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

Chung-Yi Tse* and Yujing Xu*

May 6, 2019

Abstract

Dealers in an OTC market 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: cytse@hku.hk (Tse); yujingxu@hku.hk (Xu). We would like to thank an associate editor of the journal and two referees for very helpful comments and suggestions. Comments from 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 2017 Econometric Society Asian meeting also help improve the paper. 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. 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.[1]

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.[2] Behind this overall trend, however, Adrian, Fleming, Shachar and Vogt (2017) find that individual dealers respond differently to the changing market conditions – while large dealers 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.[3] The reason for the balance sheet expansion by small dealers post-crisis is less clear though. This means that the changes in dealers' asset holdings cannot be solely attributed to regulatory changes. Instead, the differential responses by large and small dealers could be, to a certain extent, the result of the sales of assets by the former to the latter as natural responses to the changing market

[1] 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 (2017).

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

[3] 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.

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 asset supply in the market. The point of departure is that, in our model, dealers are heterogeneous in their inventory capacities. The heterogeneity can be due to risk management considerations, portfolio choices, or can result from differences in financing costs – 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.

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. In each period, investors buy from and sell to dealers in an OTC market in which only dealers who are holding at least a unit of the asset in inventory can sell to and only dealers having at least one unit of spare inventory capacity can buy from investors. Once the investor-dealer trades are completed, and only then, a perfectly competitive inter-dealer market opens, through which dealers trade to 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 inventory capacity are valued higher by a large dealer than by a small dealer. That a large dealer values the first unit of inventory more than a small dealer does is due to how the small dealer, but not the large dealer, would exhaust his entire inventory capacity in acquiring a unit of the asset and thereby would forgo any opportunity to buy from an investor in the next round of trading. That the large dealer values the last unit of inventory capacity more than a small dealer does is due to how the small dealer, but not the large dealer, may sell to an investor only by using up his only unit of inventory capacity to acquire a unit of the asset.

In other words, small dealers value a unit of the asset in between large dealers starting out with an empty inventory and large dealers starting out with one unit of inventory capacity yet to be utilized. The competitive inter-dealer market in equilibrium would then first allocate the asset to large dealers, then to small dealers if there remain units of the asset yet to be allocated to dealers, and finally to large dealers already holding a unit in inventory in case the asset supply is sufficiently large for all small dealers to already hold a unit in inventory. Such allocations are by means of small dealers selling to large dealers in a market with a small asset supply and small dealers buying from large dealers in a market with a large asset supply. Small dealers then should trade with large dealers only but not among themselves while trades between the two types of dealers tend to flow in one particular direction only in a given market.

A dealer is said to provide immediacy to another dealer if the first dealer sells to (buys from) the second dealer when it takes longest on average for the second dealer to buy (sell) the asset in the market. In our model, large dealers sell to small dealers when there is a large asset supply during which it should be easiest for small dealers themselves to buy the asset from investors. When there is a small asset supply, at which times it should be hardest for dealers to buy from investors, it is small dealers who sell to large dealers. Apparently, the large dealers in our model do not provide immediacy for small dealers. Rather, it is the small dealers who trade to provide immediacy for the large dealers, selling to (buying from) large dealers at times during which it takes a long time for the latter to buy (sell) in the market.

The last implication is consistent with the findings in Adrian et al. (2017) if the booming market pre-crisis is time during which it is hardest for dealers to find willing sellers among their customers due to the robust demand for risky investment then and the market bust post-crisis is time during which it is hardest for dealers to find willing buyers with an appetite for even relatively safe investment. In the earlier period, large dealers gain inventory from small dealers and expand their balance sheets faster. In the later period, small dealers amass inventory from large dealers and expand their balance sheets relative to those of large dealers.

It is well known that in many OTC markets, as documented in Li and Schürhoff (2014) and Hollifield, Neklyudov and Spatt (2017), a set of dealers, known as the core dealers, 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 trade with. Under this interpretation, in our model, it is the small peripheral dealers who provide immediacy for the large core dealers. The prediction seems counterintuitive. But there 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. That the two allocations coincide perhaps is not surprising given a competitive inter-dealer market. 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 (highest) level, dealers who possess inventory (spare capacity) to sell (buy) can only be few and far between. In equilibrium, the inter-dealer price 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.

Related Literature Our framework is adapted from the seminal models of OTC markets in Duffie et al. (2005) and Lagos and Rocheteau (2009). In these models, to simplify, the authors assume that 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 apparently not set up to study inter-dealer trades as the trades cannot possibly exhibit any distinctive structure.

In this paper, we extend the two aforementioned models by assuming that dealers only have periodic access to the inter-dealer market and that they possess different inventory capacities to study the direction and structure of inter-dealer trades. First of all, these changes give rise to a model in which dealers may choose to hold inventory to facilitate future trades. Dealers in Lagos, Rocheteau and Weill (2011) and Weill (2011), respectively, also have incentives to hold inventory when there is a negative shock knocking the market off the steady state and when there is a transient selling pressure in a competitive dynamic market.

Our primary contribution is an investigation of how the asset supply helps determine the direction of trades among dealers facing different inventory constraints and the implications thereof on how small dealers in the periphery of the trading network provide immediacy for large dealers in the core of the network. Our paper then contributes to the literature on how dealers specialize and form core-periphery trading networks in which the trading direction is also persistent. The literature has studied how a dealer becomes a core dealer when his search ability is high (Neklyudov (2015)), when he invests more to raise his contact rate (Farboodi, Jarosch and Shimer (2018)) or to improve his bargaining skills (Farboodi, Jarosch, Menzio and Wiriadinata (2018)), or when the dealer specializes in serving clients who trade frequently (Sambalaibat (2018)), how the usual network externality should lead to most dealers choosing to set up costly connections with just a handful of dealers (Wang (2017)), and how information imperfection should result in individual agents specializing in market making and taking on the role of core dealers (Chang and Zhang (2019)).[4]

Dealers' inventory constraints become especially relevant following the balance-sheet regulations that come into force in the aftermath of the 2008 financial crisis. Our paper adds to the growing literature on the roles of such constraints in determining the direction, structure and volume of inter-dealer trades. In particular, Dunne, Hau and Moore (2015) study the effects of the interaction of inventory constraint and adverse selection on market stability while Cimon and Garriott (2018) show how dealers are adapting to the recent regulations on inventory by shifting to an agency basis of trade. Besides, Choi and Huh (2018) find that customers, not

[4] Other explanations for why dealers trade among themselves in the literature include the amassment of cash endowment in Colliard and Demange (2017), the mitigation of information asymmetry in Glode and Opp (2016), the sharing of inventory risks among risk averse dealers in Ho and Stoll (1983), Atkeson, Eisfeldt and Weill (2015) and Üslü (2019).

dealers, are increasingly the liquidity providers post-crisis, a finding in line with our prediction that small dealers provide immediacy for large ones.

Small dealers in our model trade to provide immediacy for large dealers as they value a unit of inventory in between large dealers having an empty inventory and large dealers already holding a unit inventory. There is a subtle similarity to how investors having an intermediate valuation of the asset in Hugonnier, Lester and Weill (2018) and Shen, Wei and Yan (2018) endogenously become dealers, buying from investors with the lowest valuation and then staying on the market selling to investors with the highest valuation.[5] The difference between our model and theirs is that the difference in valuation in ours arises endogenously rather assumed and that the small dealers in our model, as intermediaries for large dealers, either just sell to or buy from their large dealer customers in a given market.

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 and compare those implications against the available empirical evidence in Section 3. In Section 4, we explore two additional implications of the model as pertaining to how dealers' inventories and the inter-dealer trading volume vary with the asset supply. In Section 5, we discuss two extensions of the model and demonstrate how the major results hold in more general settings. 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 of the extensions of the model in Section 5, the formal analysis of the constrained efficiency of equilibrium, and an additional extension that adds to the generality of our analysis.

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. An investor can hold either zero or one

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

unit of the asset at a time. A high-valuation investor derives a per period return of $v > 0$ in holding a unit of the asset, whereas low-valuation investors and dealers derive the same return normalized to zero. Investors can only buy and 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 remain as high-valuation investors but have yet to acquire a unit of the asset, they constitute the population of investor-buyers (I_B) in the market. The low-valuation investors who do own a unit of the asset become the investor-sellers (I_S) 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 as governed by a constant-returns matching function. Define market tightness

$$\theta = \frac{n^D}{n_B^I + n_S^I} \tag{1}$$

as the ratio of the measure of all dealers $n^D = n^{SD} + n^{LD}$, for n^{SD} and n^{LD} denoting the respective measures of small and large dealers, to the measure of all investors on the market, equal to the sum of the measures of investor-buyers n_B^I and investor-sellers n_S^I. Each investor is randomly matched with a dealer at probability $\eta(\theta) \in [0, 1]$, whereas a dealer is randomly matched with an investor at probability $\mu(\theta) = \eta(\theta)/\theta$. The meeting probability $\eta(\theta)$ satisfies the usual conditions:

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

Only dealers holding at least a unit of the asset in inventory can sell to the investor-buyers and only dealers having at least a unit of spare inventory capacity can buy from the investor-sellers that they meet. 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 split the surplus in equal halves. Those investor-buyers who succeed in buying a unit would start collecting the payoff v in the next period as long as they remain as high-valuation investors, whereas those investor-sellers 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. Dealers trade in the market to rebalance inventories to prepare for the next round of trading with investors. The analysis of inter-dealer trading in this market is the main focus of the paper.

At the end of the period, each high-valuation investor is hit by a liquidity shock at probability $\delta \in (0,1)$ and turns into a low-valuation investor forever thereafter.[6] Those who own a unit of the asset turn into investor-sellers in the next period and those who do not cease to remain as investor-buyers but simply exit the market.

For brevity, we restrict attention to studying steady-state equilibrium in this paper.

