Inter-Dealer Trades in OTC Markets – Who Buys and Who Sells?

Chung-Yi Tse and Yujing Xu*

June 28, 2018

Abstract

Oftentimes, dealers in an OTC market may not be able to trade with one another whenever they desire to do so, just as investors find it necessary to incur time and effort to buy and sell assets not traded in a centralized exchange. Moreover, an individual dealer can obviously only carry limited quantities of the asset over time and the inventory capacities among dealers may certainly be different. In this environment, dealers trade among themselves, whenever the opportunities arise, to rebalance inventories for facilitating the sale and purchase of the asset to and from investors. In equilibrium, the small-capacity dealers sell to the large-capacity dealers when the asset supply is at a low level but buy from them when the asset supply is at a high level. It is the small-capacity dealers who trade to provide immediacy for the large-capacity dealers – a prediction, though counterintuitive, is supported by some available empirical evidence.

Keywords: OTC Market, Inter-Dealer Trades, Dealers' Inventories
JEL classifications: D53, D85, G23

*Faculty of Business and Economics, University of Hong Kong. E-mail: tsechung@econ.hku.hk (Tse); yujingxu@hku.hk (Xu). We would like to thank Charles Leung, Dongkyu Chang, Kim-Sau Chung, Pei-yu Melody Lo, the seminar audiences of the City University of Hong Kong, Hong Kong Baptist University and the audiences of the 2017 Econometric Society Asian meeting for very helpful comments and suggestions. Xu gratefully acknowledges financial support from HK GRF grant 17517816.

1 Introduction

Many financial assets, including government and corporate bonds, asset-backed securities, and derivatives, are traded in over-the-counter (OTC) markets instead of in centralized exchanges.[1] Two distinguishing features of OTC markets are that trades are almost always intermediated by dealers of various kinds and that the dealers do not just trade with investors but also among themselves. Indeed, inter-dealer trades can account for a significant fraction of the overall transactions for a given asset.[2]

An important question on inter-dealer trades in OTC markets that seems to have attracted scant attention is how market conditions help shape the directions of trade among heterogeneous dealers.

Numerous studies have documented and modeled how dealers in OTC markets on the whole gradually raise their inventories of assets in the early to mid 2000s but begin to wind down their inventory holdings in the down market that followed the 2008 financial crisis.[3] Behind this overall trend, however, individual dealers seem to have responded differently to the changing market conditions. Adrian, Fleming, Schachar and Vogt (2017), in particular, find that while large dealers tend to expand their balance sheets much more than small dealers do pre-crisis, small dealers actually expand their balance sheets during which large dealers are downsizing their inventory holdings post-crisis. An apparent explanation for the large dealers' responses is that large dealers are overwhelmingly dealers affected by the implementation of the Volcker rule.[4] The reason for the balance sheet expansion by small dealers post-crisis, however, is less clear. This suggests that the changes in dealers' asset holdings cannot be solely due to regulatory changes. The differential responses by large and small dealers could instead be the result of the sales of assets by the former to the latter to a certain extent as natural responses to the changing market environment.

In this paper, we extend the seminal random search models of the OTC market of Duffie, Gârleanu and Pedersen (2005) and Lagos and Rocheteau (2009) to study how dealers trade with one another for managing inventory levels for trading with their customers and how the directions of trade among dealers is mainly determined by the relative asset supply in the market. The point of departure is that, in our model, (1) dealers have only imperfect access to trading with other dealers and (2) they are heterogeneous in their inventory capacities. These are arguably very plausible assumptions. First, to be sure, in reality, there is not a frictionless platform on which dealers can continuously trade among themselves in a typical OTC market, just as investors must expend time and effort in buying and selling the asset. The heterogeneity in inventory capacity among dealers can result from differences in financing costs

[1] As an example, the gross market value of global OTC derivatives totaled 38,286 billion US dollars in 2014 and 29,992 billion US dollars in 2015 (Bank for International Settlement, Semiannual OTC derivatives statistics, updated on May 4, 2016).

[2] Li and Schürhoff (2014) show that in the period covered by their data set, 16 million out of 60 million transactions in municipal bonds are inter-dealer trades. A similar percentage of inter-dealer trade is also documented in Hollifield, Neklyudov and Spatt (2016).

[3] Di Maggio, Kermani and Song (2017) and Randall (2015), among others.

[4] The so-called Volcker rule, enacted in 2010 and gradually becoming binding in the few years afterward, prohibits banking entities with access to discount window lendings and FDIC deposit insurance from engaging in proprietary trading. See Bao, O'Hara and Zhou (2016) for example.

– dealers who finance asset purchases out of retained earnings and owners' equities can face different opportunity costs of funds, whereas dealers who finance asset purchases by borrowing can be charged different risk premia. The heterogeneity can also be due to risk management considerations or portfolio choices.[5]

With imperfect access to inter-dealer trading, it becomes imperative for dealers to choose the appropriate levels of inventory holding to be able to meet the uncertain future buy and sell orders from their customers. In our model, dealers possessing different inventory capacities attain their respective optimal inventory holdings by buying and selling among themselves when the opportunities come, and that dealers of different inventory capacities play different roles in the inter-dealer market.

In particular, in our model, there is a given measure of what we call small dealers, each endowed with one unit of inventory capacity, and a given measure of what we call large dealers, each endowed with two units of inventory capacity. At the beginning of each period, investors who value the asset highly but have no asset in hand (high-valuation non-owners hereinafter) and investors who own a unit of the asset but do not value it (low-valuation owners hereinafter) enter the market to buy from and sell to dealers. Investors and dealers randomly meet in this investor-dealer market in which a given dealer can only sell to (buy from) an investor if the dealer has at least one unit of inventory (spare capacity) beforehand. Once the investor-dealer trades are completed, and only then, a perfectly competitive inter-dealer market opens, through which dealers can rebalance their inventory holdings.

Underlying most of the results in the paper is a particular ranking of the marginal benefits of inventory according to which both the first unit of inventory and the last unit of spare inventory capacity are valued higher by a large dealer than by a small dealer. The ranking is due to how the small, but not the large, dealer would exhaust his entire inventory capacity in acquiring one unit of the asset and how the large, but not the small, dealer need not fill up his entire inventory to already possess a unit of the asset for sale to investors.

The ranking implies that in equilibrium, all small dealers either tend to sell or buy in the inter-dealer market. Large dealers, on the other hand, sell as well as buy in any equilibrium. This means that small dealers should mostly trade with large dealers but not among themselves while trades between the two types of dealers tend to flow in just one direction. The direction of trade between a given pair of small and large dealers is then persistent, a prediction consistent with the findings in Li and Schürhoff (2014).

Which direction the given trade tends to take turns out to depend on the asset supply in a rather surprising manner. If large dealers are able to and indeed tend to hold a larger inventory of the asset, perhaps it seems intuitive that they should sell to small dealers in equilibrium. We find that this is the case, however, only when the asset supply is relatively abundant, at which times small dealers should find it easiest to buy the asset themselves from investors. When the asset supply is relatively meagre, at which times large dealers should find it hardest to buy the asset themselves from investors, it is the small dealers who sell to provide inventory for the large dealers. Such trading directions coincide precisely with what the findings in Adrian et al. (2017) suggest if we interpret a small asset supply in our model as times of a booming market in which the asset demand is strong relative to the asset supply and a large asset supply in

[5] In this paper, we do not attempt to model how the heterogeneity arises endogenously but instead restrict attention to exploring the consequences thereof.

our model as times of a market bust in which the asset demand is weak relative to the asset supply. In the former case, as in the up market before the financial crisis, large dealers amass inventory by buying from small dealers. In the latter case, as in the down market after the financial crisis, small dealers gain inventory by buying from large dealers.

A dealer is said to provide immediacy to another dealer if the first dealer tends to sell to (buy from) the second dealer when it takes longest on average for the second dealer to buy (sell) the asset in the market as arising from, for example, a relatively meagre (abundant) asset supply. Apparently, the large dealers in our model do not provide immediacy for small dealers. Rather, it is the small dealers who provide immediacy for the large dealers by selling to the latter when the asset supply is at a relatively low level but buying from them when the asset supply is at a relatively high level.

It is well known that in many OTC markets, as documented in Li and Schürhoff (2014) and Hollifield, Neklyudov and Spatt (2016), a set of dealers, known as the core dealers, tend to trade with all dealers in the market, whereas the rest, known as the peripheral dealers, only trade with the core dealers. Given that small dealers in our model on the whole trade with the large dealers, rather than among themselves, whereas the large dealers trade with all dealers, the two types of dealers behave similarly as the peripheral and core dealers do, respectively, identified in the empirical studies, with regard to the set of dealers they are predicted to trade with. Under this interpretation then, our model predicts that it is the small peripheral dealers who provide immediacy for the large core dealers. The prediction seems counterintuitive. But there indeed exists empirical evidence supporting it as we shall discuss in the following.

The equilibrium in our model is constrained efficient in that the allocation of inventories and spare capacities among dealers falling out from inter-dealer trades in equilibrium coincide with the planning optimum. We show that rather intuitively, the optimum allocations serve to enable investors to buy and sell the asset most rapidly. The equilibrium allocations, on the other hand, are such that inventories and spare capacities are held by dealers who value them the most for their trading needs with investors. That the two allocations coincide perhaps is not surprising but more interestingly, it suggests that for efficiency, small peripheral dealers indeed should trade to provide immediacy for the large core dealers.

In addition to implications on trading directions among dealers, a further novel result in our model is that the inter-dealer trading volume is "M-shaped" in response to changes in the asset supply – trading is most active when the asset supply is at a moderately low, but not the lowest, level and at a moderately high, but not the highest, level. Dealers trade among themselves to rebalance inventory, to which the need is greatest when either they find it hardest to acquire inventory or liquidity from investors, i.e., when the asset supply is at the lowest or the highest level. But precisely when the asset supply is at the lowest or the highest level, dealers who possess inventory (spare capacity) to sell (buy) can only be few and far between. In equilibrium, prices must then rise (fall) to dampen the demand (supply). In this way, trading is most active when the demand for and the supply of inventory are both at relatively high levels, arising from there being a moderately high or low asset supply. The inter-dealer trading volume is also non-monotonic with respect to the fraction of large dealers in the dealer population, reaching the maximum level when the fraction is at some intermediate level, whereby the dealer population is most diverse in terms of inventory capacity.

Related Literature

It has long been recognized, going back to Ho and Stoll (1983), that dealers may trade among themselves for inventory risk concerns. In these models and their modern variants, such as Atkeson, Eiseﬂdt and Weill (2015) and Üslü (2016), a risk-averse trader having a greater exposure to some risky assets sells a certain fraction of his holding to another risk-averse trader with a lesser initial exposure to the mutual benefits of both. The dealers in our model also trade to mitigate inventory risks, broadly understood. By assuming that dealers are risk neutral, we emphasize that inventory management concerns can also arise from how dealers would like to conduct business with their customers while maintaining the best overall levels of inventory and spare inventory capacity. Why dealers trade among themselves, other than for inventory risk sharing, is a topic of active ongoing research. Colliard and Demange (2017) study post-issuance intermediation chains where each dealer has limited cash endowment, so that they need to trade with one another to amass the cash endowment of a group of dealers. Glode and Opp (2016) argue that when there is an adverse selection problem, a longer intermediation chain can narrow the information gap between two successive dealers and help mitigate the problem of information asymmetry. Hugonnier, Lester and Weill (2016) and Shen, Wei and Yan (2016) assume that investors value the asset differently and show that those with intermediate valuations endogenously become dealers as they stay in the market to sell to investors with even higher valuations after buying.[6]

In addition to offering an explanation for inter-dealer trades, the primary contribution of our paper is an investigation of how the relative asset supply helps determine the direction of trades among different types of dealers and the implications thereof on who provide immediacy for whom in a simple and parsimonious model. A secondary contribution is a novel implication on how the volume of inter-dealer trades should vary with the asset supply non-monotonically. In other models of inter-dealer trades, it is not clear how a change in the asset supply may have any clear-cut implications on trading directions and how the asset supply-trading volume relation can be non-monotonic.

Our framework is adapted from the seminal models of OTC markets in Duffie, Gârleanu and Pedersen (2005) and Lagos and Rocheteau (2009), with two main differences being dealers' imperfect access to inter-dealer trading and the heterogeneity of dealers' inventory capacity – two features that make the present model more suitable for studying the inter-dealer trading relationship. In the two aforementioned papers, whenever a dealer trades with an investor, the dealer can instantaneously offset the transaction by trading in a perfectly competitive inter-dealer market that opens at all times. Such an environment, in which a dealer trades with another dealer only if and when he meets an investor, is arguably not the best environment to study inter-dealer trades as the trades have neither persistent direction nor particular structure.

A host of recent papers are motivated to explain the empirical finding that the inter-dealer market exhibits a core-periphery structure. For instance, Neklyudov (2015) proposes a random search model assuming that dealers differ in their search abilities. He shows that the dealers with higher search abilities are more interconnected and hence are in the center. We highlight another factor, inventory capacity, that can influence a dealer's position in inter-dealer trading relationships.

[6] A similar mechanism is proposed in Piazzesi and Schneider (2009) in their analysis of the housing market.

Another strand of investigation, a notable example of which is Wang (2017), models explicitly the formation of a core-periphery trading network in an environment in which a dealer can only trade with another dealer with which it has previously paid to set up an account. These models study how most dealers would choose to set up accounts with just a handful of dealers due to the usual network externality. In our model, the trading structure that emerges endogenously resembles a core-periphery trading network without any kind of network externality and where the establishment of a prior link is not needed for any two dealers to trade.

Ours is not the only model of an OTC market in which dealers hold inventory. Lagos, Rocheteau and Weill (2011) demonstrate that dealers hold inventory to speed up future trades when there is a negative shock knocking the market off the steady state, even if dealers do not value the asset. Weill (2011) shows that the same intuition also applies in a competitive dynamic market with a transient selling pressure. Dealers in our model hold inventory also to facilitate trade, as they do not have continuous access to the inter-dealer market. The difference is that they hold inventory even in the steady state and they gain by trading among themselves.

For brevity, we restrict attention to studying steady-state equilibrium in this paper. The rest of the paper is organized as follows. In Section 2, we set up the model and then study the model's equilibrium. We discuss the model's implications on trading directions between small and large dealers in Section 3 and compare those implications against the available empirical evidence. In Section 4, we discuss three extensions of the model and demonstrate how the major results hold in more general settings. We return to our basic model in Section 5 to study the model's comparative statics. Section 6 concludes with discussions on the constrained efficiency of equilibrium in particular. All proofs are relegated to the Appendix, which also includes three respective Sections for the details of one extension of the model, two Propositions on the comparative statics of prices, and the formal analysis of the constrained efficiency of equilibrium.

2 Model and Analysis

2.1 Basic Environment

Time is discrete and runs forever. Two groups of agents – investors and dealers – buy and sell an asset with supply fixed at A in an OTC market. A high-valuation investor derives a per period return of $\upsilon > 0$ in holding a unit of the asset, whereas low-valuation investors and dealers derive the same per period return normalized to zero. An individual investor can hold either zero or one unit of the asset at a time and can only buy or sell the asset through dealers of which there are two types: (1) small dealers, each of whom can hold up to one unit of the asset at a time and (2) large dealers, each of whom can hold up to two units. All agents are risk neutral and discount the future at the same factor β.

At the beginning of each period, a measure of e investors enter the market as high-valuation investors with no assets in hand. Together with the entrants in previous periods who have yet to acquire a unit of the asset, they – the high-valuation non-owners – constitute the population of investor-buyers in the market. Among investors who do own a unit of the asset, the low-valuation owners become the investor-sellers in equilibrium.

Each period is divided into two subperiods. In the first subperiod, a decentralized investor-dealer market opens in which the bilateral meetings between investors and dealers take place. We assume that investor-buyers and dealers with at least a unit of the asset for sale (dealer-sellers hereinafter) meet in one segment of the market and investor-sellers and dealers with spare inventory capacity to buy (dealer-buyers hereinafter) meet in another segment of the market. The matches in each market segment are formed in accordance with the same Mortensen-Pissarides constant-returns matching function, whereby, given market tightness $\theta \in [0, \infty)$ for the ratio of the measures of buyers to sellers for the given market segment, a seller meets a randomly selected buyer at a probability $\eta(\theta) \in [0, 1]$, while a buyer meets a randomly selected seller at the probability $\mu(\theta) = \eta(\theta)/\theta$. The meeting probability $\eta(\theta)$ satisfies the usual conditions:

$$\frac{\partial \eta}{\partial \theta} > 0; \quad \frac{\partial^2 \eta}{\partial \theta^2} < 0; \quad \lim_{\theta \to 0} \frac{\partial \eta}{\partial \theta} = 1; \quad \lim_{\theta \to \infty} \frac{\partial \eta}{\partial \theta} = 0.$$

With two market segments, there are two market tightness: (1) θ_{ID} for the ratio of the measures of investor-buyers to dealer-sellers and (2) θ_{DI} for the ratio of the measures of dealer-buyers to investor-sellers.

Prices in the investor-dealer market fall out of the bargaining between the buyers and sellers in the bilateral meetings in which the agents on the two sides are assumed to possess equal bargaining power.[7] An individual dealer may search as both a dealer-buyer and a dealer-seller in a given period but the meeting technology only allows the dealer to meet at most one investor-buyer and one investor-seller in the period.[8] At the end of the subperiod, those high-valuation non-owners who succeed in buying a unit turn into new high-valuation owners, while those low-valuation owners who succeed in selling their units leave the market for good.

In the second subperiod, a competitive inter-dealer market opens, in which dealers buy and sell as many units of the asset among themselves as they see fit at a given market price, subject to their asset holdings and spare inventory capacities. Finally, at the end of the second subperiod, each high-valuation owner, except for those who have just purchased the asset in the current period, turns into a low-valuation owner at a probability $\delta \in (0, 1)$.

An apparent alternative to our assumed meeting technology is that investors and dealers search and match in one unified market in which an investor meets a randomly selected dealer at a probability that depends on the ratio of all dealers to all investors. In this setup, there would be bilateral meetings between an investor-seller and a dealer-seller and between an investor-buyer and a dealer-buyer that cannot lead to any profitable exchanges between the agents concerned. Such no-trade meetings should not be common occurrences in reality. Hendershott and Madhavan (2015) report the increasing prevalence of electronic trading platforms for corporate bonds on which investors post their buy and sell orders. A dealer in these markets is then usually well informed of whether an investor is buying or selling before he initiates

[7] The assumption of equal bargaining power is without loss of generality and merely serves to simplify.

[8] This simplifying assumption can be understood as a discrete-time version of a continuous-time meeting process. If the time interval between two periods is small enough relative to the arrival rate of a meeting, then the probability of having more than one meeting per period approaches zero. The assumption can also be justified by a dealer's limited execution capacity in reality. We should explain in Section 4 how the major results of the analysis survive while relaxing the restriction.

contact with the investor. Our assumed meeting technology embodies the standard assumption that how many matches of one type are formed depends only on the measures of agents who may become partners in such matches, whereas in the unified market setup, the measures of other types of agents also matter and exert negative effects on the measures of the matches of the given type that would be formed, which means that for instance, an investor-buyer's matching probability decreases with the measures of dealer-buyers and investor-sellers. Our assumed meeting technology has the virtue of choosing a simple versus a complicated setting when there is no compelling reason for choosing the latter.

In the models of Duffie, Gârleanu and Pedersen (2005) and Lagos and Rocheteau (2009), dealers do not hold any inventory at all in the steady state, as they can immediately offset any transaction with investors in an inter-dealer market that opens at all times. In contrast, the dealers in our model access the inter-dealer market only after or before they meet and trade with investors. Without continuous access to trading with other dealers, a dealer in our model can sell to an investor only if the dealer is holding at least a unit of the asset beforehand and thus the dealer may find it optimal not to offload all units of the asset he acquires from investors at the first opportunity. In a similar vein, a dealer may find it optimal not to entirely fill up his inventory in the inter-dealer market, in anticipation of using the spare capacity for trading with an investor-seller if and when he meets one in the next period.

2.2 Value Functions

A small dealer, S_i, $i = 0, 1$, is either holding 0 or 1 unit of the asset at the beginning of a period when the investor-dealer market opens, whereas a large dealer, L_i, $i = 0, 1, 2$, may also be holding up to 2 units of the asset.

In the investor-dealer market, an investor-buyer meets a dealer-seller at probability $\mu(\theta_{ID})$. The dealer-seller can be an S_1, an L_1 or an L_2. Let p_{I_B,S_1}, p_{I_B,L_1}, and p_{I_B,L_2} be the respective prices at which the investor-buyer buys from these different dealers. Then, the investor-buyer has asset value,

$$U^B = \mu(\theta_{ID}) \left(\beta U_H^{ON} - \frac{n_1^{SD}}{n_S^D} p_{I_B,S_1} - \frac{n_1^{LD}}{n_S^D} p_{I_B,L_1} - \frac{n_2^{LD}}{n_S^D} p_{I_B,L_2} \right) + (1 - \mu(\theta_{ID})) \beta U^B, \quad (1)$$

where n_i^{SD} and n_i^{LD} are the respective measures of small and large dealers holding an i-unit inventory,

$$n_S^D = n_1^{SD} + n_1^{LD} + n_2^{LD}$$

the measure of all dealer-sellers, and U_H^{ON} the asset value of a high-valuation owner. In defining this value function and the ones that follow, we assume that all meetings in the investor-dealer market yield a non-negative match surplus. We show in Lemma 1 below how the assumption holds true in any equilibrium with active trading between investors and dealers.

A high-valuation owner derives a per period return υ from holding a unit of the asset and may turn into a low-valuation owner at probability δ at the end of the period. Hence,

$$U_H^{ON} = \upsilon + \beta \left(\delta U_L^{ON} + (1 - \delta) U_H^{ON} \right), \quad (2)$$

where U_L^{ON} denotes the asset value of a low-valuation owner who seeks to sell his unit of the asset. In each period, the investor-seller meets a dealer-buyer at probability $\eta(\theta_{DI})$. The dealer may be an S_0, an L_0 or an L_1.[9] Let p_{S_0,I_S}, p_{L_0,I_S}, and p_{L_1,I_S} be the respective prices at which the low-valuation investor sells to these different dealers. Hence,

$$U_L^{ON} = \eta(\theta_{DI}) \left(\frac{n_0^{SD}}{n_B^D} p_{S_0,I_S} + \frac{n_0^{LD}}{n_B^D} p_{L_0,I_S} + \frac{n_1^{LD}}{n_B^D} p_{L_1,I_S} \right) + (1 - \eta(\theta_{DI})) \beta U_L^{ON}, \tag{3}$$

where

$$n_B^D = n_0^{SD} + n_0^{LD} + n_1^{LD}$$

is the measure of all dealer-buyers.

In addition to trading with investors in the investor-dealer market in the first subperiod, dealers may also trade among themselves in the second subperiod in the competitive inter-dealer market. Write V_i^{SD} and W_i^{SD}, $i = 0, 1$, as the respective asset values of a small dealer entering the investor-dealer market in the first subperiod and the inter-dealer market in the second subperiod with an i-unit inventory. If the asset is traded in the inter-dealer market at price p,

$$W_0^{SD} = \max\left\{\beta V_0^{SD}, \beta V_1^{SD} - p\right\}, \tag{4}$$

$$W_1^{SD} = \max\left\{p + \beta V_0^{SD}, \beta V_1^{SD}\right\}, \tag{5}$$

$$V_0^{SD} = \mu(\theta_{DI}) \left(W_1^{SD} - p_{S_0,I_S}\right) + (1 - \mu(\theta_{DI})) W_0^{SD}, \tag{6}$$

$$V_1^{SD} = \eta(\theta_{ID}) \left(p_{I_B,S_1} + W_0^{SD}\right) + (1 - \eta(\theta_{ID})) W_1^{SD}. \tag{7}$$

In (4), an S_0 entering the inter-dealer market chooses between buying a unit in the market and not buying, whereas in (5), an S_1 chooses between selling the unit and not selling. Clearly, if the first dealer strictly prefers to buy where $p < \beta(V_1^{SD} - V_0^{SD})$, the second dealer must strictly prefer not to sell and vice versa. In (6), an S_0 entering the investor-dealer market meets an investor-seller at probability $\mu(\theta_{DI})$ and buys the unit from the investor at price p_{S_0,I_S}. In (7), an S_1 meets an investor-buyer at probability $\eta(\theta_{ID})$ and sells the unit to the investor at price p_{I_B,S_1}.

