highlight that for low ρ, welfare simply depends on the split of the asset between dealers (m_{do}) and customers ($s - m_{do}$). In particular, welfare increases (decreases) with m_{do} when $y_d > \hat{y}$ ($< \hat{y}$), as dealers are the better (worse) users than an *average* customer. As discussed above, when n increases, a "swelling" bottleneck of m_{do} captures the asset from the customers. Welfare losses then occur with larger n, if $y_d < \hat{y}$.

Proposition 3 (Search technology and welfare). A higher search intensity ρ always improves welfare. A larger search capacity n improves welfare when ρ is sufficiently high. However, if ρ is low enough and if $y_d < \hat{y} \in (\underline{y}_d, \overline{y}_d)$, a larger n reduces welfare.

Do trading technologies with larger search capacity n, such as the RFQ platforms, improve welfare? The proposition above highlights that the answer depends on how 'fast" the markets are (i.e., how high ρ is). Notably, the model caveats that in slow markets, there might be welfare losses due to inefficient distribution of inventory between customers and dealers, as numerically illustrated in Figure 1(d).

3.4 Inventory transparency and allocation efficiency

Our analysis so far admits generic forms of $\pi(\cdot)$, which is the probability for a searching customer's contact to be a match. Below we turn to the specific parametrization of "random matching with signals" (p. 8). That is, $\pi(\cdot)$ is now given as in Equation (1). We interpret customers' signal quality ψ as *transparency* of dealers' inventories and examine how it affects welfare.

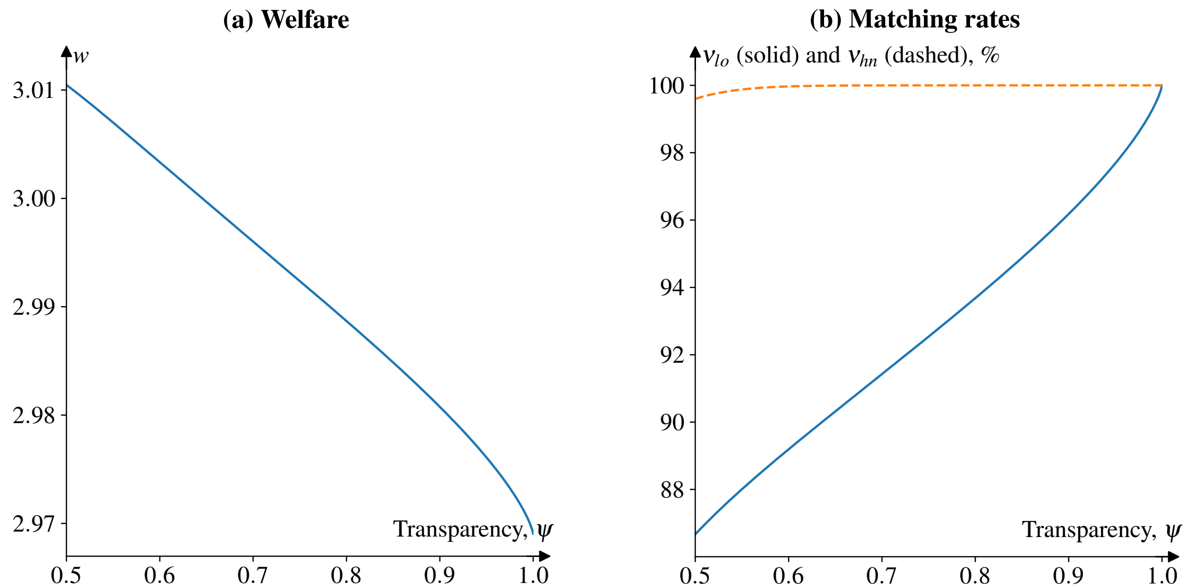

Figure 3: Effects of transparency, ψ. This figure plots how transparency ψ affects welfare in Panel (a) and the matching rates in Panel (b). The illustration is for a low level of $\rho = 0.01$. The other parameters are set at $n = 5$, $s = 0.6$, $m_d = \lambda_d = \lambda_u = r = 0.1$, $y_h = 10.0$, $y_d = 3.5$, and $y_l = 0.0$. The customers' intrinsic bargaining power q is irrelevant here, as neither the equilibrium demographics or welfare depends on it.

Proposition 4 (Transparency and welfare). Better inventory transparency ψ improves welfare only when ρ is sufficiently high. If ρ is low enough and if $y_d < \hat{y} \in (\underline{y}_d, \overline{y}_d)$, ψ reduces welfare.

Figure 3(a) illustrates a case of low ρ, where improved transparency hurts welfare. The key intuition is that the change in transparency ψ asymmetrically affects customers on the short and the long sides of the market, similar to the asymmetric effects of n in Section 3.3. A higher ψ improves matching by increasing both ν_{hn} and ν_{lo}, helping customers direct more accurately their searches to dealers with the right inventory capacity. Yet, the short side matching rate (ν_{hn}) increases much less than the long side (ν_{lo}), since it is already close to the upper bound of 100%, as illustrated in Figure 3(b). The bottleneck again emerges and possibly hurts welfare, mirroring Proposition 3.

The dissemination of post-trade information of corporate bonds via TRACE (Transaction Reporting and Compliance Engine), starting in 2002, was perhaps the most significant transparency shock in the corporate bonds market. A large volume of the literature has documented its impact on market quality, applauding the improved liquidity and the reduced trading costs; see, e.g., Bessembinder, Maxwell, and Venkataraman (2006), Edwards, Harris, and Piwowar (2007), and, Goldstein, Hotchkiss, and Sirri (2007). To the extent that post-trade transparency from TRACE also improves customers' inference about dealer inventories, our model cautions that the resulting better matching—the improved "liquidity"—not necessarily always translates to better welfare in terms of allocation (Figure 3a vs. 3b). In particular, the reduction of search frictions might only force the intermediaries to inefficiently hoard corporate bonds. Our model thus also highlights the importance of examining the impact on dealer inventories of such transparency shocks.

3.5 The endogenous split of trading gain

The analysis so far only requires the general form of expected trading gain shares, $\{\gamma_{hn}, \gamma_{lo}\}$. Under the assumed trading mechanism ("Price determination" on p. 9), such expected trading gain shares can be *endogenously* determined:

> **Corollary 1 (The split of trading gains).** Define $\gamma(\pi, n) := q + (1-q)\Big(1 - \frac{n\pi\cdot(1-\pi)^{n-1}}{1-(1-\pi)^n}\Big)$ for $\pi \in (0,1)$ and $n \in \mathbb{N}$. Then, $\gamma_{hn} = \gamma(\pi_{do}, n)$ and $\gamma_{lo} = \gamma(\pi_{dn}, n)$.

Several remarks are in order. First, as long as $n \geq 2$, $\gamma(\cdot)$ is strictly increasing in π, from $\gamma(0,n) = q$ to $\gamma(1,n) = 1$. Take, for example, an hn-buyer searching for do-dealers. With a higher π_{do}, each contacted do-dealer knows that she is more likely competing with some other do-dealers among the other $(n-1)$ contacts. Such fiercer competition gives more trading gains to the hn-buyer. Indeed, when $\pi_{do} \to 1$, the competition is perfect and hn-buyers extract full surplus with $\gamma_{hn} \to 1$. On the other extreme, if $\pi_{do} \to 0$, each do-dealer knows that she is likely the monopolist among all n contacted and therefore quotes a monopolistic price. Indeed, as $\pi_{do} \to 0$, $\gamma_{hn} \to q$, which is the baseline probability that the customer can make a TIOLIO to all dealers.

Second, when $n = 1$, $\gamma_{hn} = \gamma_{lo} = q$, as if the searching customer engages in a Nash bargaining with one matched dealer with respective bargaining power parameters q and $1 - q$. Our setup thus nests such exogenous splits of trading gains, commonly assumed in the literature (see, e.g., DGP and HLW). When $n \geq 2$, our model highlights that under SMS, the expected trading gain shares $\{\gamma_{hn}, \gamma_{lo}\}$ are endogenous, in particular, of the dealer demographics m_{do} and m_{dn}. Such an endogenous split of trading gains is a distinguishing feature of our model.

Finally, whenever $n \geq 2$, the contacted dealers compete against an *unknown number of others*, as some of the n contacted dealers might not be of the matching type. That is, every contacted matching dealer knows that there is a non-zero probability that she actually is the only match. As is known in the literature (e.g., Burdett and Judd, 1983), in this case, dealers follow mixed strategies in setting their prices. This suggests that dealers' strategic behavior can be a source of price dispersion. Even though dealers are homogeneous in our model, it still features price dispersion, a robust empirical feature of OTC markets. For example, Hendershott and Madhavan (2015) document a significant dispersion in dealers' responding quotes in corporate bond market. Hau et al. (2017) find evidence for price dispersion in foreign exchange derivatives.

4 SMS versus BB: How to search

In real-world trading, investors can choose their trading technologies. For example, while bilateral bargaining is still the dominant form of trading in corporate bonds, electronic platforms with RFQ protocols have been on the rise (O'Hara and Zhou, 2020). We consider investors' choice of "Click or Call" (Hendershott and Madhavan, 2015) in this section.

In the framework set up in Section 2, we introduce two trading technologies, BB and SMS, which differ in parameters $\{n^k, \rho^k, q^k\}$, $k \in \{\text{BB}, \text{SMS}\}$. (Some realistic parameter restrictions are imposed below.) Each customer can choose, at any point in time, which technology to use to contact dealers, if she wants to trade. All dealers can be reached either by BB or by SMS. The other model ingredients remain the same as in Section 2.

Section 4.1 analyzes how customers choose between the two technologies in a steady state equilibrium. We then examine whether SMS-like electronic trading (e.g., RFQ) can completely replace traditional bilateral bargaining. The answer is no, as Section 4.2 shows that in stress periods (e.g., after a fire sale), BB is used more often than SMS. Finally, Section 4.3 draws implications on welfare, policy, and market design.

Parameter constraints: Motivated by "calls" (BB) and "clicks" (SMS), we assume

$$n^{\mathrm{BB}} = 1, \quad n^{\mathrm{SMS}} > 2, \text{ and } \rho^{\mathrm{BB}} \leq \rho^{\mathrm{SMS}}. \tag{19}$$

In a bilateral call, a customer bargains with one dealer, hence $n^{\mathrm{BB}} = 1$. By clicking, a typical real-world RFQ protocol connects the customer to multiple dealers, at least three in most of the applications (see Remark 3), hence $n^{\mathrm{SMS}} > 2$.[9] Earlier research has shown that electronic platforms like MarketAxess can "provide considerable time savings relative to ... bilateral negotiations" (Hendershott and Madhavan, 2015); and can "improve the speed of execution" (O'Hara and Zhou, 2020), motivating our assumption of $\rho^{\mathrm{BB}} \leq \rho^{\mathrm{SMS}}$.

The probabilities to set prices in respective technology, q^{BB} and q^{SMS}, also play an important role. In most of the applications (e.g., MarketAxess), a customer using RFQ is always on the receiving end of dealers' TIOLIOs, suggesting that $q^{\mathrm{SMS}} = 0$. On the other hand, in bilateral calls, there is always room for negotiation and it is natural to expect that $q^{\mathrm{BB}} > 0$. We impose no such constraints here and proceed to examine how q^{SMS} and q^{BB} affect the customers' technology choices.

4.1 Choosing between SMS and BB

The analysis focuses on steady states, as before. There are three sets of equilibrium objects: (i) customers' optimal technology choices, (ii) demographics, and (iii) value functions. Compared to

[9] Excluding the special case of $n^{\mathrm{SMS}} = 2$ reduces the cases to consider when characterizing the equilibrium, streamlining the exposition. The full characterization for $n^{\mathrm{SMS}} \geq 2$ is provided in the proof of Proposition 5.

Section 3, the novel part is the analysis of (i), which is the focus below. The analyses of (ii) and (iii) are analogous to those in Section 3 and, hence, collated in Appendix C.2-C.3.

Recall from Section 2 that the customers can be categorized into four types, $\sigma \in \{ho, ln, hn, lo\}$. Now the customers of each type-σ can be further split into subtypes σ-BB and σ-SMS. We distinguish these two subtypes by superscripting the relevant variables with the chosen technology $k \in \{\text{BB}, \text{SMS}\}$. For example, their masses satisfy $m_\sigma^{\text{BB}} + m_\sigma^{\text{SMS}} = m_\sigma$ and they have (possibly different) value functions V_σ^{BB} and V_σ^{SMS}.

The analysis can be simplified in two ways. First, note that in a stationary equilibrium, the value functions are time-invariant. That is, if a type-σ customer prefers one technology over the other at some point of time, her technology choice will persist until her type changes (due either to a preference shock or to trading). Hence, without loss of generality, we can focus on a type-σ customer's technology choice at the moment she becomes type-σ. Second, both ho and ln customers will be bystanders in equilibrium and do not trade—a result following positive trading gains (Lemma 2). Therefore, there is no need to distinguish ln^{SMS} versus ln^{BB} or ho^{SMS} versus ho^{BB}. Only the technology choices of the trading customers, hn and lo, need to be studied below.

Denote by $\theta_\sigma \in [0, 1]$ the probability of a newborn type-σ customer choosing SMS (hence choosing BB with probability $1 - \theta_\sigma$), where $\sigma \in \{hn, lo\}$. Then

$$\theta_\sigma \begin{cases} = \mathbb{1}_{\{V_\sigma^{\text{SMS}} > V_\sigma^{\text{BB}}\}}, & \text{if } V_\sigma^{\text{SMS}} \neq V_\sigma^{\text{BB}}; \\ \in [0, 1], & \text{if } V_\sigma^{\text{SMS}} = V_\sigma^{\text{BB}}. \end{cases} \tag{20}$$

We shall focus on *symmetric* equilibria, where all customers of type σ choose the same θ_σ.

To sustain an equilibrium, the technology choices $\{\theta_{hn}, \theta_{lo}\}$ must agree with the value functions $\{V_\sigma\}$ according to Equation (20). The value functions are, in turn, chained to $\{\theta_{hn}, \theta_{lo}\}$ via many layers of endogenous variables (see Appendix C.3): the trading gain intensities $\{\zeta_\sigma\}$, the dealers' pricing, and the many demographic variables $\{m_\sigma\}$. This is a big fixed-point problem. It turns out that the equilibrium $\{\theta_{hn}, \theta_{lo}\}$ ultimately boil down to comparing the size of dealer-owners

m_{do} (and $m_{dn} = m_d - m_{do}$) with some threshold. Recalling that $\pi(\cdot)$ is monotone, the lemma below restates this comparison in terms of $\pi_{do} = \pi\left(\frac{m_{do}}{m_d}\right)$ and $\pi_{dn} = \pi\left(\frac{m_{dn}}{m_d}\right)$:

Lemma 4. If the technologies satisfy

$$\rho^{\text{SMS}} q^{\text{SMS}} n^{\text{SMS}} < \rho^{\text{BB}} q^{\text{BB}} n^{\text{BB}}, \tag{21}$$

then Equation (20) can be equivalently written as

$$\theta_{hn} \begin{cases} = \mathbb{1}_{\{\pi_{do} > \pi^*\}}, & \text{if } \pi_{do} \neq \pi^* \\ \in [0,1], & \text{if } \pi_{do} = \pi^* \end{cases} \text{and } \theta_{lo} \begin{cases} = \mathbb{1}_{\{\pi_{dn} > \pi^*\}}, & \text{if } \pi_{dn} \neq \pi^* \\ \in [0,1], & \text{if } \pi_{dn} = \pi^* \end{cases}, \tag{22}$$

where π^* uniquely solves $z^{\text{SMS}}(\pi) = z^{\text{BB}}(\pi)$, with $z^k(\cdot)$ defined in Equation (24) below for $k \in \{\text{SMS}, \text{BB}\}$. If, instead, $\rho^{\text{SMS}} q^{\text{SMS}} n^{\text{SMS}} \geq \rho^{\text{BB}} q^{\text{BB}} n^{\text{BB}}$, then $\theta_{hn} = \theta_{lo} = 1$.

Below we discuss the key steps behind the lemma. First, the value functions are pinned down by the HJB equations (40)-(43) in Appendix C.3. The proof of Lemma 4 shows that V_σ^k is a monotone function in ζ_σ^k, for $\sigma \in \{lo, hn\}$. This intuitive result says that when a searching customer chooses between SMS vs. BB, she is essentially comparing the trading gain intensities $\zeta_\sigma^{\text{SMS}}$ vs. ζ_σ^{BB}. Hence, the technology choices (20) can be equivalently written as:

$$\theta_\sigma \begin{cases} = \mathbb{1}_{\{\zeta_\sigma^{\text{SMS}} > \zeta_\sigma^{\text{BB}}\}}, & \text{if } \zeta_\sigma^{\text{SMS}} \neq \zeta_\sigma^{\text{BB}}; \\ \in [0,1], & \text{if } \zeta_\sigma^{\text{SMS}} = \zeta_\sigma^{\text{BB}}. \end{cases} \tag{23}$$

Second, analogous to $\{\zeta_{lo}, \zeta_{hn}\}$ in Section 3.2, we write $\zeta_{hn}^k = z^k(\pi_{do})$ and $\zeta_{lo}^k = z^k(\pi_{dn})$, where $z(\cdot)$ is defined for $\pi \in (0,1)$ as

$$z^k(\pi) := \rho^k \nu^k(\pi) \gamma^k(\pi) = \rho^k \cdot \left(1 - (1-\pi)^{n^k}\right)\left(q^k + (1-q^k)\left(1 - \frac{n^k \pi \cdot (1-\pi)^{n^k - 1}}{1 - (1-\pi)^{n^k}}\right)\right). \tag{24}$$

Note that the superscript k is not exponent but the technology $k \in \{\text{BB}, \text{SMS}\}$. That is, customers essentially choose $\{\theta_\sigma\}$ by examining whether and how $z^{\text{SMS}}(\pi)$ and $z^{\text{BB}}(\pi)$ cross each other.

Lemma 4 essentially characterizes such crossing. Under the condition (21), $z^{\text{SMS}}(\pi)$ crosses $z^{\text{BB}}(\pi)$ from below once at $\pi^* \in \left(0, \frac{1}{2}\right)$. That is, a hn-buyer (lo-seller) prefers BB over SMS when

$\pi_{do} < \pi^*$ ($\pi_{dn} < \pi^*$). This might come as a surprise, given that the condition (19) has guaranteed that SMS not only helps reach dealers faster but also induces more competitive quotes. Why would a customer still prefer BB?