2.2 Value Functions

A small dealer, S_i, $i = 0, 1$, is either holding 0 or 1 unit of the asset at a time. Write V_i^{SD} as the asset value of an S_i entering the investor-dealer market in the first subperiod with an i-unit inventory and W_i^{SD} the asset value of the dealer entering the inter-dealer market in the second subperiod. If the asset is traded in the inter-dealer market at price p,

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

$$W_1^{SD} = W_0^{SD} + p, \tag{3}$$

A large dealer L_i, $i = 0, 1, 2$, can hold up to two units of the asset in inventory. Then,

$$W_0^{LD} = \max\left\{\beta V_0^{LD}, \beta V_1^{LD} - p, \beta V_2^{LD} - 2p\right\}, \tag{4}$$

$$W_i^{LD} = W_0^{LD} + ip, \text{ for } i = 1, 2, \tag{5}$$

[6] As in other search-based models of OTC market that use an overlapping-generation structure (Vayanos and Wang (2007), Vayanos and Weill (2008), Afonso (2011) and Sambalaibat (2018)), all investors, irrespective of ownership status, are hit by the liquidity shock at the same probability at the end of each period.

In (2), an S_0 entering the inter-dealer market chooses between buying a unit and not buying. In (4), an L_0 entering the inter-dealer market may choose to buy up to two units. Eqs. (3) and (5) follow as a unit of the asset is worth p in the competitive inter-dealer market.

We call any dealer having at least a unit of spare inventory to buy from investors a dealer-buyer. A dealer-buyer gains from buying from an investor-seller for the amount p, for what a unit of the asset is worth in the inter-dealer market, minus the payment the dealer makes to the investor for the unit. To the investor-seller, the gain from trade is equal to the payment from the dealer net of the continuation value of being an investor-seller $\left(U_S^I\right)$. This means that the surplus of trade between an investor-seller and any dealer-buyer is simply equal to

$$z_{I_S} = p - \beta U_S^I, \tag{6}$$

Dealers having an empty inventory can only be dealer-buyers. Then, the respective asset values of an S_0 and an L_0 entering the investor-dealer market are equal to

$$V_0^{SD} = W_0^{SD} + \eta(\theta) \frac{n_S^I}{n^D} \frac{z_{I_S}}{2}, \tag{7}$$

$$V_0^{LD} = W_0^{LD} + \eta(\theta) \frac{n_S^I}{n^D} \frac{z_{I_S}}{2}, \tag{8}$$

where $\eta(\theta) \frac{n_S^I}{n^D} = \mu(\theta) \frac{n_S^I}{n_S^I + n_B^I}$ is the probability that a dealer meets an investor-seller. On the other side of the market, the asset value of an investor-seller is equal to

$$U_S^I = \eta(\theta) \frac{n_B^D}{n^D} \frac{z_{I_S}}{2} + \beta U_S^I., \tag{9}$$

where

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

is the measure of all dealer-buyers, for n_i^{SD} and n_i^{LD} denoting the respective measures of small and large dealers holding an i-unit inventory when the investor-dealer market opens.

We call any dealer who holds at least a unit of the asset in inventory for sale to investors a dealer-seller. A dealer-seller gains from selling to an investor-buyer for the payment he receives from the investor net of the value of the asset p in the inter-dealer market. To the investor-buyer, the gain from trade is equal to the capital gain of acquiring the unit minus the payment made to the dealer. The investor-buyer's capital gain, for U_H^{ON} and U_B^I denoting, respectively, the asset value of a high-valuation investor holding a unit of the asset (high-valuation owner hereafter) and the asset value of an investor-buyer, is equal to $(1-\delta)\, U_H^{ON} + \delta U_S^I - (1-\delta)\, U_B^I$. This is the case since the investor only remains a high-valuation investor in the next period absent any liquidity shock. All this means that there must be the same surplus of trade between an investor-buyer and any dealer-seller, equal to

$$z_{I_B} = \beta \left((1-\delta) \left(U_H^{ON} - U_B^I \right) + \delta U_S^I \right) - p. \tag{10}$$

Dealers holding a full inventory can only be dealer-sellers. Then, the respective asset values of an S_1 and an L_2 entering the investor-dealer market are equal to

$$V_1^{SD} = W_1^{SD} + \eta(\theta) \frac{n_B^I}{n^D} \frac{z_{I_B}}{2}, \tag{11}$$

$$V_2^{LD} = W_2^{LD} + \eta(\theta) \frac{n_B^I}{n^D} \frac{z_{I_B}}{2}, \tag{12}$$

where $\eta(\theta) \frac{n_B^I}{n^D} = \mu(\theta) \frac{n_B^I}{n_S^I + n_B^I}$ is the probability that a dealer meets an investor-buyer. On the other side of the market, the asset value of an investor-buyer is equal to

$$U_B^I = \eta(\theta) \frac{n_S^D}{n^D} \frac{z_{I_B}}{2} + (1-\delta)\, \beta U_B^I, \tag{13}$$

where

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

is the measure of all dealer-sellers.

As an L_1, in holding a unit inventory and having a unit of spare inventory capacity, is both a dealer-buyer and a dealer-seller, the dealer has asset value equal to

$$V_1^{LD} = W_1^{LD} + \eta(\theta) \frac{n_S^I}{n^D} \frac{z_{I_S}}{2} + \eta(\theta) \frac{n_B^I}{n^D} \frac{z_{I_B}}{2}. \tag{14}$$

Finally, a high-valuation owner derives a per period return υ from holding a unit of the asset and turns into a low-valuation investor *cum* investor-seller at probability δ at the end of the period in which case,

$$U_H^{ON} = \upsilon + \beta \left((1-\delta)\, U_H^{ON} + \delta U_S^I \right). \tag{15}$$

2.3 Inter-dealer Market Trades

By (2) and (3), whether an S_0 entering the inter-dealer finds it optimal to buy and whether an S_1 finds it optimal to sell depend on how the inter-dealer market price p compares with $\beta\left(V_1^{SD} - V_0^{SD}\right)$. Similarly, large dealers entering the market make their buy and sell decisions by comparing p against $\beta\left(V_1^{LD} - V_0^{LD}\right)$ and $\beta\left(V_2^{LD} - V_1^{LD}\right)$.

Proposition 1 $V_1^{LD} - V_0^{LD} \geq V_1^{SD} - V_0^{SD} \geq V_2^{LD} - V_1^{LD}$. *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 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. In our model, all dealers may each meet up to one investor, who may be a seller or a buyer, in a given round of trading. The benefit of holding a unit inventory for a large dealer versus a small dealer is flexibility: a large dealer who owns a unit can sell as well as buy, whereas a small dealer who owns a unit may only sell. This explains the first inequality of the Proposition. If a dealer earns a strictly positive surplus from buying from an investor with $z_{I_S} > 0$, the inequality is strict. There is the same cost of holding a full inventory for the two types of dealers in forgoing any opportunity to buy from an investor in the next round of trading. The small dealer, however, benefits from acquiring the unit as he may then be able to sell to an investor-buyer should he meet one in the next period. There is not any such upside to the purchase by a large dealer already holding a one-unit inventory as the dealer can sell without adding a unit to inventory. This explains the second inequality of the Proposition. If a dealer earns a strictly positive surplus from selling to an investor with $z_{I_S} > 0$, the inequality is strict.

Who buys and who sells in the inter-dealer market depend on what price clears the market, a price that must be bounded by

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

in equilibrium 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. But such a parameter configuration can at best make up a zero-measure subset of the parameter space.[7] 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 mixing probability for the mixed strategy played by the marginal buyers or sellers.

Lemma 1 *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 $z_{I_B}>0$.*

The case for $p=\beta\left(V_2^{LD}-V_1^{LD}\right)$ deserves further explanation. An L_1 is both a dealer-seller and a dealer-buyer to begin with. In filling up his inventory, the dealer turns into an L_2 while losing the status of a dealer-buyer and so must be worse off buying the unit at any positive p and is at best indifferent at $p=0$ and that dealers do not gain from buying from investors at all where $z_{I_S}=0$.[8] There must be a positive surplus in an investor-buyer trade though ($z_{I_B}>0$) since the investor, but not the dealer, is better off owning a unit than otherwise.

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 1,

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

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

[7] See the Online Appendix for the formal proof.

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

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 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

$$m_0^{LD} \geq m_1^{SD} + m_2^{LD}. \tag{16}$$

Because each L_1 is indifferent between selling and not selling, if (16) 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 multiplicity has no bearing at all on the 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. The same kind of multiplicity in the next two types of equilibrium we define in the following is similarly inconsequential.

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

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

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

$$m_1^{SD} \geq m_0^{LD} - m_2^{LD}. \tag{17}$$

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

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

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

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

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 selling in the market to meet the demand out of all S_0s and L_0s each buying one unit; i.e.,

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

In Table 1, we summarize the identities of the buyers and sellers upon entry into and the optimal inventories of large and small dealers upon exiting the inter-dealer market. Notice that in all three equilibrium types, at least a fraction of L_0s buy and at least a fraction of L_2s 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.

Each candidate equilibrium places a set of restrictions on the inventories held by small and large dealers and may obtain only if the inter-dealer market can clear at the given p. We next proceed to complete the definitions of the equilibria and derive the conditions for each type of equilibrium to hold.

Equilibrium	Buyers	Sellers	Optimal Inventory		Restrictions on
			Small dealers	Large dealers	n_i^{SD} and n_i^{LD}
Selling	L_0	$L_2^*, S_1^*, (L_1)$	0	1 and 0	$n_0^{SD} = n^{SD}$ $n_1^{SD} = n_2^{LD} = 0$
Balanced	L_0^*,	L_2^*	0 and 1	1	$n_1^{LD} = n^{LD}$ $n_0^{LD} = n_2^{LD} = 0$
Small dealers sell	(S_0)	S_1			
Small dealers buy	S_0	(S_1)			
Buying	$S_0^*, L_0^*, (L_1)$	L_2	1	1 and 2	$n_1^{SD} = n^{SD}$ $n_0^{SD} = n_0^{LD} = 0$

* inframarginal buyers and sellers who gain from trade; ( ) dealers who are indifferent to trading and do not need to trade for market clearing

Table 1: Characteristics of the three types of equilibrium

2.4 Equilibrium Conditions

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

$$n_0^{SD} + n_1^{SD} = n^{SD}, \tag{20}$$

$$n_0^{LD} + n_1^{LD} + n_2^{LD} = n^{LD}. \tag{21}$$

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

$$n_H^{ON} + n_S^I + n_1^{SD} + n_1^{LD} + 2n_2^{LD} = A, \tag{22}$$

where n_H^{ON} denotes the measure of high-valuation owners.