A large dealer can hold up to two units of the asset in inventory. The asset values, V_i^{LD} and W_i^{LD}, $i = 0, 1, 2$, satisfy, respectively,

$$W_0^{LD} = \max\left\{\beta V_0^{LD}, \beta V_1^{LD} - p, \beta V_2^{LD} - 2p\right\}, \tag{8}$$

$$W_1^{LD} = \max\left\{p + \beta V_0^{LD}, \beta V_1^{LD}, \beta V_2^{LD} - p\right\}, \tag{9}$$

$$W_2^{LD} = \max\left\{2p + \beta V_0^{LD}, p + \beta V_1^{LD}, \beta V_2^{LD}\right\}, \tag{10}$$

$$V_0^{LD} = \mu(\theta_{DI}) \left(W_1^{LD} - p_{L_0,I_S}\right) + (1 - \mu(\theta_{DI})) W_0^{LD}, \tag{11}$$

$$\begin{aligned} V_1^{LD} = \; & \mu(\theta_{DI}) (1 - \eta(\theta_{ID})) \left(W_2^{LD} - p_{L_1,I_S}\right) \\ & + (1 - \mu(\theta_{DI})) \eta(\theta_{ID}) \left(p_{I_B,L_1} + W_0^{LD}\right) \\ & + \mu(\theta_{DI}) \eta(\theta_{ID}) \left(p_{I_B,L_1} - p_{L_1,I_S} + W_1^{LD}\right) \\ & + (1 - \mu(\theta_{DI})) (1 - \eta(\theta_{ID})) W_1^{LD}, \end{aligned} \tag{12}$$

[9] In holding a unit in inventory and having one unit of spare inventory capacity, an L_1 is both a dealer-seller and a dealer-buyer in the given period.

$$V_2^{LD} = \eta\left(\theta_{ID}\right)\left(p_{I_B,L_2} + W_1^{LD}\right) + \left(1 - \eta\left(\theta_{ID}\right)\right) W_2^{LD}. \tag{13}$$

The prices for the investor-dealer trades in the above are assumed to be determined by standard Nash bargaining in which the two agents of a given trade possess equal bargaining power.[10]

2.3 Prices and Match Surpluses in the Investor-dealer Market

In defining the value functions in (1)-(3), (6) and (7), and (11)-(13), we assume that there are non-negative match surpluses in any and all meetings in the investor-dealer market. A priori this need not be true as it is not inconceivable that a given type of dealer holding a particular inventory may only find it profitable to trade with other dealers but not with investors. We verify in Lemma 1 below how our assumption holds true in any steady-state equilibrium with active trading.

Lemma 1 *In any steady-state equilibrium, the match surplus for meetings between an investor-seller and any dealer-buyer all equals to*

$$z_{I_S} = p - \beta U_L^{ON}, \tag{14}$$

while any such exchanges take place at the same price,

$$p_{S_0,I_S} = p_{L_0,I_S} = p_{L_1,I_S} = \frac{p + \beta U_L^{ON}}{2}. \tag{15}$$

The match surplus for meetings between an investor-buyer and any dealer-seller all equals to

$$z_{I_B} = \beta\left(U_H^{ON} - U^B\right) - p, \tag{16}$$

while any such exchanges take place at the same price,

$$p_{I_B,S_1} = p_{I_B,L_1} = p_{I_B,L_2} = \frac{p + \beta\left(U_H^{ON} - U^B\right)}{2}. \tag{17}$$

In a narrow sense, the Lemma holds because of a competitive inter-dealer market in which dealers buy and sell the asset at the same price p. In this case, if a dealer buying a unit from an investor finds it optimal to sell it in the inter-dealer market right after, the dealer earns a surplus equal to p minus the payment he makes to the investor. On the other hand, if the dealer finds it optimal to keep the unit for selling to other investors later on, he earns just the same surplus, as the unit would have cost him p in the inter-dealer market had he not bought it earlier from the investor. On the other side of the trade, to the investor-seller, the gain from trade is equal to the payment from the dealer net of the continuation value of being a low-valuation owner, which is otherwise independent of the identity of the counterparty of the trade. The match surplus, equal to the sum of the surpluses from trade of the two sides, is then the same across all trades between an investor-seller and any dealer-buyer as given in (14). A similar logic explains (16). In a broader sense, as we shall show in Lemma 5 when

[10] See (51)-(56) in the proof of Lemma 1 in the Appendix for the pricing equations.

we extend the analysis to allow for a frictional inter-dealer market, that any trade between a dealer and an investor-seller should yield a non-negative surplus is that had the dealer found it not optimal to buy from an investor, the dealer could only be buying the asset from another dealer for him to be an active agent in equilibrium. But then there cannot be any greater surplus for a unit to be acquired by one dealer first and then passed onto another dealer than for the unit to be bought by the latter dealer in the first instance.

2.4 Inter-dealer Market Trades

By (4) and (5), whether a small dealer entering the inter-dealer market wants to buy or sell depends on how the inter-dealer market price p compares with $\beta\left(V_1^{SD} - V_0^{SD}\right)$. Similarly, by (8)-(10), a large dealer entering the market decides to buy or sell by comparing p against $\beta\left(V_1^{LD} - V_0^{LD}\right)$ and $\beta\left(V_2^{LD} - V_1^{LD}\right)$. To proceed, we first establish that:

Proposition 1 $V_1^{LD} - V_0^{LD} \geq V_1^{SD} - V_0^{SD} \geq V_2^{LD} - V_1^{LD}$ *in any active steady-state equilibrium in which* $z_{I_S} \geq 0$ *and* $z_{I_B} \geq 0$*. The first inequality is strict if* $z_{I_S} > 0$*, whereas the second inequality is strict if* $z_{I_B} > 0$*.*

Proposition 1 says that in an active steady-state equilibrium, an L_0 has the most to gain from acquiring a unit of the asset in the inter-dealer market, followed by an S_0, whereas an L_1 has the least to gain. Intuitively, $V_1^{LD} - V_0^{LD} \geq V_1^{SD} - V_0^{SD}$ because the opportunity cost for the large dealer in utilizing his first unit of inventory capacity should be lower than the opportunity cost for the small dealer in utilizing his only unit of inventory capacity – in acquiring a unit in the inter-dealer market, the large dealer, but not the small dealer, still has spare inventory capacity to buy one more unit from an investor in the next period to capture any possible surplus of trade. If the latter surplus is strictly positive ($z_{I_S} > 0$), then a large dealer gains strictly more from the first unit of inventory than a small dealer does. When acquiring a unit in the inter-dealer market is at the expense of exhausting one's inventory capacity for both the large and small dealers, however, the small dealer should have more to gain than the large dealer $\left(V_1^{SD} - V_0^{SD} \geq V_2^{LD} - V_1^{LD}\right)$ since the large dealer holding a one-unit inventory already has a unit for sale to investors in the upcoming period, whereas the small dealer does not. If there is a strictly positive surplus in an investor-buyer-dealer-seller trade ($z_{I_B} > 0$), the small dealer gains strictly more from the last unit of inventory than the large dealer does.

The ranking in Proposition 1 implies that L_0s at least weakly prefer to buy and L_2s at least weakly prefer to sell in equilibrium. Who else will buy and sell depends on what price clears the inter-dealer market, a price that must be bounded by

$$p \in \left[\beta\left(V_2^{LD} - V_1^{LD}\right), \beta\left(V_1^{LD} - V_0^{LD}\right)\right],$$

since at any p above the upper bound of the interval, there can only be sellers and at any p below the lower bound, there can only be buyers in the market. Besides, for any p not exactly equal to β times one of the three marginal benefits in Proposition 1, any and all dealers who desire to trade either *strictly* prefer to buy or sell. In this case, the market clears only if the parameters conspire to just equate the measures of buyers and sellers. Such a parameter configuration, however, can only make up a zero-measure subset of the parameter space, the formal proof of which is postponed to Section 2.6. Equilibrium obtains in general only for p just equal to $\beta\left(V_1^{LD}-V_0^{LD}\right)$, $\beta\left(V_1^{SD}-V_0^{SD}\right)$, or $\beta\left(V_2^{LD}-V_1^{LD}\right)$, at which there is one type of dealer holding a given inventory indifferent between selling and not selling or between buying and not buying. The market may then clear at some particular mixing probability for the mixed strategy played by the marginal buyers or sellers. Furthermore, we can show that:

Lemma 2 *For $p=\beta\left(V_1^{LD}-V_0^{LD}\right)$ or $\beta\left(V_1^{SD}-V_0^{SD}\right)$, both z_{I_S} and z_{I_B}, and p itself are strictly positive, whereas for $p=\beta\left(V_2^{LD}-V_1^{LD}\right)$, z_{I_S} and p itself are equal to zero while $z_{I_B}>0$. In all cases, the candidate equilibria are active equilibria in which the gains from trade between investors and dealers are non-negative.*

The case for $p=\beta\left(V_2^{LD}-V_1^{LD}\right)$ deserves further explanation. An L_1 in the inter-dealer market choosing to buy a unit forgoes the opportunity to buy from an investor in the next period as he exhausts his entire inventory capacity in doing so. Meanwhile, given that he already possesses a unit in inventory to begin with, he can sell to an investor in the next period with or without buying a unit in the inter-dealer market. The dealer must then be worse off acquiring the unit at any positive p and may only be indifferent in doing so at $p=0$.[11] With $p=0$, a dealer is willing to buy a unit from an investor also only at a zero price, which means that there must be but a zero surplus in an investor-seller-dealer-buyer trade ($z_{I_S}=0$).[12] On the other hand, there must be a positive surplus in an investor-buyer-dealer-seller trade ($z_{I_B}>0$) since the investor, but not the dealer, is strictly better off owning a unit than not owning.

It is useful to classify equilibrium into three types, corresponding to p equal to each candidate equilibrium price.

The "Selling" Equilibrium In the Selling Equilibrium, $p=\beta\left(V_1^{LD}-V_0^{LD}\right)$. By Proposition 1 and Lemma 2,

$$p=\beta\left(V_1^{LD}-V_0^{LD}\right)>\beta\left(V_1^{SD}-V_0^{SD}\right)>\beta\left(V_2^{LD}-V_1^{LD}\right),$$

in which case no dealers strictly prefer to buy with p anchored at the highest possible marginal value of inventory. In the meantime, any dealers, large and small, with a filled inventory strictly prefer to sell. For this reason, we call this the Selling Equilibrium in which the optimal inventory of a small dealer is zero unit whereas that of a large dealer is zero or one unit. For the market to clear, a fraction or all of L_0s must buy since they are the only possible buyers. And if there are sufficiently many L_0s to meet the supply out of all S_1s and L_2s each selling one unit, the market can indeed clear. Specifically, let m_i^{SD}, $i=0,1$ and m_i^{LD}, $i=0,1,2$, be

[11] A $p=0$ results from the normalization that both low-valuation owners and dealers derive zero flow payoff from owning the asset. At a positive normalized payoff, p would become positive without affecting the qualitative results to follow.

[12] In case investors can trade among themselves, an investor-seller should always be able to sell to an investor-buyer at a positive price, given that the latter gains from holding a unit of the asset. In the present setting, the investor-seller does not have direct access to trading with an investor-buyer but can only trade with a dealer, who may possibly not gain from acquiring the unit in which case the trade can only take place at a zero price.

the respective measures of small and large dealers entering the inter-dealer market holding an i-unit inventory. The inter-dealer market can clear at $p = \beta\left(V_1^{LD} - V_0^{LD}\right)$ for

$$m_0^{LD} \geq m_1^{SD} + m_2^{LD}. \tag{18}$$

Because each L_1 is indifferent between selling and not selling, if (18) holds as a strict inequality, there is room for a fraction or all of them selling in equilibrium. In this way, there is a continuum of equilibrium, indexed by the measure of L_1 sellers, with the measure of L_0 buyers to exceed the given measure of L_1 sellers by an amount to exactly cover the supply out of the inframarginal sellers S_1s and L_2s each selling one unit. In equilibrium, since a sale by an L_1, who will become an L_0 afterward, must be matched by a purchase by an L_0, who will become an L_1 afterward, such trades merely result in those agents concerned switching identities. Hence, the possible multiplicity of equilibrium has no bearing at all on the definitions and main characteristics of the Selling Equilibrium. In particular, the conditions for the existence of the equilibrium and the allocations that follow are completely isomorphic to the multiplicity as they both are derived solely from the optimal inventories of dealers. And arguably, among the continuum, the equilibrium in which all L_1s refrain from selling is the most compelling since L_1s not only do not gain from trade at the given price but any sales by them are not needed for market clearing. Moreover, there involve the least transactions to allow all dealers to reach their respective optimal inventories with no L_1s selling. Formally, their non-participation can be justified by there being a small cost for each trade to be completed, whereby no L_1s would choose to sell if trading is costly.[13] The analysis to follow though does not hinge on restricting attention to this particular equilibrium but is entirely general unless otherwise noted. Such caveats are valid for the next two types of equilibrium we define in the following.

The "Balanced" Equilibrium In the Balanced Equilibrium, $p = \beta\left(V_1^{SD} - V_0^{SD}\right)$. By Proposition 1 and Lemma 2,

$$\beta\left(V_1^{LD} - V_0^{LD}\right) > p = \beta\left(V_1^{SD} - V_0^{SD}\right) > \beta\left(V_2^{LD} - V_1^{LD}\right),$$

from which it follows that L_0s strictly prefer to buy one unit while L_2s strictly prefer to sell one unit. We refer to this as the Balanced Equilibrium, in which the optimal inventory of a large dealer is one unit, whereas that of a small dealer is zero or one unit. For the inter-dealer market to clear, if large dealers buying (selling) outnumber large dealers selling (buying) in the market, small dealers on balance must sell (buy). That is, in case $m_0^{LD} \geq m_2^{LD}$, the Balanced Equilibrium obtains for

$$m_1^{SD} \geq m_0^{LD} - m_2^{LD}. \tag{19}$$

Otherwise ($m_2^{LD} \geq m_0^{LD}$), the market can clear at $p = \beta\left(V_1^{SD} - V_0^{SD}\right)$ for

$$m_0^{SD} \geq m_2^{LD} - m_0^{LD}. \tag{20}$$

[13] No L_0s would choose to buy too if they are obliged to pay the transaction cost. The inframarginal sellers L_2s and S_1s, however, would be ready and willing to finance the trades with them as they strictly gain from selling.

In the former case, there can be a certain measure of S_0s each choosing to buy a unit in equilibrium as long as there is a commensurate increase in the measure of S_1 sellers, whereas in the latter case, there can be a certain measure of S_1s each selling a unit likewise matched by an equal increase in the measure of S_0 buyers.

The "Buying" Equilibrium In the Buying Equilibrium, $p = \beta\left(V_2^{LD} - V_1^{LD}\right)$. By Proposition 1 and Lemma 2,

$$\beta\left(V_1^{LD} - V_0^{LD}\right) = \beta\left(V_1^{SD} - V_0^{SD}\right) > \beta\left(V_2^{LD} - V_1^{LD}\right) = p = 0,$$

from which it follows that no dealers strictly prefer to sell with p anchored at the lowest possible marginal value of inventory. Meanwhile, any dealers with an empty inventory, large and small, strictly prefer to buy. For this reason, we call this the Buying Equilibrium in which the optimal inventory of a large dealer is one or two units, whereas that of a small dealer is one unit. For the inter-dealer market to clear at the given p, there should be sufficiently many L_2s in the market to meet the demand out of all S_0s and L_0s each buying one unit; i.e.,

$$m_2^{LD} \geq m_0^{SD} + m_0^{LD}. \tag{21}$$

And if (21) holds as a strict inequality, there can be a certain measure of L_1s each buying a unit in equilibrium matched by the same increase in the measure of L_2 sellers.

Gains from trade in our model arise (1) out of dealers not having continuous access to trading among themselves and (2) from dealers having different inventory capacities. The former incentivizes dealers to manage inventories and the latter gives rise to different optimal inventories among dealers. Because L_0s value a unit of inventory the most while L_2s the least, the equilibrium price must settle at where they gain from trading with one another. There can also be mutually beneficial trades between a pair of large and small dealers when one dealer's inventory falls short of while the other dealer's inventory exceeds their respective optimal inventories. On the other hand, at a price such that it is optimal for an S_1 to sell where the receipt of the selling price p more than offsets the lost of the unit inventory for the small dealer, it cannot be optimal for an S_0 to buy since in this case, the gain from acquiring the unit inventory cannot cover the purchase price p. All this means that in equilibrium, small dealers do not gain by trading among themselves.

In Table 1, we summarize the optimal inventories of large and small dealers upon exiting and the identities of the buyers and sellers upon entry into the inter-dealer market. The first line of each cell of the "Buyers" and "Sellers" columns indicate the identities of the inframarginal buyers and sellers in the three types of equilibrium who do gain from trade. In parentheses are the identities of dealers who not only do not gain from trade but whose trades are not needed for market clearing and do not affect equilibrium allocations. Notice that in all three equilibrium types, at least a fraction of L_0s must buy and at least a fraction of L_2s must sell. The defining difference among the equilibria is the role played by small dealers. In the Selling Equilibrium, S_1s sell while S_0s stay out of the market. In the Buying Equilibrium, S_0s buy while S_1s stay out of the market. In the Balanced Equilibrium, small dealers may either tend to sell or buy, depending on whether or not the buyers among large dealers outnumber the sellers.

Equilibrium	Price	Optimum Inventory		Buyers	Sellers
		small dealers	large dealers		
Selling	$p = \beta\left(V_1^{LD} - V_0^{LD}\right)$	0	1 and 0	 L_0	L_2, S_1 (L_1)
Balanced Small dealers sell Small dealers buy	$p = \beta\left(V_1^{SD} - V_0^{SD}\right)$	0 and 1	1	L_0 (S_0) S_0	L_2 S_1 (S_1)
Buying	$p = \beta\left(V_2^{LD} - V_1^{LD}\right)$	1	1 and 2	S_0, L_0 (L_1)	 L_2

Table 1: Prices, Optimal Inventories, Buyers and Sellers in Equilibrium

A direct implication of the results in Table 1 is that while there is no persistent trading direction among large dealers since they sell as well as buy in any equilibrium, small dealers either tend to sell to or buy from large dealers in a given type of equilibrium. Now, if only one type of equilibrium can hold for a given parameter configuration – a result we will establish in Proposition 2 to follow – the direction of trade between small and large dealers is persistent. The implication is consistent with the findings in Li and Schürhoff (2014) that given that there is a directional (buy or sell) trade between two dealers in one month, the probability that the same directional trade remains in the next month is 62%.

Our major results thus far – Proposition 1 and the implications thereof – seemingly rest on a number of simplifying assumptions. First, if small dealers can only hold at most one unit of the asset in inventory, it may seem trivial that they do not gain from trading among themselves in equilibrium. An obvious question to ask is how the results may hold otherwise in case they each possess more than a unit inventory capacity. Second, an important lesson in Li and Schürhoff (2014) and the follow-up study in Henderschott, Li, Livdan and Schürhoff (2016) is that the inter-dealer market is itself a decentralized market as opposed to a Walrasian market. Finally, if large dealers can hold up to two units in inventory and may possess up to two units of spare inventory capacity, perhaps a more natural assumption is that they can meet up to two investor-buyers and two investor-sellers in each period. We will discuss how Proposition 1 and its main implications survive all three generalizations in Section 4 below.

A given type of equilibrium places a set of restrictions on the measures of dealer-buyers, dealer-sellers and the asset held by these dealers, which in turn impact on the meeting probabilities in the investor-dealer market. Then, a candidate equilibrium can be equilibrium only if these restrictions are met and where the meeting probabilities are bounded below one, in addition to the existence of some positive mixing probability for the marginal buyers' or sellers' mixed strategy which clears the inter-dealer market. We now proceed to study the underlying environment as defined by the asset supply A, the turnover rate of high-valuation owners δ, the entry rate of high-valuation non-owners e, and the measures of large and small dealers in which the restrictions of each type of equilibrium are met.

2.5 Accounting Identities, Market Tightness, and Stock-Flow Equations

If the market is populated by n^{SD} small dealers and n^{LD} large dealers, then

$$n_0^{SD} + n_1^{SD} = n^{SD}, \tag{22}$$

$$n_0^{LD} + n_1^{LD} + n_2^{LD} = n^{LD}. \tag{23}$$

The asset is in fixed supply equal to A, and hence,

$$n_H^{ON} + n_L^{ON} + n_1^{SD} + n_1^{LD} + 2n_2^{LD} = A, \tag{24}$$

where n_H^{ON} and n_L^{ON}, denote, respectively, the measures of high-valuation owners and low-valuation owners.

Let n_B^I denote the measure of high-valuation non-owners *cum* investor-buyers. Then,

$$\theta_{ID} = \frac{n_B^I}{n_S^D} = \frac{n_B^I}{n_1^{SD} + n_1^{LD} + n_2^{LD}}. \tag{25}$$

With the population of investor-sellers comprised of all low-valuation owners,

$$\theta_{DI} = \frac{n_B^D}{n_L^{ON}} = \frac{n_0^{SD} + n_0^{LD} + n_1^{LD}}{n_L^{ON}}. \tag{26}$$

In the steady state, the respective inflows and outflows of high-valuation owners, low-valuation owners, and investor-buyers are equal. Hence,

$$n_B^I \mu(\theta_{ID}) = \delta n_H^{ON}, \tag{27}$$

$$\delta n_H^{ON} = \eta(\theta_{DI}) n_L^{ON}, \tag{28}$$

$$e = n_B^I \mu(\theta_{ID}). \tag{29}$$

Not all n_i^{SD} and n_i^{LD} can be positive in a given type of equilibrium. In the Selling Equilibrium for example, by the third and fourth columns of Table 1, all small dealers exit the inter-dealer market with an empty inventory, whereas large dealers may do so with either an empty or a one-unit inventory. The restrictions on the measures of small and large dealers with various levels of inventory in the three types of equilibrium are as depicted in Table 2.

	Selling Equilibrium	Balanced Equilibrium	Buying Equilibrium
n_0^{SD}	n^{SD}	$[0, n^{SD}]$	0
n_1^{SD}	0	$[0, n^{SD}]$	n^{SD}
n_0^{LD}	$[0, n^{LD}]$	0	0
n_1^{LD}	$[0, n^{LD}]$	n^{LD}	$[0, n^{LD}]$
n_2^{LD}	0	0	$[0, n^{LD}]$

Table 2: Measures of dealers entering the investor-dealer market in equilibrium

Given the measures of dealers, n_i^{SD}, $i = 0, 1$, and n_i^{LD}, $i = 0, 1, 2$, when the investor-dealer market opens in the first subperiod, the corresponding measures of dealers leaving the market and entering the inter-dealer market in the second subperiod are given by the following.

$$m_0^{SD} = (1 - \mu(\theta_{DI}))\, n_0^{SD} + \eta(\theta_{ID})\, n_1^{SD}, \tag{30}$$

$$m_1^{SD} = \mu(\theta_{DI})\, n_0^{SD} + (1 - \eta(\theta_{ID}))\, n_1^{SD}, \tag{31}$$

$$m_0^{LD} = (1 - \mu(\theta_{DI}))\left(n_0^{LD} + \eta(\theta_{ID})\, n_1^{LD}\right), \tag{32}$$

$$m_1^{LD} = \mu(\theta_{DI})\, n_0^{LD} + [\mu(\theta_{DI})\, \eta(\theta_{ID}) + (1 - \mu(\theta_{DI}))(1 - \eta(\theta_{ID}))]\, n_1^{LD} + \eta(\theta_{ID})\, n_2^{LD}, \tag{33}$$

$$m_2^{LD} = (1 - \eta(\theta_{ID}))\left(\mu(\theta_{DI})\, n_1^{LD} + n_2^{LD}\right). \tag{34}$$

For example, (30) says that S_0s entering the inter-dealer market are among the S_0s entering the investor-dealer market who fail to buy a unit in the market and the S_1s who succeed in selling the unit they each possess.