To see the potential advantage of BB, consider for example an *hn*-buyer looking for *do*-sellers. Suppose m_{do} is very low and, hence, $\pi_{do} = \pi\left(\frac{m_{do}}{m_d}\right)$ is also very low. Then the *hn*-buyer customer finds one counterparty dealer with probability approximately $n^k \pi_{do}$—one and only one success from n^k Bernoulli draws at rate π_{do}. (As π_{do} is small, the event of finding multiple dealers is negligibly unlikely.) It follows that a successfully contacted dealer in this case knows that she is almost surely a monopolist and will always quote an ask as high as possible, leaving no trading gains to the *hn*-buyer. The customer only gets non-zero trading gains only if she can make a TIOLIO, i.e., with probability q^k. Taken together, for small π, the customers' trading gain intensity is $z^k(\pi) \approx \rho^k \cdot (n^k \pi) \cdot q^k$. Comparing BB with SMS in this case yields:

$$\lim_{\pi \downarrow 0} \frac{z^{\mathrm{BB}}(\pi)}{z^{\mathrm{SMS}}(\pi)} = \frac{\rho^{\mathrm{BB}} n^{\mathrm{BB}} q^{\mathrm{BB}}}{\rho^{\mathrm{SMS}} n^{\mathrm{SMS}} q^{\mathrm{SMS}}}.$$

The condition (21), therefore, ensures that for sufficiently small π, i.e., for relatively few counterparty dealers, BB has an advantage over SMS. In real-world trading, the condition (21) seems to hold because customers using SMS, like RFQ protocols, do not have many opportunities, if at all, to set reserve prices. That is, q^{SMS} is observed to be (close to) zero in the real world.[10]

We are now ready to state the equilibrium.

Proposition 5 (Steady state equilibrium with technology choices). A unique stationary equilibrium exists depending on the asset supply s: There exist thresholds $0 < s_{hn,0} < s_{hn,1} \leq s_{lo,1} < s_{lo,0} < 1 + m_d$ so that

[10] Complementing the condition (21), the condition (19) in turn ensures that SMS is preferred when there are sufficiently many dealer counterparties.That is, $\lim_{\pi \uparrow 1} \left(z^{\mathrm{BB}}(\pi)/z^{\mathrm{SMS}}(\pi)\right) = \rho^{\mathrm{BB}} q^{\mathrm{BB}} / \rho^{\mathrm{SMS}} \leq 1$. It is interesting to note that only q^{BB} appears but not q^{SMS} in the limit of $\pi \uparrow 1$. With $n^{\mathrm{SMS}} > 1$ and $\pi \uparrow 1$, the multiple counterparty dealers in SMS almost always engage in Bertrand competition, and the customer always gets the full trading gain, regardless of q^{SMS}. On the contrary, with $n^{\mathrm{BB}} = 1$, a customer using BB meets only one counterparty dealer, who will always set the monopolist price, leaving surplus to the customer only with probability q^{BB}.

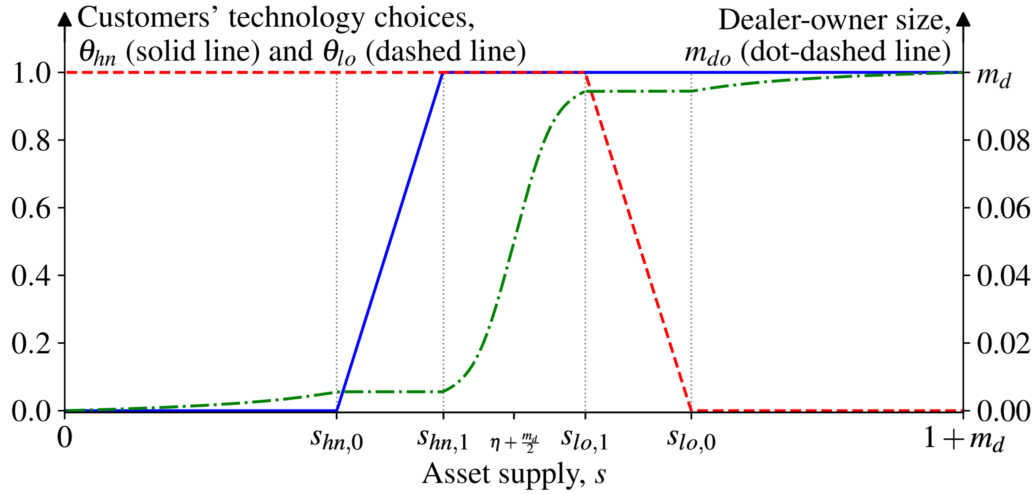

Figure 4: Equilibrium technology choice plotted against asset supply. This figure plots customers' technology choices against the asset supply s in equilibrium. The hn-buyers' choice θ_{hn} is plotted in the solid line, while the lo-sellers' choice θ_{lo} is plotted in the dashed line (the left axis). The dot-dashed line plots the population size of do-seller dealers (the right axis). The technology parameters are set at $(\rho^{\mathrm{BB}}, n^{\mathrm{BB}}, q^{\mathrm{BB}}) = (3.0, 1, 0.5)$ and $(\rho^{\mathrm{SMS}}, n^{\mathrm{SMS}}, q^{\mathrm{SMS}}) = (3.0, 5, 0.0)$. The other parameters are $\lambda_d = \lambda_u = 0.1$, $r = 0.1$, $m_d = 0.1$, $y_h = 10.0$, $y_d = 3.5$, and $y_l = 0.0$.

	(a) hn-buyers' proba -bility to use SMS, θ_{hn}	(b) lo-sellers' proba -bility to use SMS, θ_{lo}	(c) asset holding by dealers, m_{do}
(1) $0 < s \leq s_{hn,0}$	0	1	$g(0, 1, m_{do}) = s$
(2) $s_{hn,0} \leq s \leq s_{hn,1}$	$g(\theta_{hn}, 1, m_d^*) = s$	1	m_d^*
(3) $s_{hn,1} < s < s_{lo,1}$	1	1	$g(1, 1, m_{do}) = s$
(4) $s_{lo,1} \leq s \leq s_{lo,0}$	1	$g(1, \theta_{lo}, m_d - m_d^*) = s$	$m_d - m_d^*$
(5) $s_{lo,0} < s < 1 + m_d$	1	0	$g(1, 0, m_{do}) = s$

where $g(x_1, x_2, x_3) = s$ uniquely solves θ_{hn}, θ_{lo}, and m_{do} in columns (a), (b), and (c), respectively. The constant π^* is given in Lemma 4 and $m_d^* := \pi^{-1}(\pi^*)m_d$. The function $g(\cdot)$ and the thresholds $\{s_{hn,0}, s_{hn,1}, s_{lo,1}, s_{lo,0}\}$ are given in the proof.

Figure 4 illustrates the equilibrium by plotting the technology choices θ_{hn} (solid) and θ_{lo} (dashed) on the left axis and the dealer-owner population size m_{do} (dot-dashed) on the right axis. The four thresholds of $\{s_{hn,0}, s_{hn,1}, s_{lo,1}, s_{lo,0}\}$ cut the support of $s \in (0, 1 + m_d)$ into five regions on the horizontal axis. Consider the solid line, i.e., θ_{hn}, for example. When the asset supply s is extremely low, SMS is very unattractive for the hn-buyers, because they know it is very difficult to find a counterparty do-dealer (the dot-dashed line), and even if they do, they are going to be charged with a monopoly price using SMS. When s is sufficiently high, there are sufficiently many do-dealers, whose price competition makes SMS sufficiently attractive with high trading gain intensity ζ_{hn}^{SMS} for hn-buyers. As such, the solid line flattens at $\theta_{hn} = 1$ for $s > s_{hn,1}$. In between, we see θ_{hn} monotonically increases for $s_{hn,0} \leq s \leq s_{hn,1}$. Such a mixed strategy is supported by the constant $m_{do} = \pi^* m_d$ in the region—the hn-buyers are indifferent to BB and SMS. The pattern for the dashed line, i.e., θ_{lo}, is exactly the opposite, as lo-sellers seek dn-dealers, whose mass is $m_{dn} = m_d - m_{do}$.

4.2 Stress periods

O'Hara and Zhou (2020) show that after downgrade, a corporate bond's electronic (SMS) volume share falls relative to voice trading (BB). The analysis developed above provides a theoretical framework to study investors' endogenous technology choice when under such stress.[11]

One consequence of a corporate bond downgrade is that many previously buy-and-hold long-term investors now no longer wish to hold such bonds. Ambrose, Cai, and Helwege (2008) and Ellul, Jotikasthria, and Lundblad (2011) document such fire sales by insurance companies. In the context of our model, we interpret such fire selling in two different ways, (i) an exogenous increase in the total supply s of the asset—a supply shock; and / or (ii) an exogenous increase in the customers' intensity of drawing low preference λ_d—a demand shock. (A third alternative, reducing λ_u, is equivalent to increasing λ_d and is omitted for brevity.) To fit the fire-selling interpretation, we also assume that the asset is in excess supply, as defined in Lemma 3.

[11] It is worth emphasizing that following the steady state equilibrium in Section 4.1, the results in this subsection also pertain only to steady states, e.g., before and after corporate bond downgrades.

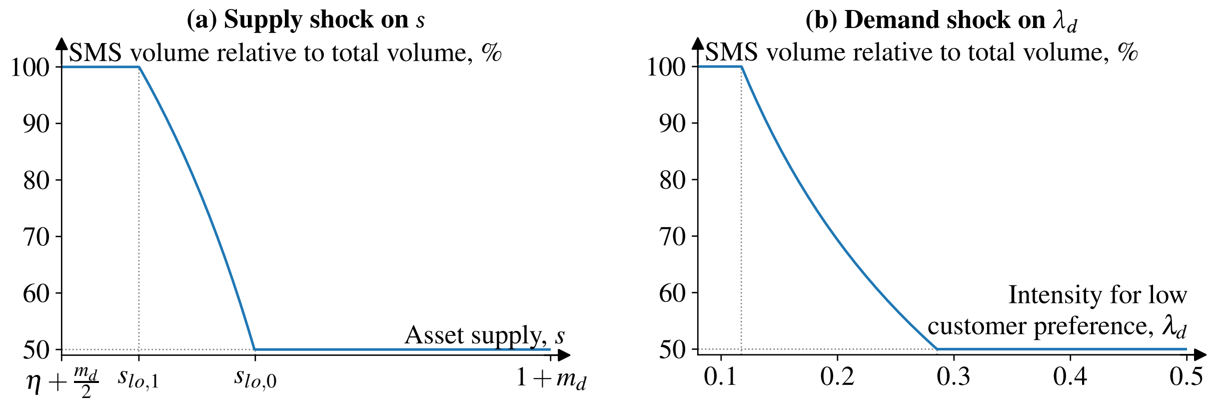

Figure 5: Usage of SMS in a stationary equilibrium after surges in supply. This figure plots the usage of SMS (in a stationary equilibrium)—SMS volume relative to total volume—when the asset supply s surges in Panel (a) and when the customers' low-valuation preference shock intensity λ_d increases in Panel (b). The technology parameters are set at $(\rho^{\text{BB}}, n^{\text{BB}}, q^{\text{BB}}) = (3.0, 1, 0.5)$ and $(\rho^{\text{SMS}}, n^{\text{SMS}}, q^{\text{SMS}}) = (3.0, 5, 0.0)$. The other parameters are $\lambda_d = 0.1$ (for Panel a), $\lambda_u = 0.1$, $r = 0.1$, $m_d = 0.1$, $s = 0.6$ (for Panel b), $y_h = 10.0$, $y_d = 3.5$, and $y_l = 0.0$.

The SMS volume share is defined as the share of the total trading volume executed using SMS:

$$\frac{\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}} + \rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}}{\left(\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}} + \rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}\right) + \left(\rho^{\text{BB}} m_{lo}^{\text{BB}} v_{lo}^{\text{BB}} + \rho^{\text{BB}} m_{hn}^{\text{BB}} v_{hn}^{\text{BB}}\right)}. \tag{25}$$

Figure 5(a) and (b) below illustrate how this SMS volume share responds to shocks in s and λ_d, respectively. In Panel (a), the volume ratio is initially flat at 100% because both lo-sellers and hn-buyers always use SMS ($\theta_{hn} = \theta_{lo} = 1.0$). As the supply s rises higher (between $s_{lo,1}$ and $s_{lo,0}$), lo-sellers start to use less SMS, resulting in the decreasing segment. As s increases further, there are no more lo-sellers using SMS—all of them use BB, while all hn-buyers use SMS. That is, $m_{lo}^{\text{SMS}} = m_{hn}^{\text{BB}} = 0$. In this case, the SMS volume ratio above reduces to

$$\frac{\rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}}{\rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}} + \rho^{\text{BB}} m_{lo}^{\text{BB}} v_{lo}^{\text{BB}}} = \frac{t}{2t} = 50\%,$$

where the second equality follows because the trading volume in this case can be written as $t = \rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}} = \rho^{\text{BB}} m_{lo}^{\text{BB}} v_{lo}^{\text{BB}}$. Overall, the SMS volume ratio drops with the decline of the SMS usage θ_{lo}, as seen before in Figure 4. The same pattern is observed from Panel (b), where we increase the customers' negative preference shock intensity λ_d, effectively reducing the demand for the asset. The proposition below summarizes the results formally.

Proposition 6 (SMS usage under stress). The usage of SMS decreases with either the asset's excess supply or with its excess demand. That is, all else equal, for $s > s_{hn,1}$ ($< s_{lo,0}$), the ratio defined in (25) weakly decreases when s increases (decreases) or when λ_d increases (decreases).

The proposition also gives the mirroring result: SMS usage also drops when the asset's excess demand exacerbates ($s < \eta + \frac{m_d}{2}$).

The key intuition for the decrease of the SMS volume share can be understood from the worsening pricing for the *lo*-sellers. As the asset supply s increases after the fire sell, there are more and more *do*-dealers, as shown in the dot-dashed line in Figure 4. This is also evidenced empirically by **?**, who show that the majority of dealers enter a positive inventory cycle upon a corporate bond's downgrade (e.g., their Figure 2C). The remaining *dn*-dealers, facing less competition, therefore, will charge worse and worse prices to the *lo*-sellers in SMS. Expecting such worsening prices from SMS, the *lo*-sellers then avoid using SMS and switch to BB. In particular, our model yields an additional prediction regarding prices in SMS and in BB under a fire sell:

Corollary 2 (Prices in SMS versus in BB under fire sell). When there is excess supply, an *lo*-seller's expected trading price using SMS worsens relative to using BB.

Therefore, one way to empirically test our theory is to compare the trading prices in BB and in SMS when the asset is under fire sell and examine if the price in SMS is worse than that in BB.

To compare, Hendershott and Madhavan (2015) also shed light on customers' choice between "call" and "click." There the key disadvantage of SMS (click) is the leakage of one's private information to the multiple contacted dealers, as opposed to the only one in BB (call). Our mechanism complements theirs by explaining the shift of trading volume to BB after shocks not affecting information asymmetry, such as corporate bond downgrades.

4.3 Efficiency and welfare

This subsection studies whether the market's equilibrium technology choices are socially optimal: Given the technologies $\{n^k, \rho^k, q^k\}$, $k \in \{\text{BB}, \text{SMS}\}$, how will a social planner choose $\{\theta_{lo}, \theta_{hn}\}$ for the customers to maximize welfare? When, if at all, will the market's equilibrium choices $\{\theta_{lo}, \theta_{hn}\}$ coincide with the planner's $\{\theta^*_{lo}, \theta^*_{hn}\}$? What are the implications for policies and market design?

It turns out that the answers critically depend on the characteristics of the asset. Among others, how quickly customers can find dealers, i.e., $\{\rho^{\text{BB}}, \rho^{\text{SMS}}\}$, matters a lot. Recall from the technology assumption (19) that $\rho^{\text{BB}} \leq \rho^{\text{SMS}}$. Therefore, we discuss the high-ρ^{BB} and the low-ρ^{SMS} cases separately below.

4.3.1 The case of high search intensity

Proposition 7 (A social planner's technology choices). When the search intensity ρ^{BB} ($\leq \rho^{\text{SMS}}$) is sufficiently high, welfare w is monotone increasing in SMS usages by both types of customers and the social planner chooses $\theta^*_{lo} = \theta^*_{hn} = 1$.

The intuition largely follows Proposition 3. When the search intensity is high, Proposition 3 shows that welfare is monotone increasing in n. As such, by assigning both $\theta^*_{lo} = \theta^*_{hn} = 1$, the planner essentially chooses n^{SMS} over n^{BB} to maximize welfare.

However, the market's technology choices do not always coincide with the planner's. This is because a searching customer cares not only about the probability of finding a counterparty dealer but also about the endogenous split of the trading gain with the dealer. Figure 6 sketches such possible discrepancies. The solid line and the dashed line plot, respectively, the market's choices of θ_{hn} and θ_{lo} against the asset supply s. (Note that the patterns are qualitatively the same as in Figure 4.) The shaded areas indicate that there is inefficiency in the market's technology choices. For example, when the excess supply s is relatively extreme, $s > s_{lo,1}$, as in fire sell (Section 4.2), the dealer sector becomes overloaded (m_{do} too large), giving lo-sellers a hard time finding dn-dealers.

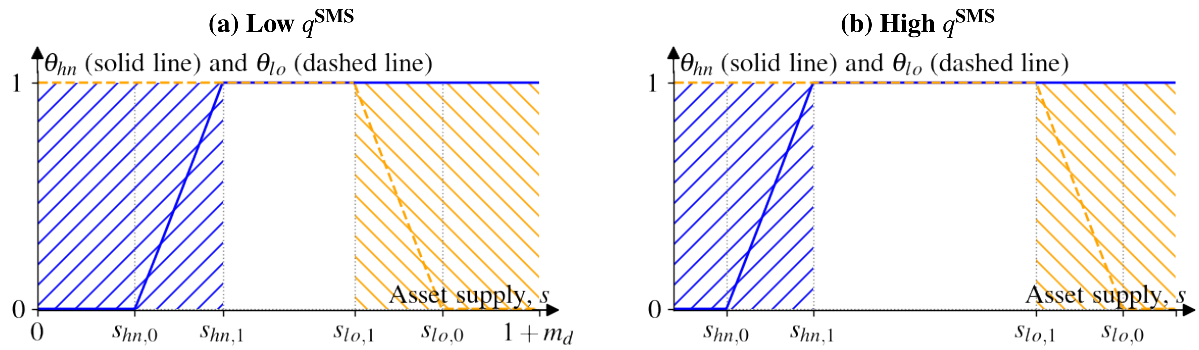

Figure 6: Market's technology choices versus a social planner's under high search intensity. This figure sketches the inefficiency due to the difference between the market's equilibrium technology choices and a social planner's when the search intensity $\rho := \min[\rho^{\text{BB}}, \rho^{\text{SMS}}]$ is high. The solid (blue) line and the dashed (orange) line are θ_{hn} and θ_{lo}, respectively, the hn-buyers' and the lo-sellers' equilibrium probabilities of using SMS. The "//"(blue) and "\\" (orange) shaded areas indicate, respectively, where θ_{hn} and θ_{lo} differ from the planner's corresponding choices θ^*_{hn} and θ^*_{lo}. Panel (a) shows the patterns for low q^{SMS}, while Panel (b) shows for a higher q^{SMS}.