In the steady state, the respective inflows and outflows of high-valuation owners, investor-sellers, and investor-buyers are equal. That is,

$$(1-\delta)\, n_B^I \eta(\theta) \frac{n_S^D}{n^D} = \delta n_H^{ON}, \tag{23}$$

$$\delta \left(n_H^{ON} + n_B^I \eta(\theta) \frac{n_S^D}{n^D} \right) = \eta(\theta) \frac{n_B^D}{n^D} n_S^I, \tag{24}$$

$$e=\left(\delta+(1-\delta)\,\eta\left(\theta\right)\frac{n_S^D}{n^D}\right)n_B^I. \tag{25}$$

Not all n_i^{SD} and n_i^{LD} can be positive in a given type of equilibrium. For example, in the Selling Equilibrium, because all small dealers exit the inter-dealer market with an empty inventory, $n_0^{SD}=n^{SD}$ and $n_1^{SD}=0$ must hold, and because large dealers may do so with either an empty or a one-unit inventory, $n_2^{LD}=0$ must hold. The restrictions on n_i^{SD} and n_i^{LD} in the three types of equilibrium are as depicted in the last column of Table 1.

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.

$$m_0^{SD}=\left(1-\eta\left(\theta\right)\frac{n_S^I}{n^D}\right)n_0^{SD}+\eta\left(\theta\right)\frac{n_B^I}{n^D}n_1^{SD}, \tag{26}$$

$$m_1^{SD}=\eta\left(\theta\right)\frac{n_S^I}{n^D}n_0^{SD}+\left(1-\eta\left(\theta\right)\frac{n_B^I}{n^D}\right)n_1^{SD}, \tag{27}$$

$$m_0^{LD}=\left(1-\eta\left(\theta\right)\frac{n_S^I}{n^D}\right)n_0^{LD}+\eta\left(\theta\right)\frac{n_B^I}{n^D}n_1^{LD}, \tag{28}$$

$$m_1^{LD}=\eta\left(\theta\right)\frac{n_S^I}{n^D}n_0^{LD}+\left(1-\eta\left(\theta\right)\frac{n_S^I}{n^D}\right)\left(1-\eta\left(\theta\right)\frac{n_B^I}{n^D}\right)n_1^{LD}+\eta\left(\theta\right)\frac{n_B^I}{n^D}n_2^{LD}, \tag{29}$$

$$m_2^{LD}=\eta\left(\theta\right)\frac{n_S^I}{n^D}n_1^{LD}+\left(1-\eta\left(\theta\right)\frac{n_B^I}{n^D}\right)n_2^{LD}. \tag{30}$$

For example, (26) 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.

2.5 Equilibrium

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_S^I and n_B^I that satisfy (20)-(25), the restrictions on n_i^{SD} and n_i^{LD} in Table 1 and the market-clearing conditions for the type of equilibrium under consideration in (16)-(19), with m_i^{SD} and m_i^{LD} given by (26)-(30).

Proposition 2 *For* $A \in \left[0, n^{LD} + \frac{e}{\delta}\right)$, *define*

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

where $\Omega_S(0) < 0$, $\partial \Omega_S(A)/\partial A > 0$, *and* $\lim_{A \to n^{LD} + \frac{e}{\delta}} \Omega_S(A) > 0$. *For* $A \in \left(n^D + \frac{e}{\delta}, \infty\right)$, *define*

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

where $\lim_{A \to n^D + \frac{e}{\delta}} \Omega_B(A) > 0$, $\partial \Omega_B(A)/\partial A < 0$ *and* $\lim_{A \to \infty} \Omega_B(A) < 0$.

(a) For $A < n^{LD} + \frac{e}{\delta}$ *and* $\Omega_S(A) < 0$, *the Selling Equilibrium holds. As A increases up to where* $\Omega_S(A) = 0$, $n_1^{LD} = n^{LD}$.

(b) Thereafter, $\Omega_S(A) > 0$, *whereupon the market turns into the Balanced Equilibrium with* $n_1^{SD} < n^{SD}/2$. *The condition* $n_1^{SD} < n^{SD}/2$ *is equivalent to* $n_B^I > n_S^I$ *and* $m_0^{LD} > m_2^{LD}$, *whereby small dealers sell to clear the inter-dealer market. For* $n^{LD} + \frac{e}{\delta} \leq A < n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta}$, *the Balanced Equilibrium with* $n_1^{SD} < n^{SD}/2$ *continues to hold until* $A = n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta}$ *at which point* $n_1^{SD} = n^{SD}/2$ *and* $m_0^{LD} = m_2^{LD}$ *so that the sales and purchases made by large dealers just clear the inter-dealer market.*

(c) For $n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta} < A \leq n^D + \frac{e}{\delta}$, *the Balanced Equilibrium with* $n_1^{SD} > n^{SD}/2$, *in which case small dealers buy to clear the inter-dealer market, first holds. For* $A > n^D + \frac{e}{\delta}$ *and* $\Omega_B(A) > 0$, *the Balanced Equilibrium with* $n_1 > n^{SD}/2$ *continues to hold. As A increases up to where* $\Omega_B(A) = 0$, $n_1^{SD} = n^{SD}$.

(d) Thereafter, $\Omega_B(A) < 0$, *whereupon the market turns into the Buying Equilibrium in which* $n_2^{LD} > 0$.

Proposition 2 essentially describes how the asset would be allocated to the two types of dealers in equilibrium. Part (a) is concerned with when the asset supply A is at the lowest level relative to the asset demand, the latter of which can be measured by the rate e at which investors enter the market, there would only be enough of the asset left, after accounting for the amount held by investors, to allow for a fraction of large dealers, who value the first unit of the asset the most, to each hold a unit in inventory, whereupon the Selling Equilibrium holds. The Selling Equilibrium then holds up to where the market has enough of the asset

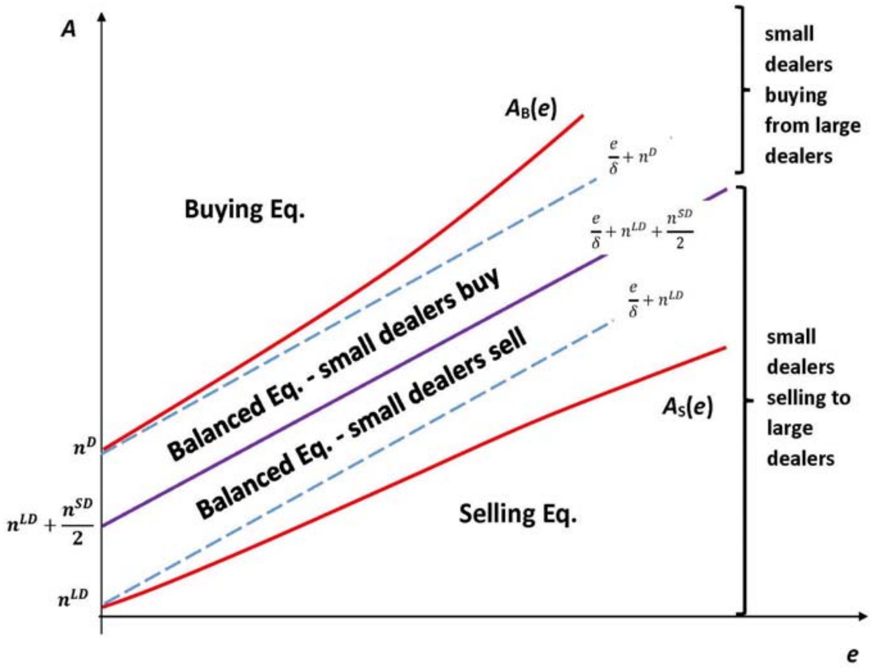

Figure 1: Equilibrium – Existence and Uniqueness

for each large dealer to hold a unit, at which point the Balanced Equilibrium comes into play as referred to by Part (b). In the Balanced Equilibrium, all large dealers and a fraction of small dealers each hold a unit in inventory. When the latter fraction is less than one-half due to a still relatively small asset supply, there are more buyers than sellers among investors so that dealers are more likely to meet the former than the latter to result in there being more L_0s, who buy in the inter-dealer market, than L_2s, who sell in the inter-dealer market, at the closing of the investor-dealer market. In this case, small dealers sell to eliminate the excess demand among large dealers in the inter-dealer market. Part (c) of the Proposition is for A large enough to allow for more than one-half of all small dealers to hold a unit inventory. With the now relatively large asset supply, there are fewer buyers than sellers among investors so that dealers are less likely to meet the former than the latter to result in there being fewer L_0s than L_2s at the closing of the market. In this case, small dealers buy to eliminate the excess supply among large dealers in the inter-dealer market. When the asset supply has risen to the level to allow for all small dealers to each hold a unit in inventory, the Balanced Equilibrium eventually gives way to the Buying Equilibrium, as referred to by Part (d) of the Proposition, in which a fraction of large dealers would hold a two-unit inventory. Figure 1 illustrates the transition from the Selling to the Balanced and then to the Buying Equilibria as A increases for each e, where $A_S(e)$ and $A_B(e)$ are the respective solutions for A to $\Omega_S(A) = 0$ and $\Omega_B(A) = 0$ for each e.