2.6 Equilibrium

Given $\{n^{SD}, n^{LD}, A, e, \delta\}$, a steady-state equilibrium consists of the respective non-negative values of n_0^{SD}, n_1^{SD}, n_0^{LD}, n_1^{LD}, n_2^{LD}, n_H^{ON}, n_L^{ON} and n_B^I that satisfy (22)-(29), the restrictions on n_i^{SD} and n_i^{LD} in Table 2 and the market-clearing conditions for the type of equilibrium under consideration in (18)-(21), with m_i^{SD} and m_i^{LD} given by (30)-(34). Write $n^D = n^{SD} + n^{LD}$ as the total measure of dealers.

Proposition 2 *The Selling Equilibrium and the Buying Equilibrium may only hold for $e < n^{LD}$. The Balanced Equilibrium may only hold for $e < n^{LD} + \frac{n^{SD}}{2}$.*

a. For $e < n^{LD}$, define

$$B_S \equiv e + \frac{n^D}{\mu^{-1}\left(\frac{e}{n^D}\right)},$$

$$B_M \equiv n^{LD} + \frac{n^D}{\mu^{-1}\left(\frac{e}{n^D}\right)},$$

$$B_L \equiv n^D + \frac{n^{LD}}{\mu^{-1}\left(\frac{e}{n^{LD}}\right)},$$

where $B_S \leq B_M \leq B_L$.

(i) for $A - e/\delta \in (B_S, B_M]$, the Selling Equilibrium holds,

(ii) for $A - e/\delta \in [B_M, B_L]$, the Balanced Equilibrium holds,

(iii) for $A - e/\delta \geq B_L$, the Buying Equilibrium holds.

b. For $e \in \left[n^{LD}, n^{LD} + \frac{n^{SD}}{2}\right)$, the Balanced Equilibrium exists if

$$A - \frac{e}{\delta} > e + \frac{n^D + n^{LD} - e}{\mu^{-1}\left(\frac{e}{n^D + n^{LD} - e}\right)} \equiv \mathcal{B}_M.$$

c. Define

$$\widetilde{B} \equiv n^{LD} + \frac{n^{SD}}{2} + \frac{n^{LD} + \frac{n^{SD}}{2}}{\mu^{-1}\left(\frac{e}{n^{LD}+\frac{n^{SD}}{2}}\right)},$$

where $B_M \leq \widetilde{B} \leq B_L$ *in case* $e < n^{LD}$ *and* $\mathcal{B}_M \leq \widetilde{B}$ *in case* $e \in \left[n^{LD}, n^{LD} + \frac{n^{SD}}{2}\right)$. *In the Balanced Equilibrium's inter-dealer market, for* $A - e/\delta < (>)\, \widetilde{B}$, *small dealers sell (buy) for market clearing.*

First, a steady-state equilibrium exists only if the necessary conditions on e, the entry rate of investors, stated at the beginning of the Proposition, are met. These conditions arise because if a measure of e investors enter the market in each period, in the steady state, there have to be the same measure of e investors succeeding in buying a unit each and the same measure of e investors succeeding in selling a unit each in the same period. All this requires that more than e dealer-sellers and more than e dealer-buyers are present in the market. In the Selling Equilibrium, only large dealers are carrying inventories for sale, whereas in the Buying Equilibrium, only large dealers possess spare capacities to buy. Then, for either type of equilibrium to exist, a necessary condition is that $e < n^{LD}$. In the Balanced Equilibrium, however, a fraction of small dealers enter the investor-dealer market with an empty inventory and a fraction enter holding a one-unit inventory. Then, there will also be small dealers among both dealer-buyers and dealer-sellers and a steady-state equilibrium may exist even for $e \geq n^{LD}$ but not for $e \geq n^{LD} + \frac{n^{SD}}{2}$ since if more than one-half of all small dealers search as buyers (sellers), then there can only be fewer than one-half searching as sellers (buyers). In Figure 1, we illustrate how a steady-state equilibrium may only exist for $e < n^{LD} + \frac{n^{SD}}{2}$ and that for $e \in \left[n^{LD}, n^{LD} + \frac{n^{SD}}{2}\right)$, the only type of equilibrium that can hold is the Balanced Equilibrium.

With e investors entering the market in each period and with each high-valuation owner turning into a low-valuation owner at probability δ, there are e/δ high-valuation owners in the steady state each holding a unit. The expression $A - e/\delta$ in Parts (a)-(c) then, on the one hand, denotes the quantity of the asset for sale held by investors and dealers altogether, and on the other hand, can be interpreted as the difference between the asset supply A and the asset demand e/δ.

Among the three types of equilibrium, the Selling Equilibrium, in which only a fraction of large dealers may hold just a one-unit inventory, involves dealers holding the least inventory, whereas the Buying Equilibrium, in which only a fraction of large dealers may still have one unit of spare inventory capacity, involves dealers holding the largest inventory. The condition in Part a(i) can be shown to be the condition for how when the dealers' inventory, held entirely by large dealers, just suffices to satisfy the demand from investor-buyers to allow for the Selling Equilibrium to hold and it holds until all large dealers are holding a unit. At this point, according to Part a(ii), the Balanced Equilibrium begins to hold and it holds until all small dealers are holding a unit in inventory as well. Thereafter, by Part a(iii), the Buying Equilibrium holds, in which all dealers hold at least a one-unit inventory and a fraction of large dealers are holding a two-unit inventory. Part (b), as in Part (a), says that a steady-state equilibrium exists once the dealers' inventory is up to the level to satisfy investors' asset

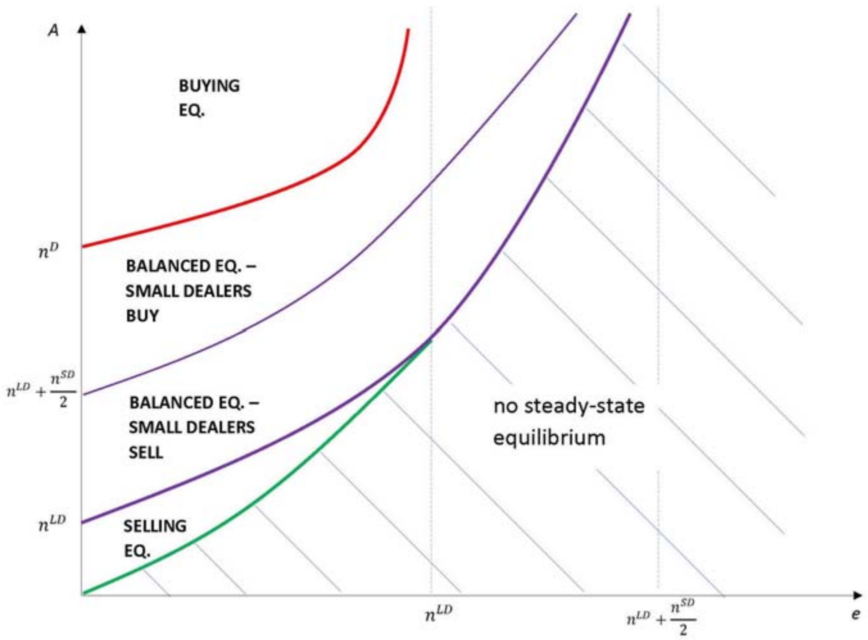

Figure 1: Equilibirum – Existence and Uniqueness

demand. Figure 1, with e on the horizontal axis measuring investors' asset demand and A on the vertical axis measuring the asset supply, illustrates the partition of e-A space for the three types of equilibrium to hold.

Parts a(iii) and (b) say that a steady-state equilibrium holds even for arbitrarily large $A - e/\delta$. In our model, as $A - e/\delta$ increases, it takes longer for each low-valuation owner to sell to a dealer in a more congested market, during which the measure of low-valuation owners and their asset holdings increase without bounds. With n_L^{ON} increasing in tandem with the asset supply, the inventory held by dealers never rises above the level that would leave them with insufficient spare inventory capacity to buy e units of the asset from investors.

In Part (c), when $A - e/\delta = \widetilde{B}$ holds exactly in the Balanced Equilibrium, $A - e/\delta$ is at which dealers buy from and sell to investors at the same probability. Then, there would be just as many L_0s and L_2s entering the inter-dealer market to result in large dealers' demand and supply being equal. For any smaller (larger) $A - e/\delta$, as in the lower (upper) subset of the e-A space for the Balanced Equilibrium to hold in Figure 1, dealers buy at a smaller (larger) probability while selling at a higher (smaller) probability in the investor-dealer market to result in fewer (more) L_2s entering the inter-dealer market to sell than L_0s entering the market to buy. Small dealers sell (buy) in equilibrium to eliminate the excess demand (supply) among large dealers.

The proof of the Proposition in the Appendix shows that the condition in Parts (a)-(c) are also the conditions for how the inter-dealer market can clear for the given type of equilibrium in (18)-(21). Recall that all dealers exit the inter-dealer market and enter the investor-dealer market with their respective optimal inventories. A given round of trading with investors afterwards would leave all dealers that have just traded weakly prefer to trade again in the inter-dealer market to restore their respective optimal inventories. Those dealers who strictly prefer to trade in one direction must, however, be less numerous than dealers who weakly prefer to trade in the other direction, meeting the requirement for market clearing, since in each type of equilibrium, there is one type of dealers indifferent between holding two levels of inventory.

So far we have only entertained the possibility that the price in the inter-dealer market be equal to one of the three marginal values of inventory in Proposition 1, the justification for which is that the market in general cannot clear at any p strictly in between any two of the marginal values. The following provides the formal result.

Lemma 3 *Equilibrium obtains for a given* $p \in \left(\beta\left(V_1^{SD}-V_0^{SD}\right), \beta\left(V_1^{LD}-V_0^{LD}\right)\right)$ *only for* $A - e/\delta = B_M$. *Equilibrium obtains for a given* $p \in \left(\beta\left(V_2^{LD}-V_1^{LD}\right), \beta\left(V_1^{SD}-V_0^{SD}\right)\right)$ *only for* $A - e/\delta = B_L$.

Together with Proposition 2, the Lemma says that at precisely the boundary between the Selling and the Balanced Equilibria, the inter-dealer market can clear at any price in between the Selling and the Balanced Equilibrium prices, whereas at precisely the boundary between the Balanced and Buying Equilibria, the inter-dealer market can clear at any price in between the Balanced and the Buying Equilibrium prices. At a given boundary, however, the allocations of the two equilibria concerned are identical and so exactly where p lies in between the two marginal values of inventory is immaterial.

2.7 The Makeup of the Dealer Population and Equilibrium Types

Proposition 2 focuses on the role the asset supply plays in determining equilibrium outcomes. To complete the analysis, we next turn our attention to the role the makeup of the dealer population plays.

To begin, if there were no large dealers, any inter-dealer trades would only be between an S_1 selling to an S_0. The equilibrium terms of trade must then be such that the two parties are indifferent between trading and not trading; i.e. $p = \beta\left(V_1^{SD}-V_0^{SD}\right)$ as in the Balanced Equilibrium. For any positive n^{LD} not up to e, by Proposition 2, there can still only be a Balanced Equilibrium in which small dealers remain indifferent between trading and not trading. But as n^{LD} rises up to and above e, the Selling and the Buying Equilibrium may begin to hold in which small dealers strictly prefer to trade.

The conditions in Proposition 2 may be manipulated to trace out the evolution of equilibrium type as n^{LD} increases from the smallest admissible value to n^D.

Corollary 1 *Holding fixed* n^D*, the Balanced Equilibrium begins to hold for*

$$n^{LD} > \eta^{-1}\left(\frac{e}{A-\frac{e}{\delta}-e}\right)\left(A-\frac{e}{\delta}-e\right)-n^D+e.$$

As n^{LD} *increases from the lower bound in the above towards* n^D*, the Balanced Equilibrium holds throughout only if* $A - e/\delta = \widehat{B}$*, where*

$$\widehat{B} \equiv n^D + \frac{n^D}{\mu^{-1}\left(\frac{e}{n^D}\right)}.$$

In general, the Balanced Equilibrium changes into the Buying Equilibrium for $A - e/\delta > \widehat{B}$ at

$$n^{LD} = \eta^{-1}\left(\frac{e}{A - \frac{e}{\delta} - n^D}\right)\left(A - \frac{e}{\delta} - n^D\right),$$

but into the Selling Equilibrium for $A - e/\delta < \widehat{B}$ at

$$n^{LD} = A - \frac{e}{\delta} - \frac{n^D}{\mu^{-1}\left(\frac{e}{n^D}\right)}.$$

When large dealers become relatively more numerous then, the Balanced Equilibrium in general must give way to either the Buying or the Selling Equilibrium. In particular, if the asset supply is relatively abundant as with $A - e/\delta > \widehat{B}$, for the inter-dealer market to clear in the Balanced Equilibrium, S_0s buy to eliminate the excess supply among large dealers. But then when n^{LD} rises up to and n^{SD} falls down to some given levels, the remaining S_0s would no longer be sufficiently numerous to fulfil such a role. In this case, equilibrium can only obtain when L_2s no longer strictly prefer to sell as when the inter-dealer market price falls from $p = \beta\left(V_1^{SD} - V_0^{SD}\right)$ to $p = \beta\left(V_2^{LD} - V_1^{LD}\right)$, at which point the Buying Equilibrium takes hold. Conversely, how the Balanced Equilibrium changes into the Selling Equilibrium when the asset supply is relatively meager as with $A - e/\delta < \widehat{B}$ is due to how not all L_0s can remain as buyers when the remaining S_1s would no longer suffice to fill the gap between the supply and demand from large dealers as those dealers make up a large enough fraction of the dealer population.

3 Direction of Trade: Small Dealers Provide Immediacy for Large Dealers

A major point of interest of our analysis is that the relative asset supply $A - e/\delta$ plays an important role in determining the direction of trades among dealers having different capacities to hold the asset – a question that cannot be answered by other models of inter-dealer trades in which the asset supply should merely scale the trading volume but not affect the direction of trades.

Before proceeding further, it is useful to note that in our model, as $A - e/\delta$ increases, dealers are holding relatively fewer units of the asset than investors do as the overall dealers' inventory holding is bounded by the level that would leave them with enough spare capacity to buy e units of the asset from investors while low-valuation owners' asset holdings are increasing without bounds. This implication is consistent with the findings in Di Maggio, Kermani and Song (2017), among others, that dealers as a whole do not step up but instead scale back their liquidity provision in the down market that follows the 2008 financial crisis.

Now, as to how $A - e/\delta$ helps determine the directions of inter-dealer trades, a direct corollary of Proposition 2 is that:

Corollary 2

(a) For $A - e/\delta \in (B_S, B_M]$, small dealers only sell to but do not buy from large dealers.

(b) For $A - e/\delta \in \left[B_M, \widetilde{B}\right)$*, small dealers sell to more than they buy from large dealers if they buy from large dealers at all.*

(c) For $A - e/\delta \in \left(\widetilde{B}, B_L\right]$*, small dealers buy from more than they sell to large dealers if they sell to large dealers at all.*

(d) For $A - e/\delta \geq B_L$*, small dealers only buy from but do not sell to large dealers.*

Parts (a) and (d) are due to how small dealers only sell in the Selling Equilibrium and only buy in the Buying Equilibrium, respectively. In the Balanced Equilibrium where $A - e/\delta < \widetilde{B}$ as for Part (b), S_1s sell to close the gap between large dealers' demand for and supply of the asset in the inter-dealer market. While a fraction of S_0s may choose to buy in equilibrium, small dealers must sell more than they buy for market clearing. Part (c) describes the mirror opposite of Part (b). In all, the above says that small dealers tend to sell to large dealers when there is a small relative asset supply and vice versa. The implication happens to coincide with what the findings in Adrian et al. (2017) suggest if a small relative asset supply in our model is taken to be a booming market in which the asset demand is strong relative to the asset supply and a large relative asset supply in our model is taken as a market bust in which the asset demand is weak relative to the asset supply. In the former case, as in the up market before the financial crisis, large dealers amass inventory by buying from small dealers. In the latter case, as in the down market after the financial crisis, small dealers gain inventory by buying from large dealers.

A priori it seems intuitive that large dealers, to the extent that they are able to and indeed tend to hold a larger inventory, should on balance sell to small dealers. In our model, however, this is the case only when the asset supply is relatively abundant – just when small dealers should find it easiest to buy the asset from investors themselves. A dealer is said to provide immediacy to another dealer if the first dealer sells the asset to (buys from) the other dealer at times when it takes longest on average for the latter to buy (sell) the asset in the market as arising, in particular, from a meagre (abundant) asset supply. Apparently, large dealers in our model do not provide immediacy for small dealers.

The proof of Proposition 2 shows that for $A - e/\delta < \widetilde{B}$, $\eta(\theta_{ID}) > \mu(\theta_{DI})$, whereby dealers are more likely to meet investor-buyers than investor-sellers. This is due to how the scarcity of the asset should give rise to more selling opportunities than buying opportunities for dealers. In our model, small dealers sell to large dealers to provide inventory for the latter to sell to investors. For $A - e/\delta > \widetilde{B}$, with $\eta(\theta_{ID}) < \mu(\theta_{DI})$, dealers need spare inventory capacity more than inventory in a market with a large asset supply as the abundance of the asset gives rise to more buying opportunities than selling opportunities for dealers. In our model, small dealers buy from large dealers to help them free up capacities to buy from investors. In all, it is the small dealers who provide immediacy for large dealers in our model.

It is well known that inter-dealer trades in many OTC markets can be characterized by a so-called core-periphery trading structure in which a set of dealers, referred to as the core dealers, are observed to trade with all dealers while the rest, referred to as the peripheral dealers, are observed to trade only with the core dealers but not among themselves (Li and Schürhoff (2014) and Hollifield, Neklyudov and Spatt (2016)). True, assuming that inter-dealer trades take place in a competitive inter-dealer market, our model has no predictions as to the exact identity of the dealer whom a given dealer tends to trade with. But we do know that all small dealers either tend to sell or buy in a given type of equilibrium in which case a given small dealer should not be buying from or selling to another small dealer.[14] On the other hand, in all three types of equilibrium in our model, there are large dealers who are selling and others who are buying in which case a given large dealer can be trading with another large dealer in one period but then with a small dealer in different time periods. In this way, the large and small dealers in our model behave similarly as the core and the peripheral dealers do, respectively, identified in the empirical studies with regard to the set of dealers they are predicted to trade with.[15]

In our model then, the large dealers can be interpreted to play the role of core dealers in a core-periphery trading network, trading with all dealers in the market while the small dealers play the role of peripheral dealers, trading only with the core dealers.[16] Under this interpretation, our model predicts that the small peripheral dealers provide immediacy for the large core dealers in general by selling to the latter when it takes a relatively long time for the core dealers to buy from investors themselves but buying from the latter when it is difficult for the core dealers themselves to sell to investors.[17]

This implication is counterintuitive but there indeed exists empirical evidence supporting it. Hollifield et al. (2015) report in their Table 6 that the percentage bid-ask spreads peripheral dealers earn when they are the first links of the intermediation chains, buying from an investor for selling to other dealers, is smaller than the percentage spreads they earn when they are the last links, buying from another dealer for selling to an investor. There are no statistically significant differences between the two spreads for core dealers, however.

In our model, when the latter part of the Balanced Equilibrium or the Buying Equilibrium holds as arising from an abundant asset supply, the small peripheral dealers are in the last link of the intermediation chains. And then with an abundant asset supply, the small peripheral dealers would be selling to investors in a slow market (small θ_{ID}) – a market in which it takes a long time on average for a dealer to sell the unit in his inventory. The intermediation services provided by small dealers for large dealers by selling to investors on their behalf in the slow market should then command a high return.

[14] In the Balanced Equilibrium, there can be small dealers buying (selling) where small dealers should be sellling for market clearing. But those sales (purchases) do not benefit any dealers and we think it is not far-fetched to presume that they should not take place and indeed they would not take place with a small trading cost in place as we argued earlier. In this case, either all small dealers sell or buy in the Balanced Equilibrium, as in the Selling and Buying Equilibria.

[15] One interpretation of a competitive market is that trades are literally centralized and that buyers and sellers do not trade with each other but instead with a Walrasian auctioneer. Of course, in this interpretation, our model cannot be taken to imply any kind of trading structure whatsoever. In our view, such an interpretation is needlessly agnostics. To us, the defining characteristic of perfect competition is that all trades take place at a price that equates demand and supply and that the Walrasian auctioneer story is but one story justifying the assumption. For more sophisticated theoretical foundations that begin with bi-lateral trades, see Gale (2000).

[16] Li and Schürhoff (2014) find that core dealers tend to hold more assets in inventory. In all three types of equilibrium in our model, the optimal inventory level for a large dealer (who is in the core) is at least weakly higher than that of a small dealer (who is in the periphery).

[17] In other network-theoretic models of inter-dealer trades with a core-periphery structure, the core dealers are either identified with dealers that sell to and thus provide inventory for peripheral dealers (Farboodi (2014) and Zhong (2014)), dealers that generally provide immediacy for peripheral dealers by virtue of being more connected to other dealers (Wang (2017)), or dealers that simply tend to trade more frequently (Neklyudov (2015) and Hugonniery et al. (2016)).

On the other hand, when the Selling Equilibrium or the earlier part of the Balanced Equilibrium holds as arising from a meager asset supply, small dealers are in the first link of the intermediation chains. With the meager asset supply, the small peripheral dealers would be buying from investors in a tight market (large θ_{DI}),[18] a market populated by a large number of dealer-buyers versus a small number of investor-sellers. In the tight market, where the competition among dealers can be intense, there should only be a small return earned by a dealer from intermediating the sale by an investor since the dealer has a relatively weak bargaining position vis-à-vis an investor having plenty of other meeting opportunities.

Averaging over markets with various levels of asset supply, small dealers in our model indeed should earn a higher markup when they buy from other dealers for selling to investors than when they buy from investors for selling to other dealers, just as what Hollifield et al. (2015) find in their empirical analysis. To confirm this conjecture, we calculate and then plot the markup

$$\rho_{DB} \equiv \frac{p - p_{I_S}}{p_{I_S}}$$

small dealers earn by intermediating the sale by investors in the Selling Equilibrium and the earlier part of the Balanced Equilibrium and the markup

$$\rho_{DS} \equiv \frac{p_{I_B} - p}{p}$$

they earn by intermediating the sale by other dealers in the latter part of the Balanced Equilibrium against A in Figure 2.[19] When the Buying Equilibrium holds, $\rho_{DS} = \infty$ with $p = 0$, and so is left out of the plot. By changing the normalization that low-valuation investors and dealers place a positive value, instead of zero, on holding a unit of the asset, p would stay positive in any equilibrium while the basic forces in our model should still contrive to give rise to a relatively large ρ_{DS} in the Buying Equilibrium.

The tendencies that ρ_{DB} should be low for small A and ρ_{DS} should be high for large A arise out of the influences of market tightness on agents' bargaining positions and should be common to most models of frictional OTC markets.[20] Models that are built to be consistent with core dealers providing immediacy for small dealers would then likely to predict just the opposite of the findings in Hollifield et al. (2015).

Last but not least, we should mention that our model can also be consistent with the finding in Hollifield et al. (2015) that the central dealers do not tend to earn higher or lower percentage markups between buying from and selling to customers in a dealer chain. The

[18] We formally show that both θ_{ID} and θ_{DI} are (weakly) decreasing and continuous in A in Proposition 4a below.