Then lo-sellers become less willing to use SMS (θ_{lo} decreases with s) because in SMS their trading gains are too low. The same holds when $s < s_{hn,1}$ (extreme excess demand).

Since the planner wants to encourage SMS usage, a simple, welfare-improving market design mandate seems to readily follow: Let customers indicate their reservation values when searching dealers via SMS. In the model, such a design translates to an increase in q^{SMS}. By Lemma 4, when q^{SMS} is large enough, such that the inequality (21) flips, the customers endogenously choose SMS efficiently: $\theta^{\text{SMS}}_{hn} = 1 = \theta^*_{hn}$ and $\theta^{\text{SMS}}_{lo} = 1 = \theta^*_{lo}$. Indeed, this is what we find by contrasting Figure 6(a) and (b): The shaded area of the market's inefficient technology adoption is reduced from (a) of low q^{SMS} to (b) high q^{SMS}.

In practice, however, customers are almost always on the receiving end of TIOLIOs on electronic platforms; i.e., $q^{\text{SMS}} = 0$. We argue that one reason behind such an inefficient design is the dealers' incentive to participate. For example, when q^{SMS} becomes large, close to one, the dealers get a vanishing share of trading gains. Therefore, to the extent that the dealers have certain influence on the design of trading protocols on the electronic platforms, they would avoid a high q^{SMS}, or perhaps none at all, to let customers make TIOLIOs. Even if the dealers are independent of the trading protocol design, the platform operator will have to incentivize dealers' participation, without which the platform will not run, by, e.g., setting a low q^{SMS}.

4.3.2 The case of low search intensity

The case of low search intensity is more nuanced. The planner's choices in addition depend on the comparison between dealers' instantaneous utility y_d and an average customer's $\hat{y}$:

> **Proposition 8 (A social planner's technology choices).** When the search intensity ρ^{SMS} ($\geq \rho^{\text{BB}}$) is sufficiently low, the social planner chooses $\theta^*_{lo} = 1 - \theta^*_{hn} = \mathbb{1}_{\{y_d > \hat{y}\}}$ to maximize welfare.

To see why, recall from Equation (17): $w = \frac{\hat{y}}{r}(s - m_{do}) + \frac{y_d}{r} m_{do}$. Thus, the planner wants to maximize (minimize) m_{do}, i.e., to shift as much asset holding as possible to dealers (customers), if and only if $y_d > \hat{y}$ ($y_d < \hat{y}$). To do so, the planner will polarize $\{\theta^*_{lo}, \theta^*_{hn}\}$ because they affect m_{do} in opposite directions (Lemma D.1): If more lo-sellers use SMS, dn-dealers get to buy more often, increasing m_{do}; but if more hn-buyers use SMS, more do-dealers get to sell their assets, decreasing m_{do}. As a result, the planner sets $\theta^*_{lo} = \mathbb{1}_{\{y_d > \hat{y}\}}$ and $\theta^*_{hn} = 1 - \theta^*_{lo}$.

Figure 7(a) sketches the proposition for the case of $y_d < \hat{y}$, in which the planner wants to allocate the asset to the customers as much as possible, thus assigning $\theta^*_{hn} = 1$ and $\theta^*_{lo} = 0$. This is against the lo-sellers' wish, as they want to sell the asset to the dealers. As a result, the market's technology choices are efficient (coinciding with the planner's) only when the asset is in extreme supply, i.e., when $s > s_{lo,0}$. Panel (b), flipping Panel (a), sketches the case of $y_d > \hat{y}$.

The patterns shown in Figure 7 caveat that the intuition regarding welfare and market design obtained from the high-ρ case does not carry through when the matching of the asset is intrinsically slow. For example, compared to corporate bonds, whose matching on MarketAxess take only a few minutes (Hendershott and Madhavan, 2015), collateralized loan obligations (CLOs) trade much more slowly, taking days as the B/OWIC run through emails require considerably longer time to

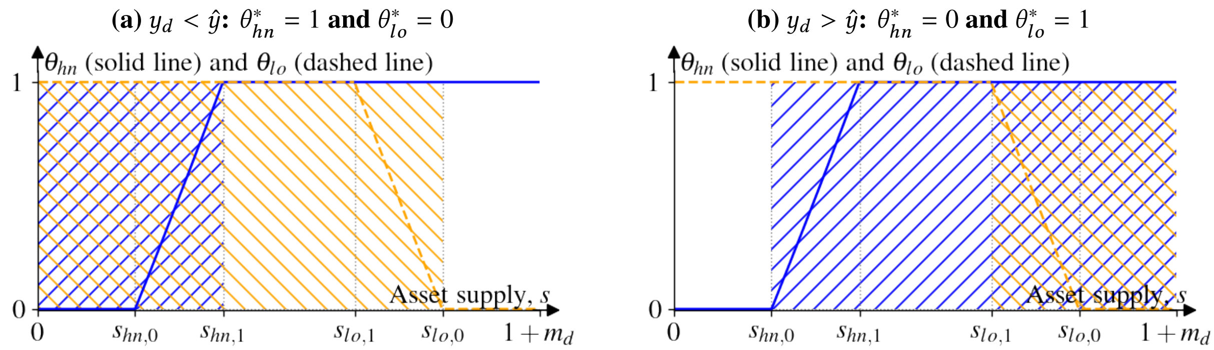

Figure 7: Market's technology choices versus a social planner's under low search intensity. This figure sketches the inefficiency due to the difference between the market's equilibrium technology choices and a social planner's when the search intensity $\rho := \min[\rho^{BB}, \rho^{SMS}]$ is low. The solid (blue) line and the dashed (orange) line are θ_{hn} and θ_{lo}, respectively, hn-buyers' and lo-sellers' equilibrium probabilities of using SMS. The "//"(blue) and "\\" (orange) shaded areas indicate, respectively, where θ_{hn} and θ_{lo} deviate from the planner's corresponding choices θ^*_{hn} and θ^*_{lo}. Panel (a) shows the pattern for the case of $y_d < \hat{y}$, in which case $\theta^*_{hn} = 1$ and $\theta^*_{lo} = 0$, and Panel (b) the opposite, in which case $\theta^*_{hn} = 0$ and $\theta^*_{lo} = 1$.

organize (Hendershott et al., 2020). For such slow-moving assets, the planner always wants some customers to use BB to prevent the asset from being held inefficiently in the wrong hands.

5 Conclusion

This paper studies "simultaneous multilateral searching" (SMS), which has been popularized in practice recently through trading protocols like "Request-for-Quote" (RFQ) in OTC markets. The idea is that a searching customer can reach out to multiple dealers simultaneously, solicit quotes from them, and then trade with the one offering the best quote. This search mechanism differs from the conventional "bilateral bargaining" (BB), in which a searching customer meets a single dealer and negotiates the terms of trade.

A steady state equilibrium is characterized in an extension of the standard search framework. The key insight revealed is that the split of the trading gain between a searching and a quoting investor is an endogenous equilibrium outcome, as opposed to the exogenous split (à la Nash) in the literature assuming BB. In addition, two search parameters, the intensity and the capacity, are analyzed in terms of their contrasting welfare implications. A novel bottleneck effect, arising from (and only from) the search capacity, is shown to hinder the efficient asset allocation and might possibly hurt welfare. Such a bottleneck might arise also from transparency of dealer inventories.

Allowing customers to endogenously choose between SMS and BB, the model finds an intrinsic hindrance in the adoption of SMS and further suggests potential inefficiency in terms of asset allocation. The model underscores channels through which both regulation and market design can affect customers' search preferences and, ultimately, the asset allocation efficiency.

Appendix

A Comparison with directed search

This appendix analyzes a model of OTC trading with "directed search" (DS), where dealers continuously post quotes and, after observing them, each customer directs her search to one chosen dealer. This is a realistic feature of corporate bonds trading, as dealers sometimes broadcast their indicative bids and asks to customers on electronic platforms (Section III.B of Bessembinder, Spatt, and Venkataraman, 2020). The similarity of this paper and the DS literature is that both allow endogenous trading gain shares accruing to customers and dealers. See, e.g., Guerrieri, Shimer, and Wright (2010), Lester, Rocheteau, and Weill (2015), Chang (2018) and, for a review, Wright et al. (2020). The purpose is to compare the findings of DS model in this appendix with those of SMS from Section 3. The key result is that the steady state equilibrium under DS can be proxied by the SMS equilibrium either (i) when the dealers inventory transparency ψ is sufficiently higher; or (ii) when the search capacity n is sufficiently large.

A.1 A model of directed search

The model setup follows Section 2, except for the parts of "search" and "price determination," which are modified as follows: All dealer owners (type do) constantly post ask quotes, while all dealer non-owners (type dn) post bid quotes. Customers observe these quotes and direct their searches to chosen dealers at independent Poisson processes with the same intensity ρ. Customers and their chosen dealers meet pairwise and once met, the pair exchanges the asset at the dealer quoted price.

Denote by v_{hn} the probability for a *hn*-*do* match and by v_{lo} for a *lo*-*dn* match. Assume that

$$v_{hn} = 1 \text{ if } m_{do} > 0 \text{ and } v_{lo} = 1 \text{ if } m_{dn} > 0. \tag{26}$$

This is because at any instant, there is only an infinitesimal measure of customers searching ($\rho m_{hn}\mathrm{d}t$ buyers and $\rho m_{lo}\mathrm{d}t$ sellers), which is negligible compared to the vast mass of quoting dealers (m_{do} and m_{dn}, if positive). Effectively, the assumption (26) lets the $\mathrm{d}t$-measure customers find dealers with certainty as long as the measures of counterparties are strictly positive. In the case when some dealer masses become zero, i.e., $m_{do} \to 0$ or $m_{dn} \to 0$, the matching probabilities will be solved endogenously.

We shall look for a steady-state, symmetric dealer pricing equilibrium, characterized by the following time-invariant variables: (i) the demographics $\{m_{ho}, m_{hn}, m_{lo}, m_{ln}, m_{do}, m_{dn}\}$; (ii) the matching probabilities $\{v_{hn}, v_{lo}\}$ in case the corresponding dealer mass is zero; and (iii) the symmetric ask and bid, p_a and p_b, by the *do*- and *dn*-dealers, respectively. Note that since the quotes are symmetric, the customers randomly direct their searches to all counterparty dealers.

A.1.1 Demographics and matching probabilities

The six demographic variables $\{m_{ho}, m_{hn}, m_{lo}, m_{ln}, m_{do}, m_{dn}\}$ and the two matching probabilities $\{v_{hn}, v_{lo}\}$ can be pinned down by the six equations (2)-(7), which hold as before, plus the two conditions given in (26):

Proposition 9 (Demographics under DS). In a steady-state, the matching probabilities satisfy

$$v_{hn} = \min\left[1, \frac{m_{lo}}{m_{hn}}\right] \text{ and } v_{lo} = \min\left[1, \frac{m_{hn}}{m_{lo}}\right], \tag{27}$$

where the ratios $\frac{m_{lo}}{m_{hn}}$ and $\frac{m_{hn}}{m_{lo}}$ depend on the asset supply s:

$$\operatorname{sign}\left[\frac{m_{lo}}{m_{hn}} - 1\right] = -\mathbb{1}_{\{0<s<\eta\}} + \mathbb{1}_{\{\eta+m_d<s<1+m_d\}}. \tag{28}$$

Given the $\{v_{hn}, v_{lo}\}$ above, the steady-state demographics exist and are the unique solution to the linear equation system (2)-(7).

The key insight from (27) is that the matching probabilities $\{v_{hn}, v_{lo}\}$ only depend on the relative sizes of *hn*-buyers and *lo*-sellers in the market. Perhaps surprisingly, the dealer sizes m_{do} and m_{dn} do not show up, as if the customers are directly searching for counterparties, skipping the dealers. To see why, recall that (2)-(7) also imply the dealer stationarity condition (8): $\rho m_{hn} v_{hn} = \rho m_{lo} v_{lo}$, which balances the asset inflow to and the outflow from the dealers (otherwise the dealer sizes will change overtime). Note that it can be equivalently interpreted as the *hn*-buyers are directly matched with the lo-sellers. That is, the dealers are merely passing the asset from the sellers to the buyers, not affecting the matching probabilities $\{\nu_{hn}, \nu_{lo}\}$.

One might wonder why ν_{hn} and ν_{lo} can be less than one: Since there is always only a $\mathrm{d}t$ amount—zero measure—of customers searching for dealers, by assumption (26), should not the matching probabilities always be one? It turns out that $\nu_{hn} < 1$ and $\nu_{lo} < 1$ precisely when the measure of the corresponding dealers is also zero. Consider, for example, the case of $m_{do} = 0$, which implies that $m_{dn} = m_d - m_{do} > 0$. Then, by assumption (26), $\nu_{lo} = 1$ and there is always an amount of $\rho m_{lo}\nu_{lo}\mathrm{d}t$ trades occurring between dn-buyers and lo-sellers. These trades then create a "transient" fringe of do-sellers (of the same magnitude of $\mathrm{d}t$) who then quote asks to the searching hn-buyers. The dealer stationarity condition (8) above then requires:

$$\rho m_{hn}\nu_{hn} = \rho m_{lo}\nu_{lo} \implies \nu_{hn} = \frac{m_{lo}\nu_{lo}}{m_{hn}} = \frac{m_{lo}}{m_{hn}}.$$

Note that it must be $\nu_{hn} < 1$ in this case (as verified in the proof of Proposition 9); that is, *all* such transient do-sellers are sought after by hn-buyers. Otherwise, some of the do-sellers will accumulate overtime, making the dealer masses nonstationary.

A.1.2 Value functions and prices

The value functions can be found using the same set of HJB equations (10)-(15) as in Section 3.2. Hence, Proposition 1 continues to hold for directed search. In particular, the condition for positive trading gains remains the same. The only differences are in the trading gain intensities, $\{\zeta_\sigma\}$.

We first consider the interior equilibrium where $m_{do} \in (0, m_d)$. In this case, there is perfect competition among dealers and so $\zeta_{dn} = \zeta_{do} = 0$. To see why, consider, for example, the do-sellers' quote p_a to the hn-buyers. Since the searching hn-buyers' measure is vanishingly small ($\rho m_{hn}\mathrm{d}t$), the standard Bertrand price competition applies to the do-sellers, giving $p_a = R_d$ as the only symmetric price in equilibrium. Similar argument implies that in this case the dn-buyers quote $p_b = R_d$. Summing up, the dealers quotes are given by $p_a = p_b = R_d$, and the trading gain intensities are given by $\zeta_{lo} = \rho\nu_{lo}$, $\zeta_{hn} = \rho\nu_{hn}$, and $\zeta_{do} = \zeta_{dn} = 0$.

Next, in the corner equilibrium of $m_{do} = 0$, any do-dealer exists only transiently: Whenever a dn-dealer has bought the asset (from an lo-seller), he immediately trades again with an hn-buyer, becoming dn-dealer again. Therefore, $\Delta_{hd} = 0$ in this case. This follows from Equation (14) if one divides it by ζ_{do} and sets $\zeta_{do} = \infty$ (because a do-dealer trades immediately without waiting). Then $p_a = R_h = R_d$. The Bertrand competition argument for $m_d > 0$ measure of dn-buyers still implies that $p_b = R_d$. Likewise, in the corner equilibrium of $m_{dn} = 0$, we have $\Delta_{dl} = 0$ and $p_a = p_b = R_d$.

A.2 Comparing DS with SMS

In DS, a customer can first observe all dealers' quotes and then direct her search to a chosen dealer. In SMS, a customer can reach only n dealers (where n is set by the platform like RFQ), not knowing the types of the dealers, only with noisy signals of quality ψ. This appendix concludes with the following convergence result:

> **Proposition 10 (Convergence of SMS to DS).** The equilibrium demographics, value functions, and prices in the SMS model converge to those in DS either (i) when $\psi \to 1$ under the "random matching with signal" specification with $n \geq 2$; or (ii) when $n \to \infty$ under the general specification.

Intuitively, one should expect as the search capacity $n \to \infty$, the customer can almost surely find at least one counterparty dealer. Likewise, if the signal quality (dealer inventory transparency) $\psi \to 1$, the customer can always direct her quote to the correct dealers. Therefore, in both limits of SMS, the customers' searches converge to those in DS.

B Price dispersion with homogenous dealers

This section studies the dealers pricing strategies under SMS. Despite the homogeneity of dealers, in equilibrium, there still is price dispersion in their quotes. Consider a customer just contacted n dealers via SMS. First, there is probability q that she can make a TIOLIO to the dealers. In this case, it is optimal for her to set the price of the TIOLIO at the dealers' reservation value, i.e., $p = R_d$.

Second, there is probability $1 - q$ that the n dealers independently quote to the customer. For concreteness, suppose the customer is an *hn*-buyer. In this case, a quoting dealer must be a *do*-seller and he would like to capture the full surplus by setting $p \uparrow R_h$. However, he faces potential competition from the other $(n - 1)$ dealers, as their asking quotes might be lower than his. Yet not all of the other $(n - 1)$ dealers are necessarily also *do*-sellers. The quoting *do*-seller therefore engages in a price competition with *unknown number of competitors*.

Such price competition differs from the standard Bertrand price competition, in which every *do*-seller quotes his reservation price of R_d and the *hn*-buyer gets the full surplus Δ_{hd}. Instead, every *do*-seller has an incentive to charge a higher price, $R_d + \alpha\Delta_{hd}$ for some $\alpha \in [0, 1]$. (When $\alpha = 1$, $R_d + \alpha\Delta_{hd} = R_h$, which is the *hn*-buyer's reservation value.) This is because he might actually be the only *do*-seller among the n contacted dealers, in which case his quote is the only price available to the contacting *hn*-buyer. As long as $\alpha \leq 1$, the buyer will accept it[12] and the

[12] To see this, note that by accepting an offer $p = R_d + \alpha\Delta_{hd}$, the customer-buyer becomes *ho*-bystander and gets a

dealer can pocket the difference $\alpha\Delta_{hd}$ as his profit. In a Nash equilibrium, however, the fraction α cannot be deterministic, as the undercutting argument of Bertrand competition will lead to $\alpha \downarrow 0$. Yet, it would be strictly better off to quote some $\alpha > 0$ if all the potential competitors were to quote $\alpha \downarrow 0$. The heuristic discussion above is formalized in the proof of the following proposition.