It turns out that the equations in (26)-(30) linking the measures of dealers in the inter-dealer to the investor-dealer markets, together with the restrictions on n_i^{SD} and n_i^{LD} in Table 1, suffice to guarantee that the market-clearing conditions in (16)-(19) would be satisfied. 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.

3 Who Buys and Who Sells?

3.1 Trading Direction

A major point of interest of our analysis is that the asset supply A, relative to the demand as measured by e, plays an important role in determining the direction of trades among dealers having different capacities to hold the asset. A direct corollary of Proposition 2 is that:

Corollary 1

(a) For $A \leq A_S(e)$, small dealers only sell to but do not buy from large dealers.

(b) For $A \in \left(A_S(e), n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta}\right)$, small dealers sell to more than they buy from large dealers if they buy from large dealers at all.

(c) For $A \in \left(n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta}, A_B(e)\right)$, small dealers buy from more than they sell to large dealers if they sell to large dealers at all.

(d) For $A \geq A_B(e)$, small dealers only buy from but do not sell to large dealers.

Parts (a) and (d), respectively, are due to how small dealers only sell in the Selling Equilibrium and only buy in the Buying Equilibrium. In the Balanced Equilibrium where $A < n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta}$ 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 buy given that small dealers are indifferent between holding an empty or a one-unit inventory, small dealers must sell more than they buy for market clearing. Part (c) describes the mirror opposite of Part (b). In Figure 1, for each e, small dealers on balance sell to large dealers for A up to the demarcation between the first and the second halves of the Balanced Equilibrium but buy from large dealers for any larger A.

Because large dealers value the first unit of inventory more than small dealers, for an asset supply at a relatively low level, the competitive inter-dealer market would allocate the asset to large dealers as far as possible. This allocation takes place through small dealers selling to large dealers. Because small dealers value the last unit of inventory more than large dealers, for an asset supply at a relatively high level, the competitive inter-dealer market would first fill up the inventories of small dealers. This allocation takes place through small dealers buying from large dealers.

3.2 Small Dealers Provide Immediacy for 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 for another dealer if the 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. Apparently, large dealers in our model do not provide immediacy for small dealers.

Proposition 3 *The probability that a dealer meets an investor-buyer $\eta(\theta)\frac{n_B^I}{n^D}$ is everywhere decreasing and the probability that a dealer meets an investor-seller $\eta(\theta)\frac{n_S^I}{n^D}$ is everywhere increasing in A.*

The Proposition, together with Proposition 2(b)-(c), implies that in the Selling Equilibrium and the first half of the Balanced Equilibrium, $\eta(\theta)\frac{n_B^I}{n^D} > \eta(\theta)\frac{n_S^I}{n^D}$, whereas in the second half of the Balanced Equilibrium and in the Buying Equilibrium, $\eta(\theta)\frac{n_B^I}{n^D} < \eta(\theta)\frac{n_S^I}{n^D}$. All this means that for a relatively small asset supply, with which the scarcity of the asset gives rise to more opportunities for dealers to meet investor-buyers than sellers, small dealers sell to large dealers to provide inventory for the latter to sell to investors. For a relatively large asset supply, with $\eta(\theta)\frac{n_B^I}{n^D} > \eta(\theta)\frac{n_S^I}{n^D}$, dealers need spare inventory capacity more than inventory as the abundance of the asset gives rise to more opportunities for dealers to meet investor-sellers than buyers. In this case, small dealers buy from large dealers to help them free up capacities to buy from investors. In our model, it is the small dealers who provide immediacy for large dealers.

The last implication is consistent with the findings in Adrian et al. (2017) if the booming market pre-crisis is time during which it is hardest for dealers to find willing sellers among their customers due to the robust demand for risky investment then and the market bust post-crisis is time during which it is hardest for dealers to find willing buyers with an appetite for even relatively safe investment. In the earlier period, large dealers gain inventory from small dealers and expand their balance sheets faster than the latter. In the later period, small dealers amass inventory from large dealers and expand their balance sheets relative to those of large dealers.

A further implication pertaining to trading direction is that the direction of trade between a given pair of dealers tends to be persistent – while there is no persistent trading direction among large dealers since they sell as well as buy in any equilibrium, in a given type of equilibrium, small dealers either tend to sell to or buy from large dealers. 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%.

3.3 Core-Periphery Trading Network

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, trade with all dealers while the rest, referred to as the peripheral dealers, trade only with the core dealers but not among themselves (Li and Schürhoff (2014) and Hollifield et al. (2017)).

In our model, 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.[9] In all three types of equilibrium, 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 trade with.[10] Hence, the large dealers in our model can be interpreted to play the role of core dealers in a core-periphery trading network, trading with all dealers while the small dealers play the role of peripheral dealers, trading only with the core dealers.[11] Under this interpretation, our model says that the peripheral dealers trade to provide immediacy for the 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.[12]

[9] In the Balanced Equilibrium, there can be small dealers buying (selling) when small dealers should be selling (buying) 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 or if there is a small information asymmetry as we shall argue below in Section 4. In this case, either all small dealers sell or buy in the Balanced Equilibrium, as in the Selling and Buying Equilibria.

[10] 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 bilateral trades, see Gale (2000).

[11] 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).

[12] 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 (2017) and Zhong (2016)), dealers that generally provide immediacy for peripheral dealers by virtue of being more

This implication is counterintuitive but there 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 chain, buying from an investor for selling to other dealers, are smaller than the percentage spreads they earn when they are the last links, buying from another dealer for selling to an investor. Meanwhile, there are no statistically significant differences between the two spreads for core dealers.

In the Selling Equilibrium and the earlier part of the Balanced Equilibrium of our model, which obtains with a relatively small asset supply, small dealers are the first links of the intermediation chain as sellers in the inter-dealer market. If there is only a relatively small asset supply, the small peripheral dealers would be buying from investors in 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.

In the latter part of the Balanced Equilibrium and the Buying Equilibrium, which obtains with a relatively large asset supply, the small peripheral dealers are the last links of the intermediation chain as buyers in the inter-dealer market. If there is a relatively large asset supply, the small peripheral dealers would be selling to investors in a slow market in which it takes a long time on average for a dealer to sell a unit. The intermediation services provided by small dealers for large dealers by selling to investors on their behalf should then command a high return. Thus, in our model, as what Hollifield et al. (2015) find in their empirical analysis:

Conjecture 1 *Small dealers should earn a smaller markup* $\rho_{DB} \equiv \frac{p-p_{I_S}}{p_{I_S}}$ *when they are the first links of the intermediation chain, buying from investors at price* p_{I_S} *for selling to other dealers than the markup* $\rho_{DS} \equiv \frac{p_{I_B}-p}{p}$ *they earn when they are the last links, buying from other dealers for selling to investors at price* p_{I_B}*.*[13]

connected to other dealers (Wang (2017)), or dealers that simply tend to trade more frequently (Neklyudov (2019), Hugonnier et al. (2018) and Sambalaibat (2018)).

[13]Because there is the same surplus of trade between an investor-seller and any dealer-buyer, all dealers buy from investors at the same investor-sell price p_{I_S}. Because there is the same surplus of trade between an

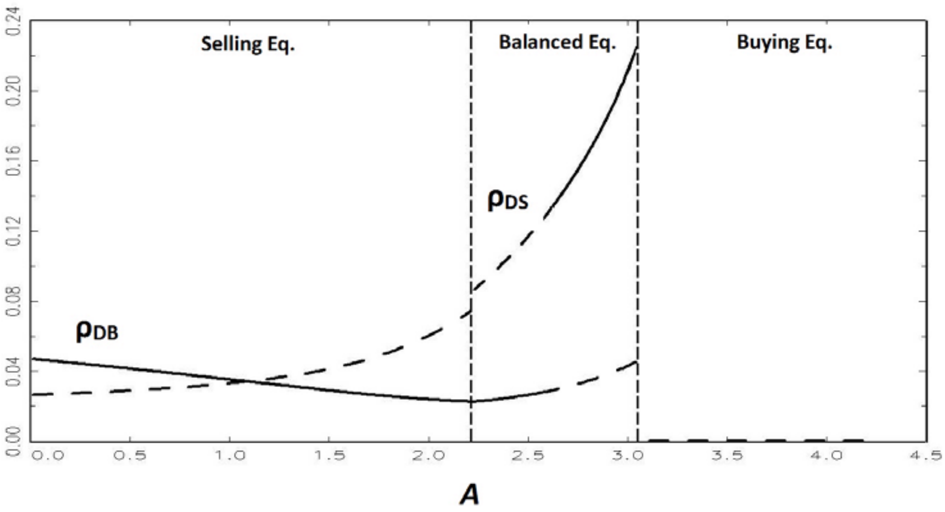

Figure 2: Dealers' Markups

To confirm the conjecture, we calculate and then plot the two markups against A in Figure 2.[14] 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 large ρ_{DS} in the Buying Equilibrium. In Figure 2, small dealers earn the markup ρ_{DB} as the first links of the intermediation chain in the Selling and the first half of the Balanced Equilibria and the markup ρ_{DS} as the last links in the second half of the Balanced and the Buying Equilibria. Indeed, the first markup (the solid portion of the ρ_{DB} curve) is everywhere below the second (the solid portion of the ρ_{DS} curve) over the relevant ranges of A.

Last but not least, 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 large core dealers in our model

investor-buyer and any dealer-seller, all dealers sell to investors at the same investor-buy price p_{I_B}.

[14] The example assumes $\eta(\theta) = 1 - e^{-\theta}$, $e = 0.15$, $\delta = 0.08$, $\beta = 0.98$, $n^{LD} = 0.5$, $n^{SD} = 0.5$ and A varying from 0.1 and up. The equations for all endogenous variables can be found in the proofs of Lemma 1 and Proposition 2. The equations for p_{I_S} and p_{I_B} are derived and presented in the Online Appendix.

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 2, 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 for smaller As in the Selling Equilibrium. Averaging over the A for assets included in a given sample, there can be no statistically significant differences between the two observed markups large dealers earn.