[19] The numerical analyses assume $\eta(\theta) = 1 - e^{-\theta}$, $e = 0.8$, $\delta = 0.1$, $\beta = 0.95$, $n^{LD} = 0.5$, $n^{SD} = 1$ and A varying from 9.86 – the lower bound for the Selling Equilibrium to hold – and up. The equations for p, θ_{DI}, and θ_{ID} for the Selling Equilibrium are given by (73), (82), and (84), respectively. The equations for p, θ_{ID}, and θ_{DI} in the Balanced Equilibrium are given by (75), (95), and (96), respectively. The equations for p_{I_B} and p_{I_S} for the Selling and Balanced Equilibria are stated in Lemma A1 in the Appendix. The discontinuities in ρ_{DS} and ρ_{DB} are due to p changing its anchor from one to another indifference conditions as one equilibrium type changes to another. See Proposition 4c below.

[20] And the predictions indeed are seen to be consistent with the empirical results reported in Friewald and Nagler (2017), among others.

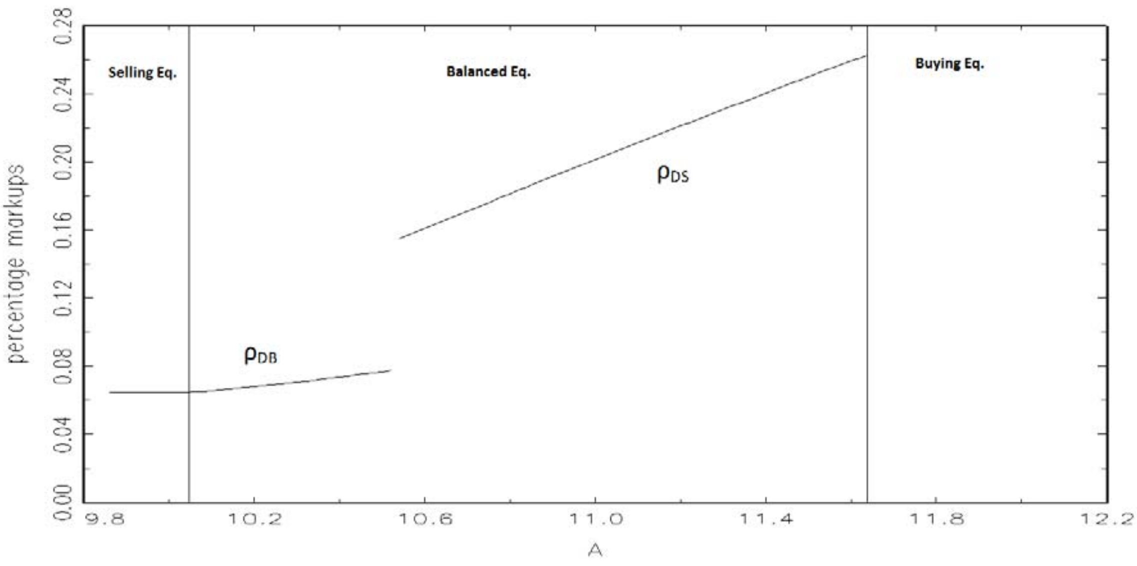

Figure 2: Small dealers' percentage bid-ask spreads

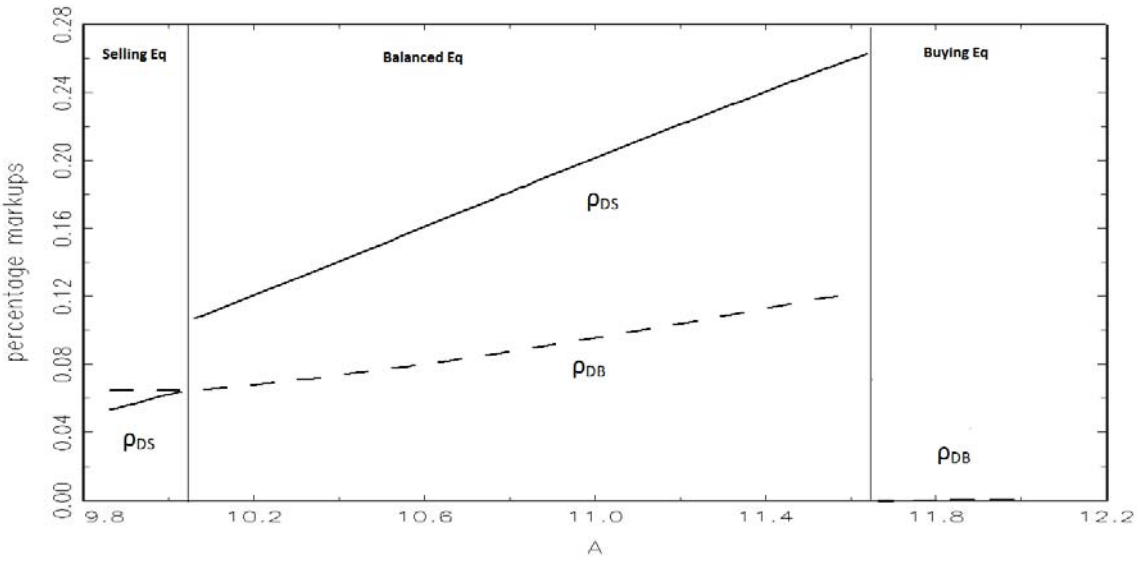

Figure 3: Large dealers' percentrage bid-ask spreads

large core dealers in our model are dealer-buyers as well as dealer-sellers in all three types of equilibrium. For each A then, there are large dealers earning ρ_{DS} as well as ρ_{DB}. In Figure 3, we plot ρ_{DS} and ρ_{DB} for all levels of A. Although large dealers earn a higher ρ_{DS} than ρ_{DB} in the Balanced and in the Buying Equilibria, they earn a higher ρ_{DB} instead in the Selling Equilibrium. Averaging over the A for assets included in a given sample, there can be no statistically significant differences between the observed ρ_{DS} and ρ_{DB} for the large dealers.

4 Robustness of the Equilibrium Features

The model we studied in this paper is admittedly a very special model, with numerous important simplifying assumptions. In this section, we explain how our major results regarding trading directions should survive three generalizations that are most warranted. Notwithstanding the discussions below falling short of full-fledged formal analyses in various occasions for the interest of brevity, the claims not derived from formal proofs in the following all appear to be intuitive extensions of the corresponding results in the main model. All lemmas and Proposition 3 below are formal results though the proofs of which are in the Appendix.

4.1 Larger Inventory Capacity for Small Dealers

Perhaps it seems trivial that, in our model, small dealers, in having just one unit of inventory capacity, never gain from trading among themselves. The question then is if and how the trading directions in Corollary 2 survive the generalization where small dealers each possess more than a unit of inventory capacity and thereby may gain by trading with one another.

Consider, in particular, that the small dealers each possess a two-unit inventory capacity, while the large dealers each possess a three-unit inventory capacity. These larger capacities are relevant only if a dealer may buy and sell up to two units of the asset in a period. The simplest extension is to assume that there are two types of investors – small and large, where the former, comprising a fraction α of the investor population, may each hold either zero or one unit, whereas the latter, comprising the rest of the investor population, may each hold either zero or two units, and that dealers meet investors randomly independent of dealers' types. In this environment, a dealer-seller (-buyer) holding a one-unit inventory (spare capacity) can only sell to (buy from) small investors who buy (sell) one unit of the asset at a time, whereas a dealer-seller (-buyer) holding at least a two-unit inventory (spare capacity) can sell to (buy from) large investors, who buy (sell) two units at a time, as well as small investors.

Now, a ranking of the marginal benefits of inventory similar to that in Proposition 1 should remain – an additional unit of the asset should be valued higher by a large dealer than by a small dealer at the same initial level of inventory for the two dealers since the former would retain a greater spare capacity for future buying needs than the latter in using up a unit of capacity for acquiring the unit. On the other hand, the large dealer should value an additional unit of the asset less than the small dealer when they start with the same spare capacity, as the former has a larger initial inventory than the latter beforehand. The ranking of the marginal benefits of inventory in Proposition 1 may then be generalized to

$$V_1^{LD} - V_0^{LD} \geq V_1^{SD} - V_0^{SD} \geq V_2^{LD} - V_1^{LD} \geq V_2^{SD} - V_1^{SD} \geq V_3^{LD} - V_2^{LD},$$

where the inter-dealer market clears in general only at p equal to β times one of the above marginal values. We next proceed to inquire how the trading directions between small and large dealers remain persistent while small dealers trade to provide immediacy for large dealers. For the ease of exposition and brevity and without loss of generality, we assume, for the following, that dealers who do not gain from trade and whose trades are not required for market clearing do not trade in the inter-dealer market.

Case 1a $p = \beta\left(V_1^{LD} - V_0^{LD}\right)$ The buyers in the inter-dealer market are comprised of a fraction of L_0s and the sellers are S_1s, S_2s, L_2s, and L_3s.

Case 1b $p = \beta\left(V_1^{SD} - V_0^{SD}\right)$ **with a fraction of S_1s selling in the inter-dealer market** The buyers in the inter-dealer market are all of L_0s and the sellers are a fraction of S_1s, and all of S_2s, L_2s and L_3s.

For p to settle at the highest or the second highest marginal values of inventory, the two cases above should hold for the smallest A. Given that all small dealers who trade in the inter-dealer market (S_1s and S_2s) sell and they sell to L_0s, small dealers trade to provide immediacy for large dealers and the trading direction between small and large dealers is persistent.

Case 2 $p = \beta\left(V_1^{SD} - V_0^{SD}\right)$ **with a fraction of S_0s buying in the inter-dealer market** The buyers in the inter-dealer market are all of L_0s and a fraction of S_0s, and the sellers are L_2s, S_2s and L_3s. When this type of equilibrium first starts to hold, the fraction of S_0s who buy is arbitrarily close to zero. When this equilibrium turns into the equilibrium at $p = \beta\left(V_2^{LD} - V_1^{LD}\right)$ so that all S_0s are buying, as we shall demonstrate below in the next case, there would be as many S_0s as S_2s. In between, we conjecture that there remains fewer S_0 buyers than S_2 sellers. Then, on balance, small dealers are selling to and thus are still providing immediacy for large dealers at a p that should hold for relatively small A. The trading direction, though not perfectly, is largely persistent.

Case 3 $p = \beta\left(V_2^{LD} - V_1^{LD}\right)$ The buyers in the inter-dealer market are all of L_0s, S_0s, and possibly a fraction of L_1s. The sellers are S_2s, L_3s, and possibly a fraction of L_2s. Given that all small dealers leave the inter-dealer market with one unit of inventory whereas large dealers do so with either one or two units of inventory, when the investor-dealer market opens, all dealers are dealer-sellers as well as dealer-buyers. All this can be shown to imply that:

Lemma 4 *(a) $\eta(\theta_{ID}) = \mu(\theta_{DI})$, (b) $m_0^{SD} = m_2^{SD}$, and (c) $m_0^{LD} = m_3^{LD}$.*

Part (a) of the Lemma says that it is equally likely for dealers to meet investor-buyers and investor-sellers in which case a unit of inventory would not be any more or less useful than a unit of inventory capacity to dealers. There should then be no particular role to be played by small dealers in providing immediacy for large dealers. Parts (b) and (c) together say that what follows next is that there are equal measures of inframarginal buyers $\left(m_0^{SD} + m_0^{LD}\right)$ and sellers $\left(m_2^{SD} + m_3^{LD}\right)$, whereby the indifferent traders L_1s and L_2s do not trade in the inter-dealer market. This equilibrium then corresponds to the mid-point of the Balanced Equilibrium in the basic model at which $\eta(\theta_{ID}) = \mu(\theta_{DI})$ and that the indifferent traders S_0s and S_1s do not trade. Here, with small dealers selling and buying in the inter-dealer market equally numerous, they indeed do not tend to provide immediacy for large dealers. For $\eta(\theta_{ID}) = \mu(\theta_{DI})$ exactly, A must be at just one particular level, as in the basic model.

Case 4 $p = \beta\left(V_2^{SD} - V_1^{SD}\right)$ **with a fraction of S_2s selling in the inter-dealer market** This is the mirror opposite of case 2. Small dealers buy from more than they sell to large dealers, helping large dealers free up inventory capacities and providing immediacy for them on balance with p settling at a relatively low level as arising from relatively large A.

Case 5 $p = \beta\left(V_2^{SD} - V_1^{SD}\right)$ **with a fraction of S_1s buying in the inter-dealer market or** $p = \beta\left(V_3^{LD} - V_2^{LD}\right)$ The is the mirror opposite of case 1. Small dealers buy from large dealers only, providing immediacy for large dealers, with p settling at the lowest levels as arising from the largest A.

The above suggests that the result that small dealers provide immediacy should also generalize to where there are more than two inventory capacities, as similar mechanisms should be operative to give rise to smaller-capacity dealers selling (buying) the asset to (from) larger-capacity dealers when the asset supply is small (large).

4.2 Frictional inter-dealer market

In reality, the inter-dealer market is better described as a decentralized market as suggested by the findings in Li and Schürhoff (2014) and Henderschott, Li, Livdan and Schürhoff (2016), where it takes time and effort for a dealer to find a counterparty to trade with, in which case dealers, by all means, have incentives to manage inventory for future trading needs. By assuming dealers only have periodic, instead of continuous, access to the competitive inter-dealer market, the dealers in our model likewise have incentives to manage inventory. Where the incentives are similar, many features of the equilibrium in the present model should survive in an arguably richer model of a frictional inter-dealer market.

In the following, we report the results of our analysis of a model of a frictional inter-dealer market, which is otherwise identical to the main model of the paper, except that the model is set in continuous time as it is a more convenient setting to analyze a model in which both the investor-dealer and the inter-dealer markets are decentralized. In the revised model, we continue to assume that the search and matching in the investor-dealer market takes place in two market segments, with respective market tightness θ_{ID} and θ_{DI}. The inter-dealer market is frictional, however, in which a given dealer meets another randomly selected dealer at a fixed rate α per unit of time, and where the terms of trade between two dealers are determined by Nash Bargaining, as are prices in the investor-dealer market. The value functions and equilibrium conditions are presented in Appendix 7.1. The pricing equations are standard and are omitted for brevity.

We first verify that:

Lemma 5 *Any investor-dealer match yields a non-negative surplus in any equilibrium in which both small and large dealers are active.*

This Lemma generalizes Lemma 1 for a competitive inter-dealer market. The proof of the Lemma proceeds with the idea that first, if a given dealer-seller DS chooses not to sell to investor-buyers (IB), he must then sell to other dealers, say dealer d, for otherwise the unit will never be passed on to an IB, in which case it would never be optimal for dealer DS to acquire the unit in the first place. Now, if it is optimal for dealer d to sell to an IB, it must be optimal for dealer DS to sell to the IB as well – there cannot be any greater surplus of trade for the unit to pass to another dealer before the unit is sold to an IB. A similar argument explains how a given dealer possessing spare capacity must find it optimal to buy from an investor-seller should the opportunity arises if the dealer should stay active in equilibrium at all. Given the Lemma, we then proceed to show that the counterpart to Proposition 1 for the competitive inter-dealer market holds.[21]

Proposition 3 $V_1^{LD} - V_0^{LD} \geq V_1^{SD} - V_0^{SD} \geq V_2^{LD} - V_1^{LD}$ *in any active equilibrium. The two equalities are strict unless the surplus for the IB-L_1 match,*

$$z_{I_B,L_1} \equiv U_H^{ON} - U^B - \left(V_1^{LD} - V_0^{LD}\right).$$

is equal to zero.

In our baseline model, we assume that inter-dealer trading opportunities arrive each time after dealers have traded with investors. The question then is how the ranking of the marginal values of inventory in Proposition 1 may survive in a more general model in which the inter-dealer market may only open every n periods for any $n \geq 1$. Now, with a frictional inter-dealer market, a dealer indeed may only meet another dealer for an inter-dealer trade after he has traded with investors for any number of times. The ranking in Proposition 1 remains intact.

With the same ranking of the marginal values of inventory and if the inequalities are strict, among the respective surpluses for the possible inter-dealer trades, only

$$z_{L_0,S_1} \equiv V_1^{LD} - V_0^{LD} + V_0^{SD} - V_1^{SD} > 0,$$

$$z_{S_0,L_2} \equiv V_1^{SD} - V_0^{SD} + V_1^{LD} - V_2^{LD} > 0,$$

$$z_{L_0,L_2} \equiv V_1^{LD} - V_0^{LD} + V_1^{LD} - V_2^{LD} > 0.$$

whereas all other bilateral trades between dealers either yield a negative surplus or just a zero surplus for trades that merely involve two dealers switching identities. Hence, exchanges between an L_0 and an S_1, between an S_0 and an L_2, and between an L_0 and an L_2 exhaust all profitable exchanges among dealers, as in the competitive inter-dealer market model.

In a competitive inter-dealer market, we have shown that, in Corollary 2, small dealers on balance sell to (buy from) large dealers when the asset supply is relatively meagre (abundant). In a frictional inter-dealer market, both L_0-S_1 trades (small dealers' sales to large dealers) and S_0-L_2 trades (small dealers' purchases from large dealers), each yielding a positive surplus, would take place in equilibrium for all levels of asset supply. Even so, the Corollary can

[21] All notations for the revised model have the same meanings as for the main model.

remain valid if the L_0-S_1 trades happen to dominate the S_0-L_2 trades when the asset supply is relatively meagre and vice versa.[22]

In our numerical analyses,[23] we find that the volume of L_0-S_1 trades, given by[24]

$$SD_s = n_1^{SD} \alpha \frac{n_0^{LD}}{n^D} \tag{35}$$

is decreasing in A, whereas the volume of S_0-L_2 trades, given by

$$SD_b = n_0^{SD} \alpha \frac{n_2^{LD}}{n^D} \tag{36}$$

is increasing in A as shown in Figure 4, and that SD_s dwarfs (is dwarfed by) SD_b for small (large) A.[25] Hence, even though small dealers do buy from and sell to large dealers for all levels of A, they overwhelmingly sell to provide inventory for large dealers for small A but buy from them to provide capacity for large dealers for large A just as in the competitive inter-dealer market model.

The first order effect of an increase in A is that there should be more dealers searching with a full inventory (larger n_1^{SD} and n_2^{LD}) and fewer dealers searching with an empty inventory (smaller n_0^{SD} and n_0^{LD}) for the additional units of the asset to be in circulation. A priori then, by (35) and (36), respectively, it seems that either SD_s or SD_b may go up or down with A.

There are also second order effects arising from inter-dealer trading, however. Specifically, with more dealers searching with a full inventory who may sell to dealers searching with an empty inventory, inter-dealer trading should further reinforce the decline in the measures of the latter group of dealers. Since L_0s buy when they meet S_1s as well as L_2s, whereas S_0s buy only when they meet L_2s, this second order effect would be felt more on n_0^{LD} than on n_0^{SD}. Other things equal, there are stronger forces for SD_s to go down with A than for SD_b.

On the other hand, with fewer dealers searching with an empty inventory who may buy from dealers searching with a full inventory, inter-dealer trading should similarly further reinforce the increase in the measures of the latter group of dealers. Since L_2s sell when they meet L_0s as well as S_0s, while S_1s sell only when they meet L_0s, this second order effect should be felt more on n_2^{LD} than on n_1^{SD}. Other things equal, there are stronger forces for SD_b to go up with A than for SD_s. All together then, where there are stronger forces for SD_s to go down and stronger forces for SD_b to go up with A, on balance, SD_s should only become decreasing and SD_b become increasing in A, as depicted in Figure 4.

[22] The equilibrium conditions turn out to be highly nonlinear in equilibrium objects, whereby it does not seem possible to analytically solve the model and to derive conditions on model fundamentals for the existence and uniqueness of equilibrium. Assuming a Walrasian inter-dealer market simplifies considerably and enables us to derive a rich set of analytical results.

[23] In the numerical example, we assume $\eta(\theta) = \theta^{0.5}$, $\alpha = 1$, $e = 0.1$, $d = 0.05$, and $\{n^{SD}, n^{LD}\} = \{0.6, 0.1\}$. A unique steady-state equilibrium exists for $A \in [2.77, 3.61]$.

[24] An S_1 meets another dealer at the rate α and a fraction n_0^{LD}/n^D of those dealers are L_0. The equation for SD_b is constructed similarly.

[25] These results are from assuming a large n^{SD} $(= 0.6)$ relative to n^{LD} $(= 0.1)$. As checks for robustness, we find the same qualitative results hold with $\{n^{SD}, n^{LD}\} = \{0.1, 0.6\}$ and $\{0.35, 0.35\}$.

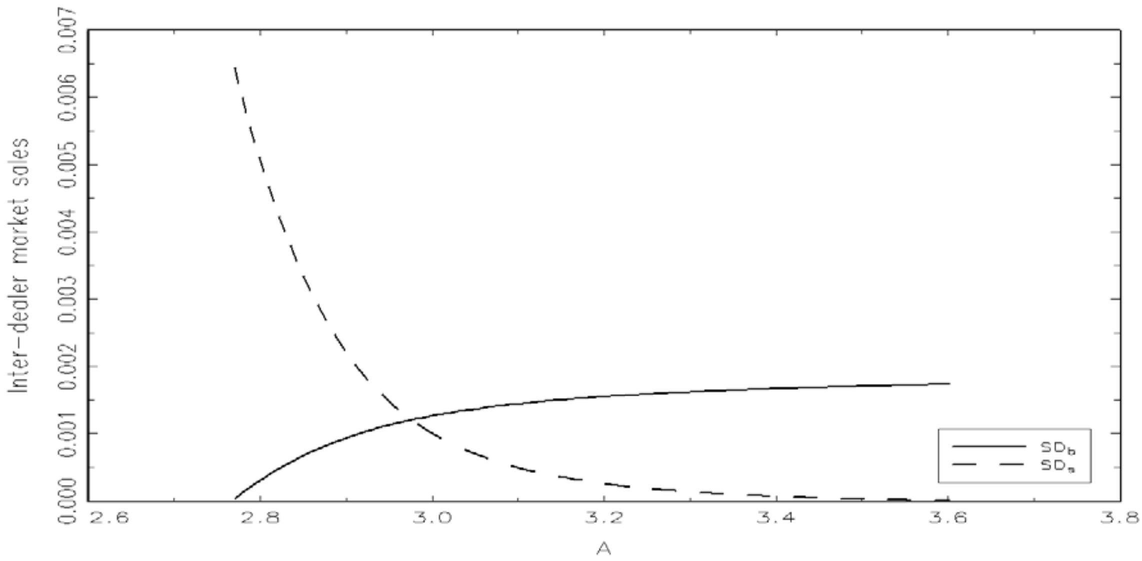

Figure 4: Small-large dealers trade volume

4.3 Matching Opportunity

If each large dealer can hold up to two units in inventory and may possess up to two units of spare inventory capacity, perhaps a more natural assumption is that they can meet up to two investor-buyers and two investor-sellers in each period. We shall demonstrate below how the ranking of the marginal benefits of inventory in Proposition 1 can be left intact.

First, if a large dealer has up to two matching opportunities with investor-buyers and with investor-sellers, respectively, a reasonable matching technology should be such that

1. the probability that a large dealer meets at least one investor-buyer is weakly higher than the probability that a small dealer meets one investor-buyer,

2. the probability that a large dealer meets at least one investor-seller is weakly higher than the probability that a small dealer meets one investor-seller.

If, in addition, the matching technology exhibits diminishing returns in the sense that

3. the probability that a large dealer meets two investor-buyers is weakly lower than the probability that a small dealer meets one investor-buyer,

4. the probability that a large dealer meets two investor-sellers is weakly lower than the probability that a small dealer meets one investor-seller,

then the ranking in Proposition 1 remains.[26]

[26] These assumptions are easily satisfied if the two matching outcomes for the large dealer are independent

The arguments are as follows. First, consider the costs and benefits of acquiring the first unit of inventory in the inter-dealer market for the two types of dealers. Filling up the first unit of capacity in the inter-dealer market is costly to a small dealer as long as he shall meet one investor-seller in the next period but is costly to a large dealer only if he meets two investor-sellers in the next period since if a large dealer meets only one investor-seller, he still has capacity to buy. By (4), the expected cost is higher for the small dealer. The expected benefit is higher for the large dealer – if (1) holds, the large dealer can sell the unit with weakly higher probability. Then, $V_1^{LD} - V_0^{LD} \geq V_1^{SD} - V_0^{SD}$ should follow. The inequality should be strict if either one of the relation in (1) or (4) is strict.