Proposition 11 (Dealers' equilibrium quoting). Suppose a customer contacts n (≥ 1) dealer(s). With probability $1-q$, each dealer independently makes a TIOLIO to the customer. Within symmetric strategies, there is a unique mixed-strategy equilibrium for the dealers. Define $F(x;\pi,n) := \frac{1}{\pi} - \left(\frac{1}{\pi} - 1\right)x^{-\frac{1}{n-1}}$ for $(1-\pi)^{n-1} \leq x \leq 1$, $\pi \in (0,1)$, and $n \in \mathbb{N}$. Then,

- a do-seller asks $R_l + \alpha\Delta_{hd}$, where α is random with c.d.f. $F(\alpha;\pi_{do},n)$; and
- a dn-buyer bids $R_h - \beta\Delta_{dl}$, where β is random with c.d.f. $F(\beta;\pi_{dn},n)$.

Note that when $n=1$, the c.d.f. $F(\cdot)$ becomes degenerate with $F(x) = \mathbb{1}_{\{x\geq 1\}}$.

The proposition above implies that a quoting do-seller expects a trading price of $R_l + \bar{\alpha}\Delta_{hd}$ and a quoting dn-buyer expects $R_h - \bar{\beta}\Delta_{dl}$, where

$$\bar{\alpha} := \mathbb{E}[\alpha] = (1-\pi_{do})^{n-1} \quad \text{and} \quad \bar{\beta} := \mathbb{E}[\beta] = (1-\pi_{dn})^{n-1}. \tag{29}$$

To see this, consider a quoting do-seller and note that under the mixed-strategy equilibrium, he must be indifferent across all possible $\alpha \in [0,1]$. In particular, the only situation for quoting $\alpha = 1$ to "win" is that there are no other competing do-sellers; that is, with probability $(1-\pi_{do})^{n-1}$. Therefore, when contacted, a quoting do-seller expects a profit of $\bar{\alpha}\Delta$, where $\bar{\alpha}$ can be interpreted as his expected trading gain share. Likewise, a quoting dn-buyer expects $\bar{\beta}\Delta$.

Proposition 11 characterizes a contacted dealer's quoting strategy. From a contacting customer's perspective, however, the expected trading price has a different distribution, because she can pick the best quote and because there might be no quote if none of the contacted dealers is of the matching type. Consider a contacting hn-buyer for example. He contacts n dealers knowing that the number of counterparties he will actually find, N_{do}, is random and follows a binomial distribution with n draws and success rate π_{do}. Each of these N_{do} dealers then quotes a random price according to $F(\alpha;\pi_{do},n)$, following Proposition 11. (The hn-buyer can safely ignore the other $n - N_{do}$ dealers' quotes, as they also want to buy.) The contacting hn-buyer then picks the lowest ask among the N_{do} available quotes. Conditional on the realization $N_{do} \geq 1$, the c.d.f. of this minimum ask is $1-(1-F(\alpha;\cdot))^{N_{do}-1}$. (When $N_{do} = 0$, the hn-buyer finds no ask quote and there is no trade.) Averaging across all possible $N_{do} \in \{1,\dots,n\}$, the corollary below gives the expectation of this minimum ask quote.

continuation value of $V_{ho} - p$. If instead he rejects the offer, his value remains as V_{hn}. This customer-buyer will accept the offer as long as $V_{ho} - p \geq V_{hn}$, a condition equivalent to $\alpha \leq 1$.

Proposition 12 (Trading prices). Define $G(\pi, n) := \frac{n\pi\cdot(1-\pi)^{n-1}}{1-(1-\pi)^n}$ for some $\pi \in (0,1)$ and $n \in \mathbb{N}$. Then, with probability q, a searching customer sets the price equal to the dealers' reservation value R_d; and with probability $1-q$,

- an hn-buyer expects ask quotes from do-dealer(s) with probability $(1-(1-\pi_{do})^n)$ and the expected best ask, is $R_d + \bar{A}\Delta_{hd}$, where $\bar{A} = G(\pi_{do}, n)$; and
- an lo-seller expects bid quotes from dn-dealer(s) with probability $(1-(1-\pi_{dn})^n)$ and the expected best bid, is $R_d - \bar{B}\Delta_{dl}$, where $\bar{B} = G(\pi_{dn}, n)$.

The above expected quotes, $R_d + \bar{A}\Delta_{hd}$ and $R_d - \bar{B}\Delta_{dl}$, are also the average trading prices in buyer- and seller-initiated trades, respectively. Note that when $n = 1$, $\bar{A} = \bar{B} = 1$ for all $\pi \in (0,1)$.

Several features of the equilibrium pricing above are worth highlighting. Proposition 11 and 12 imply that there is price dispersion in equilibrium. Such dispersion is due to the unknown number of competitors, an intrinsic feature of SMS: The contacted dealers' types are unknown to each other. In the current stylized model, such types boil down to the dealers' inventory holdings (do vs. dn). In real-world trading, agents' other characteristics (like risk-aversion, patience, wealth, and relationship with customers, etc.) can enrich their possible types. As long as such a friction remains, price dispersion will be a robust feature in equilibrium. The models by Duffie, Dworczak, and Zhu (2017) and Lester et al. (2018) also feature similar price setting mechanisms. The key novelty here is that such price dispersion is endogenously parametrized by the equilibrium demographics of counterparties, through $\pi_{do} = \pi\left(\frac{m_{do}}{m_d}\right)$ and $\pi_{dn} = \pi\left(\frac{m_{do}}{m_d}\right)$.

C Additional useful results

C.1 Demographics with one search technology

While the system (2)-(7) has only two zero-flow conditions (Equations 6 and 7), the stationarity of all other types of agents is also implied. Apart from the dealer stationarity (8), $-(5) - (6)$ gives $v_{lo}m_{lo}\rho - m_{ln}\lambda_u + m_{hn}\lambda_d = 0$, ensuring that the net flow in and out of ln-bystanders is zero. Likewise, $(5) - (7)$ gives $v_{hn}m_{hn}\rho - m_{ho}\lambda_d + m_{lo}\lambda_u = 0$, ensuring that the net flow of ho-bystanders is zero.

We also derive some useful expressions for the customer masses. Equations (2), (3), and (5) together imply the stable fractions of the high-type and the low-type customers:

$$m_{ho} + m_{hn} = \frac{\lambda_u}{\lambda_d + \lambda_u} =: \eta \text{ and } m_{lo} + m_{ln} = \frac{\lambda_u}{\lambda_d + \lambda_u} = 1 - \eta. \tag{30}$$

Then combining the market clearing condition (2) and the lo-seller net flow (6), we obtain

$$m_{lo} = (1-\eta)(s - m_{do}) - \frac{t}{\lambda_u + \lambda_d}, \tag{31}$$

which intuitively says that the stationary mass of lo-sellers is a fraction $(1-\eta)$ of the residual asset supply $(s - m_{do})$ available to customers, less a term $t/(\lambda_u + \lambda_d)$ due to their active trading. Combining (2) and (6) gives

$$m_{hn} = \eta \cdot (1 + m_{do} - s) - \frac{t}{\lambda_u + \lambda_d}. \tag{32}$$

Note that $1 + m_{do} - s$, which is the total mass of non-owner customers in this economy. That is, the stationary mass of hn-buyers is the high-type fraction η of all non-owner customers, less the same term due to trading. The above expressions are in fact generic in the search literature. For example, if, as in DGP, customers find each other at intensity ρ without dealers, then Equations (31) and (32) still hold with $m_{do} = 0$ and $t = 2\rho m_{hn} m_{lo}$.

C.2 Demographics with two search technologies

There are six customer population sizes: $\{m_{ho}, m_{ln}, m_{hn}^{\text{SMS}}, m_{hn}^{\text{BB}}, m_{lo}^{\text{SMS}}, m_{lo}^{\text{BB}}\}$; in addition, there are two types of dealers, $\{m_{do}, m_{dn}\}$. For notation simplicity, write

$$m_{hn} = m_{hn}^{\text{SMS}} + m_{hn}^{\text{BB}}; \quad \text{and} \quad m_{lo} = m_{lo}^{\text{SMS}} + m_{lo}^{\text{BB}}.$$

Then the four (aggregate) customer masses, $\{m_{ho}, m_{ln}, m_{hn}, m_{lo}\}$, must satisfy the conditions (2)-(5) in Section 3.1. The other four conditions are analogous to the stationarity conditions (6) and (7):

$$\text{net flow of } lo\text{-sellers using SMS:} \quad -v_{lo}^{\text{SMS}} m_{lo}^{\text{SMS}} \rho^{\text{SMS}} - \lambda_u m_{lo}^{\text{SMS}} + \theta_{lo} \lambda_d m_{ho} = 0 \tag{33}$$

$$\text{net flow of } lo\text{-sellers using BB:} \quad -v_{lo}^{\text{BB}} m_{lo}^{\text{BB}} \rho^{\text{BB}} - \lambda_u m_{lo}^{\text{BB}} + (1-\theta_{lo}) \lambda_d m_{ho} = 0 \tag{34}$$

$$\text{net flow of } hn\text{-buyers using SMS:} \quad -v_{hn}^{\text{SMS}} m_{hn}^{\text{SMS}} \rho^{\text{SMS}} - \lambda_d m_{hn}^{\text{SMS}} + \theta_{hn} \lambda_u m_{ln} = 0 \tag{35}$$

$$\text{net flow of } hn\text{-buyers using BB:} \quad -v_{hn}^{\text{BB}} m_{hn}^{\text{BB}} \rho^{\text{BB}} - \lambda_d m_{hn}^{\text{BB}} + (1-\theta_{hn}) \lambda_u m_{ln} = 0 \tag{36}$$

where $v_{lo}^k = 1 - \left(1 - \frac{m_{dn}}{m_d}\right)^{n^k}$ and $v_{hn}^k = 1 - \left(1 - \frac{m_{do}}{m_d}\right)^{n^k}$ are the probabilities for a customer to find at least one counterparty dealer using technology $k \in \{\text{BB}, \text{SMS}\}$. Compared to Equations (6) and (7) in Section 3.1, the key differences are (i) that every variable here is technology-dependent and superscripted with $k \in \{\text{BB}, \text{SMS}\}$; and (ii) that only a fraction of θ_σ of the newborn σ-customer use SMS, while the rest $(1-\theta_\sigma)$ use BB, where $\sigma \in \{hn, lo\}$.

The conditions (2)-(5) and (33)-(36) exactly pin down the eight demographic variables:

Lemma 5 (Stationary demographics with technology choice). Given the customers' technology choices $\{\theta_{lo}, \theta_{hn}\} \in [0,1]^2$, Equations (2)-(5) and (33)-(36) uniquely pin down the demographics $\{m_{ho}, m_{ln}, m_{hn}^{\text{SMS}}, m_{hn}^{\text{BB}}, m_{lo}^{\text{SMS}}, m_{lo}^{\text{BB}}\} \in [0,1]^6$ and $\{m_{do}, m_{dn}\} \in (0, m_d)^2$.

The resulting expressions are similar to those implied by Lemma 1. In particular, (33) + (34) − (35) − (36) + (5) gives the trading volume expression

$$t := \sum_k \nu_{lo}^k m_{lo}^k \rho^k = \sum_k \nu_{hn}^k m_{hn}^k \rho^k, \tag{37}$$

ensuring the stationarity of both dealer types. The h- and l-type customer stationarity (30) also holds the same, and so do the expressions for the total size of trading customers $m_{lo} = \sum_k m_{lo}^k$ and $m_{hn} = \sum_k m_{hn}^k$. The stationarity of all other types of agents are also ensured: For example, −(5) − (33) − (34) gives $-\lambda_u m_{ln} + \sum_k \left(\nu_{lo}^k m_{lo}^k \rho^k + \lambda_d m_{hn}^k\right) = 0$, which ensures the stationarity of ln-bystanders. Likewise, (5) − (35) − (36) gives $-\lambda_d m_{ho} + \sum_k \left(\nu_{hn}^k m_{hn}^k \rho^k + \lambda_u m_{ln}^k\right) = 0$, which ensures the stationarity of ho-bystanders.

C.3 Value functions with two search technologies

Given the technology choices $\{\theta_\sigma\}$, hence also the demographics (Lemma 5), the value functions for all six agent types can be derived analogously to those in Equations (10)-(15). For example, the value functions of an ho-bystander and an ln-bystander must satisfy the HJB equations

$$y_h + \lambda_d \cdot \left(\max\left[V_{lo}^{\text{SMS}}, V_{lo}^{\text{BB}}\right] - V_{ho}\right) - rV_{ho} = 0; \tag{38}$$

$$\lambda_u \cdot \left(\max\left[V_{ho}^{\text{SMS}}, V_{ho}^{\text{BB}}\right] - V_{ln}\right) - rV_{ln} = 0. \tag{39}$$

Compared with Equations (10) and (11), the only difference is that upon a preference shock, a newborn trading customer can choose which technology to use, hence the term of $\max\left[V_\sigma^{\text{SMS}}, V_\sigma^{\text{BB}}\right]$ in the above HJBs ($\sigma \in \{lo, hn\}$).

The HJB equations for the trading agents are also similar to before:

(40) HJB of lo-sellers using technology k: $$y_l + \lambda_u \cdot (V_{ho} - V_{lo}^k) - rV_{lo}^k + \zeta_{lo}^k \Delta_{dl}^k = 0;$$

(41) HJB of hn-buyers using technology k: $$\lambda_d \cdot (V_{ln} - V_{hn}^k) - rV_{hn}^k + \zeta_{hn}^k \Delta_{hd}^k = 0;$$

(42) HJB of do-dealers: $$y_d - rV_{do} + \sum_k \zeta_{do}^k \Delta_{hd}^k = 0;$$

(43) HJB of dn-dealers: $$-rV_{dn} + \sum_k \zeta_{dn}^k \Delta_{dl}^k = 0.$$

Compared to Equations (12)-(15), the only difference is that the trading gains $\{\Delta_{hd}, \Delta_{dl}\}$ and the trading gain intensities $\{\zeta_{lo}, \zeta_{hn}, \zeta_{do}, \zeta_{dn}\}$ are technology specific, superscripted with $k \in \{\text{BB}, \text{SMS}\}$. For completeness, we derive these expressions below.

Using technology k, an lo-seller's reservation value is $R_l^k := V_{lo}^k - V_{ln}$, and that for an hn-buyer is $R_h^k := V_{ho} - V_{hn}^k$. A dealer's reservation value is the same $R_d := V_{do} - V_{dn}$ as before. Then, depending the customer's technology k, the trading gain between an hn-buyer and a do-seller is $\Delta_{hd}^k := R_h^k - R_d$ and that between a dn-buyer and an lo-seller is $\Delta_{dl}^k := R_d - R_l^k$. By Proposition 12, the dealers' respective average ask and bid are:

$$\bar{A}^k = \frac{n^k \frac{m_{do}}{m_d}\left(1 - \frac{m_{do}}{m_d}\right)^{n^k - 1}}{1 - \left(1 - \frac{m_{do}}{m_d}\right)^{n^k}} \text{ and } \bar{B}^k = \frac{n^k \frac{m_{dn}}{m_d}\left(1 - \frac{m_{dn}}{m_d}\right)^{n^k - 1}}{1 - \left(1 - \frac{m_{dn}}{m_d}\right)^{n^k}}.$$

Thus, an hn-buyer expects $\zeta_{hn}^k \Delta_{hd}^k$, while a do-dealer expects $\zeta_{do}^k \Delta_{hd}^k$, where the respective trading gain intensities are

$$\zeta_{hn}^k = \rho^k v_{hn}^k \cdot \left(q^k + (1 - q^k)(1 - \bar{A}^k)\right) \text{ and } \zeta_{do}^k = \frac{m_{hn}^k \rho^k v_{hn}^k}{m_{do}}(1 - q^k)\bar{A}^k.$$

Analogously, an lo-seller expects $\zeta_{lo}^k \Delta_{dl}^k$, while a dn-dealer expects $\zeta_{dn} \Delta_{dl}^k$, with intensities

$$\zeta_{lo}^k = \rho^k v_{lo}^k \cdot \left(q^k + (1 - q^k)(1 - \bar{B}^k)\right) \text{ and } \zeta_{dn}^k = \frac{m_{hn}^k \rho^k v_{lo}^k}{m_{dn}}(1 - q^k)\bar{B}^k.$$

Corollary 3 (Positive trading gains). When $\overline{y}_d \geq y_d \geq \underline{y}_d$ as defined in Proposition (1), there is strictly positive gains from trade, i.e., $R_d \in (R_l^k, R_h^k)$, for both $k \in \{\text{BB}, \text{SMS}\}$.

The corollary ensures that the same condition for y_d as before is sufficient to guarantee positive trading gains regardless of the equilibrium technology choices $\{\theta_\sigma\}$. Note that the positive trading gains also ensures that both the $ho-$ and ln-customers do stay out of trading, following Lemma 2.

D Collection of proofs

Lemmas 1 and 5

Proof. The proof considers the general case of Lemma 5 with arbitrary θ_{hn} and θ_{lo}. Lemma 1 is then just a special case of $\theta_{hn} = \theta_{lo} = 1$. Where convenient, we will occasionally write $\theta_\sigma^{\text{SMS}} = 1 - \theta_\sigma^{\text{BB}} = \theta_\sigma$. The idea is to first express all other unknowns as monotone functions of m_{do}. The existence and the uniqueness then follow as long as the solution to m_{do} exists and is unique.