4 Other Implications

4.1 Shocks to Asset Supply

A further implication of Proposition 2 is that:

Corollary 2 *In the new steady state after a positive shock to the asset supply, only large dealers expand inventory at low levels of asset supply while the Selling Equilibrium holds, only small dealers expand inventory at intermediate levels of asset supply while the Balanced Equilibrium holds, and again only large dealers expand inventory at high levels of asset supply while the Buying Equilibrium holds.*

The Corollary holds because in the Selling Equilibrium, all small dealers just hold an empty inventory, while in the Balanced Equilibrium, all large dealers hold one and only one unit of inventory, and finally in the Buying Equilibrium, all small dealers are already holding a full inventory. A readily testable implication of the Corollary is that if it is only large dealers who are holding any inventory in a given market, they would be the only dealers who would expand inventory in response to any moderate asset supply increase. In a slack market where all dealers are holding inventory, again they would be the only dealers who would expand inventory in response to any asset supply increase. On the other hand, in a moderate market with some, but not all small dealers holding inventory, small dealers would be the only ones who would expand inventory.

4.2 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 would not take place if there exists a small cost for each trade to be completed or a small information asymmetry.[15] In studying the model's implications on trading volumes, there is good reason to focus solely on the particular 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,

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

$$TV = \begin{cases} m_0^{LD} & A \leq n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta} \\ m_2^{LD} & A \geq n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta} \end{cases},$$

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

Proposition 4 *The inter-dealer market trading volume changes non-monotonically with A:*

(a) first increasing in the Selling Equilibrium,

(b) becomes decreasing in the Balanced Equilibrium for $A \leq n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta}$ *where small dealers sell,*

(c) then turns increasing again in the Balanced Equilibrium for $A \geq n^{LD} + \frac{n^{SD}}{2} + \frac{e}{\delta}$ *where small dealers buy,*

(d) and eventually becomes decreasing in the Buying Equilibrium.

That is, the trading volume peaks at where the Selling Equilibrium turns into the Balanced Equilibrium and when the Balanced Equilibrium turns into the Buying Equilibrium.

The Proposition says the trading volume is M-shaped and that the inter-dealer market is most active when the asset supply is at a relatively low level but not at the lowest level and

[15] For example, Bethune, Sultanum and Trachter (2016) show how private information in an OTC market hinders trades with gains from trade below the information rent.

at a relatively high level but not at the highest level. The intuition lies in how the measures of buyers and sellers vary with the asset supply.[16] Inter-dealer trading should be most active when both buyers and sellers abound. For the smallest asset supply, there can only be few dealers possessing any inventory to sell in the inter-dealer market; for the largest asset supply, there can only be few dealers having any spare capacity to buy in the market. For intermediate levels of A, the inter-dealer market price would settle at a moderate level, equal to the Balanced Equilibrium price, at which small dealers, who do not gain from inter-dealer trades at the given price, trade only for eliminating any excess demand or supply among large dealers. Thus, there cannot be much inter-dealer trading when the asset supply is either lowest, highest, or right in between.

5 Extension and Robustness

Our major results – Proposition 1 and the implications thereof – seemingly rest on a number of simplifying assumptions that may appear ad hoc. In the following, we discuss how the Proposition and its implications survive two generalizations that seem most warranted. In Appendix 7.2, we discuss an additional extension that further adds to the generality of the analysis.

5.1 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 an equally plausible assumption is that they can meet up to two investors 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 investors, a reasonable matching technology should be such that

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

[16] The same intuition is known in the literature to explain bell-shaped transaction volumes. See, for example, Kiyotaki and Wright (1989) and Hugonnier et al. (2018).

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

2. the probability that a large dealer meets two investor-sellers (-buyers) is weakly lower than the probability that a small dealer meets one investor-seller (-buyer),

then the ranking in Proposition 1 remains.

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 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 a large dealer meeting only one investor-seller still has capacity to buy from the investor. By (2), 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 at a 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 (2) is strict.

Next, consider the costs and benefits of using up the last unit of spare capacity for the two types of dealers. Exhausting one's capacity is costly to a dealer as long as the dealer shall meet one or more investor-seller in the next period. By (1), 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 (2) 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 (1) or (2) is strict.

As an example of how Assumptions (1) and (2) above can be satisfied, suppose that each large dealer participates as two independent agents in the matching process and thereby earns two independent matching opportunities. In this case, market tightness is given by $\theta = \frac{n^{SD}+2n^{LD}}{n_S^I+n_B^I}$. Notice that

$$1 - \left(1 - \mu\left(\theta\right)\frac{n_S^I}{n_B^I + n_S^I}\right)^2 > \mu\left(\theta\right)\frac{n_S^I}{n_B^I + n_S^I} > \left(\mu\left(\theta\right)\frac{n_S^I}{n_B^I + n_S^I}\right)^2, \tag{31}$$

where the first term is the probability that a large dealer meets at least one investor-seller, the second term the probability that a small dealer meets one investor-seller and the last term the probability that a large dealer meets two investor-sellers. Both Assumptions (1) and (2) hold true. In general, the two inequalities above would only fail to hold if the two matching outcomes for the large dealer are perfectly and positively correlated events – the large dealer either meets two investor-sellers at the same probability that a small dealer meets one investor-seller or not a single investor-seller at all, in which case all three terms in (31) become one and the same.[17]

5.2 Frictional Inter-dealer Market

There exists ample evidence that the inter-dealer market is better described as a decentralized market (Li and Schürhoff (2019) and Henderschott, Li, Livdan and Schürhoff (2017)), 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 as dealers in our model, who only have periodic access to the competitive inter-dealer market, do. Where the incentives are similar, the major implications of our model should survive in a 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 but is otherwise identical to the main model of the paper. The model is set in continuous time as it is a more convenient setting for a model in which both the investor-dealer and the inter-dealer markets are decentralized. With a frictional inter-dealer market, in addition to meeting an investor-buyer at the rate $\eta(\theta)\frac{n_B^I}{n^D}$ and an investor-seller at the rate $\eta(\theta)\frac{n_S^I}{n^D}$, a dealer meets another randomly selected dealer at a fixed rate α per unit of time. 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. All notations for the revised model have the same meanings as for the main model.

Proposition 5 $V_1^{LD} - V_0^{LD} \geq V_1^{SD} - V_0^{SD} \geq V_2^{LD} - V_1^{LD}$ *in any steady-state equilibrium involving trades between investors and dealers. The two equalities are strict unless the surplus for the* I_B*-*L_1 *match* $z_{I_B,L_1} \equiv U_H^{ON} - U_B^I - \left(V_1^{LD} - V_0^{LD}\right)$ *is equal to zero.*

[17] It is perhaps of interest to note that in a model where large dealers have two meeting opportunities, the inter-dealer price in the Buying Equilibrium should stay positive even if dealers and low-valuation investors do not value the asset.

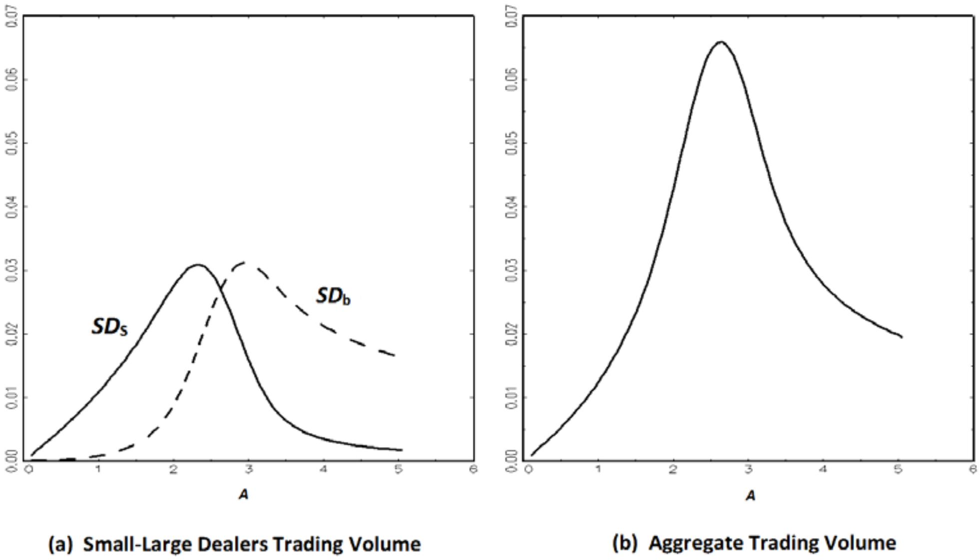

Figure 3: Inter-Dealer Market Trading Volume

There is thus the same ranking of the marginal benefits of inventory as in Proposition 1, which again implies that profitable bilateral trades between two dealers only include those between an L_0 and an S_1 (by the first inequality in the Proposition), between an S_0 and an L_2 (by the second inequality), and between an L_0 and an L_2 (by the two inequalities together).

In the competitive inter-dealer market of the main model, small dealers on balance sell to (buy from) large dealers when the asset supply is meagre (abundant). Below, we set out to verify that the same tendency remains in the frictional inter-dealer market, despite small dealers selling to and buying from large dealers at all levels of asset supply in the market, where there are positive surpluses from all such trades.

In our numerical analyses,[18] we find that the volume of small dealers' sales to large dealers

[18] It does not seem possible to analytically solve the model and to derive conditions on model fundamentals for the existence and uniqueness of equilibrium from the equilibrium conditions as detailed in Appendix 7.1. Assuming a Walrasian inter-dealer market simplifies considerably through the restrictions on n_i^{SD} and n_i^{LD} in Table 1 and thereby enables us to derive a rich set of analytical results. The numerical example uses the same parameterization used in the plots in Figure 2, except for the $\eta(\theta)$ function, which is now assumed to be $\eta(\theta) = \theta^{0.5}$. The change is necessitated by $\eta(\theta)$ being a meeting rate function instead of a meeting probability function as in the main model.

and the volume of small dealers' purchases from large dealers, given by, respectively,[19]

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

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

are both bell-shaped functions of A as shown in Figure 3(a). For a small A, there can only be few dealers holding a full inventory (S_1s and L_2s) and selling in the inter-dealer market and for a large A, there can only be few dealers searching with an empty inventory (S_0s and L_0s) and buying in the market. In either case, SD_s and SD_b must only be at relatively low levels. At intermediate levels of A, where both buyers and sellers in the inter-dealer market become fairly numerous, SD_s and SD_b attain their respective maximum.