Next, consider the costs and benefits of utilizing the last unit of spare capacity for the two types of dealers. Exhausting one's capacity is costly to a dealer, large or small, as long as the dealer shall meet one or more investor-seller in the next period. By (2), the expected cost is higher for the large dealer. A small dealer benefits from the additional unit of inventory if he meets one investor-buyer while a large dealer benefits only if he meets as many as two investor-buyers. If (3) holds, the expected benefit is higher for the small dealer. Then, $V_1^{SD} - V_0^{SD} \geq V_2^{LD} - V_1^{LD}$ should follow. The inequality should be strict if either one of the relation in (2) or (3) is strict.

5 Comparative Statics

Having shown how the major results of the model hold in more general settings, we now return to the basic model and study the model's comparative statics.

5.1 Asset Supply

Market Tightness and Turnover In Section 3, we remarked that dealers should find it easier to buy but more difficult to sell in a market with more abundant asset supply. We state the formal results in the following proposition.

Proposition 4a *(i) For $e < n^{LD}$, as A increases from $B_L + e/\delta$ at which the Selling Equilibrium first holds, $\partial\theta_{DI}/\partial A = 0$ and $\partial\theta_{ID}/\partial A < 0$. Once A reaches $B_M + e/\delta$ at which the Balanced Equilibrium begins to hold, $\partial\theta_{DI}/\partial A < 0$ and $\partial\theta_{ID}/\partial A < 0$. Finally, when A rises up to and above $B_L + e/\delta$ at which the Buying Equilibrium holds, $\partial\theta_{DI}/\partial A < 0$ and $\partial\theta_{ID}/\partial A = 0$. In the transition from one equilibrium type to another, θ_{DI} and θ_{ID} are continuous. (ii) For $e \in \left[n^{LD}, n^{LD} + \frac{n^{SD}}{2}\right)$ and that $A > \mathcal{B}_M + e/\delta$ at which the Balanced Equilibrium holds, $\partial\theta_{DI}/\partial A < 0$ and $\partial\theta_{ID}/\partial A < 0$.*

Proposition 4a implies that indeed, as A rises, a dealer-buyer meets an investor-seller at a (weakly) higher probability $\mu(\theta_{DI})$ while a dealer-seller meets an investor-buyer at a (weakly)

events. In this case, the probability that a large dealer meets at least one investor-buyer is

$$1 - (1 - \eta(\theta_{ID}))^2 = \eta(\theta_{ID})(2 - \eta(\theta_{ID})) > \eta(\theta_{ID}),$$

where the far-right term is the probability that the small dealer meets one investor-buyer. On the other hand, the probability that a large dealer meets as many as two investor-buyers is $\eta(\theta_{ID})^2 < \eta(\theta_{ID})$.

lower probability $\eta(\theta_{ID})$.

Inter-dealer Trading Volume Strictly speaking, for a given type of equilibrium, the inter-dealer market trading volume is indeterminate given the existence of a continuum of equilibrium. All but one equilibrium among the continuum though involve trades between two dealers both not benefiting from trade at all but only result in the two dealers switching identities. Such trades are spurious and probably should not take place and indeed would not take place at all if trading is costly as we argued before. In studying the model's implications on trading volumes, there is good reason for us to focus solely on the equilibrium with the least trades, where there is at least one dealer benefiting from any given exchange. In this case, trades are driven solely by the inframarginal buyers' or sellers' desire to rebalance inventories. The trading volume (TV) in the Selling, Balanced, and the Buying Equilibria are then given by, respectively,

$$TV = m_1^{SD} + m_2^{LD},$$

$$TV = \begin{cases} m_0^{LD} & A \leq \widetilde{B} + \frac{e}{\delta} \\ m_2^{LD} & A \geq \widetilde{B} + \frac{e}{\delta} \end{cases},$$

$$TV = m_0^{SD} + m_0^{LD}.$$

Proposition 4b *The inter-dealer market trading volume changes non-monotonically with A, as depicted in the table below.*

Selling Equilibrium	Balanced Equilibrium		Buying Equilibrium
	$A \leq \widetilde{B} + \frac{e}{\delta}$ small dealers sell	$A \geq \widetilde{B} + \frac{e}{\delta}$ small dealers buy	
$\frac{\partial TV}{\partial A} > 0$	$\frac{\partial TV}{\partial A} < 0$	$\frac{\partial TV}{\partial A} > 0$	$\frac{\partial TV}{\partial A} < 0$

For $e < n^{LD}$, the trading volume changes continuously as one equilibrium type changes to another, peaking at $TV = e\left(1 - \frac{e}{n^D}\right)$, when the Selling Equilibrium turns into the Balanced Equilibrium and when the Balanced Equilibrium turns into the Buying Equilibrium.

The inter-dealer market is most active then when the asset supply is at a relatively low level but not at the lowest level or at a relatively high level but not at the highest level. The prediction is in stark contrast to what we would obtain in maybe any other models of inter-dealer trades, where a higher asset supply should merely scale up the trading volume. The non-monotonicity in our model is due to the interactions of changing market tightness and optimal dealers' inventories as brought about by variations in A. In particular, say, to begin with, the asset supply is at the lowest level for which very few dealers can buy from investors in a very tight market, whereby, trades in the inter-dealer market, driven by the successful dealer-buyers' need to dispose the inventories they acquire from investors, can only be few and far between. As the asset supply rises, successful dealer-buyers become more numerous, whereby inter-dealer trades begin to pick up. When the asset supply, still at a relatively low level, reaches the point at which the Balanced Equilibrium takes hold, trading in the inter-dealer market changes into being driven by the successful dealer-sellers replenishing their inventories in the market. As A increases further, there are fewer successful dealer-sellers and such needs to replenish inventories weaken. How inter-dealer trades are least numerous when A is at the highest level, begin to pick up as A first goes down, but then declines once A falls down to a low enough level can be understood similarly.

Inter-dealer Trading Prices Proposition 4a shows that dealers find it easier to buy from and harder to sell to investors as A increases. If dealers who have bought from investors tend to sell and dealers who have sold to investors tend to buy afterwards in the inter-dealer market, an increase in A should be followed by an increase in supply and a decline in demand in the inter-dealer market and a concomitant decline in the inter-dealer market price. Besides, when one equilibrium type turns into another, the inter-dealer market price p changes its anchor from one indifference condition to another. As such, a minute change in the asset supply can cause a catastrophic change in p.

Proposition 4c *(i) For $e < n^{LD}$, as A increases from $B_L + e/\delta$ at which the Selling Equilibrium first holds, p is continuously decreasing in A. Once A reaches $B_M + e/\delta$ at which the Balanced Equilibrium begins to hold, there will be a discrete fall in p, followed by further continuous decreases as A increases further. Finally, when A rises up to and above $B_L + e/\delta$ at which the Buying Equilibrium holds, there will be another discrete fall in p all the way to zero. (ii) For $e \in \left[n^{LD}, n^{LD} + \frac{n^{SD}}{2}\right)$ and that $A > \mathcal{B}_M + e/\delta$ at which the Balanced Equilibrium begins to hold, p is continuously decreasing in A throughout.*

5.2 Measure of Large Dealers n^{LD}

In the comparative statics exercises below, we hold constant the total measure of dealers and vary the measure of large dealers. To the extent that there is least diversity among dealers when all dealers are small dealers and when all dealers are large dealers, the comparative statics with respect to n^{LD} we present below can also be interpreted as the impacts of the diversity of dealers in the dealer population on equilibrium outcomes.

Market Tightness and Turnover If more of the dealers are large dealers possessing a two-unit inventory capacity, there will be a greater overall inventory capacity among dealers. Then, first of all, there will tend to be more dealer-buyers. Furthermore, when more dealers are buying from investors, dealers' overall inventory holding tends to increase as well, giving rise to there being more dealer-sellers.

Proposition 5a *Holding fixed n^D, as n^{LD} increases from the smallest admissible value for which the Balanced Equilibrium holds, $\partial\theta_{DI}/\partial n^{LD} > 0$ and $\partial\theta_{ID}/\partial n^{LD} < 0$. If and when the Balanced Equilibrium gives way to the Buying Equilibrium, $\partial\theta_{DI}/\partial n^{LD} > 0$ and $\partial\theta_{ID}/\partial n^{LD} =$*

0. *On the other hand, if and when the Balanced Equilibrium gives way to the Selling Equilibrium,* $\partial\theta_{DI}/\partial n^{LD} = 0$ *and* $\partial\theta_{ID}/\partial n^{LD} = 0$*. The two market tightness are continuous at the point at which the Balanced Equilibrium turns into either the Buying or the Selling Equilibrium.*

The substantive implication of Proposition 5a is that the investor-sellers' matching rate $\eta(\theta_{DI})$ and the investor-buyers' matching rate $\mu(\theta_{ID})$ are both weakly increasing in n^{LD}. In this way, a market with relatively more large dealers functions better at transferring units of the asset from low- to high-valuation investors.

Inter-dealer Trading Volume While the Balanced Equilibrium holds, the trading volume in the inter-dealer market is given by $\max\{m_0^{LD}, m_2^{LD}\}$. Other things equal, there would simply be more dealers searching as L_0s and L_2s as large dealers make up a greater fraction of the dealer population. Meanwhile, with more dealer-buyers and dealer-sellers searching for investors to trade with, each dealer may only buy or sell at a lower probability to result in more large dealers remaining as L_0s and L_2s at the closing of each round of investor-dealer trades.

While the Buying Equilibrium holds, the trading volume equals $m_0^{SD} + m_0^{LD}$. As small dealers are replaced one-for-one by large dealers, dealers leaving the investor-dealer market with an empty inventory should fall in numbers since the large (but not the small) dealers may replenish any inventories they sell to investors in the same period of time by buying from other investors. When the Selling Equilibrium holds, a similar mechanism is at work to result in TV becoming decreasing in n^{LD}.

Proposition 5b *Holding* n^D *fixed, the trading volume in the inter-dealer market is increasing in* n^{LD} *while the Balanced Equilibrium holds. Once the Buying or the Selling Equilibrium takes hold, the trading volume becomes decreasing in* n^{LD}*.* TV *is continuous at where the Balanced Equilibrium turns into either the Buying or the Selling Equilibrium and reaches the highest level equal to* $e\left(1 - \frac{e}{n^D}\right)$ *at the point of transition.*

Proposition 5b shows that for any level of asset supply, the inter-dealer market is least active when there is little diversity in the dealer population with n^{LD} either at the lowest or at the highest level. With more diversity as when n^{LD} is at some intermediate level, the market becomes more active. Trading in the inter-dealer market in our model then is driven more by the heterogeneity of dealers than by some dealers possessing more than a unit of inventory capacity.

Inter-dealer Trading Prices

Proposition 5c *Holding fixed* n^D*, as* n^{LD} *rises, if and when the Balanced Equilibrium gives way to the Buying Equilibrium,* p *falls by a discrete amount down to zero; if and when the Balanced Equilibrium gives way to the Selling Equilibrium,* p *jumps up by a discrete amount and stays at a given level thereafter for all* n^{LD}*.*

During which the Balanced Equilibrium holds, we find that, through a set of numerical analysis,[27] p is decreasing in n^{LD}, both when the Balanced Equilibrium will turn into the Selling Equilibrium and when it will turn into the Buying Equilibrium. A corollary is that p can be non-monotonic with respect to increase in n^{LD} when the Balanced Equilibrium turns into the Selling Equilibrium.

Investor-Dealer Market Prices For brevity, in Propositions 4c and 5c, we have not extended the analysis to also checking how prices in the investor-dealer market may vary with A and n^{LD}. In Propositions A1 and A2 and the ensuing discussions in Appendix 7.2, we show that the dealers' ask and bid prices in the market do turn out to vary with A and n^{LD} in the just the same ways that the inter-dealer market price does.

6 Conclusion

In this paper, by means of a parsimonious extension of the standard asset market search model, we study inter-dealer trades among heterogeneous dealers in OTC markets motivated by inventory risk concerns. We depart from earlier inventory risk models by all assuming that all traders are risk neutral. Even so, the dealers benefit from trading among one another to eliminate the risks of carrying an insufficient inventory and an insufficient spare inventory capacity for their trading needs with investors.

The model yields a rich set of testable implications. First and foremost, our analysis shows that the apparently obvious notion that dealers who are able to and indeed tend to hold a larger inventory should provide inventory for dealers having a smaller inventory capacity only holds up when inventory is relatively abundant. But at such times, those smaller-capacity dealers should need the inventory the least. In contrast, in our model, it is the latter group of dealers who trade to provide immediacy for the larger-capacity dealers, selling to them when they need inventory the most and buying from them when they need the spare capacity the most. If the large dealers are interpreted as the core dealers and the small dealers are interpreted as the peripheral dealers in a core-periphery trading network, our analysis suggests that the peripheral dealers trade to provide immediacy for the core dealers, contrary to the common perception of the roles the two types of dealers should play in inter-dealer trading.

We show in Appendix 7.3 that such features of equilibrium inter-dealer trades actually help attain constrained efficiency. In the planning optimum, inventories are allocated to dealers to enable high-valuation investors to acquire the asset most rapidly and to enable units of the asset to be transferred from low-valuation investors to dealers the quickest, thereby facilitating the eventual sales to the high-valuation investors. In the competitive inter-dealer market, inventories and spare capacities are allocated to dealers who value them the most – the very dealers who have the best use of them for trading with investors. Perhaps not surprisingly, the equilibrium allocations coincide with the constrained optimum allocations. More interestingly, our analysis suggests that for efficiency, the small peripheral dealers should indeed trade to provide immediacy for the large core dealers.

[27] For brevity, the details of the numerical analysis, which are available upon request from the authors, are omitted.

7 Appendix

7.1 The Frictional Inter-Dealer Market Model

The model is set in continuous time in which a dealer meets another randomly selected dealers at the rate α per time unit. All other notations have the same meanings as for the main model.

7.1.1 Value Functions

To define the value functions, we rule out a priori any exchanges between two dealers that merely result in the two dealers concerned switching states as such exchanges cannot give rise to a positive surplus.

Small dealers An S_0, who can only buy, meets an investor-seller at the rate $\mu(\theta_{DI})$ and another dealer at the rate α. Among all dealers that the S_0 may meet, there can be a potentially profitable exchange only if the counterparty is an L_1 or an L_2. By Lemma 5, all investor-dealer trades yield non-negative surpluses. Then,

$$\begin{aligned} rV_0^{SD} &= \mu(\theta_{DI})\left(V_1^{SD} - V_0^{SD} - p_{S_0,I_S}\right) + \alpha\left\{\frac{n_1^{LD}}{n^D}\max\left\{-p_{S_0,L_1} + V_1^{SD} - V_0^{SD}, 0\right\}\right. \\ &\quad \left. + \frac{n_2^{LD}}{n^D}\max\left\{-p_{S_0,L_2} + V_1^{SD} - V_0^{SD}, 0\right\}\right\}. \end{aligned}$$

An S_1, who can only sell, meets an investor-buyer at the rate $\eta(\theta_{ID})$. The S_1 may also sell to an L_0 or an L_1. Then,

$$\begin{aligned} rV_1^{SD} &= \eta(\theta_{ID})\left(p_{I_B,S_1} + V_0^{SD} - V_1^{SD}\right) + \alpha\left\{\frac{n_0^{LD}}{n^D}\max\left\{p_{L_0,S_1} + V_0^{SD} - V_1^{SD}, 0\right\}\right. \\ &\quad \left. + \frac{n_1^{LD}}{n^D}\max\left\{p_{L_1,S_1} + V_1^{SD} - V_0^{SD}, 0\right\}\right\}. \end{aligned}$$

Large dealers An L_0 may buy from an investor-seller, an S_1, or an L_2. Then,

$$\begin{aligned} rV_0^{LD} &= \mu(\theta_{DI})\left(V_1^{LD} - V_0^{LD} - p_{L_0,I_S}\right) + \alpha\left\{\frac{n_1^{SD}}{n^D}\max\left\{-p_{L_0,S_1} + V_1^{LD} - V_0^{LD}, 0\right\}\right. \\ &\quad \left. + \frac{n_2^{LD}}{n^D}\max\left\{-p_{L_0,L_2} + V_1^{LD} - V_0^{LD}, 0\right\}\right\}. \end{aligned}$$

An L_1 may buy from an investor-seller and sell to an investor-buyer. Among dealers, he may sell to an S_0, buy from an S_1, and either buy from or sell to another L_1. Then,

$$\begin{aligned} rV_1^{LD} &= \mu(\theta_{DI})\left(V_2^{LD} - V_1^{LD} - p_{L_1,I_S}\right) + \eta(\theta_{ID})\left(p_{I_B,L_1} + V_0^{LD} - V_1^{LD}\right) \\ &\quad + \alpha\left\{\frac{n_0^{SD}}{n^D}\max\left\{p_{S_0,L_1} + V_0^{LD} - V_1^{LD}, 0\right\} + \frac{n_1^{SD}}{n^D}\max\left\{-p_{L_1,S_1} + V_2^{LD} - V_1^{LD}, 0\right\}\right. \\ &\quad \left. + \frac{n_1^{LD}}{n^D}\max\left\{-p_{L_1,L_1} + V_2^{LD} - V_1^{LD}, p_{L_1,L_1} + V_0^{LD} - V_1^{LD}, 0\right\}\right\}. \end{aligned}$$

An L_2, who may only sell, meets an investor-buyer at the rate $\eta(\theta_{ID})$. Among dealers, he may sell to an S_0 or an L_0. Then,

$$rV_2^{LD} = \eta(\theta_{ID})\left(p_{I_B,L_2} + V_1^{LD} - V_2^{LD}\right) + \alpha\left\{\frac{n_0^{SD}}{n^D}\max\left\{p_{S_0,L_2} + V_1^{LD} - V_2^{LD}, 0\right\} + \frac{n_0^{LD}}{n^D}\max\left\{p_{L_0,L_2} + V_1^{LD} - V_2^{LD}, 0\right\}\right\}.$$

Investors An investor-buyer may buy from an S_1, an L_1, or an L_2. Then,

$$rU^B = \mu(\theta_{ID})\left(U_H^{ON} - U^B - \frac{n_1^{SD}}{n_S^D}p_{I_B,S_1} - \frac{n_1^{LD}}{n_S^D}p_{I_B,L_1} - \frac{n_2^{LD}}{n_S^D}p_{I_B,L_2}\right),$$

where

$$rU_H^{ON} = \upsilon + \delta\left(U_L^{ON} - U_H^{ON}\right).$$

An investor-seller may sell to an S_0, an L_0, or an L_1. Then,

$$rU_L^{ON} = \eta(\theta_{DI})\left(\frac{n_1^{SD}}{n_B^D}p_{S_0,I_S} + \frac{n_0^{LD}}{n_B^D}p_{L_0,I_S} + \frac{n_1^{LD}}{n_B^D}p_{L_1,I_S} - U_L^{ON}\right).$$

7.1.2 Equilibrium Conditions

To begin, since, by Proposition 3, among all dealers, S_1s only sell to L_0s whereas S_0s only buy from L_2s, in the steady state in which $\dot{n}_0^{SD} = \dot{n}_1^{SD} = 0$,

$$n_1^{SD}\left(\eta(\theta_{ID}) + \alpha\frac{n_0^{LD}}{n^D}\right) = n_0^{SD}\left(\mu(\theta_{ID}) + \alpha\frac{n_2^{LD}}{n^D}\right). \quad (37)$$

Also, where L_1s do not trade in the inter-dealer market and that L_0s buy from S_1s and L_2s, whereas L_2s sell to S_0s and L_0s, the equations for $\dot{n}_0^{LD} = 0$ and $\dot{n}_2^{LD} = 0$ specialize to, respectively,

$$n_1^{LD}\eta(\theta_{ID}) = n_0^{LD}\left(\mu(\theta_{DI}) + \alpha\frac{n_1^{SD} + n_2^{LD}}{n^D}\right), \quad (38)$$

$$n_1^{LD}\mu(\theta_{DI}) = n_2^{LD}\left(\eta(\theta_{ID}) + \alpha\frac{n_0^{SD} + n_0^{LD}}{n^D}\right). \quad (39)$$

Given $\left\{n^{SD}, n^{LD}, A, e, \delta\right\}$, a steady-state equilibrium consists of the respective non-negative values of n_0^{SD}, n_1^{SD}, n_0^{LD}, n_1^{LD}, n_2^{LD}, n_H^{ON}, n_L^{ON} and n_B^I that satisfy the accounting identities in (22)-(24) and the steady-state conditions for n_H^{ON}, n_L^{ON} and n_B^I in (27)-(29), which are common for both the competitive and frictional inter-dealer market models, and (37)-(39) above.

First, by (24),

$$n_1^{SD} = A - \frac{e}{\delta} - \frac{e}{\eta(\theta_{DI})} - n_1^{LD} - 2n_2^{LD}, \quad (40)$$

from which it follows that

$$n_1^{LD} + n_1^{SD} + n_2^{LD} = A - \frac{e}{\delta} - \frac{e}{\eta(\theta_{DI})} - n_2^{LD}.$$

Substitute the equation into (25) to yield

$$n_2^{LD} = A - \frac{e}{\delta} - \frac{e}{\eta(\theta_{DI})} - \frac{e}{\eta(\theta_{ID})}. \tag{41}$$

Next, by (22), (23), (40), and (41),

$$n_0^{SD} + n_0^{LD} + n_1^{LD} = n^D - \frac{e}{\eta(\theta_{ID})} + n_1^{LD}.$$

Substitute the equation into (26) to yield

$$n_1^{LD} = \frac{e}{\mu(\theta_{DI})} + \frac{e}{\eta(\theta_{ID})} - n^D. \tag{42}$$

It is then straightforward to derive the following,

$$n_0^{SD} = A - \frac{e}{\delta} - n^{LD} - \frac{e}{\eta(\theta_{DI})} - \frac{e}{\eta(\theta_{ID})} + \frac{e}{\mu(\theta_{DI})}, \tag{43}$$

$$n_1^{SD} = n^D - A + \frac{e}{\delta} + \frac{e}{\eta(\theta_{DI})} + \frac{e}{\eta(\theta_{ID})} - \frac{e}{\mu(\theta_{DI})}, \tag{44}$$

$$n_0^{LD} = n^{LD} + n^D - A + \frac{e}{\delta} - \frac{e}{\mu(\theta_{DI})} + \frac{e}{\eta(\theta_{DI})}. \tag{45}$$

Any $\{\theta_{ID}, \theta_{DI}\}$ pair satisfying (37) and (38) with the measures of dealers given by (40)-(45), where the resulting $n_0^{SD}, n_1^{SD} \in [0, n^{SD}]$ and $n_0^{LD}, n_1^{LD}, n_2^{LD} \in [0, n^{LD}]$, is a steady-state equilibrium.