To begin with, add (33) and (34) to get

$$\text{(44)} \qquad \text{all } lo\text{-seller stationarity:} \qquad -\lambda_u m_{lo} + \lambda_d m_{ho} - \rho m_{lo} \nu_{lo} = 0,$$

where $m_{lo} := \sum_k m_{lo}^k$ is the total lo-seller mass, $\rho := \max\left[\rho^{\text{SMS}}, \rho^{\text{BB}}\right]$, $\nu_{lo} := \frac{1}{\rho m_{lo}} \sum_k \rho^k m_{lo}^k \nu_{lo}^k$ is the (weighted) average matching rate for an lo-seller, and $m_{ho} := \sum_k m_{ho}^k$ is the total ho-bystander mass. Similarly, adding (35) and (36) yields

$$\text{(45)} \qquad \text{all } hn\text{-buyer stationarity:} \qquad -\lambda_d m_{hn} + \lambda_u m_{ln} - \rho m_{hn} \nu_{hn} = 0,$$

where $m_{hn} := \sum_k m_{hn}^k$, $\nu_{hn} := \frac{1}{\rho m_{hn}} \sum_k \rho^k m_{hn}^k \nu_{hn}^k$, and $m_{ln} := \sum_k m_{ln}^k$. Taking $\{\nu_{lo}, \nu_{hn}\}$ as given, Equations (30), (44), and (45) form a linear system of the four masses $\{m_{ho}, m_{ln}, m_{hn}, m_{lo}\}$, which have the unique solution of

$$\text{(46)} \qquad \begin{aligned} m_{ho} &= \eta \frac{\lambda_u \nu_{hn} + \rho \nu_{lo} \nu_{hn}}{\lambda_u \nu_{hn} + \lambda_d \nu_{lo} + \rho \nu_{hn} \nu_{lo}}; \quad m_{ln} = (1-\eta) \frac{\lambda_d \nu_{lo} + \rho \nu_{lo} \nu_{hn}}{\lambda_u \nu_{hn} + \lambda_d \nu_{lo} + \rho \nu_{hn} \nu_{lo}}; \\ m_{hn} &= (1-\eta) \frac{\lambda_u \nu_{lo}}{\lambda_u \nu_{hn} + \lambda_d \nu_{lo} + \rho \nu_{hn} \nu_{lo}}; \quad m_{lo} = \eta \frac{\lambda_d \nu_{hn}}{\lambda_u \nu_{hn} + \lambda_d \nu_{lo} + \rho \nu_{hn} \nu_{lo}}. \end{aligned}$$

Plug in the expressions of m_{ho} and $m_{lo} = \sum_k m_{lo}^k$ into the market clearing condition (2) to get

$$\text{(47)} \qquad \eta \frac{(\lambda_u + \lambda_d) \nu_{hn} + \rho \nu_{lo} \nu_{hn}}{\lambda_u \nu_{hn} + \lambda_d \nu_{lo} + \rho \nu_{lo} \nu_{hn}} + m_{do} - s = 0.$$

This is an equation with unknowns $\{m_{do}, \nu_{hn}, \nu_{lo}\}$. It remains to express ν_{hn} and ν_{lo} as (monotone) functions of m_{do}.

Consider ν_{lo} for example. Note that (33) and (34) imply that

$$\text{(48)} \qquad m_{lo}^k = \frac{\lambda_d m_{ho} \theta_{lo}^k}{\lambda_u + \rho^k \nu_{lo}^k}$$

where $\theta_{lo}^{\text{BB}} := 1 - \theta_{lo}$ and $\theta_{lo}^{\text{SMS}} := \theta_{lo}$. Hence, from the earlier definition,

$$\text{(49)} \qquad \nu_{lo} = \frac{\sum_k \rho^k m_{lo}^k \nu_{lo}^k}{\rho m_{lo}} = \frac{\sum_k \rho^k m_{lo}^k \nu_{lo}^k}{\rho \sum_k m_{lo}^k} = \frac{\sum_k \frac{\rho^k \theta_{lo}^k \nu_{lo}^k}{\lambda_u + \rho^k \nu_{lo}^k}}{\rho \sum_k \frac{\theta_{lo}^k}{\lambda_u + \rho^k \nu_{lo}^k}},$$

which is monotone increasing in both ν_{lo}^k for $k \in \{BB, SMS\}$. Recall from the definition $\nu_{lo}^k := 1 - (1 - \pi_{dn})^{n^k}$ that both ν_{lo}^k are monotone decreasing in m_{do}. Therefore, so is ν_{lo}. In the same way, both ν_{hn}^k are monotone increasing in m_{do} and so is ν_{hn}.

Now return to Equation (47). Since both ν_{lo} and ν_{hn} can be expressed as a unique function in m_{do}, (47) is an equation of a single unknown m_{do}. To prove the existence of the solution, consider the limits of the support of $m_{do} \in [0, m_d]$. As $m_{do} \downarrow 0$, both $v^k_{lo} \uparrow 1$ while both $v^k_{hn} \downarrow 0$, and as a result, $v_{lo} \uparrow 1$ and $v_{hn} \downarrow 0$. The left-hand side of (47), therefore, reaches $-s < 0$. Reversely, as $m_{do} \uparrow m_d$, $v_{lo} \downarrow 0$ and $v_{hn} \uparrow 1$, the left-hand side of (47) reaches $1 + m_d - s > 0$ (as it is assumed that $0 < s < 1 + m_d$). Therefore, by continuity, the solution to m_{do} always exists.

To prove uniqueness, examine the derivative of the left-hand side of (47) with respect to m_{do}:

$$-\eta\lambda_d \frac{(\lambda_u + \lambda_d + \rho v_{hn})v_{hn}}{(\lambda_u v_{hn} + \lambda_d v_{lo} + \rho v_{hn} v_{lo})^2} \frac{\partial v_{lo}}{\partial m_{do}} + \eta\lambda_d \frac{(\lambda_u + \lambda_d + \rho v_{lo})v_{lo}}{(\lambda_u v_{hn} + \lambda_d v_{lo} + \rho v_{hn} v_{lo})^2} \frac{\partial v_{hn}}{\partial m_{do}} + 1 > 0, \tag{50}$$

where the inequality holds because v_{lo} decreases, while v_{hn} increases, in m_{do}. That is, the left-hand side of (47) is strictly monotone increasing in m_{do}. Hence, there exists one and only one m_{do} that solves (47). Therefore, the demographics equation system always has a unique solution. □

Lemma 2

Proof. If one did switch to trading, her expected trading price p would fall between the reservation values. For example, if an *ho*-customer were to sell, she would get a price between $R_l = V_{lo} - V_{ln} \leq p \leq V_{do} - V_{dn} = R_d$ and continue with V_{hn}. Given the strictly positive trading gains, we have $R_d < R_h = V_{ho} - V_{hn}$, implying $V_{ho} > V_{hn} + p$, and the *ho*-customer never wants to sell. The same holds for an *lo*-customer. They are really bystanders. □

Lemma 3

Proof. Calculate the difference between m_{hn} and m_{lo} using the expressions (32) and (31) to get

$$m_{hn} - m_{lo} = \eta + m_{do} - s = \eta\lambda_d \cdot \frac{v_{lo} - v_{hn}}{\lambda_u v_{hn} + \lambda_d v_{lo} + \rho v_{hn} v_{lo}}, \tag{51}$$

where the last equality follows Equation (47). Therefore, $\text{sign}[m_{hn} - m_{lo}] = \text{sign}[v_{lo} - v_{hn}]$. Recall that $v_{lo} = 1 - (1 - \pi_{dn})^n$ and $v_{hn} = 1 - (1 - \pi_{do})^n$, from which it follows that $v_{lo} > v_{hn}$ if and only if $m_{dn} > m_{do}$. Given that $m_{dn} + m_{do} = m_d$, therefore, $m_{hn} > m_{lo}$ if and only if $m_{do} < m_d/2$. Use again $m_{hn} - m_{lo} = \eta + m_{do} - s$, which is negative if and only if $s > \eta + m_{do} > \eta + m_d/2$. □

Lemma 4

Proof. Consider **first** the case of $\rho^{\text{SMS}} q^{\text{SMS}} n^{\text{SMS}} < \rho^{\text{BB}} q^{\text{BB}} n^{\text{BB}}$. The proof first establishes the single-crossing of $z^{\text{SMS}}(\pi)$ and $z^{\text{BB}}(\pi)$ at some $\pi^* \in (0, 1)$. The general idea is to characterize the shapes of $z^{\text{BB}}(\pi)$ and $z^{\text{SMS}}(\pi)$. In particular, it will be shown that z^{BB} is linearly increasing in π, while z^{SMS} is sigmoid-shaped in π, starting below z^{BB} for sufficiently small π; and the two satisfy $z^{\mathrm{BB}}(0) = z^{\mathrm{SMS}}(0) = 0$ and $z^{\mathrm{BB}}(1) < z^{\mathrm{SMS}}(1)$. Therefore, there is always one and only one intersection point $\pi^* \in (0, 1)$.

Consider z^{BB} first. With $n^{\mathrm{BB}} = 1$, $z^{\mathrm{BB}} = q^{\mathrm{BB}}\rho^{\mathrm{BB}}\pi$, which is linearly increasing from 0 at $\pi = 0$ to $q^{\mathrm{BB}}\rho^{\mathrm{BB}}$ at $\pi = 1$. Next, consider z^{SMS}. For notation simplicity, the superscripts SMS on n, ρ, and q are omitted when there is no confusion. With $n = n^{\mathrm{SMS}} > 1$, $z^{\mathrm{SMS}} = \big(1 - (1 - \pi)^{n-1}(1 - \pi + (1 - q)n\pi)\big)\rho$, whose first-order derivative with respect to π is $\frac{\partial z^{\mathrm{SMS}}}{\partial \pi} = -n\rho(1 - \pi)^{n-2}(\pi(1 - n) + q(\pi n - 1))$, which is positive. To see why, note that the bracketed term, $\pi(1 - n) + q(\pi n - 1)$ is linear in π and is negative for both $\pi = 0$ and $\pi = 1$ and so it is negative for all π. Thus, $z^{\mathrm{SMS}}(\pi)$ is strictly monotone increasing on $\pi \in (0, 1)$. Its second-order derivative with respect to π is $\frac{\partial^2 z^{\mathrm{SMS}}}{\partial \pi^2} = (n - 1)n\rho(1 - \pi)^{n-3}(\pi - n\pi + (\pi n - 2)q + 1)$, which is positive if and only if $\pi < \frac{1-2q}{n-1-nq}$. Note that $\frac{1-2q}{n-1-nq} > 0$, because $\rho^{\mathrm{SMS}}q^{\mathrm{SMS}}n^{\mathrm{SMS}} < \rho^{\mathrm{BB}}q^{\mathrm{BB}}n^{\mathrm{BB}}$ implies $q = q^{\mathrm{SMS}} < 1/n^{\mathrm{SMS}} \leq 1/2$. Summarizing the above, $z^{\mathrm{SMS}}(\cdot)$ is sigmoid-shaped for $\pi > 0$: it is monotone increasing, initially convex, but eventually concave.

Now note that in the lower end, $z^{\mathrm{SMS}}|_{\pi\downarrow 0} = z^{\mathrm{BB}}|_{\pi\downarrow 0} = 0$. Further, the slope of $z(\cdot)$ satisfies $\lim_{\pi\downarrow 0} \frac{\mathrm{d}z}{\mathrm{d}\pi} = n\rho q$. Therefore, the assumption $\rho^{\mathrm{SMS}}q^{\mathrm{SMS}}n^{\mathrm{SMS}} < \rho^{\mathrm{BB}}q^{\mathrm{BB}}n^{\mathrm{BB}}$ ensures that for π sufficiently small, $z^{\mathrm{SMS}} < z^{\mathrm{BB}}$. On the upper end of $\pi \uparrow 1$, $z^{\mathrm{SMS}} \to \rho^{\mathrm{SMS}} \geq \rho^{\mathrm{BB}} \geq q^{\mathrm{BB}}\rho^{\mathrm{BB}}$, where the first inequality follows (19) and the second follows $q^{\mathrm{BB}} \in [0, 1]$. That is, z^{SMS} exceeds z^{BB} eventually. Therefore, there exists a unique $\pi^* \in (0, 1)$ at which $z^{\mathrm{SMS}}(\pi^*) = z^{\mathrm{BB}}(\pi^*)$.

Next, it is clear that V^k_σ is monotone increasing in ζ^k_σ, where $k \in \{\mathrm{BB}, \mathrm{SMS}\}$ and $\sigma \in \{lo, hn\}$. Hence, comparing the value functions is equivalent to comparing the trading gain intensities $\{\zeta^k_\sigma\}$; i.e., the technology choice (20) is equivalent to (23). With the single-crossing property established above, it then follows that the comparison of the $\{\zeta^k_\sigma\}$ is equivalent to (22).

Finally, consider the case of $\rho^{\mathrm{SMS}}q^{\mathrm{SMS}}n^{\mathrm{SMS}} \geq \rho^{\mathrm{BB}}q^{\mathrm{BB}}n^{\mathrm{BB}}$. The only change is that the slope of $z^k(\pi)$ at the lower end now is higher for SMS than for BB. Thus, the only intersection possible is at $\pi = 0$, i.e., $z^{\mathrm{SMS}} > z^{\mathrm{BB}}$ for all $\pi \in (0, 1)$, i.e., SMS is always preferred and, hence, $\theta_{hn} = \theta_{lo} = 1$. □

Lemma D.1

> **Lemma D.1.** Write the left-hand side of Equation (47) as a function of $g(\theta_{lo}, \theta_{hn}, m_{do}, s)$. Then (1) $\frac{\partial g}{\partial m_{do}} > 0$, (2) $\frac{\partial g}{\partial \theta_{lo}} < 0$, (3) $\frac{\partial g}{\partial \theta_{hn}} > 0$, and (4) $\frac{\partial g}{\partial s} < 0$. In particular, (5) $m_{do} \downarrow 0$ when $s \downarrow 0$ and $m_{do} \uparrow 1 + m_d$ when $s \uparrow 1 + m_d$ regardless of θ_{lo} and θ_{hn}. Finally, (6) $\frac{\partial m_{do}}{\partial \theta_{lo}} > 0 > \frac{\partial m_{do}}{\partial \theta_{hn}}$.

Proof. (1) $\frac{\partial g}{\partial m_{do}}$ has been evaluated in (50) in the proof of Lemma 5. (2) Note that θ_{lo} affects $g(\cdot)$ only through v_{lo}, which is given by (49). Carefully simplifying, it can be found that $\frac{\partial v_{lo}}{\partial \theta_{lo}} = \frac{(v^{\mathrm{SMS}}_{lo} - v^{\mathrm{BB}}_{lo})(\lambda_u + v^{\mathrm{SMS}}_{lo})(\lambda_u + v^{\mathrm{BB}}_{lo})}{(\lambda_u + (1-\theta_{lo})v^{\mathrm{SMS}}_{lo} + \theta_{lo}v^{\mathrm{BB}}_{lo})^2} > 0$ where the inequality holds because $v^{\mathrm{SMS}}_{lo} > v^{\mathrm{BB}}_{lo}$ always holds (with

$\rho^{\text{SMS}} \geq \rho^{\text{BB}}$ and $n^{\text{SMS}} > n^{\text{BB}} = 1$). The partial derivative of $g(\cdot)$ with respect to ν_{lo} is $\frac{\partial g}{\partial \nu_{lo}} = -\frac{(\lambda_d+\lambda_u+\nu_{hn})\lambda_d \nu_{hn}}{(\lambda_u \nu_{hn}+(\lambda_d+\nu_{hn})\nu_{lo})^2} < 0$. Therefore, by chain rule, $\frac{\partial g}{\partial \theta_{lo}} < 0$. (3) can be proved similarly by showing that $\frac{\partial \nu_{hn}}{\partial \theta_{hn}} > 0$ and that $\frac{\partial g}{\partial \nu_{hn}} > 0$. The details are omitted for brevity. (4) is straightforward as $\frac{\partial g}{\partial s} = -1$. (5) By implicit function theorem, $g(\cdot) = 0$ implies that m_{do} strictly increases in s; see (1) and (4) above. The limit values as $s \downarrow 0$ or $s \uparrow 1$ can then be easily verified, regardless of θ_{lo} and θ_{hn}. (6) directly follows the implicit function theorem by (1)-(3). □

Lemma D.2

> **Lemma D.2.** When there is excess supply, both the search intensity ρ and the capacity n increase m_{do} and reduce m_{dn}, but their effects on customers' matching rates are different: A higher ρ increases ν_{hn} but decreases ν_{lo}, while a larger n increases both ν_{hn} and ν_{lo}.

Proof. The key equation is (47) in the proof of Lemma 1. Define the left-hand side as $f(m_{do}, \rho, n)$. Recall that Equation (50) has shown that $\frac{\partial f}{\partial m_{do}} > 0$. In addition, simple calculus gives $\text{sign}\left[\frac{\partial f}{\partial \rho}\right] = \text{sign}[\nu_{lo} - \nu_{hn}]$. Since excess supply is assumed, i.e., $m_{hn} < m_{lo}$, Equation (51) gives $\nu_{lo} < \nu_{hn}$. Hence, $\frac{\mathrm{d}m_{do}}{\mathrm{d}\rho} = -\frac{\partial f}{\partial \rho} / \frac{\partial f}{\partial m_{do}} > 0$, i.e., a higher ρ increases m_{do} and, because $m_{dn} = m_d - m_{do}$, decreases m_{dn}. It then also follows that a higher ρ increases $\nu_{hn} = 1 - (1 - \pi_{do})^n$ but decreases $\nu_{lo} = 1 - (1 - \pi_{dn})^n$.

Consider the effect of a larger n next. In that case it is more convenient to work with an equivalent version of (47):

$$\eta(1 + m_{do} - s)\left(\frac{\lambda_d + \lambda_u}{\nu_{lo}} + \rho\right) - (1 - \eta)(s - m_{do})\left(\frac{\lambda_d + \lambda_u}{\nu_{hn}} + \rho\right) = 0. \tag{52}$$

Define the left-hand side of (52) as $g(m_{do}, \rho, n)$. Since $\frac{\partial g}{\partial m_{do}} > 0$ we have $\text{sign}\left[\frac{\mathrm{d}m_{do}}{\mathrm{d}n}\right] = -\text{sign}\left[\frac{\partial g}{\partial n}\right]$ and it remains to sign $\frac{\partial g}{\partial n}$. Taking π_{do} and π_{dn} as given, then $\text{sign}\left[\frac{\partial g}{\partial n}\right] = \text{sign}[hh(1 - \pi_{dn}) - hh(1 - \pi_{do})]$, where

$$hh(x) := (\rho(1 - x^n) + \lambda_d + \lambda_u)^{-1} \frac{\log x}{x^{-n} - 1}.$$

Further, one can show that $hh(x)$ is decreasing in x. Then, because $\pi_{do} > \pi_{dn}$, $hh(1 - \pi_{dn}) < hh(1 - \pi_{do})$. It follows that $\frac{\mathrm{d}m_{do}}{\mathrm{d}n} > 0$, i.e., a higher n increases m_{do} and, hence, decreases m_{dn}.

Since higher n increases m_{do} it follows immediately that higher n increases $\nu_{hn} = 1 - (1 - \pi_{do})^n$. To see the effect of n on ν_{lo} we first prove that trading volume is increasing in n. The trading volume can be written as $t = \rho m_{hn} \nu_{hn}$ (Equation 8). Equation (32) gives another link between t and m_{hn}.

Combining the two gives

$$t = \frac{(1 + m_{do} - s)\lambda_u \rho}{(\lambda_d + \lambda_u)v_{hn}^{-1} + \rho},$$

which is increasing in n both directly and through the dependence of m_{do} on n. Thus, trading volume increases in n. Now, writing $t = \rho m_{lo} v_{lo}$ and using (31) we obtain $v_{lo} = \frac{\lambda_d+\lambda_u}{\frac{\lambda_d \rho(s-m_{do})}{t} - \rho}$ from which it follows that v_{lo} increases in n. □

Lemma D.3

Lemma D.3. When there is excess supply, the transparency ψ increase m_{do} and reduce m_{dn}.