The Figure also shows that $SD_s > SD_b$ for smaller As and vice versa.[20] Hence, even though small dealers do buy from and sell to large dealers for all levels of A, they sell to more than they buy from large dealers when the asset supply is at a low level but buy from more than they sell to the latter when the asset supply is at a high level, just as in the competitive inter-dealer market model. This happens because small dealers only trade with large dealers in the frictional inter-dealer market also. In particular, given that S_1s only sell to L_0s whereas L_2s sell to S_0s in addition to L_0s, the ratio n_1^{SD}/n_2^{LD} is largest when S_0s are most plentiful since in this case there would be fewest dealers remaining as L_2s relative to S_1s. In turn, S_0s are most plentiful amidst a small asset supply. And so in this environment, SD_s tends to exceed SD_b. On the other hand, given that S_0s only buy from L_2s, whereas L_0s buy from S_1s in addition to L_2s, the ratio n_0^{LD}/n_0^{SD} is smallest when S_1s are most plentiful since in this case there would be fewest dealers remaining as L_0s relative to S_0s. In turn, S_1s are most plentiful amidst a large asset supply for the additional units of the asset to be in circulation. And so in this environment, SD_s tends to fall below SD_b.

In view of Figure 3(a), in which both SD_s and SD_b peak near each other around some intermediate level of asset supply, the aggregate trading volume, equal to the sum of SD_s, SD_b,

[19]The equation for SD_s obtains as 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.

[20]These results are from assuming $n^{SD} = n^{LD}$. As checks for robustness, we find the same qualitative results hold with $n^{SD} > n^{LD}$ and $n^{SD} < n^{LD}$

and the volume of L_0-L_2 trades $\left(\alpha \frac{n_0^{LD} n_2^{LD}}{n^D}\right)$, should likewise be single-peaked, as depicted in Figure 3(b). In the our main model, the aggregate trading volume, however, is M-shaped.

This distinction between the two models suggests an empirical test for the efficiency of a given inter-dealer market. If the competitive inter-dealer market is efficient – a result we present formally in Appendix 7.3 and discuss in the Concluding Section of the paper – aggregate trading volume should not reach the maximum level at some intermediate level of asset supply. In the Balanced Equilibrium of the competitive inter-dealer market, large dealers should only trade with small dealers to the extent that the market cannot clear otherwise. In the frictional inter-dealer market, not all possible trades between an L_0 and an L_2, which yield the highest surplus among all inter-dealer trades, can take place and a unit of the asset may only be passed on from an L_2 to an L_0 over time through a small dealer. Such roundabout trades are most numerous at an intermediate asset supply where both S_0s and S_1s tend to be relatively abound. The abundance of inter-dealer trades in a given market at some intermediate asset supply can thus be a sign of search frictions giving rise to a myriad of roundabout trades that would not have taken place otherwise.

Furthermore, the frictional inter-dealer market model, but not the competitive inter-dealer market model, may also be used to address questions of what determines the differential markups dealers earn since in a frictional inter-dealer market, the terms of trade between an investor and a dealer would differ among dealers of different inventory capacities and actual inventory holdings, in addition to inter-dealer prices depending on who buys and who sells.

6 Concluding Remarks

First and foremost, our analysis in this paper shows that small dealers tend to sell to large dealers when large dealers need inventory the most and buy from large dealers when large dealers need 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 conception 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.[21] More interestingly, this means that for efficiency, the small peripheral dealers should trade to provide immediacy for the large core dealers.

[21] If investors' portfolio decisions are on the intensive margin where they choose how many units of the asset to hold, see Lagos and Rocheteau (2006) for an example, trades in the investor-dealer markets are inefficient due to a holdup problem introduced by bargaining. This inefficiency could affect the constrained efficiency of inter-dealer trades. But the conclusion that the provision of immediacy by small for large dealers is good for speeding up trades should stand.

7 Appendix

7.1 The Frictional Inter-Dealer Market Model

The model is set in continuous time in which all agents discount the future at the rate r and 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 $\eta(\theta)\frac{n_S^I}{n^D}$ 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. If all investor-dealer trades yield non-negative surpluses, as we verify in the proof of Proposition 5,

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

where $p_{b,s}$ denotes the terms of exchange between buyer b and seller s. An S_1, who can only sell, meets an investor-buyer at the rate $\eta(\theta)\frac{n_B^I}{n^D}$. The S_1 may also sell to an L_0 or an L_1. Then,

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

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

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

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,

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

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,

$$\begin{aligned} rV_2^{LD} &= \eta(\theta)\frac{n_B^I}{n^D}\left(p_{I_B,L_2}+V_1^{LD}-V_2^{LD}\right)+\alpha\left\{\frac{n_0^{SD}}{n^D}\max\left\{p_{S_0,L_2}+V_1^{LD}-V_2^{LD},0\right\}\right. \\ &\left.+\frac{n_0^{LD}}{n^D}\max\left\{p_{L_0,L_2}+V_1^{LD}-V_2^{LD},0\right\}\right\}. \end{aligned}$$

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

$$rU_B^I=\eta(\theta)\frac{n_S^D}{n^D}\left(U_H^{ON}-U_B^I-\frac{n_1^{SD}}{n_S^D}p_{I_B,S_1}-\frac{n_1^{LD}}{n_S^D}p_{I_B,L_1}-\frac{n_2^{LD}}{n_S^D}p_{I_B,L_2}\right)-\delta U_B^I,$$

where

$$rU_H^{ON}=v+\delta\left(U_S^I-U_H^{ON}\right).$$

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

$$rU_S^I=\eta(\theta)\frac{n_B^D}{n^D}\left(\frac{n_1^{SD}}{n_B^D}p_{S_0,I_S}+\frac{n_0^{LD}}{n_B^D}p_{L_0,I_S}+\frac{n_1^{LD}}{n_B^D}p_{L_1,I_S}-U_S^I\right).$$

7.1.2 Equilibrium Conditions

To begin, since, by Proposition 5, among all dealers, S_1s only sell to L_0s whereas S_0s only buy from L_2s, in the steady state in which $\dot{n}_0^{SD}=\dot{n}_1^{SD}=0$,

$$n_1^{SD}\left(\eta(\theta)n_B^I+\alpha n_0^{LD}\right)=n_0^{SD}\left(\eta(\theta)n_S^I+\alpha n_2^{LD}\right). \tag{32}$$

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

$$n_1^{LD}\eta(\theta)n_B^I=\left(n^{LD}-n_1^{LD}-n_2^{LD}\right)\left(\eta(\theta)n_S^I+\alpha n_1^{SD}+\alpha n_2^{LD}\right), \tag{33}$$

The continuous time versions of the respective steady-state conditions for $\dot{n}_H^{ON} = 0$, $\dot{n}_S^I = 0$, and $\dot{n}_B^I = 0$ are given by

$$n_B^I \eta(\theta) \frac{n_S^D}{n^D} = \delta n_H^{ON}, \tag{34}$$

$$\delta n_H^{ON} = \eta(\theta) \frac{n_B^D}{n^D} n_S^I, \tag{35}$$

$$e = \left(\delta + \eta(\theta) \frac{n_S^D}{n^D} \right) n_B^I, \tag{36}$$

whereas the adding up constraint (22) remains valid.

To begin, we first use (34) and (36) to solve for

$$n_H^{ON} = \frac{e}{\delta} - n_B^I, \tag{37}$$

while rewriting (36) as

$$n_S^D = \frac{e - \delta n_B^I}{\eta(\theta) n_B^I} n^D. \tag{38}$$

Then, by (34) and (35),

$$n_B^D = \frac{e - \delta n_B^I}{\eta(\theta) n_S^I} n^D. \tag{39}$$

Substituting (37) and (38) into (22) and by virtue of $n_S^D = n_1^{SD} + n_1^{LD} + n_2^{LD}$,

$$n_2^{LD} = A - \frac{e}{\delta} - n_S^I + n_B^I - \frac{e - \delta n_B^I}{\eta(\theta) n_B^I} n^D. \tag{40}$$

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

$$\begin{aligned} n_1^{SD} &= n^D - n_2^{LD} - n_B^D \\ &= n^D - A + \frac{e}{\delta} + n_S^I - n_B^I + \frac{e - \delta n_B^I}{\eta(\theta)} n^D \left(\frac{1}{n_B^I} - \frac{1}{n_S^I} \right), \end{aligned} \tag{41}$$

where the second line is from substituting from (39) and (40). Where $n_0^{SD} = n^{SD} - n_1^{SD}$,

$$n_0^{SD} = -n^{LD} + A - \frac{e}{\delta} - n_S^I + n_B^I - \frac{e - \delta n_B^I}{\eta(\theta)} n^D \left(\frac{1}{n_B^I} - \frac{1}{n_S^I} \right).$$

Finally, substituting (37), (40), and (41) into (22),

$$n_1^{LD} = -n^D + \frac{e - \delta n_B^I}{\eta(\theta)} n^D \left(\frac{1}{n_S^I} + \frac{1}{n_B^I} \right).$$

To complete the characterization, we bring in the definition of market tightness from (1). Then (32) and (33) are in terms of n_B^I and n_S^I only.

7.2 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 1 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 investors who may each hold zero or one unit and large investors who may each hold 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, whereas a dealer-seller (-buyer) holding at least a two-unit inventory (spare capacity) can sell to (buy from) large investors as well as small investors.