7.2 Dealers' Bid and Ask Prices

Lemma A1 *In all three types of equilibrium, the dealers' bid price; i.e., the price at which investors sell to dealers is given by*

$$p_{I_S} = \frac{1 - \beta + \beta\eta(\theta_{DI})}{2(1-\beta) + \beta\eta(\theta_{DI})} p, \tag{46}$$

whereas the dealers' ask prices; i.e., the prices at which investors buy from dealers in the Selling, Balanced, and Buying Equilibria are given by, respectively,

$$p_{I_B} = \left(1 + \frac{1-\beta}{\beta\eta(\theta_{ID})}\right) p, \tag{47}$$

$$p_{I_B} = \left(1 + \frac{\left(1 - \beta + \beta\frac{\mu(\theta_{DI})}{2}\right)(4(1-\beta) + 2\beta\eta(\theta_{DI})) - \beta^2\mu(\theta_{DI})\eta(\theta_{DI})}{\beta\eta(\theta_{ID})(4(1-\beta) + 2\beta\eta(\theta_{DI}))}\right) p, \tag{48}$$

$$p_{I_B} = \frac{\beta(1-\beta)\upsilon}{(1-\beta+\beta\delta)(2(1-\beta)+\mu(\theta_{ID})\beta)}. \tag{49}$$

Proposition A1 *For $e \leq n^{LD}$, as A increases from $B_L + e/\delta$ at which the Selling Equilibrium first holds, p_{I_S} and p_{I_B} are continuously decreasing in A. Once A reaches $B_M + e/\delta$ at which the Balanced Equilibrium begins to hold, there will be discrete falls in the two prices. While the Balanced Equilibrium holds, p_{I_S} is continuously decreasing in A. And then finally, when A rises up to and above $B_L + e/\delta$, at which the Buying Equilibrium holds, there will be further discrete falls in the two prices – p_{I_S} all the way to zero and p_{I_B} to some positive value. Thereafter, the two prices do not vary with A any longer. For $e \in \left(n^{LD}, n^{LD} + \frac{n^{SD}}{2}\right]$ and that $A > B_M + e/\delta$ at which the Balanced Equilibrium begins to hold, p_{I_S} is likewise continuously decreasing in A.*

The Proposition leaves out how p_{I_B} may vary with A in the Balanced Equilibrium as it does not seem possible to sign $\partial p_{I_B}/\partial A$ in said equilibrium. Our numerical analyses do reveal, however, that p_{I_B} does decline in A while the Balanced Equilibrium holds,[28] just as p and p_{I_S} do.

For the smallest admissible n^{LD}, the market starts off in a Balanced Equilibrium, in which p, as our numerical analyses in the main text indicate, tends to decline with increases in n^{LD}. In the same numerical analyzes, we find that p_{I_S} and p_{I_B} follow the same tendency.

Proposition A2 *Holding fixed n^D, as n^{LD} rises, if and when the Balanced Equilibrium gives way to the Buying Equilibrium, both p_{I_S} and p_{I_B} fall by some discrete amount – p_{I_S} to zero and p_{I_B} to some positive value; if and when the Balanced Equilibrium gives way to the Selling Equilibrium, both p_{I_S} and p_{I_B} jump up by some discrete amount. In either the Buying or the Selling Equilibrium, the two prices do not vary with n^{LD}.*

7.3 Efficient Decentralized Market Trades

A social planner maximizes the discounted flow payoffs for investors over time from the ownership of the asset given by,

$$W = \max\left\{\sum_{t=0}^{\infty} \beta^t n_H^{ON}(t)\,\upsilon\right\}, \tag{50}$$

subject to the same search and matching frictions that agents in the model face.

A priori, the equilibrium trades in the frictional investor-dealer market are constrained efficient where any trades with a positive surplus, but only such trades, will take place with the terms of trade in the bilateral meetings reached via Nash Bargaining. Specifically, any investor-buyer and dealer-seller trade is efficient with the former, but not the latter, deriving the flow payoff υ in holding a unit of the asset. But then a dealer-seller becomes a dealer-seller

[28] Under the same parameter configurations as for the numerical analyses preceding Proposition 4c, except that n^{LD} is fixed at 0.846.

in the first place only by acquiring the asset from an investor-seller. Then, any and all trades between an investor-seller and a dealer-buyer are also efficient.

This means that it suffices for us to ask how the planner may wish to allocate units of inventory among the dealers in each period after the investor-dealer trades are completed and whether the allocation coincides with the allocation that falls out from the inter-dealer market in equilibrium.

Lemma A2 *In the steady state of the planner's solution, units of inventory not held by investors are allocated to dealers to maximize the measures of dealer-sellers (dealers who hold inventory) and dealer-buyers (dealers who possess spare capacity). To maximize the measure of dealer-sellers, first allocate one unit each to either small or large dealers, and then allocate any remaining inventory to the large dealers. To maximize the measure of dealer-buyers, first allocate one unit each to large dealers, and then allocate any remaining inventory to either large or small dealers. The two objectives are then attained simultaneously by allocating inventory in the following order: (1) one unit each to large dealers; (2) if there remains any inventory, then one unit each to small dealers; (3) if there remains any inventory, one more unit each to large dealers.*

Proposition A3 *In the steady state, the allocations from the decentralized market trades coincide with the planning optimum.*

7.4 Proofs of Lemmas and Propositions

Proof of Lemma 1 Assuming equal bargaining power, the respective prices an investor-buyer pays to an S_1, an L_1, and an L_2 satisfy,

$$\beta\left(U_H^{ON} - U^B\right) - p_{I_B,S_1} = W_0^{SD} - W_1^{SD} + p_{I_B,S_1}, \tag{51}$$

$$\beta\left(U_H^{ON} - U^B\right) - p_{I_B,L_1} = W_0^{LD} - W_1^{LD} + p_{I_B,L_1}, \tag{52}$$

$$\beta\left(U_H^{ON} - U^B\right) - p_{I_B,L_2} = W_1^{LD} - W_2^{LD} + p_{I_B,L_2}. \tag{53}$$

On the other hand, the respective prices an investor-seller receives from selling to an S_0, an L_0, and an L_1 satisfy,

$$p_{S_0,I_S} - \beta U_L^{ON} = W_1^{SD} - W_0^{SD} - p_{S_0,I_S}, \tag{54}$$

$$p_{L_0,I_S} - \beta U_L^{ON} = W_1^{LD} - W_0^{LD} - p_{L_0,I_S}, \tag{55}$$

$$p_{L_1,I_S} - \beta U_L^{ON} = W_2^{LD} - W_1^{LD} - p_{L_1,I_S}. \tag{56}$$

Notice that

$$W_1^{SD} = W_0^{SD} + p, \tag{57}$$

$$W_0^{LD} = W_1^{LD} - p, \tag{58}$$

$$W_2^{LD} = W_1^{LD} + p. \tag{59}$$

The lemma then follows from (51)-(56).

Proof of Proposition 1 Substitute (15) and (17) and (57)-(59) into the value functions (6) and (7) and (11)-(13),

$$V_0^{SD} = W_0^{SD} + \frac{\mu(\theta_{DI})}{2}\left(p - \beta U_L^{ON}\right), \tag{60}$$

$$V_1^{SD} = W_0^{SD} + \left(1 - \frac{\eta(\theta_{ID})}{2}\right)p + \frac{\eta(\theta_{ID})}{2}\beta\left(U_H^{ON} - U^B\right), \tag{61}$$

$$V_0^{LD} = W_1^{LD} - \left(1 - \frac{\mu(\theta_{DI})}{2}\right)p - \frac{\mu(\theta_{DI})}{2}\beta U_L^{ON}, \tag{62}$$

$$V_1^{LD} = W_1^{LD} + \frac{\mu(\theta_{DI}) - \eta(\theta_{ID})}{2}p - \frac{\mu(\theta_{DI})}{2}\beta U_L^{ON} + \frac{\eta(\theta_{ID})}{2}\beta\left(U_H^{ON} - U^B\right), \tag{63}$$

$$V_2^{LD} = W_1^{LD} + \left(1 - \frac{\eta(\theta_{ID})}{2}\right)p + \frac{\eta(\theta_{ID})}{2}\beta\left(U_H^{ON} - U^B\right). \tag{64}$$

We can then calculate

$$\left(V_1^{LD} - V_0^{LD}\right) - \left(V_1^{SD} - V_0^{SD}\right) = \frac{\mu(\theta_{DI})}{2}\left(p - \beta U_L^{ON}\right), \tag{65}$$

$$\left(V_1^{SD} - V_0^{SD}\right) - \left(V_2^{LD} - V_1^{LD}\right) = \frac{\eta(\theta_{ID})}{2}\left(\beta\left(U_H^{ON} - U^B\right) - p\right). \tag{66}$$

Notice that the terms inside the brackets in (65) and (66) denote, respectively, the surpluses of trade between an investor-seller and any dealer-buyer and between an investor-buyer and any dealer-seller in (14) and (16). If either of the two is negative, there cannot be any trade in equilibrium between investors and dealers in the steady state.

Proof of Lemma 2 Substitute (15) into (3) and rearrange,

$$U_L^{ON} = \frac{\frac{\eta(\theta_{DI})}{2}}{1 - \beta + \beta\frac{\eta(\theta_{DI})}{2}}p. \tag{67}$$

Substitute the equation into (2) and rearrange,

$$U_H^{ON} = \frac{\left(1 - \beta + \frac{\eta(\theta_{DI})}{2}\beta\right)\upsilon + \beta\delta\frac{\eta(\theta_{DI})}{2}p}{(1 - \beta + \beta\delta)\left(1 - \beta + \frac{\eta(\theta_{DI})}{2}\beta\right)} \tag{68}$$

Substitute (17) into (1) and rearrange,

$$U^B = \frac{\frac{\mu(\theta_{ID})}{2}\left(\beta U_H^{ON} - p\right)}{1 - \beta + \beta\frac{\mu(\theta_{ID})}{2}}. \tag{69}$$

Substituting from (68),

$$U^B = \frac{\mu(\theta_{ID})}{2}\frac{\beta\left(1 - \beta + \frac{\eta(\theta_{DI})}{2}\beta\right)\upsilon - (1 - \beta)\left(1 - \beta + \beta\delta + \frac{\eta(\theta_{DI})}{2}\beta\right)p}{(1 - \beta + \beta\delta)\left(1 - \beta + \beta\frac{\mu(\theta_{ID})}{2}\right)\left(1 - \beta + \frac{\eta(\theta_{DI})}{2}\beta\right)}. \tag{70}$$

Then, by (68) and (70),

$$U_H^{ON} - U^B = \frac{\left(\frac{\mu(\theta_{ID})}{2}(1-\beta+\beta\delta)\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)+\beta\delta\frac{\eta(\theta_{DI})}{2}(1-\beta)\right)p+\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)(1-\beta)\upsilon}{(1-\beta+\beta\delta)\left(1-\beta+\frac{\mu(\theta_{ID})}{2}\beta\right)\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)} \tag{71}$$

Set $p = \beta\left(V_1^{LD} - V_0^{LD}\right)$ and by (62) and (63),

$$p = \frac{\beta\frac{\eta(\theta_{ID})}{2}}{1-\beta+\beta\frac{\eta(\theta_{ID})}{2}}\beta\left(U_H^{ON} - U^B\right). \tag{72}$$

Then use (71) to obtain

$$p = \frac{\beta^2\frac{\eta(\theta_{ID})}{2}\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)\upsilon}{\left((1-\beta+\beta\delta)\left(1-\beta+\beta\frac{\mu(\theta_{ID})}{2}\right)+\beta\frac{\eta(\theta_{ID})}{2}(1-\beta)\right)\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)+\delta\beta^2\frac{\eta(\theta_{ID})}{2}(1-\beta)}. \tag{73}$$

Given the positivity of p in (73) and by (67) and (72),

$$0 < \beta U_L^{ON} < p < \beta\left(U_H^{ON} - U^B\right).$$

Next, set $p = \beta\left(V_1^{SD} - V_0^{SD}\right)$ and by (60) and (61),

$$p = \frac{\frac{\eta(\theta_{ID})}{2}\beta^2\left(U_H^{ON} - U^B\right) + \frac{\mu(\theta_{DI})}{2}\beta^2 U_L^{ON}}{1-\beta+\beta\left(\frac{\eta(\theta_{ID})}{2}+\frac{\mu(\theta_{DI})}{2}\right)}. \tag{74}$$

Then use (67) and (71) to obtain

$$p = \frac{\frac{\eta(\theta_{ID})}{2}\beta^2\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)\upsilon}{\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta+\beta\frac{\mu(\theta_{DI})}{2}\right)\left(1-\beta+\frac{\mu(\theta_{ID})}{2}\beta\right)(1-\beta+\beta\delta)+\left(1-\beta+\beta\delta+\frac{\eta(\theta_{DI})}{2}\beta\right)\beta\frac{\eta(\theta_{ID})}{2}(1-\beta)}. \tag{75}$$

Given the positivity of p in (75) and by (67) and (74),

$$0 < \beta U_L^{ON} < p < \beta\left(U_H^{ON} - U^B\right).$$

Finally, set $p = \beta\left(V_2^{LD} - V_1^{LD}\right)$ and by (63) and (64),

$$p = \frac{\beta\frac{\mu(\theta_{DI})}{2}}{1-\beta+\beta\frac{\mu(\theta_{DI})}{2}}\beta U_L^{ON}. \tag{76}$$

With (67),

$$p = \beta U_L^{ON} = 0.$$

Next, by (71),

$$U_H^{ON} - U^B = \frac{(1-\beta)\upsilon}{(1-\beta+\beta\delta)\left(1-\beta+\frac{\mu(\theta_{ID})}{2}\beta\right)} > 0. \tag{77}$$

Thus,

$$0 = \beta U_L^{ON} = p < \beta\left(U_H^{ON} - U^B\right).$$

Proof of Proposition 2 Before proceeding to prove the Proposition, it is useful to establish the following.

Remark 1 For $x \leq 1$, $x < \eta^{-1}(x)$.
Proof. Given $\mu(x) = \eta(x)/x$ and that $\mu(x) < 1$, $\eta(x) < x$. And then for $x \leq 1$, the last condition implies $x < \eta^{-1}(x)$.

Remark 2 For $x \geq 1$, $x > \mu^{-1}\left(\frac{1}{x}\right)$.
Proof. Given that $\eta(x) = x\mu(x)$ and that $\eta(x) < 1$, $\mu(x) < \frac{1}{x}$. And then for $x \geq 1$, the last condition implies $x > \mu^{-1}\left(\frac{1}{x}\right)$.

Remark 3 For $e \leq n$, $\frac{n}{\mu^{-1}\left(\frac{e}{n}\right)}$ is decreasing in n.
Proof. By differentiation.

We can manipulate (27)-(29) and obtain,

$$n_H^{ON} = \frac{e}{\delta}, \tag{78}$$

$$n_L^{ON} = \frac{e}{\eta(\theta_{DI})}, \tag{79}$$

$$n_B^{I} = \frac{e}{\mu(\theta_{ID})}. \tag{80}$$

Selling Equilibrium In the Selling Equilibrium, $n_2^{LD} = n_1^{SD} = 0$ and $n_0^{SD} = n^{SD}$. Then, together with (79) and (80), the two market tightness equations, (25) and (26), specialize to, respectively,

$$\eta(\theta_{ID}) = \frac{e}{n_1^{LD}}, \tag{81}$$

$$\mu(\theta_{DI}) = \frac{e}{n^{D}}. \tag{82}$$

By (24), (78), (79), and that $n_1^{SD} = n_2^{LD} = 0$ in the Selling Equilibrium,

$$n_1^{LD} = A - \frac{e}{\delta} - \frac{e}{\eta(\theta_{DI})}. \tag{83}$$

Substitute the equation into (81) and rearrange,

$$\eta(\theta_{ID}) = \frac{\eta(\theta_{DI})e}{(A - \frac{e}{\delta})\eta(\theta_{DI}) - e}. \tag{84}$$

Once θ_{DI} is known from (82), the above uniquely gives θ_{ID}. For θ_{DI} from (82) to be a valid equilibrium, it has to be such that the resulting: (a) $\eta(\theta_{ID}) \in (0,1)$, as given by (84) and (b) $n_1^{LD} \in (e, n^{LD}]$, as given by (83). For (b) to be satisfied, $e < n^{LD}$ must hold.

By (84), for $\eta(\theta_{ID}) \in (0,1)$,

$$\frac{e}{\eta(\theta_{DI})} < A - \frac{e}{\delta} - e. \tag{RS.1}$$

Substituting from (82) and rearranging, $(RS.1)$ holds if

$$A - \frac{e}{\delta} > e + \frac{n^D}{\mu^{-1}\left(\frac{e}{n^D}\right)} = B_S. \tag{AS.1}$$

Given that $\eta(\theta_{ID}) < 1$, $n_1^{LD} > e$ holds for sure. By (83), for $n_1^{LD} \leq n^{LD}$,

$$\frac{e}{\eta\left(\theta_{DI}\right)} \geq A - \frac{e}{\delta} - n^{LD}. \tag{RS.2}$$

Given θ_{DI} from (82), the above becomes,

$$A - \frac{e}{\delta} \leq n^{LD} + \frac{n^D}{\mu^{-1}\left(\frac{e}{n^D}\right)} = B_M. \tag{AS.2}$$

Notice that for $(RS.1)$ and $(RS.2)$ to be satisfied at the same time, it has to be such that $e < n^{LD}$, in which case (82) and the two conditions $(AS.1)$ and $(AS.2)$ are guaranteed to be well-defined.

Substituting (31), (32), and (34) into the inter-dealer market equilibrium condition (18),

$$\left(n^{SD} + n_1^{LD}\right)\mu(\theta_{DI}) \leq n_0^{LD}\left(1 - \mu(\theta_{DI})\right) + n_1^{LD}\eta(\theta_{ID}).$$

And then with (81)-(84), the condition can be shown to simplify to $(RS.2)$.

To sum up, the Selling Equilibrium holds if and only if $(AS.1)$ and $(AS.2)$ hold; i.e., $A - \frac{e}{\delta} \in (B_S, B_M]$, in addition to $e < n^{LD}$.

Buying Equilibrium In the Buying Equilibrium, $n_0^{LD} = n_0^{SD} = 0$ and $n_1^{SD} = n^{SD}$. Then, together with (79) and (80), the two market tightness equations, (25) and (26), specialize to, respectively,

$$\eta\left(\theta_{ID}\right) = \frac{e}{n^D}, \tag{86}$$

$$\mu\left(\theta_{DI}\right) = \frac{e}{n_1^{LD}}. \tag{87}$$

By (24), (78), (79), and that $n_0^{SD} = n_0^{LD} = 0$ in the Buying Equilibrium,

$$n_1^{LD} = n^{SD} + 2n^{LD} - A + \frac{e}{\delta} + \frac{e}{\eta(\theta_{DI})}. \tag{88}$$

Substitute the equation into (87),

$$e(\theta_{DI} - 1) - \eta(\theta_{DI})\left(n^{SD} + 2n^{LD} - A + \frac{e}{\delta}\right) = 0, \tag{89}$$

which is an equation in θ_{DI} alone. It is straightforward to verify that there is a unique positive solution of θ_{DI} to the equation, and that the LHS is positive (negative) for θ_{DI} above (below) the solution of the equation. For the solution to be a valid equilibrium, it has to be such that the resulting $n_1^{LD} \in (e, n^{LD}]$, as given by (88). Then $e < n^{LD}$ must be satisfied.

Rearranging (88), $n_1^{LD} \leq n^{LD}$ if and only if

$$\frac{e}{\eta(\theta_{DI})} \leq A - \frac{e}{\delta} - n^D. \tag{RB.1}$$

A necessary condition for the equation to hold is that

$$A - \frac{e}{\delta} - n^D \geq e. \tag{AB1.a}$$

Then, $(RB.1)$ holds if the LHS of (89) is non-positive when evaluated at $\theta_{DI} = \eta^{-1}\left(\frac{e}{A-\frac{e}{\delta}-n^D}\right)$. Where $e < n^{LD}$, which is necessary for the RHS of (87) and also guarantees the RHS of (86) to be bounded below one, the condition reads

$$A - \frac{e}{\delta} \geq n^D + \frac{n^{LD}}{\mu^{-1}\left(\frac{e}{n^{LD}}\right)} = B_L, \tag{AB1.b}$$

which subsumes $(AB1.a)$.

Rearranging (88), $n_1^{LD} > e$ if and only if

$$A - \frac{e}{\delta} + e - n^{SD} - 2n^{LD} < \frac{e}{\eta(\theta_{DI})} \tag{RB.2}$$

Hence, if

$$A - \frac{e}{\delta} - n^{SD} - 2n^{LD} \leq 0,$$

then $(RB.2)$ holds for sure. Otherwise, the condition holds if the LHS of (89) is positive when evaluated at $\theta_{DI} = \eta^{-1}\left(\frac{e}{A-\frac{e}{\delta}+e-n^{SD}-2n^{LD}}\right)$; i.e.,

$$\eta^{-1}\left(\frac{e}{A - \frac{e}{\delta} + e - n^{SD} - 2n^{LD}}\right) > \frac{e}{A - \frac{e}{\delta} + e - n^{SD} - 2n^{LD}}.$$

But the condition is guaranteed to hold by Remark 1.

Substituting (30), (32), and (34) into the inter-dealer market equilibrium condition (21),

$$\left(n^{SD} + n_1^{LD}\right)\eta(\theta_{ID}) \leq n_1^{LD}\mu(\theta_{DI}) + n_2^{LD}\left(1 - \eta(\theta_{ID})\right).$$

Then, by (86)-(89), the condition can be shown to simplify to $(RB.1)$.

To sum up, the Buying Equilibrium holds if and only if $(AB.1b)$ holds; i.e., $A - \frac{e}{\delta} \geq B_L$, in addition to $e < n^{LD}$.

Balanced Equilibrium In the Balanced Equilibrium, $n_0^{LD} = n_2^{LD} = 0$ and $n_1^{LD} = n^{LD}$. Then, together with (79) and (80), the two market tightness equations, (25) and (26), specialize to, respectively,

$$\eta\left(\theta_{ID}\right) = \frac{e}{n_1^{SD} + n^{LD}}, \tag{91}$$

$$\mu\left(\theta_{DI}\right) = \frac{e}{n_0^{SD} + n^{LD}}. \tag{92}$$

By (24), (78), and (79), and that $n_0^{LD} = n_2^{LD} = 0$ in the Balanced Equilibrium,

$$n_1^{SD} = A - n^{LD} - \frac{e}{\delta} - \frac{e}{\eta(\theta_{DI})}, \tag{93}$$

and therefore

$$n_0^{SD} = n^D - A + \frac{e}{\delta} + \frac{e}{\eta(\theta_{DI})}. \tag{94}$$

Substituting (93) and (94) into (91) and (92), respectively, and rearranging,

$$\eta(\theta_{ID}) = \frac{\eta(\theta_{DI})\, e}{(A - e/\delta)\,\eta(\theta_{DI}) - e}, \tag{95}$$

$$e(\theta_{DI} - 1) - \eta(\theta_{DI})\left(n^D + n^{LD} - A + \frac{e}{\delta}\right) = 0, \tag{96}$$

which are respectively the same equations that give θ_{ID} in the Selling Equilibrium in (84) and θ_{DI} in the Buying Equilibrium in (89).

For now, we restrict attention to where $e < \frac{n^D + n^{LD}}{2}$. Later on, we will verify that the condition is necessary for the existence of the Balanced Equilibrium. For the solution of (96) to be a valid equilibrium, it has to be such that (a) n_0^{SD} as given by (94) satisfies $n_0^{SD} \in \left[0, n^{SD}\right]$ for $e < n^{LD}$ and $n_0^{SD} \in \left(e - n^{LD}, n^D - e\right)$ for $e \in \left[n^{LD}, \frac{n^D + n^{LD}}{2}\right)$ and (b) $\eta(\theta_{ID}) \in (0,1)$, as given by (95).