Proof. Consider (52). Define the left-hand side of (52) as $g(m_{do}, \psi)$. It is straightforward to show that $\text{sign}(\partial g/\partial \psi) = \text{sign}\left[\frac{\partial}{\partial \psi}\log\left(\frac{\lambda_d+\lambda_u}{v_{lo}} + \rho\right) - \frac{\partial}{\partial \psi}\log\left(\frac{\lambda_d+\lambda_u}{v_{hn}} + \rho\right)\right]$. Denoting $x := m_{do}/m_d$, and $v(\pi) := 1 - (1-\pi)^n$ we can express $v_{hn} = v(\pi(x;\psi))$ and $v_{lo} = v(\pi(1-x;\psi))$. For $x > 1/2$ (which holds in the excess supply case) one can show that $\frac{\partial \pi(x,\psi)}{\partial \psi} < \frac{\partial \pi(1-x,\psi)}{\partial \psi}$. Additionaly, $\pi(x;\psi) > \pi(1-x;\psi)$. Then, an explicit calculation of $\frac{\partial}{\partial \psi}\log\left(\frac{\lambda_d+\lambda_u}{v(\pi(x;\psi))} + \rho\right)$ implies that $\frac{\partial}{\partial \psi}\log\left(\frac{\lambda_d+\lambda_u}{v(\pi(x;\psi))} + \rho\right) > \frac{\partial}{\partial \psi}\log\left(\frac{\lambda_d+\lambda_u}{v(\pi(1-x;\psi))} + \rho\right)$. It then follows that $\frac{\partial g}{\partial \psi} < 0$ and by implicit function theorem $\frac{\mathrm{d}g}{\mathrm{d}\psi} > 0$. □

Lemma D.3

Lemma D.4. The functions $z^{\text{SMS}}(\pi)$ and $z^{\text{BB}}(\pi)$ in Lemma 4 cross at $\pi^* > \frac{1}{2}$ if and only if $n^{\text{SMS}} = 2$ and $\frac{2q^{\text{SMS}}\rho^{\text{SMS}} - q^{\text{BB}}\rho^{\text{BB}}}{(2q^{\text{SMS}}-1)\rho^{\text{SMS}}} > \frac{1}{2}$.

Proof. Consider first the case $n^{\text{SMS}} > 2$. To establish that $\pi^* \le \frac{1}{2}$, note that $z^{\text{SMS}}(\pi)$ is monotone increasing in n^{SMS} and in q^{SMS}. Therefore, fixing $z^{\text{BB}}(\pi) = q^{\text{BB}}\rho^{\text{BB}}\pi$, the intersection π^* must be higher as n^{SMS} and q^{SMS} reduce. Likewise, fixing $z^{\text{SMS}}(\pi)$, π^* must be higher when the product of $q^{\text{BB}}\rho^{\text{BB}}$ increases. Since $q^{\text{BB}} \in [0,1]$ and $\rho^{\text{BB}} \le \rho^{\text{SMS}}$, the maximum of this product is $q^{\text{BB}}\rho^{\text{BB}} \le \rho^{\text{SMS}}$. Therefore, the maximum π^* is the solution to $z^{\text{SMS}}(\pi; n^{\text{SMS}} = 3, q^{\text{SMS}} = 0) - \rho^{\text{SMS}}\pi = 0$. Solving this equation gives the unique interior solution of $\pi^* = \frac{1}{2}$. For the case $n^{\text{SMS}} = 2$, we plug $n^{\text{SMS}} = 2$ into $z^{\text{SMS}}(\pi)$ and solve for π^* in (a linear equation) $z^{\text{BB}}(\pi) = q^{\text{BB}}\rho^{\text{BB}}\pi$. Doing so yields $\pi^* = \frac{2q^{\text{SMS}}\rho^{\text{SMS}} - q^{BB}\rho^{BB}}{(2q^{\text{SMS}}-1)\rho^{\text{SMS}}}$. □

Proposition 1

Proof. Note that the trading gain is $\Delta = R_{hn} - R_{lo} = (V_{ho} - V_{hn}) - (V_{lo} - V_{ln})$, a linear combination of the four unknown value functions. The four equations (10)-(13), therefore, is a linear equation system that uniquely pins down the four unknowns.

It only remains to prove that the trading gains are strictly positive when $\underline{y}_d \le y_d \le \overline{y}_d$. Difference Equation (10) and (13) to get $0 = y_h - rR_h - \zeta_{hn}\Delta_{hd} - \lambda_d \cdot (R_h - R_l)$. Similarly, difference Equation (12) and (11) to get $0 = y_l - rR_l + \zeta_{lo}\Delta_{dl} + \lambda_u \cdot (R_h - R_l)$. Finally, difference the two dealers' HJB equations, (14) and (15), to get $y_d - rR_d + \zeta_{do}\Delta_{hd} - \zeta_{dn}\Delta_{dl}$. Note that $\Delta_{hd} = R_h - R_d$ and $\Delta_{dl} = R_d - R_l$. Therefore, taking the $\{\zeta\}$ as given, the above form a 3-equation-3-unknown linear system, from which the reservation values $\{R_h, R_d, R_l\}$ can be uniquely solved. The resulting expressions are complicated and omitted here, but it is straightforward verify that they are all monotone increasing in y_d. (Note that the trading gain intensities $\{\zeta\}$ are independent of y_d.) Therefore, one can find the upper and the lower thresholds by solving $\overline{y}_d'$ explicitly from $R_h = R_d$ and $\underline{y}_d'$ from $R_d = R_l$:

$$\overline{y}_d' := y_l + (y_h - y_l)\frac{\zeta_{dn} + \zeta_{lo} + \lambda_u + r}{\zeta_{lo} + \lambda_d + \lambda_u + r} \quad \text{and} \quad \underline{y}_d' := y_h - (y_h - y_l)\frac{\zeta_{do} + \zeta_{hn} + \lambda_d + r}{\zeta_{hn} + \lambda_d + \lambda_u + r}.$$

The above thresholds are still endogenous of $\{\zeta\}$. To obtain the thresholds composed of exogenous parameters, note that $\overline{y}_d'$ is increasing in both ζ_{dn} and ζ_{lo}, that $\underline{y}_d'$ is decreasing in both ζ_{do} and ζ_{hn}, and that $\{\zeta\} \ge 0$. Therefore,

$$\begin{aligned}
\overline{y}_d' &\ge y_l + (y_h - y_l)\frac{\lambda_u + r}{\lambda_d + \lambda_u + r} = y_h - (y_h - y_l)\frac{\lambda_d}{\lambda_d + \lambda_u + r} =: \overline{y}_d; \\
\underline{y}_d' &\le y_h - (y_h - y_l)\frac{\lambda_d + r}{\lambda_d + \lambda_u + r} = y_l + (y_h - y_l)\frac{\lambda_u}{\lambda_d + \lambda_u + r} =: \underline{y}_d.
\end{aligned}$$

Clearly, $\overline{y}_d > \underline{y}_d$. As such, $\underline{y}_d \le y_d \le \overline{y}_d$ is sufficient to ensure $R_l < R_d < R_h$. □

Proposition 2

Proof. The effects of ρ and n are proved separately below. For concreteness, assume that the asset is in excess supply. (The case of excess demand is symmetric and omitted.)

A higher search intensity ρ: The trading volume can be written as $t = \rho m_{hn} v_{hn}$ (Equation 8). Equation (32) gives another link between t and m_{hn}. Combining the two gives

$$t = \frac{(1 + m_{do} - s)\lambda_u \rho}{(\lambda_d + \lambda_u)v_{hn}^{-1} + \rho}, \tag{53}$$

which is increasing in ρ and in m_{do} (note that v_{hn} is also increasing in m_{do}). Lemma D.2 has shown that a higher ρ increases m_{do} (given excess supply). Therefore, the volume increases with ρ. It is then also clear from (31) that m_{lo} decreases. Finally, $m_{hn} = \frac{v_{lo}}{v_{hn}} m_{lo}$ by (8). The ratio $\frac{v_{lo}}{v_{hn}} = \frac{1-\pi_{do}^n}{1-(1-\pi_{do})^n}$. Simply computing the derivative with respect to π_{do} can show that the ratio decreases with π_{do}. That is, a higher ρ, increasing m_{do} and π_{do}, results in a lower m_{hn} as well.

A larger search capacity n: Lemma D.2 has shown that a larger n also increases m_{do} (given excess supply). Note that since $v_{hn} = 1-(1-\pi(m_{do}/m_d))^n$, $\frac{\partial v_{hn}}{\partial m_{do}} > 0$ and $\frac{\partial v_{hn}}{\partial n} > 0$. From the same expression of t above, therefore, n also increases trading volume. Again, from Equation (31), it is clear that m_{lo}, the long-side, then decreases with n.

The effect on $m_{hn} = \frac{v_{lo}}{v_{hn}} m_{lo}$, the short-side, is more complicated, because now n also affects the ratio $\frac{v_{lo}}{v_{hn}}$. To prove the statement, instead, it is easier to turn to the following equivalent expression:

$$m_{hn}(m_{do}, n) := \frac{t}{\rho v_{hn}} = \frac{(1+m_{do}-s)\lambda_u}{\lambda_d + \lambda_u + \rho(1-(1-\pi(m_{do}/m_d))^n)}, \tag{54}$$

where the second equality follows Equation (32). It is straightforward to find that $\lim_{\rho\to 0} \frac{\partial m_{hn}}{\partial n} = 0$; and $\lim_{\rho\to 0} \frac{\partial m_{hn}}{\partial m_{do}} = \frac{\lambda_u}{\lambda_d+\lambda_u} > 0$. Directly computation using (52) implies $\lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}n} > 0$. Therefore, $\lim_{n\to\infty} \frac{\mathrm{d}m_{hn}}{\mathrm{d}n} = \lim_{n\to\infty} \left(\frac{\partial m_{hn}}{\partial n} + \frac{\partial m_{hn}}{\partial m_{do}} \frac{\mathrm{d}m_{do}}{\mathrm{d}n} \right) > 0$. □

Proposition 3

Proof. The proof considers the changes in ρ and in n separately. Only the case of excess supply, i.e., $s > \eta + m_d/2$, is analyzed (and the case of excess demand is analogous and is omitted).

When ρ increases: Recall welfare is $w = (y_h m_{ho} + y_d m_{do} + y_l m_{lo})/r$. By market clearing (2), substitute $m_{lo} = s - m_{ho} - m_{do}$ in the above welfare expression to get $w = (y_l s + (y_h - y_l) m_{ho} + (y_d - y_l) m_{do})/r$. By Lemma D.2, m_{do} increases with ρ. By Proposition 2, m_{hn} and m_{lo} decrease with ρ. That is, $m_{ho} = \eta - m_{hn}$ increases with ρ. Note that $y_d \in [\underline{y}_d, \overline{y}_d]$ is assumed to ensure positive trading gains (Proposition 1) and we also have that $y_l < \underline{y}_d < \overline{y}_d < y_h$. It then follows that $y_d \in (y_l, y_h)$. Therefore, welfare is increasing with ρ.

When n increases: Welfare can be written as $w = (y_l s + (y_d - y_l) m_{do} + (y_h - y_l)(\eta - m_{hn}))/r$. The effect of n goes through m_{do} and m_{hn}, which are linked through the trading volume definition of $t = \rho m_{hn} v_{hn}$. In the proof of Proposition 2, it has been shown that m_{hn} can be written as a function of m_{do} and n; see Equation (54). Applying the chain rule yields

$$\frac{\mathrm{d}m_{hn}}{\mathrm{d}n} = \frac{\partial m_{hn}}{\partial n} + \frac{\partial m_{hn}}{\partial m_{do}} \frac{\mathrm{d}m_{do}}{\mathrm{d}n}. \tag{55}$$

Combining the above, one can see that

$$\frac{\mathrm{d}w}{\mathrm{d}n} = \frac{1}{r}\left((y_d - y_l)\frac{\mathrm{d}m_{do}}{\mathrm{d}n} - (y_h - y_l)\frac{\mathrm{d}m_{hn}}{\mathrm{d}n} \right) = \frac{1}{r}\left(\left((y_d - y_l) - (y_h - y_l)\frac{\partial m_{hn}}{\partial m_{do}} \right) \frac{\mathrm{d}m_{do}}{\mathrm{d}n} - (y_h - y_l)\frac{\partial m_{hn}}{\partial n} \right).$$

Therefore, three derivatives of $\frac{\partial m_{hn}}{\partial n}$, $\frac{\partial m_{hn}}{\partial m_{do}}$, and $\frac{\mathrm{d}m_{do}}{\mathrm{d}n}$ need to be evaluated under $\rho \to 0$ and under $\rho \to \infty$.

Consider first the case of $\rho \to 0$**.** Directly computing the first partial derivative yields

$$\frac{\partial m_{hn}}{\partial n} = \frac{\lambda_u(1 + m_{do} - s)\left(1 - \frac{m_{do}}{m_d}\right)^n \rho \log\left(1 - \frac{m_{do}}{m_d}\right)}{\left(\lambda_d + \lambda_u + \rho\left(1 - \left(1 - \frac{m_{do}}{m_d}\right)^n\right)\right)^2}, \tag{56}$$

from which it follows that $\lim_{\rho\to 0} \frac{\partial m_{hn}}{\partial n} = 0$. Also, $\lim_{\rho\to 0} \frac{\partial m_{hn}}{\partial m_{do}} = \frac{\lambda_u}{\lambda_d+\lambda_u} = \eta$. Hence, $\lim_{\rho\to 0} \frac{\mathrm{d}m_{hn}}{\mathrm{d}n} = \eta \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}n}$. Therefore, $\lim_{\rho\to 0} \frac{\mathrm{d}w}{\mathrm{d}n} = \frac{1}{r}\left((y_d - y_l) \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}n} - (y_h - y_l)\eta \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}n}\right) = \frac{1}{r}(y_d - \hat{y}) \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}n}/r$, where $\hat{y} := \eta y_h + (1 - \eta)y_l$. Note that $\lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}n} > 0$ because (i) from (47), $\lim_{\rho\to 0} m_{do} \in (0, m_d)$; and (ii) given the excess supply, m_{do} increases in n (Lemma D.2). Therefore, $\mathrm{sign}\left[\lim_{\rho\to 0} \frac{\mathrm{d}w}{\mathrm{d}n}\right] = \mathrm{sign}[y_d - \hat{y}]$, proving the statement.

Next, consider the case of $\rho \to \infty$**.** Note that $\frac{\mathrm{d}m_{do}}{\mathrm{d}n} > 0$ (Lemma D.2). Then signing $\frac{\mathrm{d}w}{\mathrm{d}n}$ in this case is equivalent to

$$\mathrm{sign}\left[\lim_{\rho\to\infty} \frac{\mathrm{d}w}{\mathrm{d}n}\right] = \mathrm{sign}\left[\frac{y_d - y_l}{y_h - y_l} - \lim_{\rho\to\infty}\left(\frac{\frac{dm_{hn}}{dn}}{\frac{dm_{do}}{dn}}\right)\right].$$

Next, we use (56) to calculate $\frac{\partial m_{hn}}{\partial n} \leq 0$. Equation (56) then implies that $\lim_{\rho\to\infty} \frac{\frac{dm_{hn}}{dn}}{\frac{dm_{do}}{dn}} \leq \lim_{\rho\to\infty} \frac{\partial m_{hn}}{\partial m_{do}} = 0$. This implies that $\lim_{\rho\to\infty} \frac{\mathrm{d}w}{\mathrm{d}n} > 0$. □

Proposition 4

Proof. Similar to the proof of Proposition 3 one can see that

$$\frac{\mathrm{d}w}{\mathrm{d}\psi} = \frac{1}{r}\left((y_d - y_l)\frac{\mathrm{d}m_{do}}{\mathrm{d}\psi} - (y_h - y_l)\frac{\mathrm{d}m_{hn}}{\mathrm{d}\psi}\right) = \frac{1}{r}\left(\left((y_d - y_l) - (y_h - y_l)\frac{\partial m_{hn}}{\partial m_{do}}\right)\frac{\mathrm{d}m_{do}}{\mathrm{d}\psi} - (y_h - y_l)\frac{\partial m_{hn}}{\partial \psi}\right).$$

Therefore, three derivatives of $\frac{\partial m_{hn}}{\partial \psi}$, $\frac{\partial m_{hn}}{\partial m_{do}}$, and $\frac{\mathrm{d}m_{do}}{\mathrm{d}\psi}$ need to be evaluated under $\rho \to 0$ and under $\rho \to \infty$.

Consider first the case of $\rho \to 0$**.** Directly computing the first partial derivative yields we get that $\lim_{\rho\to 0} \frac{\partial m_{hn}}{\partial \psi} = 0$. Also, $\lim_{\rho\to 0} \frac{\partial m_{hn}}{\partial m_{do}} = \frac{\lambda_u}{\lambda_d+\lambda_u} = \eta$. Hence, $\lim_{\rho\to 0} \frac{\mathrm{d}m_{hn}}{\mathrm{d}\psi} = \eta \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}\psi}$. Therefore, $\lim_{\rho\to 0} \frac{\mathrm{d}w}{\mathrm{d}\psi} = \frac{1}{r}\left((y_d - y_l) \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}\psi} - (y_h - y_l)\eta \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}\psi}\right) = \frac{1}{r}(y_d - \hat{y}) \lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}\psi}/r$, where $\hat{y} := \eta y_h + (1 - \eta)y_l$. Note that $\lim_{\rho\to 0} \frac{\mathrm{d}m_{do}}{\mathrm{d}\psi} > 0$. Therefore, $\mathrm{sign}\left[\lim_{\rho\to 0} \frac{\mathrm{d}w}{\mathrm{d}\psi}\right] = \mathrm{sign}[y_d - \hat{y}]$, proving the statement.