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. 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_1**s 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_0**s 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 2 *At where* $p = \beta\left(V_2^{LD} - V_1^{LD}\right)$*,* $m_0^{SD} = m_2^{SD}$*.*

Proof. Use superscripts SI and LI, respectively, for measures of small and large investors. Because all dealers are dealer-sellers as well as dealer-buyers and small investors trade with all dealers, in the steady state,

$$n_B^{SI}\eta(\theta) = n_S^{SI}\eta(\theta),$$

which implies that $n_B^{SI} = n_S^{SI}$. Because all small dealers enter the investor-dealer market holding a unit inventory and that they only trade with small investors,

$$m_0^{SD} = \eta(\theta)\frac{n_B^{SI}}{n^D}n^{SD},$$

$$m_2^{SD} = \eta(\theta)\frac{n_S^{SI}}{n^D}n^{SD},$$

from which we obtain $m_0^{SD} = m_2^{SD}$. ■

With $m_0^{SD} = m_2^{SD}$, the buyers and sellers among small dealers in the inter-dealer market are equally numerous in this equilibrium so that they neither provide inventory nor capacity for large dealers.

Case 4 $p = \beta\left(V_2^{SD} - V_1^{SD}\right)$ **with a fraction of** S_2**s 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_1**s 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).

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,

$$W = \max\left\{\sum_{t=0}^{\infty} \beta^t n_H^{ON}(t)\,\upsilon\right\}, \tag{42}$$

subject to the same search and matching frictions that agents 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. Specifically, any trade between an investor-buyer and a dealer-seller 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 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.

Proposition 6 *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 and dealer-buyers. 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 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.*

Proof. The controls of the planning problem (42) are $\{n_0^{SD}(t), n_1^{SD}(t), n_0^{LD}(t), n_1^{LD}(t), n_2^{LD}(t)\}$ and the state variables are $\left\{n_H^{ON}(t), n_S^I(t), n_B^I(t)\right\}$, where the equations of motion are given by,

$$n_H^{ON}(t+1) - n_H^{ON}(t) = -\delta n_H^{ON}(t) + n_B^I(t)\,\eta(\theta(t))\,\frac{n_S^D(t)}{n^D},$$

$$n_S^I(t+1) - n_S^I(t) = \delta n_H^{ON}(t) - n_S^I(t)\,\eta(\theta(t))\,\frac{n_B^D(t)}{n^D},$$

$$n_B^I(t+1) - n_B^I(t) = e - \left(\delta + \eta(\theta(t))\,\frac{n_S^D(t)}{n^D}\right) n_B^I(t).$$

Given the definitions of $\theta(t)$ in (1), the adding up constraints in (20)-(22) can be summarized by the following two equations:

$$\eta(\theta(t))\,\frac{n_S^D(t)}{n^D} = \eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{n^D - n_0^{SD}(t) - n_0^{LD}(t)}{n^D},$$

$$\eta(\theta(t))\,\frac{n_B^D(t)}{n^D} = \eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{n^D + n^{LD} - A - n_0^{LD}(t) + n_S^I(t) + n_H^{ON}(t)}{n^D}.$$

Write $N_S^D(t)$ for $\eta(\theta(t))\,\frac{n_S^D(t)}{n^D}$ and $N_B^D(t)$ for $\eta(\theta(t))\,\frac{n_B^D(t)}{n^D}$. In the above, a pair of $\left\{n_0^{SD}(t), n_0^{LD}(t)\right\}$ uniquely determines the pair $\left\{N_S^D(t), N_B^D(t)\right\}$. This means that the controls can be stated in terms of $\left\{N_S^D(t), N_B^D(t)\right\}$, whereby the admissible values are given by

$$N_S^D(t) \in \left[\eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{\underline{n}_S^D\left(n_H^{ON}(t), n_S^I(t)\right)}{n^D}, \eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{\overline{n}_S^D\left(n_H^{ON}(t), n_S^I(t)\right)}{n^D}\right],$$

$$N_B^D(t) \in \left[\eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{\underline{n}_B^D\left(n_H^{ON}(t), n_S^I(t)\right)}{n^D}, \eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{\overline{n}_B^D\left(n_H^{ON}(t), n_S^I(t)\right)}{n^D}\right],$$

with $\overline{n}_S^D$ and $\underline{n}_S^D$ denoting, respectively, the largest and smallest possible n_S^D and $\overline{n}_B^D$ and $\underline{n}_B^D$ denoting, respectively, the largest and smallest n_B^D, given state variables $n_H^{ON}(t)$ and $n_S^I(t)$. Let $A^D(t) = A - n_H^{ON}(t) - n_S^I(t)$ be the inventory of asset to be held by dealers. Note that:

(1) To attain $\overline{n}_S^D$, first allocate one unit each of $A^D(t)$ to either small or large dealers, and then allocate one more unit each to large dealers if $A^D(t) > n^D$.

(2) To attain $\underline{n}_S^D$, first allocate two units each of $A^D(t)$ to large dealers, and then allocate one unit each to small dealers if $A^D(t) > 2n^{LD}$.

(3) To attain $\overline{n}_B^D$, first allocate one unit each of $A^D(t)$ to large dealers, and then allocate one unit each to either large or small dealers if $A^D(t) > n^{LD}$.

(4) To attain $\underline{n}_B^D$, first allocate one unit each of $A^D(t)$ to small dealers, and then allocate two units each to large dealers if $A^D(t) > n^{SD}$.

To proceed, write (42) as

$$W\left(n_H^{ON}(t), n_S^I(t), n_B^I(t)\right) = \max_{N_S^D(t), N_B^D(t)} \left\{n_H^{ON}(t)\,\upsilon + \beta W\left(n_H^{ON}(t+1), n_S^I(t+1), n_B^I(t+1)\right)\right\},$$

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 $N_S^D(t)$ and $N_B^D(t)$. 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 $N_S^D(t)$ ($\underline{N}_S^D(t)$ and $\overline{N}_S^D(t)$) and the lower and upper bounds of $N_B^D(t)$ ($\underline{N}_B^D(t)$ and $\overline{N}_B^D(t)$).

Restricting attention to the steady state, we omit all time indices in the following. The first order conditions for $N_S^D(t)$ and $N_B^D(t)$ are then given by, respectively,

$$\beta n_B^I(W_1 - W_3) + \lambda_1 - \lambda_2 = 0, \tag{48}$$

$$-\beta n_S^I W_2 + \lambda_3 - \lambda_4 = 0. \tag{49}$$

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

$$W_1 = \upsilon + \beta(1-\delta)W_1 + \beta\delta W_2 - \lambda_1 \frac{\partial \underline{N}_S^D}{\partial n_H^{ON}} + \lambda_2 \frac{\partial \overline{N}_S^D}{\partial n_H^{ON}} - \lambda_3 \frac{\partial \underline{N}_B^D}{\partial n_H^{ON}} + \lambda_4 \frac{\partial \overline{N}_B^D}{\partial n_H^{ON}}, \tag{50}$$

$$W_2 = \beta(1 - N_B^D)W_2 - \lambda_1 \frac{\partial \underline{N}_S^D}{\partial n_S^I} + \lambda_2 \frac{\partial \overline{N}_S^D}{\partial n_S^I} - \lambda_3 \frac{\partial \underline{N}_B^D}{\partial n_S^I} + \lambda_4 \frac{\partial \overline{N}_B^D}{\partial n_S^I}, \tag{51}$$

$$W_3 = \beta N_S^D W_1 + \beta\left(1 - \delta - N_S^D\right) W_3 - \lambda_1 \frac{\partial \underline{N}_S^D}{\partial n_B^I} + \lambda_2 \frac{\partial \overline{N}_S^D}{\partial n_B^I} - \lambda_3 \frac{\partial \underline{N}_B^D}{\partial n_B^I} + \lambda_4 \frac{\partial \overline{N}_B^D}{\partial n_B^I}. \tag{52}$$

We first show that λ_3 must equal to 0. Suppose otherwise. Then, by the definitions 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 and so equations (48)-(52) reduce to four equations, (49)-(51) with only three unknowns, λ_3, W_1 and W_2. The set of λ_3, W_1 and W_2 that satisfy all three equations is of measure zero. Therefore, λ_3 must equal to 0.

Next, we show that λ_1 must equal to 0. Suppose otherwise. Then, $\lambda_2 = \lambda_3 = \lambda_4 = 0$ and so (49) says that $W_2 = 0$, which further implies that $\lambda_1 = 0$ from (51). This is a contradiction and so $\lambda_1 = 0$.

Next, we show that $\lambda_2 > 0$. Suppose otherwise. Together with the fact that $\lambda_1 = 0$, (48) becomes $W_1 = W_3$. We already know that $\lambda_3 = 0$. Then, (48)-(52) reduce to four equations (49)-(52) in three unknowns, λ_4, W_1 and W_2. We reach the desired contradiction.

Finally, we show that $\lambda_4 > 0$. Suppose otherwise. Then, by equation (49), $W_2 = 0$. Plug it into equation (51), we must have $\lambda_2 = 0$, which contradicts our previous conclusion that $\lambda_2 > 0$.