By (94), where $e < n^{LD}$, for $n_0^{SD} \geq 0$,

$$\frac{e}{\eta(\theta_{DI})} \geq A - \frac{e}{\delta} - n^D \tag{RBA.1}$$

has to hold. The condition is guaranteed to hold if

$$A - \frac{e}{\delta} - n^D \leq e. \tag{ABA.1a}$$

Otherwise, $(RBA.1)$ can only hold if the LHS of (96) is non-negative when evaluated at $\theta_{DI} = \eta^{-1}\left(\frac{e}{A - \frac{e}{\delta} - n^D}\right)$; i.e.,

$$A - \frac{e}{\delta} \leq n^D + \frac{n^{LD}}{\mu^{-1}\left(\frac{e}{n^{LD}}\right)} = B_L. \tag{ABA.1b}$$

Note that $(RBA.1)$ holds if either $(ABA.1a)$ or $(ABA.1b)$ is satisfied. Given that $\frac{n^{LD}}{\mu^{-1}\left(\frac{e}{n^{LD}}\right)} < e$ by Remark 2, however, the latter condition subsumes the former one to begin with. Next, for $n_0^{SD} \leq n^{SD}$, by (94),

$$\frac{e}{\eta(\theta_{DI})} \leq A - \frac{e}{\delta} - n^{LD}. \tag{RBA.2}$$

The condition holds if the LHS of (96) is non-positive when evaluated at $\theta_{DI} = \eta^{-1}\left(\frac{e}{A - \frac{e}{\delta} - n^{LD}}\right)$; i.e.,

$$A - \frac{e}{\delta} \geq n^{LD} + \frac{n^D}{\mu^{-1}\left(\frac{e}{n^D}\right)} = B_M. \tag{ABA.2}$$

Where $e \in \left[n^{LD}, \frac{n^D+n^{LD}}{2}\right)$, for $n_0^{SD} > e - n^{LD}$,

$$\frac{e}{\eta\left(\theta_{DI}\right)} > A - \frac{e}{\delta} - n^D - n^{LD} + e \tag{RBA.3}$$

has to hold. The condition is guaranteed to hold if

$$A - \frac{e}{\delta} - n^D - n^{LD} \leq 0.$$

Otherwise, $(RBA.3)$ can only hold if the LHS of (96) is positive when evaluated at $\theta_{DI} = \eta^{-1}\left(\frac{e}{A-\frac{e}{\delta}-n^D-n^{LD}+e}\right)$; i.e.,

$$\eta^{-1}\left(\frac{e}{A-\frac{e}{\delta}-n^D-n^{LD}+e}\right) - \frac{e}{A-\frac{e}{\delta}-n^D-n^{LD}+e} > 0.$$

This inequality is met for sure by Remark 1, as $A-\frac{e}{\delta}-n^D-n^{LD} > 0$ implies $\frac{e}{A-\frac{e}{\delta}-n^D-n^{LD}+e} < 1$. Next, for $n_0^{SD} < n^D - e$, by (94),

$$\frac{e}{\eta\left(\theta_{DI}\right)} < A - \frac{e}{\delta} - e, \tag{RBA.4}$$

By (95), the condition for $\eta\left(\theta_{ID}\right) \in (0,1)$ is the same condition as $(RBA.4)$. The condition holds if the LHS of (96) is negative when evaluated at $\theta_{DI} = \eta^{-1}\left(\frac{e}{A-\frac{e}{\delta}-e}\right)$; i.e.,

$$\frac{n^D+n^{LD}-e}{A-\frac{e}{\delta}-e} - \eta^{-1}\left(\frac{e}{A-\frac{e}{\delta}-e}\right) > 0. \tag{97}$$

The condition can only hold for $e < \frac{n^D+n^{LD}}{2}$, justifying our previous claim that the Balanced Equilibrium can only hold for e bounded from below the given value, because, otherwise,

$$\frac{n^D+n^{LD}-e}{A-\frac{e}{\delta}-e} \leq \frac{e}{A-\frac{e}{\delta}-e} < \eta^{-1}\left(\frac{e}{A-e/\delta-e}\right),$$

where the last inequality is by Remark 1. Given $e < \frac{n^D+n^{LD}}{2}$, (97) is equivalent to

$$A - \frac{e}{\delta} > e + \frac{n^D+n^{LD}-e}{\mu^{-1}\left(\frac{e}{n^D+n^{LD}-e}\right)} = \mathcal{B}_M. \tag{ABA.4}$$

The condition for there to be more sellers than buyers among large dealers in the inter-dealer market $m_2^{LD} \geq m_0^{LD}$, by (32) and (34), and $n_0^{LD} = n_2^{LD} = 0$ in the Balanced Equilibrium, simplifies to

$$\mu\left(\theta_{DI}\right) \geq \eta\left(\theta_{ID}\right).$$

By (95),

$$\mu\left(\theta_{DI}\right) - \eta\left(\theta_{ID}\right) = \mu\left(\theta_{DI}\right)\frac{\left(A-e/\delta\right)\eta\left(\theta_{DI}\right) - e - e\theta_{DI}}{\left(A-e/\delta\right)\eta\left(\theta_{DI}\right) - e}.$$

The denominator is guaranteed positive for $\eta(\theta_{ID}) \in [0,1]$. The expression then has the same sign as the numerator; i.e.,

$$\mu(\theta_{DI}) - \eta(\theta_{ID}) \geq 0 \Leftrightarrow \eta(\theta_{DI}) \left(A - \frac{e}{\delta}\right) - e - e\theta_{DI} \geq 0.$$

Rewrite (96) as

$$\eta(\theta_{DI}) \left(A - \frac{e}{\delta}\right) - e - e\theta_{DI} = \eta(\theta_{DI}) \left(n^D + n^{LD}\right) - 2e\theta_{DI}. \tag{99}$$

We seek conditions on how the two sides of the equation meet at a non-negative value.

Properties of $\eta(\theta)\left(A - \frac{e}{\delta}\right) - e - e\theta$

1. equal to $-e$ at $\theta = 0$,
2. tends to negative infinity as $\theta \to \infty$,
3. given condition $(ABA.4)$, so that $A - \frac{e}{\delta} - e > 0$, is inverted-U,
4. if $\max\left\{\eta(\theta)\left(A - \frac{e}{\delta}\right) - e - e\theta\right\} > 0$, rises above zero for a range of θ.

Properties of $\eta(\theta)\left(n^D + n^{LD}\right) - 2e\theta$

1. equal to 0 at $\theta = 0$,
2. tends to negative infinity as $\theta \to \infty$.
3. For $e < \frac{n^D + n^{LD}}{2}$, is inverted-U.

Given these properties of the two sides of (99), the RHS is greater than the LHS before the two sides meet, whereas the LHS is less than the RHS thereafter. Then, if at where the RHS vanishes, i.e.,

$$\theta = \mu^{-1}\left(\frac{2e}{n^D + n^{LD}}\right),$$

the LHS is non-negative; i.e.,

$$A - \frac{e}{\delta} \geq \frac{n^D + n^{LD}}{2} + \frac{\frac{n^D + n^{LD}}{2}}{\mu^{-1}\left(\frac{2e}{n^D + n^{LD}}\right)} = \widetilde{B}, \tag{ABA.S}$$

then the meeting point is where the two sides are non-negative.

If $(ABA.S)$ holds, the relevant inter-dealer market equilibrium condition is (20), which becomes

$$n^{LD}\left(\mu(\theta_{DI}) - \eta(\theta_{ID})\right) \leq \left(n_0^{SD}\left(1 - \mu(\theta_{DI})\right) + n_1^{SD}\eta(\theta_{ID})\right),$$

after substituting in (30), (32), and (34). By (93)-(95), the condition becomes

$$\left(n^D - A + \frac{e}{\delta} + \frac{e}{\eta(\theta_{DI})}\right)\left(1 - \mu(\theta_{DI})\right) + \left(\frac{e}{\mu(\theta_{DI})} - n^{LD}\right)\mu(\theta_{DI}) \geq 0. \tag{RBA.5}$$

Rewrite (96) as

$$\frac{e}{\mu(\theta_{DI})} - n^{LD} = n^D - A + \frac{e}{\delta} + \frac{e}{\eta(\theta_{DI})}$$

The LHS of $(RBA.5)$ is a weighted average of the two terms in this equation. Thus, if the equation holds where the two sides are non-negative, $(RBA.5)$ must hold. In turn, in case $e < n^{LD}$ and if $(RBA.1)$ holds, under which $n_0^{SD} \geq 0$, and in case $e \in \left[n^{LD}, \frac{n^D+n^{LD}}{2}\right)$ and if $(RBA.3)$ holds, under which $n_0^{SD} > e - n^{LD}$, the RHS of the equation is guaranteed non-negative.

If $(ABA.S)$ holds in reverse, the relevant inter-dealer market equilibrium condition is (19), which becomes

$$n^{LD}\left(\eta\left(\theta_{ID}\right) - \mu\left(\theta_{DI}\right)\right) \leq n_0^{SD}\mu\left(\theta_{DI}\right) + n_1^{SD}\left(1 - \eta\left(\theta_{ID}\right)\right),$$

after substituting in (31), (32), and (34). By (93)-(95), the condition becomes

$$\left(A - n^{LD} - \frac{e}{\delta} - \frac{e}{\eta\left(\theta_{DI}\right)}\right)\left(1 - \mu\left(\theta_{DI}\right)\right) + \left(n^D - \frac{e}{\mu\left(\theta_{DI}\right)}\right)\mu\left(\theta_{DI}\right) \geq 0. \quad \text{(RBA.6)}$$

Rewrite (96) as

$$A - n^{LD} - \frac{e}{\delta} - \frac{e}{\eta\left(\theta_{DI}\right)} = n^D - \frac{e}{\mu\left(\theta_{DI}\right)}.$$

The LHS of $(RBA.6)$ is a weighted average of the two terms in this equation. Thus if the equation holds at the point where the two sides are non-negative, $(RBA.6)$ must hold. In turn, in case $e < n^{LD}$ and if $(RBA.2)$ holds, under which $n_0^{SD} \leq n^{SD}$ and in case $e \in \left[n^{LD}, \frac{n^D+n^{LD}}{2}\right)$ and if $(RBA.4)$ holds, under which $n_0^{SD} < n^D - e$, the LHS of the equation is guaranteed non-negative.

Notice that in case $e < n^{LD}$, $(RBA.2)$ is a more stringent condition than $(RBA.4)$. Then, for θ_{ID} and θ_{DI} defined by (95) and (96) to be a valid Balanced Equilibrium, it suffices that $(ABA.1b)$ and $(ABA.2)$ hold; i.e., $A - \frac{e}{\delta} \in [B_M, B_L]$. Otherwise for $e \in \left[n^{LD}, \frac{n^D+n^{LD}}{2}\right)$, the equilibrium holds under $(ABA.4)$; i.e., $A - \frac{e}{\delta} > \mathcal{B}_M$. In either case, for $A - \frac{e}{\delta} \leq \widetilde{B}$, small dealers sell in equilibrium; otherwise small dealers buy.

Ranking of the Bounds That $B_S \leq B_M$ follows from $e \leq n^{LD}$, whereas that $B_M \leq \widetilde{B} \leq B_L$ follows from $n^{LD} \leq n^D$ and Remark 3. That $\mathcal{B}_M \leq \widetilde{B}$ follows from $e \leq n^{LD} + \frac{n^{SD}}{2}$ and Remark 3.

Proof of Lemma 3 Consider an inter-dealer market price such that

$$\beta\left(V_1^{LD} - V_0^{LD}\right) > p > \beta\left(V_1^{SD} - V_0^{SD}\right) > \beta\left(V_2^{LD} - V_1^{LD}\right).$$

The buyers are L_0s, whereas the sellers are S_1s and L_2s. This means that $n_1^{LD} = n^{LD}$ and $n_0^{SD} = n^{SD}$, $n_0^{LD} = n_2^{LD} = n_1^{SD} = 0$. In this case.

$$\theta_{DI} = \frac{n_B^D}{n_L^{ON}} = \frac{n^D}{n_L^{ON}},$$

By (79),

$$\theta_{DI} = \frac{n^D}{\frac{e}{\eta(\theta_{DI})}},$$

whereby,

$$\theta_{DI} = \mu^{-1}\left(\frac{e}{n^D}\right).$$

Substitute $n_L^{ON} = \frac{n^D}{\theta_{DI}}$, $n_H^{ON} = \frac{e}{\delta}$, $n_1^{LD} = n^{LD}$ and $n_1^{SD} = n_2^{LD} = 0$ into (24). The condition becomes the first condition in the Lemma.

Consider a price such that

$$\beta\left(V_1^{LD} - V_0^{LD}\right) > \beta\left(V_1^{SD} - V_0^{SD}\right) > p > \beta\left(V_2^{LD} - V_1^{LD}\right).$$

Following a similar procedure, we show that this price can be an equilibrium only where the second condition of the Lemma is met.

Proof of Corollary 1 The condition $e \in \left[n^{LD}, n^{LD} + \frac{n^{SD}}{2}\right)$ is equivalent to $n^{LD} \in \left(2e - n^D, e\right]$. For such n^{LD}, the Balanced Equilibrium indeed holds if the condition in Proposition 2(b) is met, which can be rewritten as the first condition of the Proposition. Notice that the RHS of the condition is greater than $2e - n^D$ by Remark 1 in the proof of Proposition 2, meaning that any n^{LD} that satisfies the condition exceeds $2e - n^D$. Now when n^{LD} rises up to e, Proposition 2(a) applies. At $n^{LD} = e$, $B_L \to \infty$ and $B_M = B_S$, in which case the Balanced Equilibrium continues to hold. At $n^{LD} = n^D$,

$$B_M = \widetilde{B} = B_L = \widehat{B},$$

in which case the Balanced Equilibrium still holds only if $A - e/\delta = \widehat{B}$. Otherwise, for $A - e/\delta < (>)\, \widehat{B}$, the Selling (Buying) Equilibrium holds. In general, as n^{LD} increases from e to n^D, B_L falls from infinity, whereas B_M and $\widetilde{B}$ go up and diverge from B_S. Eventually the three bounds converge to $\widehat{B}$. Then, for $A - e/\delta < (>)\, \widehat{B}$. the Balanced Equilibrium must turn into the Selling (Buying) Equilibrium at some $n^{LD} \in \left(e, n^D\right)$. The cutoff values are from Proposition 2(a).

Proof of Corollary 2 The Corollary follows directly from Proposition 2 and the characteristics of the three equilibria in Table 1.

Proof of Lemma 4 First, that all dealers are dealer-buyers and dealer-sellers and that αe measures of small investors buy and sell in each period imply that

$$\alpha e = n^D \eta(\theta_{ID})\alpha,$$

$$\alpha e = n^D \mu(\theta_{DI})\alpha.$$

The two conditions combine to yield $\eta(\theta_{ID}) = \mu(\theta_{DI})$, which in turn implies that $m_0^{SD} = m_2^{SD}$, given that

$$m_0^{SD} = \eta(\theta_{ID})\alpha\left(1 - \mu(\theta_{DI})\alpha\right)n^{SD},$$
$$m_2^{SD} = \mu(\theta_{DI})\alpha\left(1 - \eta(\theta_{ID})\alpha\right)n^{SD}.$$

Next, with $(1-\alpha)e$ measure of large investors buying from large dealers holding a two-unit inventory as well as $(1-\alpha)e$ measure of large investors selling to large dealers holding a one-unit inventory,

$$(1-\alpha)e = n_2^{LD}\eta(\theta_{ID})(1-\alpha),$$
$$(1-\alpha)e = n_1^{LD}\mu(\theta_{DI})(1-\alpha).$$

Given that $\eta(\theta_{ID}) = \mu(\theta_{DI})$, the above imply $n_1^{LD} = n_2^{LD} = n^{LD}/2$. To follow then is $m_0^{LD} = m_3^{LD}$ given that

$$m_0^{LD} = n_1^{LD}\eta(\theta_{ID})\alpha(1-\mu(\theta_{DI})) + n_2^{LD}\eta(\theta_{ID})(1-\alpha)(1-\alpha\mu(\theta_{DI})),$$
$$m_3^{LD} = n_1^{LD}\mu(\theta_{DI})(1-\alpha)(1-\alpha\eta(\theta_{ID})) + n_2^{LD}\mu(\theta_{DI})\alpha(1-\eta(\theta_{ID})).$$

Proof of Lemma 5 The surpluses of the possible trades are as follows.

$$z_{I_B,S_1} = U_H^{ON} - U^B - \left(V_1^{SD} - V_0^{SD}\right),$$
$$z_{I_B,L_i} = U_H^{ON} - U^B - \left(V_i^{LD} - V_{i-1}^{LD}\right) \text{ for } i = 1,2,$$
$$z_{S_0,I_S} = V_1^{SD} - V_0^{SD} - U_L^{ON},$$
$$z_{L_i,I_S} = V_{i+1}^{LD} - V_i^{LD} - U_L^{ON} \text{ for } i = 0,1,$$
$$z_{S_0,L_i} = V_1^{SD} - V_0^{SD} - \left(V_i^{LD} - V_{i-1}^{LD}\right) \text{ for } i = 1,2,$$
$$z_{L_i,S_1} = V_{i+1}^{LD} - V_i^{LD} - \left(V_1^{SD} - V_0^{SD}\right) \text{ for } i = 0,1,$$
$$z_{L_0,L_2} = V_1^{LD} - V_0^{LD} - \left(V_2^{LD} - V_1^{LD}\right),$$
$$z_{L_1,L_1} = V_2^{LD} - V_1^{LD} - \left(V_1^{LD} - V_0^{LD}\right).$$

An S_1 can sell to an IB (investor-buyer), an L_0, or an L_1. He will not sell to an IB only if selling to other dealers yields a strictly larger surplus; i.e.,

$$\max\{z_{L_0,S_1}, z_{L_1,S_1}\} > z_{I_B,S_1}.$$

Expanding the expressions for the zs,

$$\max\left\{V_1^{LD} - V_0^{LD}, V_2^{LD} - V_1^{LD}\right\} > U_H^{ON} - U^B.$$

Subtracting $U_H^{ON} - U^{IB}$ from the two sides of the condition

$$\max\left\{V_1^{LD} - V_0^{LD} - \left(U_H^{ON} - U^B\right), V_2^{LD} - V_1^{LD} - \left(U_H^{ON} - U^B\right)\right\} > 0.$$

The two terms inside the max operator are simply the negatives of z_{I_B,L_1} and z_{I_B,L_2}, respectively. Then, the condition becomes

$$\max\{-z_{I_B,L_1}, -z_{I_B,L_2}\} > 0 \Leftrightarrow \min\{z_{I_B,L_1}, z_{I_B,L_2}\} < 0.$$

All this implies that if one type of dealer-seller finds it optimal not to sell to investor-buyers, then only one type of dealer-seller may find it optimal to do so. In any active steady-state equilibrium, indeed at least one type of dealer-seller must do so.

Now, suppose only S_1s sell to IB where

$$z_{I_B,S_1} = U_H^{ON} - U^B - \left(V_1^{SD} - V_0^{SD}\right) \geq 0. \tag{103}$$

An L_1 may then only sell to an S_0 or another L_1. Selling to an S_0 is optimal if

$$z_{S_0,L_1} = V_1^{SD} - V_0^{SD} - \left(V_1^{LD} - V_0^{LD}\right) \geq 0.$$

But if the condition holds,

$$z_{I_B,L_1} = U_H^{ON} - U^B - \left(V_1^{LD} - V_0^{LD}\right) \geq 0$$

must hold given (103). The hypothesis that only S_1 sell to IB then implies that selling to another L_1 must be optimal for the L_1 (otherwise the L_1 has no one to sell to), where

$$z_{L_1,L_1} = V_2^{LD} - V_1^{LD} - \left(V_1^{LD} - V_0^{LD}\right) \geq 0. \tag{104}$$

An L_2 may sell to an S_0 or an L_0 if selling to an IB is not optimal. Selling to an S_0 is optimal if

$$z_{S_0,L_2} = V_1^{SD} - V_0^{SD} - \left(V_2^{LD} - V_1^{LD}\right) \geq 0.$$

But if the condition holds,

$$z_{I_B,L_2} = U_H^{ON} - U^B - \left(V_2^{LD} - V_1^{LD}\right) \geq 0$$

must hold given (103). The hypothesis that only S_1 sell to IB then implies that selling to an L_0 must be optimal for the L_2, where

$$z_{L_0,L_2} = V_1^{LD} - V_0^{LD} - \left(V_2^{LD} - V_1^{LD}\right) \geq 0. \tag{105}$$

The two conditions, (104) and (105), together imply that

$$V_1^{LD} - V_0^{LD} = V_2^{LD} - V_1^{LD}.$$

Thus, if neither L_1s nor L_2s find it optimal to sell to investor-buyers or to small dealers, large dealers do not gain by selling and buying among themselves either. They must then be inactive in equilibrium.

Next, suppose only L_1s sell to investor-buyers, where

$$z_{I_B,L_1} = U_H^{ON} - U^B - \left(V_1^{LD} - V_0^{LD}\right) \geq 0. \tag{106}$$

An S_1 may sell to an L_0 or to an L_1 if not selling to an investor-buyer. If the first sale is optimal, it must be optimal for the S_1 to sell to an IB as well given (106). The hypothesis that only L_1 sells to investor-buyers then requires that it is optimal for an S_1 to sell to an L_1 where

$$z_{L_1,S_1} = V_2^{LD} - V_1^{LD} - \left(V_1^{SD} - V_0^{SD}\right) \geq 0. \tag{107}$$

An L_2 may sell to an L_0 or to an S_0. If the first sale is optimal, it must be optimal for the L_2 to sell to an IB as well given (106). The condition for the second sale to be optimal is that

$$z_{S_0,L_2} = V_1^{SD} - V_0^{SD} - \left(V_2^{LD} - V_1^{LD}\right) \geq 0. \tag{108}$$

The two conditions, (107) and (108), together imply that

$$V_1^{SD} - V_0^{SD} = V_2^{LD} - V_1^{LD}.$$

Thus, if neither S_1s nor L_2s find it optimal to sell to investor-buyers, S_1s only sell to L_1s, where such trades do not yield any surplus. This implies that small dealers must be inactive in equilibrium.

The case for where only L_2s sell to investor-buyers can be shown in a similar way to imply that small dealers must be inactive in equilibrium.

The proof that in any equilibrium in which both small and large dealers are active, investor-sellers must sell to all three types of dealer-buyers can be constructed similarly.

Proof of Proposition 3 With Nash Bargaining and each agent in a match entitled to one-half of the match's surplus, we can rewrite dealers' value functions as follows.

$$rV_0^{SD} = \mu\left(\theta_{DI}\right)\frac{z_{S_0,I_S}}{2} + \alpha\left\{\frac{n_1^{LD}}{2n^D}\max\{z_{S_0,L_1},0\} + \frac{n_2^{LD}}{2n^D}\max\{z_{S_0,L_2},0\}\right\},$$

$$rV_1^{SD} = \eta\left(\theta_{ID}\right)\frac{z_{I_B,S_1}}{2} + \alpha\left\{\frac{n_0^{LD}}{2n^D}\max\{z_{L_0,S_1},0\} + \frac{n_1^{LD}}{2n^D}\max\{z_{L_1,S_1},0\}\right\},$$

$$rV_0^{LD} = \mu\left(\theta_{DI}\right)\frac{z_{L_0,I_S}}{2} + \alpha\left\{\frac{n_1^{SD}}{2n^D}\max\{z_{L_0,S_1},0\} + \frac{n_2^{LD}}{2n^D}\max\{z_{L_0,L_2},0\}\right\},$$

$$\begin{aligned} rV_1^{LD} = {} & \mu\left(\theta_{DI}\right)\frac{z_{L_1,I_S}}{2} + \eta\left(\theta_{ID}\right)\frac{z_{I_B,L_1}}{2} + \\ & \alpha\left\{\frac{n_0^{SD}}{2n^D}\max\{z_{S_0,L_1},0\} + \frac{n_1^{SD}}{2n^D}\max\{z_{L_1,S_1},0\} + \frac{n_1^{LD}}{2n^D}\max\{z_{L_1,L_1},0\}\right\}, \end{aligned}$$

$$rV_2^{LD} = \eta\left(\theta_{ID}\right)\frac{z_{I_B,L_2}}{2} + \alpha\left\{\frac{n_0^{SD}}{2n^D}\max\{z_{S_0,L_2},0\} + \frac{n_0^{LD}}{2n^D}\max\{z_{L_0,L_2},0\}\right\}.$$

Suppose $V_1^{SD} - V_0^{SD} > V_1^{LD} - V_0^{LD}$. Then $z_{I_B,S_1} < z_{I_B,L_1}$ and $z_{S_0,I_S} > z_{L_0,I_S}$. Together with the fact that $z_{L_1,I_S} \geq 0$, this implies that

$$\eta\left(\theta_{ID}\right)\frac{z_{I_B,S_1}}{2} - \mu\left(\theta_{DI}\right)\frac{z_{S_0,I_S}}{2} < \mu\left(\theta_{DI}\right)\frac{z_{L_1,I_S}}{2} + \eta\left(\theta_{ID}\right)\frac{z_{I_B,L_1}}{2} - \mu\left(\theta_{DI}\right)\frac{z_{L_0,I_S}}{2}.$$

Also, $V_1^{SD} - V_0^{SD} > V_1^{LD} - V_0^{LD}$ implies that $z_{S_0,L_1} > 0 > z_{L_0,S_1}$, $z_{S_0,L_2} > z_{L_0,L_2}$, and $z_{L_1,S_1} < z_{L_1,L_1}$. This means

$$\frac{n_1^{LD}}{2n^D}\max\{z_{S_0,L_1},0\} + \frac{n_2^{LD}}{2n^D}\max\{z_{S_0,L_2},0\} > \frac{n_1^{SD}}{2n^D}\max\{z_{L_0,S_1},0\} + \frac{n_2^{LD}}{2n^D}\max\{z_{L_0,L_2},0\}$$

and

$$\frac{n_0^{LD}}{2n^D}\max\{z_{L_0,S_1},0\}+\frac{n_1^{LD}}{2n^D}\max\{z_{L_1,S_1},0\}$$
$$<\frac{n_0^{SD}}{2n^D}\max\{z_{S_0,L_1},0\}+\frac{n_1^{SD}}{2n^D}\max\{z_{L_1,S_1},0\}+\frac{n_1^{LD}}{2n^D}\max\{z_{L_1,L_1},0\}$$

The above three inequalities together imply that $V_1^{SD}-V_0^{SD}<V_1^{LD}-V_0^{LD}$. This is a contradiction.