Next, consider the case of $\rho \to \infty$**.**] Note that $\frac{\mathrm{d}m_{do}}{\mathrm{d}\psi} > 0$ (Lemma D.3). Then signing $\frac{\mathrm{d}w}{\mathrm{d}\psi}$ in this case is equivalent to

$$\operatorname{sign}\left[\lim_{\rho\to\infty}\frac{\mathrm{d}w}{\mathrm{d}\psi}\right]=\operatorname{sign}\left[\frac{y_d-y_l}{y_h-y_l}-\lim_{\rho\to\infty}\left(\frac{\frac{dm_{hn}}{d\psi}}{\frac{dm_{do}}{d\psi}}\right)\right].$$

Next, it follows from from (54) that $\frac{\partial m_{hn}}{\partial n}\leq 0$. Implicit function theorem applied to (54) then implies that $\lim_{\rho\to\infty}\left(\frac{dm_{hn}}{d\psi}/\frac{dm_{do}}{d\psi}\right)\leq\lim_{\rho\to\infty}\frac{\partial m_{hn}}{\partial m_{do}}=0$. This implies that $\lim_{\rho\to\infty}\frac{\mathrm{d}w}{\mathrm{d}\psi}>0$. □

Proposition 5

Proposition 5 only characterizes the equilibrium for the case of $n=n^{\mathrm{SMS}}>2$, which guarantees that $\pi^*\leq\frac{1}{2}$ by Lemma D.4. Before proceeding to the proof, we first add the case of $n=2$:

> **Proposition (Equilibrium technology choices when $n=2$).** If $\pi^*\leq\frac{1}{2}$, Proposition 5 holds. If $\pi^*>1/2$, a unique stationary equilibrium exists depending on the asset supply s: There exist thresholds $0<\hat{s}_{hn,0}<\hat{s}_{hn,1}\leq\hat{s}_{lo,1}<\hat{s}_{lo,0}<1+m_d$ so that
>
> | | (a) hn-buyers' proba -bility to use SMS, θ_{hn} | (b) lo-sellers' proba -bility to use SMS, θ_{lo} | (c) asset holding by dealers, m_{do} |
> |---|---|---|---|
> | (1) $0<s\leq\hat{s}_{hn,0}$ | 0 | 1 | $g(0,1,m_{do})=s$ |
> | (2) $\hat{s}_{hn,0}\leq s\leq\hat{s}_{hn,1}$ | 0 | $g(1,\theta_{lo},m_d-m_d^*)=s$ | $m_d-m_d^*$ |
> | (3) $\hat{s}_{hn,1}<s<\hat{s}_{lo,1}$ | 0 | 0 | $g(0,0,m_{do})=s$ |
> | (4) $\hat{s}_{lo,1}\leq s\leq\hat{s}_{lo,0}$ | $g(\theta_{hn},1,m_d^*)=s$ | 0 | m_d^* |
> | (5) $\hat{s}_{lo,0}<s<1+m_d$ | 1 | 0 | $g(1,0,m_{do})=s$ |
>
> where $g(x_1,x_2,x_3)=s$ uniquely solves θ_{hn}, θ_{lo}, and m_{do} in columns (a), (b), and (c), respectively. The constant π^* is given in Lemma 4 and $m_d^*:=\pi^{-1}(\pi^*)m_d$. The function $g(\cdot)$ and the the thresholds $\{\hat{s}_{hn,0},\hat{s}_{hn,1},\hat{s}_{lo,1},\hat{s}_{lo,0}\}$ are given in the proof.

Proof. To begin with, note that both v_{lo} and v_{hn} are only functions of θ_{lo}^k and θ_{hn}^k, respectively; see, e.g., Equation (49). Following Lemma D.1, Equation (47) can be written as $g(\theta_{lo},\theta_{hn},m_{do};s)=0$. For the case of $\pi^*\leq\frac{1}{2}$, define the four thresholds $\{s_{hn,0},s_{hn,1},s_{lo,1},s_{lo,0}\}$ to be the respective unique solution to $g(\cdot;s)=0$ for $\{\theta_{lo},\theta_{hn},m_{do}\}\in\{\{1,0,\pi^{-1}(\pi^*)m_d\},\{1,1,\pi^{-1}(\pi^*)m_d\},\{1,1,(1-\pi^{-1}(\pi^*))m_d\},\{0,1,(1-\pi^{-1}(\pi^*))m_d\}\}$. For the case of $\pi^*>\frac{1}{2}$, likewise, the four thresholds $\{\hat{s}_{hn,0},\hat{s}_{hn,1},\hat{s}_{lo,1},\hat{s}_{lo,0}\}$ are defined as the respective unique solution to $g(\cdot;s)=0$ for $\{\theta_{lo},\theta_{hn},m_{do}\}\in\{\{1,0,(1-\pi^{-1}(\pi^*))m_d\},\{0,0,(1-\pi^{-1}(\pi^*))m_d\},\{0,0,\pi^{-1}(\pi^*)m_d\},\{0,1,\pi^{-1}(\pi^*)m_d\}\}$. In either case, it is easy to see that the four thresholds indeed exist according to the respective definition. In particular, the monotonicity shown in Lemma D.1 guarantees the sorting of these thresholds.

To complete the proof, for each region of s, the stated values of $\{\theta_{lo}, \theta_{hn}, m_{do}\}$ are first verified to indeed sustain an equilibrium and then shown to be unique in that region. Only the case of $\pi^* \leq \frac{1}{2}$ is discussed below for brevity.

Region 1: $0 < s < s_{hn,0}$. With $\{\theta_{lo}, \theta_{hn}\} = \{1, 0\}$, m_{do} is uniquely pinned down by Equation (47). Since $s < s_{hn,0}$, Lemma D.1 implies that $\pi_{do} < \pi^*$. Hence, by Lemma 4, $\zeta_{hn}^{\mathrm{SMS}} < \zeta_{hn}^{\mathrm{BB}}$ but $\zeta_{lo}^{\mathrm{SMS}} > \zeta_{lo}^{\mathrm{BB}}$ and, indeed, $\{\theta_{lo}, \theta_{hn}\} = \{1, 0\}$ sustains an equilibrium.

There are three possible deviations. First, suppose instead that :w $\{\theta_{lo}, \theta_{hn}\} \in (0, 1) \times (0, 1)$. This would require both hn-buyers and lo-sellers be indifferent between the two technologies. That is, $\pi_{do} = \pi_{dn} = \pi^*$ must hold, but this cannot be true because $\pi_{do} < \pi^*$ in this region. Second, suppose $\theta_{lo} = \theta_{hn} = 0$. But by Lemma D.1, this reduction in θ_{lo} would only reduce m_{do} (for a fixed s) and increase m_{dn}, making lo-sellers prefer SMS more, hence inconsistent with $\zeta_{lo}^{\mathrm{SMS}} < \zeta_{lo}^{\mathrm{BB}}$ as implied by $\theta_{lo} = 0$. Third, suppose $\theta_{lo} = \theta_{hn} = 1$. Likewise, this increase in θ_{hn} would decrease m_{do}, inconsistent with hn-buyers' switch from BB to SMS as a lower m_{do} would only strengthen $\zeta_{hn}^{\mathrm{SMS}} < \zeta_{hn}^{\mathrm{BB}}$. Since none of these alternative values of θ_{lo} and θ_{hn} can sustain the equilibrium, in this range of s, the only possible equilibrium is $\{\theta_{lo}, \theta_{hn}\} = \{1, 0\}$.

Region 2: $s_{hn,0} \leq s \leq s_{hn,1}$. With $\{\theta_{lo}, m_{do}\} = \{1, \pi^{-1}(\pi^*)m_d\}$ in this region, $g(\cdot; s) = 0$ uniquely solves $\theta_{hn} \in [0, 1]$. This is indeed an equilibrium because at $\pi_{do} = \pi^*$, hn-buyers are indifferent between SMS and BB and, hence, any $\theta_{hn} \in [0, 1]$ is admissible. On the other hand, $\pi_{do} = \pi^* < \frac{1}{2}$ implies that $m_{do} < m_d/2$ (recall that $\pi(x) \geq x$ by assumption) and so $m_{dn} > m_d/2$. It then follows that $\pi_{dn} > m_{dn}/m_d > \frac{1}{2} > \pi^*$ (because $\pi^* < 1/2$). Therefore, $\zeta_{lo}^{\mathrm{SMS}} > \zeta_{lo}^{\mathrm{BB}}$ by Lemma 4 and $\theta_{lo} = 1$ is sustained.

To rule out other equilibria, consider alternative values. Suppose $\pi_{do} > \pi^*$, implying $\theta_{hn} = 1$. Recall that $s = s_{hn,1}$ is the unique solution to $g(\cdot; s) = 0$ when $\theta_{lo} = \theta_{hn} = 1$ and $m_{do} = \pi^{-1}(\pi^*)m_d$. The monotonicity in Lemma D.1 would then require $s > s_{hn,1}$, out of this region. Suppose instead $\pi_{do} < \pi^*$, implying $\theta_{hn} = 0$. Then, similarly, the monotonicity in Lemma D.1 would require $s < s_{hn,0}$, again out of this region. Finally, suppose $\pi_{do} = \pi^*$ but $\theta_{lo} < 1$. Then $\pi_{do} = \pi^* < \frac{1}{2}$ implies that $m_{do} < m_d/2$ and so $m_{dn} > m_d/2$. It then follows that $\pi_{dn} > m_{dn}/m_d > \frac{1}{2} > \pi^*$, implying $\theta_{lo} = 1$, a contradiction.

Region 3: $s_{hn,1} < s < s_{lo,1}$. When $\theta_{lo} = \theta_{hn} = 1$, $s_{hn,1} < s < s_{lo,1}$ ensures that m_{do} as solved from $g(\cdot; s) = 0$ satisfies $\pi^{-1}(\pi^*)m_d < m_{do} < \pi^{-1}(1 - \pi^*)m_d$; and, hence, $\pi_{dn} > \pi^*$. That is, $\zeta^{\mathrm{SMS}} > \zeta^{\mathrm{BB}}$ for both hn and lo, which indeed guarantee that $\theta_{lo} = \theta_{hn} = 1$ as an equilibrium.

Again, consider other values for $\{\theta_{lo}, \theta_{hn}\}$. First, $\{\theta_{lo}, \theta_{hn}\} \in (0, 1)^2$ cannot be an equilibrium for the same reason as explained in Region 1. Second, suppose $\{\theta_{lo}, \theta_{hn}\} = \{1, 0\}$. By Lemma D.1, this reduction in θ_{hn} would result in an increase in m_{do}, but such an increase would only make SMS more attractive for hn-buyers, contradicting the reduction of θ_{hn}. Third, suppose $\{\theta_{lo}, \theta_{hn}\} = \{0, 1\}$. Then similarly by Lemma D.1, this reduction in θ_{lo} would result in a decrease in m_{do} or an increase in m_{dn}, but such an increase would only make SMS more attractive for lo-sellers, contradicting the reduction of θ_{lo}.

Region 4: $s_{lo,1} \leq s \leq s_{lo,0}$**.** This region mirrors Region 2 and the proof is omitted for brevity.

Region 5: $s_{lo,0} < s < 1+m_d$**.** This region mirrors Region 1 and the proof is omitted for brevity. □

Proposition 6

Proof. We consider the case of $s > s_{hn,1}$ and prove that the ratio defined in (25) weakly decreases in s. The volume share ratio, VS, in this region can be written as

$$\frac{\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}} + \rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}}{\left(\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}} + \rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}\right) + \left(\rho^{\text{BB}} m_{lo}^{\text{BB}} v_{lo}^{\text{BB}} + \rho^{\text{BB}} m_{hn}^{\text{BB}} v_{hn}^{\text{BB}}\right)} = \frac{1}{2} + \frac{1}{2} \frac{m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}}}{m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}}.$$

This is because in the considered region, $\theta_{hn} = 1$. Then the dealer stationarity (37) reduces to

$$\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}} + \rho^{\text{BB}} m_{lo}^{\text{BB}} v_{lo}^{\text{BB}} = \rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}. \tag{57}$$

We consider three cases next:

- $s < s_{lo,1}$. In this case, $\theta_{lo} = 1$, which means that the dealer stationarity condition (37) writes as $\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}} = \rho^{\text{SMS}} m_{hn}^{\text{SMS}} v_{hn}^{\text{SMS}}$ implying $VS = 1$.
- $s > s_{lo,0}$. In this case, $\theta_{lo} = 0$, implying $VS = 1/2$.
- $s_{lo,1} \leq s \leq s_{lo,0}$. In this case, m_{do} is a constant, invariant of s, and so both v_{lo}^{BB} and v_{lo}^{SMS} are constants as well. Then $\text{sign}\frac{\mathrm{d}VS}{\mathrm{d}s} = \text{sign}\frac{\mathrm{d}}{\mathrm{d}s}\left(m_{lo}^{\text{SMS}}/m_{hn}^{\text{SMS}}\right)$. Using again (57),

$$\frac{m_{lo}^{\text{SMS}}}{m_{hn}^{\text{SMS}}} = \frac{\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{hn}^{\text{SMS}}}{\rho^{\text{SMS}} m_{lo}^{\text{SMS}} v_{lo}^{\text{SMS}} + \rho^{\text{BB}} m_{lo}^{\text{BB}} v_{lo}^{\text{BB}}} = \frac{\rho^{\text{SMS}} v_{hn}^{\text{SMS}}}{\rho^{\text{SMS}} v_{lo}^{\text{SMS}} + \rho^{\text{BB}} v_{lo}^{\text{BB}} \left(\frac{m_{lo}^{\text{BB}}}{m_{lo}^{\text{SMS}}}\right)}$$

  Using the stationarity conditions (33) and (34),

$$\frac{m_{lo}^{\text{SMS}}}{m_{lo}^{\text{BB}}} = \frac{\lambda_u + \rho^{\text{BB}} v_{lo}^{\text{BB}}}{\lambda_u + \rho^{\text{SMS}} v_{lo}^{\text{SMS}}} \frac{\theta_{lo}}{1 - \theta_{lo}},$$

  increasing in θ_{lo}, which is the only variable endogenous of s. Proposition 5 has shown that in this range, θ_{lo} decreases with s. Therefore, by chain rule, $\text{sign}\frac{\mathrm{d}V}{\mathrm{d}s} < 0$.

Combining the three cases completes the proof for the claims regarding s. To prove the claims regarding λ_d, note that from Equation (47), cateris paribus, the left-hand side is monotone increasing in λ_d (the excess supply implies $v_{hn} > v_{lo}$; see Equation (51)) but decreasing in s. Therefore, increases in s are equivalent to those in λ_d. Hence, all results about s above also hold for λ_d. $\square$

Proposition 7 and 8

Proof. Welfare can be written as $w = \frac{1}{r}(y_l s + (y_d - y_l)m_{do} + (y_h - y_l)(\eta - m_{hn}))$. Consider a small change in either $\theta \in \{\theta_{hn}, \theta_{lo}\}$. We then have

$$\operatorname{sign}\left[\frac{\mathrm{d}w}{\mathrm{d}\theta}\right] = \operatorname{sign}\left[(y_d - y_l)\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta} - (y_h - y_l)\frac{\mathrm{d}m_{hn}}{\mathrm{d}\theta}\right].$$

Moreover, following $m_{ho} + m_{hn} = \eta$ and using the expressions (31) and (32), we have

$$\frac{\mathrm{d}m_{hn}}{\mathrm{d}\theta} = -\frac{\mathrm{d}m_{hn}}{\mathrm{d}\theta} = \eta\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta} - \frac{1}{\lambda_u + \lambda_d}\frac{\mathrm{d}t}{\mathrm{d}\theta}. \tag{58}$$

Combining the above two, we get

$$\operatorname{sign}\left[\frac{\mathrm{d}w}{\mathrm{d}\theta}\right] = \operatorname{sign}\left[(y_d - \hat{y})\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta} + \frac{y_h - y_l}{\lambda_u + \lambda_d}\frac{\mathrm{d}t}{\mathrm{d}\theta}\right], \tag{59}$$

where $\hat{y} := \eta y_h + (1-\eta)y_l$. The derivative of $\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}$ can be signed by the implicit function theorem using the results from Lemma D.1: $\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta_{lo}} > 0$ and $\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta_{hn}} < 0$. To see how volume t changes with respect to θ, recall from Equations (31) and (32) and use $t = \rho m_{lo} v_{lo} = \rho m_{hn} v_{hn}$ to get

$$t = \frac{\lambda_d(s - m_{do})}{\frac{\lambda_d + \lambda_u}{\rho v_{lo}} + 1} \text{ and } t = \frac{\lambda_u(1 + m_{do} - s)}{\frac{\lambda_d + \lambda_u}{\rho v_{hn}} + 1}. \tag{60}$$

Note that θ_{hn} in the first expression only affects t through m_{do}. Therefore, t is increasing in θ_{lo}. Likewise, θ_{lo} affects t in the second expression only through m_{do}. Hence, t is also increasing in θ_{hn}. That is, $\frac{\mathrm{d}t}{\mathrm{d}\theta} > 0$ for either $\theta \in \{\theta_{lo}, \theta_{hn}\}$.

The case of sufficiently high ρ, i.e., $\rho := \min[\rho^{\mathbf{BB}}, \rho^{\mathbf{SMS}}] \to \infty$: Since $\frac{\mathrm{d}t}{\mathrm{d}\theta} > 0$,

$$\operatorname{sign}\left[\lim_{\rho\to\infty}\frac{\mathrm{d}w}{\mathrm{d}\theta}\right] = \operatorname{sign}\left[(y_d - \hat{y})\lim_{\rho\to\infty}\left(\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}\Big/\frac{\mathrm{d}t}{\mathrm{d}\theta}\right) + \frac{y_h - y_l}{\lambda_u + \lambda_d}\right].$$

Hence, one needs to find $\lim_{\rho\to\infty}\left(\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}/\frac{\mathrm{d}t}{\mathrm{d}\theta}\right)$.

Consider first $\theta = \theta_{lo}$. Then differentiate the second expression of t in (60) with respect to $\theta = \theta_{lo}$, to get $\left(\frac{\lambda_d + \lambda_u}{v_{hn}} + \rho\right)\frac{\mathrm{d}t}{\mathrm{d}\theta} = \rho\lambda_u\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}$. Hence, $\lim_{\rho\to\infty}\left(\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}/\frac{\mathrm{d}t}{\mathrm{d}\theta}\right) = \frac{1}{\lambda_u}$. (Note that $v_{hn}^k = 1 - (1 - m_{do}/m_d)^{n^k}$ is always nonzero, because $m_{do} > m_d/2$ in the case of excess supply.) Then $\operatorname{sign}\left[\lim_{\rho\to\infty}\frac{\mathrm{d}w}{\mathrm{d}\theta}\right] = \operatorname{sign}\left[\frac{y_d - \hat{y}}{\lambda_u} + \frac{y_h - y_l}{\lambda_u + \lambda_d}\right] = \operatorname{sign}[y_d - \hat{y} - (y_h - y_l)\eta] = \operatorname{sign}[y_d - y_l] > 0$. (Recall that $y_d \in (\overline{y}_d', \overline{y}_d') \subset (y_l, y_h)$ by Corollary 3).

Consider $\theta = \theta_{hn}$. Then differentiate the first expression of t in (60) with respect to $\theta = \theta_{hn}$. As $\rho \to \infty$, the limit of m_{do} may be binding at m_d, resulting in $v_{lo} \to 0$. If it is not binding, i.e., if $\lim_{\rho\to\infty} m_{do} < m_d$, then $v_{lo}^k > 0$ and $\lim_{\rho\to\infty}\left(\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}/\frac{\mathrm{d}t}{\mathrm{d}\theta}\right) = -\frac{1}{\lambda_d}$. If it is binding, i.e., $m_{do} \to m_d$ and $v_{lo} \to 0$, however ρv_{lo} has a strictly positive limit as folows from (47). Then, one can again derive $\lim_{\rho\to\infty}\left(\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}/\frac{\mathrm{d}t}{\mathrm{d}\theta}\right) > -\frac{1}{\lambda_d}$. Therefore, $\mathrm{sign}\left[\lim_{\rho\to\infty}\frac{\mathrm{d}w}{\mathrm{d}\theta}\right] > \mathrm{sign}\left[-\frac{y_d-\hat{y}}{\lambda_d} + \frac{y_h-y_l}{\lambda_u+\lambda_d}\right] = \mathrm{sign}[-y_d + \hat{y} - (y_h - y_l)(1-\eta)] = \mathrm{sign}[y_h - y_d] > 0$.

The case of sufficiently low ρ, i.e., $\rho := \max\{\rho^{\mathbf{BB}}, \rho^{\mathbf{SMS}}\} \to 0$: For either $\theta \in \{\theta_{lo}, \theta_{hn}\}$, directly calculating $\frac{\mathrm{d}t}{\mathrm{d}\theta}$ from (60) and taking the limit yield $\lim_{\rho\to 0}\frac{\mathrm{d}t}{\mathrm{d}\theta} = 0$. Yet, $\lim_{\rho\to 0}\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta} \neq 0$, which follows by taking the limit in the calculations of Lemma D.1. Hence, $\lim_{\rho\to 0}\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta_{lo}} > 0$ and $\lim_{\rho\to 0}\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta_{hn}} < 0$ remain. Therefore, $\lim_{\rho\to 0}\mathrm{sign}\left[\frac{\mathrm{d}w}{\mathrm{d}\theta}\right] = \mathrm{sign}\left[(y_d - \hat{y})\lim_{\rho\to 0}\frac{\mathrm{d}m_{do}}{\mathrm{d}\theta}\right]$, proving the statement made in the proposition. □

Proposition 9

Proof. The expressions in (27) follow the dealer stationarity condition (8). Note that the assumption (26) can be equivalently written as$(v_{hn} - 1)m_{do} = 0$ and $(v_{lo} - 1)m_{dn} = 0$, which, together with (2)-(7), give three sets of solutions to the six demographic variables and the two matching probabilities. The three solutions one-to-one map into the three regions of the asset supply s: $0 < \eta < \eta + m_d < 1 + m_d$. The equations (27) and (28) can be easily verified using the three regions of the solution. □

Proposition 10

Proof. Consider first demographics. Recall that $v_{hn} = v(\pi_{do}; n) = 1 - (1 - \pi_{do})^n$. Whenever $m_{hn} > 0$, $\pi_{do} \in (0, 1)$ and in the limit of $n \to \infty$, $v_{hn} \to 1$. When $m_{hn} = 0$, then $\pi_{do} = 0$ and $v_{hn} = 0$. The same holds for the limit of $\psi \to 1$ (with $n \geq 2$), under the specific functional form of $\pi(\cdot; \psi)$. The limits for v_{lo} follow analogously. Therefore, Equation (26) holds in both limits, which then proves the convergence of all demographic variables together with the demographic conditions (2)-(7).

For value functions and prices, in both limit as $n \to \infty$ and $\psi \to 1$ (with $n \geq 2$), we have perfect competition among dealers, whenever $m_{do} \in (0, m_d)$. Thus, $p_a = p_b = R_d$ and the value function system is the same as in DS. In the case of $m_{do} = 0$ ($m_{do} = m_d$) we have $\Delta_{hd} = 0$ ($\Delta_{dl} = 0$) in the two limits ($n \to \infty$ and $\psi \to 1$) and so, again, prices and values are as in DS. □

Proposition 11

Proof. The proof only focuses on a contacted *do*-seller's symmetric quoting strategy. The same analysis applies to *dn*-buyers and is omitted. Consider first the trivial case of $n = 1$. A contacted *do*-seller then knows that he is the only one quoting. It is then trivial that with probability $(1 - q)$, he will quote the highest possible ask price, i.e., the *hn*-buyer's reservation value $R_h = R_d + \Delta_{hd}$. This can be viewed as a degenerate mixed strategy with c.d.f. $F(\alpha)$ converging to a unity probability mass at $\alpha = 1$ as stated in the proposition.

Next consider $n \geq 2$. Given the reservation values, it suffices to restrict the ask quote within $[R_d, R_h]$. Without loss of generality, a *do*-seller's strategy can be written as $R_d + \alpha\Delta_{hd}$ by choosing $\alpha \in [0, 1]$. Suppose α has a c.d.f. $F(\alpha)$ with possible realizations $[0, 1]$ (some of which might have zero probability mass). The following four steps pin down the specific form of $F(\cdot)$ so that it sustains a symmetric equilibrium.

Step 1: There are no probability masses in the support of $F(\cdot)$. If at $\alpha^* \in (0, 1]$ there is some non-zero probability mass, any *do*-seller has an incentive to deviate to quoting with the same probability mass but at a level infinitesimally smaller than α^*. This way, he converts the strictly positive probability of tying with others at α^* to winning over others. (The undercut costs no expected revenue as it is infinitesimally small.) If at $\alpha^* = 0$ there is non-zero probability mass, again, any *do*-seller will deviate, this time to an α just slightly above zero. This is because allocating probability mass at zero brings zero expected profit. Deviating to a slightly positive α, therefore, brings strictly positive expected profit. Taken together, there cannot be any probability mass in $\alpha \in [0, 1]$. Note that any pure symmetric-strategy equilibria are ruled out.

Step 2: The support of $F(\cdot)$ is connected. The support is not connected if there is $(\alpha_1, \alpha_2) \subset [0, 1]$ on which there is zero probability assigned and there is probability density on α_1. If this is the case, then any *do*-seller will deviate by moving the probability density on α_1 to any $\alpha \in (\alpha_1, \alpha_2)$. Such a deviation is strictly more profitable because doing so does not affect the probability of winning (if one wins at bidding α_1, he also wins at any $\alpha > \alpha_1$) and because $\alpha > \alpha_1$ is selling at a higher price.

Step 3: The upper bound of the support of $F(\cdot)$ is 1. The logic follows Step 2. Suppose the upper bound is $\alpha^* < 1$. Then, allocating the probability density at α^* to 1 is a profitable deviation: It does not affect the probability of winning and upon winning sells at a higher price.

Step 4: Deriving the c.d.f. $F(\cdot)$. Suppose all other *do*-sellers, when contacted, quote according to some same distribution $F(\cdot)$. Consider a specific seller called i. Quoting $R_d + \alpha\Delta_{hd}$, i gets to trade with the searching buyer if, and only if, such a quote is the best that the buyer receives. The buyer examines all quotes received. For each of the $n - 1$ contacts, with probability $1 - \pi_{do}$ the dealer is not a *do*-seller and in this case i's quote beats the no-quote. With probability π_{do}, the contacted investor is indeed another lo-seller, who quotes at α'. Then, only with probability $\mathbb{P}(\alpha < \alpha') = 1 - F(\alpha)$ will i's quote win. Taken together, for each of the $n-1$ potential competitor, i wins with probability $(1-\pi_{do}) + \pi_{do}(1-F(\alpha))$, and he needs to win all these $n-1$ times to capture the trading gain of $\alpha\Delta_{hd}$. That is, i expects a profit of $(1-\pi_{do}F(\alpha))^{n-1}\alpha\Delta_{hd}$. In particular, at the highest possible $\alpha = 1$, the above expected profit simplifies to $(1-\pi_{do})^{n-1}\Delta_{hd}$, because $F(1) = 1$. In a mixed-strategy equilibrium, i must be indifferent of quoting any values of α in the support. Equating the two expressions above and solving for $F(\cdot)$, one obtains the c.d.f. stated in the proposition. It can then be easily solved that the lower bound of the support must be at $(1-\pi_{do})^{n-1}$, where $F(\cdot)$ reaches zero. This completes the proof. □

Proposition 12

Proof. Consider a searching hn-buyer, for example. He contacts n dealers but knows that the number of counterparties he will actually find, N, is a random variable that follows a binomial distribution with n draws and success rate π_{do}. Each of these N counterparties then quotes a random price according to $F(\alpha; \pi_{do}, n)$, as stated in Proposition 11. The searching buyer chooses the lowest ask across the N available quotes. The c.d.f. of this minimum α is $1 - (1 - F(\alpha; \cdot))^{N-1}$ for $N \geq 1$. Since the probability of $N \geq 1$ is $(1 - (1-\pi_{do})^n)$, one obtains the the conditional c.d.f., as stated in the corollary. The same applies to a searching lo-seller. □

Corollary 1

Proof. Proposition 12 shows how the trading gains are split between one contacting customer and n *potential* counterparty dealers. Recall that with probability q, the customer is able to capture the full trading gain. Therefore, conditional on finding at least one dealer of her matching type, an hn-buyer expects a profit of $q\Delta_{hd} + (1-q)(R_h - (R_d + \bar{A}\Delta_{hd})) = (q + (1-q)(1-\bar{A}))\Delta_{hd}$, while an lo-seller expects $q\Delta_{dl} + (1-q)((R_d - \bar{A}\Delta_{hd}) - R_d) = (q + (1-q)(1-\bar{B}))\Delta_{dl}$. Substituting in $\bar{A}$ and $\bar{B}$ gives the stated $\gamma(\cdot)$ expression. □

Corollary 2

Proof. Since $n^{\text{BB}} = 1 < n^{\text{SMS}}$, below the notation n, without the superscript, indicates n^{SMS}. Proposition 12 gives the expression of $\bar{B}^k$ for $k \in \{\text{BB}, \text{SMS}\}$. In particular, $\bar{B}^{\text{BB}} = 1$, and for SMS, $\bar{B}^{\text{SMS}} = \frac{n \cdot (1-\pi_{do})\pi_{do}^{n-1}}{1-(1-\pi_{do})^k}$. Then $\bar{B}^{\text{SMS}}/\bar{B}^{\text{BB}} = \bar{B}^{\text{SMS}}$. By Lemma D.1, m_{do} is weakly increasing with s and hence so does π_{do}, thus proving the claim. To prove the claims regarding λ_d, note that from

Equation (47), cateris paribus, the left-hand side is monotone increasing in λ_d (the excess supply implies $v_{hn} > v_{lo}$; see Equation 51) but decreasing in s. Hence, all results about s hold for λ_d. □

Corollary 3

Proof. In equilibrium, the trading customers either have a strict preference for one of the technology or are indifferent. Consider *lo*-sellers, for example. If the preference is strict, then only one of the two HJBs in (40) is relevant; and if indifference, then the two HJBs reduce to the same one. The same holds for *hn*-buyers in their two HJBs (41). Likewise, the max[·] operator in Equations (40) and (41) can be dropped in equilibrium. Hence, defining $V_{lo} = \max_k[\{V^k_{lo}\}]$ and $V_{hn} = \max_k[\{V^k_{hn}\}]$, the HJB equations (38)-(43) can be reduced to the exactly the same set of (10)-(15) as if there is only one technology. Therefore, solving the same equation system, Proposition 1 holds. □

References

Ambrose, Brent W., Nianyun (Kelly) Cai, and Jean Helwege. 2008. "Forced Selling of Fallen Angels." *The Journal of Fixed Income* 18 (1):72–85.

An, Yu. 2020. "Competing with Inventory in Dealership Markets." Working paper.

Babus, Ana and Cecilia Parlatore. 2017. "Strategic Fragmented Markets." Working paper.

Bessembinder, Hendrik, William Maxwell, and Kumar Venkataraman. 2006. "Market Transparency, Liquidity Externalities, and Institutional Trading Costs in Corporate Bonds." *Journal of Financial Economics* 82 (2):251–288.

Bessembinder, Hendrik, Chester Spatt, and Kumar Venkataraman. 2020. "A Survey of the Microstructure of Fixed-Income Markets." *Journal of Financial and Quantitative Analysis* 55 (1):1–45.

Burdett, Kenneth and Kenneth L. Judd. 1983. "Equilibrium Price Dispersion." *Econometrica* 51 (4):955–969.

Butters, Gerard R. 1977. "Equilibrium Distributions of Sales and Advertising Prices." *The Review of Economic Studies* 44 (3):465–491.

Chang, Briana. 2018. "Adverse Selection and Liquidity Distortion." *The Review of Economic Studies* 85 (1):275–306.

Chowdhry, Bhagwan and Vikram Nanda. 1991. "Multimarket Trading and Market Liquidity." *Review of Financial Studies* 4 (3):483–511.

Duffie, Darrell, Piotr Dworczak, and Haoxiang Zhu. 2017. "Benchmarks in Search Markets." *The Journal of Finance* 72 (5):1983–2044.

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

Duffie, Darrell, Lei Qiao, and Yeneng Sun. 2019. “Continuous-time Random Matching.” Working paper.

Duffie, Darrell and Yeneng Sun. 2007. “Existence of Independent Random Matching.” *The Annals of Applied Probability* 17 (1):386–419.

———. 2012. “The exact law of large numbers for independent random matching.” *Journal of Economic Theory* 147:1105–1139.

Dugast, Jérôme, Semih Üslü, and Pierre-Olivier Weill. 2019. “A Theory of Participation in OTC and Centralized Markets.” Working paper.

Edwards, Amy K., Lawrence E. Harris, and Michael S. Piwowar. 2007. “Corporate Bond Market Transparency and Transaction Costs.” *The Journal of Finance* 62 (3):1421–1452.

Ellul, Andrew, Chotibhak Jotikasthria, and Christian T. Lundblad. 2011. “Regulatory pressure and fire sales in the corporate bond market.” *Journal of Financial Economics* 101 (3):596–620.

Fermanian, Jean-David, Olivier Guéant, and Jiang Pu. 2017. “The behavior of dealers and clients on the European corporate bond market: the case of Multi-Dealer-to-Client platforms.” Working paper. URL https://arxiv.org/abs/1511.07773.

Glode, Vincent and Christian C. Opp. 2019. “Over-the-Counter versus Limit-Order Markets: The Role of Traders’ Expertise.” *The Review of Financial Studies* Forthcoming.

Goldstein, Michael A., Edith S. Hotchkiss, and Erik R. Sirri. 2007. “Transparency and Liquidity: A Controlled Experiment on Corporate Bonds.” *Review of Financial Studies* 20 (2):235–274.

Guerrieri, Veronica, Robert Shimer, and Randall Wright. 2010. “Adverse Selection in Competitive Search Equilibrium.” *Econometrica* 78 (6):1823–1862.

Hau, Harald, Peter Hoffmann, Sam Langfield, and Yannick Timmer. 2017. “Discriminatory Pricing of Over-the-Counter Derivatives.” Working paper.

Hendershott, Terrence, Dan Li, Dmitry Livdan, and Norman Schürhoff. 2017. “Relationship Trading in OTC Markets.” Working paper.

———. 2020. “Fake Liquidity?” Working paper.

Hendershott, Terrence and Ananth Madhavan. 2015. “Click or Call? Auction versus Search in the Over-the-Counter Market.” *The Journal of Finance* 70 (1):419–447.

Hugonnier, Julien, Benjamin Lester, and Pierre-Olivier Weill. 2020. “Frictional Intermediation in Over-the-Counter Markets.” *The Review of Economics Studies* 87 (3):1432–1469.

Jovanovic, Boyan and Albert J. Menkveld. 2015. “Dispersion and Skewness of Bid Prices.” Working paper.

Klemperer, Paul. 1999. “Auction theory: A guide to the literature.” *Journal of Economic Surveys* 13 (3):227–286.

Lagos, Ricardo and Guillaume Rocheteau. 2009. "Liquidity in Asset Markets with Search Frictions." *Econometrica* 77 (2):403–426.

Lagos, Ricardo, Guillaume Rocheteau, and Pierre-Olivier Weill. 2011. "Crises and Liquidity in Over-the-Counter Markets." *Journal of Economic Theory* 146 (6):2169–2205.

Lee, Tommy and Chaojun Wang. 2019. "Why Trade Over-the-Counter? When Investors Want Price Discrimination." Working paper.

Lester, Benjamin, Guillaume Rocheteau, and Pierre-Olivier Weill. 2015. "Competing for Order Flow in OTC Markets." *Journal of Money, Credit and Banking* 47 (S2):77–126.

Lester, Benjamin, Ali Shourideh, Venky Venkateswaran, and Ariel Zetlin-Jones. 2018. "Market-making with Search and Information Frictions." Working paper.

Liu, Ying, Sebastian Vogel, and Yuan Zhang. 2017. "Electronic Trading in OTC Markets vs. Centralized Exchange." Working paper.

Nash, John F. 1950. "The Bargaining Problem." *Econometrica* 18 (2):155–162.

O'Hara, Maureen and Xing Zhou. 2020. "The Electronic Evolution of Corporate Bond Dealers." *Journal of Financial Economics* Forthcoming.

Pagano, Marco. 1989. "Trading Volume and Asset Liquidity." *The Quarterly Journal of Economics* 104 (2):255–274.

Riggs, Lynn, Esen Onur, David Reiffen, and Haoxiang Zhu. 2019. "Swap Trading after Dodd-Frank: Evidence from Index CDS." *Journal of Financial Economics* Forthcoming.

Saar, Gideon, Jian Sun, Ron Yang, and Haoxiang Zhu. 2019. "From Market Making to Matchmaking: Does Bank Regulation Harm Market Liquidity?" Working paper.

Shi, Shouyong. 2019. "Sequentially Mixed Search and Equilibrium Price Dispersion." Working paper.

Sun, Yeneng. 2006. "The exact law of large numbers via Fubini extension and characterization of insurable risks." *Journal of Economic Theory* 126:31–69.

Üslü, Semih. 2019. "Pricing and Liquidity in Decentralized Asset Markets." *Econometrica* Forthcoming.

Varian, Hal R. 1980. "A Model of Sales." *American Economic Review* 70 (4):651–659.

Vayanos, Dimitri and Pierre-Olivier Weill. 2008. "A Search-Based Theory of the On-the-Run Phenomenon." *The Journal of Finance* 63 (3):1361–1398.

Vogel, Sebastian. 2019. "When to Introduce Electronic Trading Platforms in Over-the-Counter Markets?" Working paper.

Wang, Chaojun. 2017. "Core-Periphery Trading Networks." Working paper.

Weill, Pierre-Olivier. 2007. "Leaning against the Wind." *Review of Economic Studies* 74:1329–1354.

Wright, Randall, Philipp Kircher, Benoît Julien, and Veronica Guerrieri. 2020. "Directed Search and Competitive Search: A Guided Tour." *Journal of Economic Literature* Forthcoming.

Yueshen, Bart Zhou. 2017. "Uncertain Market Making." Working paper.

Zhu, Haoxiang. 2012. "Finding a Good Price in Opaque Over-the-Counter Markets." *The Review of Financial Studies* 25 (4):1255–1285.

List of Figures