Now suppose $V_2^{LD}-V_1^{LD}>V_1^{SD}-V_0^{SD}$. Similarly, we can show that this implies $z_{L_1,I_S}>z_{S_0,I_S}$, $z_{I_B,L_2}<z_{I_B,S_1}$, $z_{S_0,L_2}<0<z_{L_1,S_1}$ and $z_{L_0,L_2}<z_{L_0,S_1}$. These inequalities in turn imply that $V_2^{LD}-V_1^{LD}<V_1^{SD}-V_0^{SD}$. This is a contradiction.

Given that we have shown $V_1^{LD}-V_0^{LD}\geq V_1^{SD}-V_0^{SD}\geq V_2^{LD}-V_1^{LD}$, it is straightforward to verify that the two equalities hold are strict unless $z_{I_B,L_1}=0$.

Proof of Proposition 4a In the Selling Equilibrium, θ_{DI} is implicitly given by (82), in which A is absent. By (84), θ_{ID} is decreasing in A given that θ_{DI} does not vary with A.

In the Balanced Equilibrium, θ_{DI} is implicitly given by (96), the solution to which is at a point where the LHS of the equation is increasing. In the meantime, the LHS of the equation is increasing in A. Then, $\partial\theta_{DI}/\partial A<0$. To evaluate the effect of A on θ_{ID}, first rewrite (96) as

$$A-\frac{e}{\delta}=n^D+n^{LD}-\frac{e\left(\theta_{DI}-1\right)}{\eta\left(\theta_{DI}\right)}.$$

Then substitute the equation into (95) to yield

$$\eta\left(\theta_{ID}\right)=\frac{\eta\left(\theta_{DI}\right)e}{\left(n^D+n^{LD}\right)\eta\left(\theta_{DI}\right)-e\theta_{DI}},$$

the RHS of which is increasing in θ_{DI} due to the concavity of η. Then, θ_{ID} must be decreasing in A.

In the Buying Equilibrium, θ_{ID} is implicitly given by (86), in which A is absent, whereas θ_{DI} is given by the same equation that defines θ_{DI} in the Balanced Equilibrium.

The continuity can be established by verifying that the equations for θ_{DI} and θ_{ID} for one equilibrium type coincide with another at each of the two cutoff values of $A-e/\delta$.

Proof of Proposition 4b By (31), (34), the restrictions in the first column of Table 2, and (82), in the Selling Equilibrium,

$$TV=\frac{e}{n^D}\left(n^D-n^{LD}+e\left(\frac{1}{\eta\left(\theta_{ID}\right)}-1\right)\right). \tag{115}$$

The result of the Proposition then follows, given that, by Proposition 4a, θ_{ID} is decreasing in A in the Selling Equilibrium. By (34) and (32) and the restrictions in the second column of Table 2, in the Balanced Equilibrium,

$$TV=\begin{cases} n^{LD}\eta\left(\theta_{ID}\right)\left(1-\mu(\theta_{DI})\right) & A\leq\widetilde{B}+\frac{e}{\delta} \\ n^{LD}\mu\left(\theta_{DI}\right)\left(1-\eta(\theta_{ID})\right) & A>\widetilde{B}+\frac{e}{\delta}\end{cases}. \tag{116}$$

The result of the Proposition then follows given that, by Proposition 4a, both θ_{ID} and θ_{ID} are decreasing in A in the Selling Equilibrium. By (30), (32), the restrictions in the third column of Table 2, and (86), in the Buying Equilibrium,

$$TV = \frac{e}{n^D}\left(n^D - n^{LD} + e\left(\frac{1}{\mu(\theta_{DI})} - 1\right)\right). \tag{117}$$

The result of the Proposition then follows given that, by Proposition 4a, θ_{DI} is decreasing in A in the Buying Equilibrium.

Evaluate (115) and the first line of (116) at where $A = B_M + e/\delta$ and (86) yields the same value of $e\left(1 - \frac{e}{n^D}\right)$. Evaluate the second line of (116) and (117) at where $A = B_L + e/\delta$ and (87) yields the same value of $e\left(1 - \frac{e}{n^D}\right)$. This proves continuity.

Proof of Proposition 4c In the Selling Equilibrium, p is given by (73), which is increasing in θ_{ID} and θ_{DI}. Given that in the Selling Equilibrium, θ_{ID} is decreasing in A but θ_{DI} is independent of A, p must be decreasing in A. That there is a discrete fall in p as the Selling Equilibrium turns into the Balanced Equilibrium can be established by showing that the denominator of (75), which gives p in the Balanced Equilibrium, is larger than that of (73) at any θ_{ID} and θ_{DI}. Moreover, by (75), p in the Balanced Equilibrium is also increasing in θ_{ID} and θ_{DI}, both of which are decreasing in A. Finally, that there is a discrete fall in p as the Balanced Equilibrium turns into the Buying Equilibrium can be established by noting that the numerator of (75) always stays strictly positive.

Proof of Proposition 5a In the Balanced Equilibrium, θ_{DI} is given by the solution to (96), whereas θ_{ID} can be recovered from (95) once θ_{DI} is known from the former equation. In the Buying Equilibrium, θ_{DI} and θ_{ID} are given by the solutions to (89) and (86), respectively. In the Selling Equilibrium, θ_{DI} is given by the solution to (82), whereas θ_{ID} can be recovered from (84) once θ_{DI} is known from the former equation. The comparative steady states followed straightforwardly from these equations. Just as in the proof of Proposition 3a, the continuity can be established by verifying that the equations for θ_{DI} and θ_{ID} for one equilibrium type coincide with another at each of the two cutoff values of $A - e/\delta$.

Proof of Proposition 5b In the Selling Equilibrium, TV, given by (115) is decreasing in n^{LD}, given that θ_{ID} is independent of n^{LD} in the Selling Equilibrium. In the Buying equilibrium, TV, given by (117) can be shown to be decreasing in n^{LD} with θ_{DI} given in (89).

In the Balanced Equilibrium, TV is given by either the first or the second line of (116). To show that both expressions are increasing in n^{LD}, we begin with noting that A^D, by (78) and (79), in the first instance is given by

$$A^D = A - \frac{e}{\delta} - \frac{e}{\eta(\theta_{DI})},$$

But by (96),

$$A - \frac{e}{\delta} - \frac{e}{\eta(\theta_{DI})} = n^{LD} + n^D - \frac{e}{\mu(\theta_{DI})}$$

Because θ_{DI} increases with n^{LD} in the Balanced Equilibrium, the LHS strictly increases with n^{LD}, and so dealers hold more inventory in total. The RHS, however, can only rise by less than the increase in n^{LD}. Given that in the Balanced Equilibrium, $A^D = n_1^{SD} + n^{LD}$, a larger n^{LD} must be accompanied by a smaller n_1^{SD}. Also, according to (94), there would also have to be a smaller n_0^{SD}. In the investor-dealer market, both dealer-sellers and dealer-buyers execute e trades in the steady state; i.e.,

$$(n^{LD} + n_1^{SD})\eta(\theta_{ID}) = (n^{LD} + n_0^{SD})\mu(\theta_{DI}) = e$$

Because $n_1^{SD}\eta(\theta_{ID})$ and $n_0^{SD}\mu(\theta_{DI})$ strictly decrease with n^{LD}, $n^{LD}\eta(\theta_{ID})$ and $n^{LD}\mu(\theta_{DI})$ must be strictly increasing in n^{LD}. This implies that both the first and the second lines of (116) are strictly increasing in n^{LD}.

The proof of continuity is as in Proposition 3c.

Proof of Proposition 5c In the Selling Equilibrium, p is given by (73), which does not directly depend on n^{LD}, given θ_{ID} and θ_{DI}. But then, the two market tightness in the Selling Equilibrium do not vary with n^{LD}. The proof for the jump in p that occurs when the Balanced Equilibrium gives way to the Buying or the Selling Equilibrium follow from Proposition 4b.

Proof of Lemma A1 The equation for p_{I_S} is from combining (15) and (67). The equations for p_{I_B} are from combining (17) and (72) for the Selling Equilibrium, (17), (67), and (74) for the Balanced Equilibrium, and (17), (77), and $p = 0$ for the Buying Equilibrium.

Proof of Propositions A1 and A2 By (46), given p, p_{I_S} depends only on and is increasing in θ_{DI}. In the Selling and Balanced Equilibria, $\partial\theta_{DI}/\partial A = 0$ and $\partial\theta_{DI}/\partial A < 0$, respectively. Then, $\partial p_{I_S}/\partial A$ has the same negative sign as $\partial p/\partial A$ in the two types of equilibrium. Next, in the Selling Equilibrium, $\partial\theta_{DI}/\partial n^{LD} = 0$, from which it follows that $\partial p_{I_S}/\partial n^{LD}$ has the same zero value as $\partial p/\partial n^{LD}$. In the Buying Equilibrium, by Lemma A1, $p_{I_S} = p = 0$. The discrete changes in p_{I_S} at where one equilibrium type changes to another follows from the discrete changes in p.

By combining (47) and (73), in the Selling Equilibrium,

$$p_{I_B} = \frac{\frac{1}{2}\beta\left(1-\beta+\beta\eta\left(\theta_{ID}\right)\right)\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)\upsilon}{\left(1-\beta+\left(\frac{\eta(\theta_{ID})}{2}+\frac{\mu(\theta_{ID})}{2}\right)\beta\right)\left(1-\beta+\frac{\eta(\theta_{DI})}{2}\beta\right)(1-\beta+\beta\delta)-\delta\beta^3\frac{\eta(\theta_{ID})}{2}\frac{\eta(\theta_{DI})}{2}},$$

where $\partial p_{I_B}/\partial\theta_{ID} > 0$. Then, given $\partial\theta_{ID}/\partial A < 0$, it follows that $\partial p_{I_B}/\partial A < 0$. Meanwhile, $\partial p_{I_B}/\partial n^{LD} = 0$ holds given $\partial\theta_{ID}/\partial n^{LD} = \partial\theta_{DI}/\partial n^{LD} = 0$ in the Selling Equilibrium. That p_{I_B} in the Buying Equilibrium, given by (49), does not vary with n^{LD} follows from $\partial\theta_{ID}/\partial n^{LD} = 0$ in said equilibrium. The discrete changes in p_{I_B} at which one equilibrium type changes to another can be verified by checking how, given θ_{ID} and θ_{DI}, p_{I_B} in (47) exceeds p_{I_B} in (48), which in turn exceeds p_{I_B} in (49).

Proof of Lemma A2 The state variables of the planning problem (50) are $\{n_H^{ON}(t), n_L^{ON}(t), n_B^I(t)\}$, the initial conditions are $\{n_H^{ON}(0), n_L^{ON}(0), n_B^I(0)\} = \{\widehat{n}_H^{ON}, \widehat{n}_0^{ON}, \widehat{n}_B^I\}$, the controls are $\{n_0^{SD}(t), n_1^{SD}(t), n_0^{LD}(t), n_1^{LD}(t), n_2^{LD}(t)\}$ and the equations of motion are, respectively,

$$n_H^{ON}(t+1) - n_H^{ON}(t) = -\delta n_H^{ON}(t) + n_B^I(t)\,\mu(\theta_{ID}[t]),$$
$$n_L^{ON}(t+1) - n_L^{ON}(t) = \delta n_H^{ON}(t) - n_L^{ON}(t)\,\eta(\theta_{DI}[t]),$$
$$n_B^I(t+1) - n_B^I(t) = e - n_B^I(t)\,\mu(\theta_{ID}[t]).$$

The constraints are given in (22)-(26) that hold at each moment in time, which can be summarized by the following two equations:

$$\theta_{ID}(t) = \frac{n_B^I(t)}{n^D - n_0^{SD}(t) - n_0^{LD}(t)},$$

$$\theta_{DI}(t) = \frac{n^D + n^{LD} - A - n_0^{LD}(t) + n_L^{ON}(t) + n_H^{ON}(t)}{n_L^{ON}(t)}.$$

In the above, a pair of $\{n_0^{SD}(t), n_0^{LD}(t)\}$ uniquely determines the pair $\{\theta_{ID}(t), \theta_{DI}(t)\}$. This means that the controls can be stated in terms of the two market tightness only, whereby the admissible values are given by

$$\theta_{ID}(t) \in \left[\frac{n_B^I(t)}{\overline{n}_S^D(n_H^{ON}(t), n_L^{ON}(t))}, \frac{n_B^I(t)}{\underline{n}_S^D(n_H^{ON}(t), n_L^{ON}(t))}\right],$$

$$\theta_{DI}(t) \in \left[\frac{\underline{n}_B^D(n_H^{ON}(t), n_L^{ON}(t))}{n_L^{ON}(t)}, \frac{\overline{n}_B^D(n_H^{ON}(t), n_L^{ON}(t))}{n_L^{ON}(t)}\right],$$

with $\overline{n}_S^D$ and $\underline{n}_S^D$ denoting, respectively, the largest and smallest possible measures of dealer-sellers and $\overline{n}_B^D$ and $\underline{n}_B^D$ denoting, respectively, the largest and smallest possible measures of dealer-buyers, given state variables $n_H^{ON}(t)$ and $n_L^{ON}(t)$. Note that:
(1) To attain $\overline{n}_S^D$, first allocate one unit each of the assets to be held by dealers $(A - n_H^{ON}(t) - n_L^{ON}(t))$ to either small or large dealers, and then allocate one more unit each to large dealers if $A - n_H^{ON}(t) - n_L^{ON}(t) > n^D$.
(2) To attain $\underline{n}_S^D$, first allocate two units each of the assets to be held by dealers to large dealers, and then allocate one unit each to small dealers if $A - n_H^{ON}(t) - n_L^{ON}(t) > 2n^{LD}$.
(3) To attain $\overline{n}_B^D$, first allocate one unit each of the assets to be held by dealers to large dealers, and then allocate one unit each to either large or small dealers if $A - n_H^{ON}(t) - n_L^{ON}(t) > n^{LD}$.
(4) To attain $\underline{n}_B^D$, first allocate one unit each of the assets to be held by dealers to small dealers, and then allocate two units each to large dealers if $A - n_H^{ON}(t) - n_L^{ON}(t) > n^{SD}$.

To proceed, write (50) as

$$W\left(n_H^{ON}(t), n_L^{ON}(t), n_B^I(t)\right) = \max_{\theta_{ID}(t),\theta_{DI}(t)} \left\{n_H^{ON}(t)\,\upsilon_H + \beta W\left(n_H^{ON}(t+1), n_L^{ON}(t+1), n_B^I(t+1)\right)\right\},$$

in which the state variables for $t+1$ can be recovered from the equations of motions. There are four constraints corresponding to the four bounds of market tightness. Let $\lambda_1(t)$, $\lambda_2(t)$,

$\lambda_3(t)$ and $\lambda_4(t)$ be the respective Lagrange multipliers of the lower and upper bounds of $\theta_{ID}(t)$ and the lower and upper bounds of $\theta_{DI}(t)$.

Restricting attention to the steady state, we omit all time indices in the following. The first order conditions for $\theta_{ID}(t)$ and $\theta_{DI}(t)$ are then given by, respectively,

$$\beta n_B^I \mu'(\theta_{ID})(W_1 - W_3) + \lambda_1 - \lambda_2 = 0, \tag{127}$$

$$-\beta n_L^{ON} \eta'(\theta_{DI}) W_2 + \lambda_3 - \lambda_4 = 0. \tag{128}$$

In addition, there are three envelope conditions, one for each state variable:

$$\begin{aligned} W_1 =& \upsilon_H + \beta(1-\delta)W_1 + \beta\delta W_2 - \lambda_1 \frac{\partial(n_B^I/\overline{n}_S^D)}{\partial n_H^{ON}} + \lambda_2 \frac{\partial(n_B^I/\underline{n}_S^D)}{\partial n_H^{ON}} \\ &- \lambda_3 \frac{\partial(\underline{n}_B^D/n_L^{ON})}{\partial n_H^{ON}} + \lambda_4 \frac{\partial(\overline{n}_B^D/n_L^{ON})}{\partial n_H^{ON}} \end{aligned} \tag{129}$$

$$\begin{aligned} W_2 =& \beta(1-\eta(\theta_{DI}))W_2 - \lambda_1 \frac{\partial(n_B^I/\overline{n}_S^D)}{\partial n_L^{ON}} + \lambda_2 \frac{\partial(n_B^I/\underline{n}_S^D)}{\partial n_L^{ON}} \\ &- \lambda_3 \frac{\partial(\underline{n}_B^D/n_L^{ON})}{\partial n_L^{ON}} + \lambda_4 \frac{\partial(\overline{n}_B^D/n_L^{ON})}{\partial n_L^{ON}} \end{aligned} \tag{130}$$

$$W_3 = \beta\mu(\theta_{ID})(W_1 - W_3) - \frac{\lambda_1}{\overline{n}_S^D} + \frac{\lambda_2}{\underline{n}_S^D} \tag{131}$$

We first show that λ_3 must equal to 0. Suppose otherwise. By the definition of $\overline{n}_S^D$, $\underline{n}_S^D$, $\overline{n}_B^D$ and $\underline{n}_B^D$, λ_1, λ_2 and λ_4 must all equal to 0. Then equations (127) to (131) reduce to three equations, (128), (129) and (130) with only two unknowns, λ_3 and W_2. The set of λ_3 and W_2 that satisfy all three equations is of measure zero. Therefore, λ_3 must equal to 0.

By the same argument, we can show that λ_2 must equal to 0.

Next, we prove that $\lambda_1 > 0$. Suppose otherwise. Together with the fact that $\lambda_2 = 0$, this implies $W_1 = W_3 = 0$. We already know $\lambda_3 = 0$. Then equations (127) to (131) reduce to three equations, (128), (129) and (130) with only two unknowns, λ_4 and W_2. We have reached the desired contradiction.

Finally, we prove that $\lambda_4 > 0$. Suppose otherwise. Then by equation (128), $W_2 = 0$. Plug it into equation (130), we must have $\lambda_1 = 0$, which contradicts our previous conclusion that $\lambda_1 > 0$.

To summarize, we have shown that $\theta_{ID} = \frac{n_B^I}{\overline{n}_S^D(n_H^{ON}, n_L^{ON})}$ and $\theta_{DI} = \frac{\overline{n}_B^D(n_H^{ON}, n_L^{ON})}{n_L^{ON}}$. In other words, for efficiency, we should allocate the assets held by dealers to maximize the measure of dealers holding inventory and the measure of dealers having spare capacity: first allocate one unit each to large dealers; if $A - n_L^{ON} - n_H^{ON} > n^{LD}$, then allocate one unit each to small dealers; if $A - n_L^{ON} - n_H^{ON} > n^D$, then allocate one more unit each to large dealers.

Proof of Proposition A3 The allocations as described in Lemma A2 are the same as the allocations as described in the discussions following Proposition 2.

References

[1] Adrian, T., M. Fleming, O. Stackman and E. Vogt, 2017, "Market Liquidity after the Financial Crisis," *Federal Reserve Bank of New York Staff Reports* no. 796.

[2] Atkeson, A.G., A.L. Eisfeldt and P.-O. Weill, 2015, "Entry and Exit in OTC Derivatives Market," *Econometrica* 83, 2231-2292.

[3] Bao, J., M. O'Hara and A. Zhou, 2016, "The Volcker Rule and Market-Making in Times of Stress," *Finance and Economics Discussion Series* 2016-102. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2016.102.

[4] Colliard, J.-E. and G. Demange, 2017, "Cash Providers: Asset Dissemination over Intermediation Chains," mimeo.

[5] Di Maggio, M., A. Kermani and Z. Song, 2017, "The value of trading relations in turbulent times," *Journal of Financial Economics* 124, 266-284.

[6] Duffie, D., N. Gârleanu and L.H. Pedersen, 2005, "Over-the-Counter Markets," *Econometrica* 73, 1815-47.

[7] Friewald, N. and F. Nagler, 2017, "Over-the-Counter-Market Frictions and Yield Spread Changes," mimeo.

[8] Farboodi, M., 2014, "Intermediation and Voluntary Exposure to Counterparty Risk," mimeo.

[9] Gale, D., 2000, *Strategic Foundations of General Equilibrium – Dynamic Matching and Bargaining Games*, Cambridge University Press.

[10] Glode, V. and C. Opp, 2016, "Asymmetric Information and Intermediation Chains," *American Economic Review* 106, 2699-2721.

[11] Hendershott, T., D. Li, D. Livdan and N. Schürhoff, 2016, " Relationship Trading in OTC markets," mimeo.

[12] Hendershott, T. and A. Madhavan, 2015, "Click or Call? Auction Versus Search in the Over-the-Counter-Market," *Journal of Finance* 70, 419-447.

[13] Ho, T., and H. R. Stoll, 1983, "The Dynamics of Dealer Markets Under Competition," *Journal of Finance* 38, 1053–1074.

[14] Hollifield, B., A. Neklyudov and C. S. Spatt, 2016, "Bid-Ask Spreads, Trading Networks and the Pricing of Securitizations," *Review of Financial Studies*, forthcoming.

[15] Hugonnier J., B. Lester and P.-O. Weill, 2016, "Heterogeneity in Decentralized Asset Markets," mimeo.

[16] Li, D and N. Schürhoff, 2014, "Dealer Network," Swiss Finance Research Institute Research Paper, 14-50.

[17] Lagos, R. and G. Rocheteau, 2009, "Liquidity in Asset markets with Search Frictions," *Econometrica* 77, 403-26.

[18] Lagos, R., G. Rocheteau and P.-O. Weill, 2011, "Crisis and Liquidity in Over-the-Counter Markets," *Journal of Economic Theory* 146, 2169-2205.

[19] Neklyudov, A. V., 2015, "Bid-Ask Spreads and the Over-the-Counter Interdealer Markets: Core and Peripheral Dealers," mimeo.

[20] Piazzesi, M. and M. Schneider, 2009, "Momentum Traders in the Housing Market: Survey Evidence and a Search Model," *American Economic Review Papers and Proceedings* 99, 406-411.

[21] Randall, O., 2015, "Pricing and Liquidity in Over-the-Counter Markets," mimeo.

[22] Shen, J., B. Wei and H. Yan, 2016, "Financial Intermediation Chain in an OTC Market," mimeo.

[23] Üslü, S., 2016, "Pricing and Liquidity in Decentralized Asset Markets," mimeo.

[24] Wang, C., 2017, "Core-Periphery Trading Networks," mimeo.

[25] Weill, P.-O., 2011, "Liquidity Provision in Capacity Constrained Markets," *Macroeconomic Dynamic* 15, 119-144.

[26] Zhong, Z., 2014, "The Risk Sharing Benefit versus the Collateral Cost: The Formation of the Inter-Dealer Network in Over-the-Counter Trading," mimeo.