To summarize, we have shown that $N_S^D(t) = \eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{\overline{n}_S^D\left(n_H^{ON}(t), n_S^I(t)\right)}{n^D}$ and $N_B^D(t) = \eta\left(\frac{n^D}{n_S^I(t)+n_B^I(t)}\right)\frac{\overline{n}_B^D\left(n_H^{ON}(t), n_S^I(t)\right)}{n^D}$. In other words, for efficiency, we should allocate the assets held by dealers to maximize the measure of dealer-sellers and dealer-buyers: first allocate one unit each to large dealers; if $A^D > n^{LD}$, then allocate one unit each to small dealers; if $A^D > n^D$, then allocate one more unit each to large dealers. ■

Corollary 3 *In the steady state, the allocations from the decentralized market trades coincide with the planning optimum.*

Proof. The allocations as described in Proposition 6 are the same as the allocations as described in the discussions following Proposition 2. ■

7.4 Proofs of Lemmas and Propositions

Proof of Proposition 1 Given (3),

$$V_1^{SD} = W_0^{SD} + p + \eta(\theta)\frac{n_B^I}{n^D}\frac{z_{I_B}}{2}. \tag{53}$$

Given (5),

$$V_0^{LD} = W_1^{LD} - p + \eta(\theta)\frac{n_S^I}{n^D}\frac{z_{I_S}}{2}, \tag{54}$$

$$V_2^{LD} = W_1^{LD} + p + \eta(\theta)\frac{n_B^I}{n^D}\frac{z_{I_B}}{2}. \tag{55}$$

We can then calculate

$$V_1^{SD} - V_0^{SD} = p + \eta(\theta)\frac{n_B^I}{n^D}\frac{z_{I_B}}{2} - \eta(\theta)\frac{n_S^I}{n^D}\frac{z_{I_S}}{2}, \tag{56}$$

$$V_1^{LD} - V_0^{LD} = \eta(\theta)\frac{n_B^I}{n^D}\frac{z_{I_B}}{2} + p, \tag{57}$$

$$V_2^{LD} - V_1^{LD} = p - \eta(\theta)\frac{n_S^I}{n^D}\frac{z_{I_S}}{2}, \tag{58}$$

$$\left(V_1^{LD} - V_0^{LD}\right) - \left(V_1^{SD} - V_0^{SD}\right) = \eta(\theta)\frac{n_S^I}{n^D}\frac{z_{I_S}}{2}, \tag{59}$$

$$\left(V_1^{SD} - V_0^{SD}\right) - \left(V_2^{LD} - V_1^{LD}\right) = \eta(\theta)\frac{n_B^I}{n^D}\frac{z_{I_B}}{2}. \tag{60}$$

Notice that the two surpluses z_{I_B} and z_{I_S} can at worst be equal to zero in equilibrium. This proves the Proposition.

Proof of Lemma 1 Substitute (6) into (9) and rearrange,

$$U_S^I = \frac{\frac{\eta(\theta)n_B^D}{2n^D}}{1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta}p. \tag{61}$$

Substitute the equation into (15) and rearrange,

$$U_H^{ON} = \frac{\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\upsilon+\beta\delta\frac{\eta(\theta)n_B^D}{2n^D}p}{(1-\beta+\beta\delta)\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)}. \tag{62}$$

Substitute (10) into (13) and rearrange,

$$U_B^I = \frac{\frac{\eta(\theta)n_S^D}{2n^D}\left(\beta\left((1-\delta)U_H^{ON}+\delta U_S^I\right)-p\right)}{1-\left(1-\frac{\eta(\theta)n_S^D}{2n^D}\right)(1-\delta)\beta}. \tag{63}$$

Then, by (61) , (62) and (63),

$$(1-\delta)\left(U_H^{ON}-U_B^I\right)+\delta U_S^I = \frac{(1-\delta)\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\upsilon+\left(\delta n_B^D+\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)(1-\delta)n_S^D\right)\frac{\eta(\theta)}{2n^D}p}{\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\left(1-\left(1-\frac{\eta(\theta)n_S^D}{2n^D}\right)(1-\delta)\beta\right)}. \tag{64}$$

Set $p=\beta\left(V_1^{LD}-V_0^{LD}\right)$ and by (57),

$$p = \frac{\beta\frac{\eta(\theta)n_B^I}{2n^D}}{1-\beta+\frac{\eta(\theta)n_B^I}{2n^D}\beta}\beta\left(\delta U_S^I+(1-\delta)\left(U_H^{ON}-U_B^I\right)\right). \tag{65}$$

Then, use (64) to obtain

$$p = \frac{\frac{\eta(\theta)n_B^I}{2n^D}\beta^2(1-\delta)\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\upsilon}{(1-\beta)\left[\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\left(1-(1-\delta)\beta\left(1-\frac{\eta(\theta)n_S^D}{2n^D}\right)\right)+\frac{\eta(\theta)n_B^I}{2n^D}\beta\left(1-(1-\delta)\beta\left(1-\frac{\eta(\theta)n_B^D}{2n^D}\right)\right)\right]}. \tag{66}$$

Given the positivity of p in (66) and by (61) and (65),

$$0<\beta U_S^I<p<\beta\left(\delta U_S^I+(1-\delta)\left(U_H^{ON}-U_B^I\right)\right), \tag{67}$$

from which it follows that $z_{I_S}>0$ and $z_{I_B}>0$.

Next, set $p=\beta\left(V_1^{SD}-V_0^{SD}\right)$ and by (56),

$$p = \frac{\beta^2\frac{\eta(\theta)n_B^I}{2n^D}\left((1-\delta)\left(U_H^{ON}-U_B^I\right)+\delta U_S^I\right)+\frac{\eta(\theta)n_S^I}{2n^D}\beta^2U_S^I}{1-\beta+\beta\frac{\eta(\theta)}{2n^D}\left(n_B^I+n_S^I\right)}. \tag{68}$$

Then use (61) and (64) to obtain

$$p = \frac{\frac{\eta(\theta)n_B^I}{2n^D}\beta^2(1-\delta)\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\upsilon}{(1-\beta)\left[\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta+\frac{\eta(\theta)n_S^I}{2n^D}\beta\right)\left(1-(1-\delta)\beta\left(1-\frac{\eta(\theta)n_S^D}{2n^D}\right)\right)+\frac{\eta(\theta)n_B^I}{2n^D}\beta\left(1-(1-\delta)\beta\left(1-\frac{\eta(\theta)n_B^D}{2n^D}\right)\right)\right]}. \tag{69}$$

Given the positivity of p in (69) and by (61) and (68), the same conclusion as in (67) obtains.

Finally, set $p = \beta\left(V_2^{LD} - V_1^{LD}\right)$ and by (58),

$$p = \frac{\beta\frac{\eta(\theta)n_S^I}{2n^D}}{1-\beta+\beta\frac{\eta(\theta)n_S^I}{2n^D}}\beta U_S^I. \tag{70}$$

With (61),

$$p = \beta U_S^I = 0,$$

which also implies that $z_{I_S} = 0$. Next, by (64),

$$(1-\delta)\left(U_H^{ON} - U_B^I\right) + \delta U_S^I = \frac{(1-\delta)\,\upsilon}{1-\left(1-\frac{\eta(\theta)n_S^D}{2n^D}\right)(1-\delta)\,\beta} > 0. \tag{71}$$

And so,

$$0 = \beta U_S^I = p < \beta\left(U_H^{ON} - U_B^I\right) + \delta U_S^I,$$

which implies that $z_{I_B} > 0$.

Proof of Proposition 2 Solve (23)-(25) for

$$n_H^{ON} = \frac{(1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}}{\delta+(1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}}\frac{e}{\delta}, \tag{72}$$

$$n_S^I = \frac{\frac{n_S^D}{n_B^D}e}{\delta+(1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}}, \tag{73}$$

$$n_B^I = \frac{e}{\delta+(1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}}. \tag{74}$$

Substituting the above into (1) and (22) gives

$$\theta = \left(\delta+(1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}\right)\frac{n_B^D}{n_B^D+n_S^D}\frac{n^D}{e}, \tag{75}$$

$$\frac{\frac{1-\delta}{\delta}\eta(\theta)\,\frac{1}{n^D}+\frac{1}{n_B^D}}{\delta+(1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}}n_S^D e + n_1^{SD} + n_1^{LD} + 2n_2^{LD} = A. \tag{76}$$

Selling Equilibrium In the Selling Equilibrium, by Table 1, $n_0^{SD} = n^{SD}$, $n_1^{SD} = 0$, and $n_2^{LD} = 0$. Then, $n_S^D = n_1^{LD}$ and $n_B^D = n^D$, in which case (75) and (76) specialize to, respectively,

$$\theta = \left(\delta + (1-\delta)\,\eta(\theta)\,\frac{n_1^{LD}}{n^D}\right)\frac{n^D}{n_1^{LD}+n^D}\frac{n^D}{e}, \tag{77}$$

$$\frac{\frac{1-\delta}{\delta}\eta(\theta)+1}{\delta + (1-\delta)\,\eta(\theta)\,\frac{n_1^{LD}}{n^D}}\frac{n_1^{LD}}{n^D}e + n_1^{LD} = A. \tag{78}$$

Solve (78) for

$$\eta(\theta) = \frac{\left(A - n_1^{LD}\right)\delta - \frac{n_1^{LD}}{n^D}e}{\left(\frac{e}{\delta} + n_1^{LD} - A\right)(1-\delta)\frac{n_1^{LD}}{n^D}}. \tag{79}$$

Evaluate both sides of (77) by means of η, equate the RHS of the resulting expression to the RHS of (79) and then use $\mu = \eta(\theta)/\theta$ to obtain

$$\frac{\delta}{1-\delta}\frac{\left(A - n_1^{LD} - \frac{n_1^{LD}}{n^D}\frac{e}{\delta}\right)\left(n_1^{LD}+n^D\right)}{n_1^{LD}\left(n^D - n_1^{LD}\right)} = \mu\left(\frac{n^D - n_1^{LD}}{\frac{e}{\delta} + n_1^{LD} - A}\frac{n^D}{n_1^{LD}+n^D}\right). \tag{80}$$

Any solution for $n_1^{LD} \in \left[0, n^{LD}\right]$ to (80) is a Selling Equilibrium.

The RHS of (80) is well-defined and increasing for $n_1^{LD} \in \left[\max\left\{A - \frac{e}{\delta}, 0\right\}, n^{LD}\right]$ whereas the LHS is decreasing over the same range if

$$\left(A - \frac{e}{\delta} - n^D\right) n_1^{LD}\left(n_1^{LD} + n^D\right) - \left(A - n_1^{LD} - \frac{n_1^{LD}}{n^D}\frac{e}{\delta}\right) n^D\left(n^D - n_1^{LD}\right) < 0,$$

which is guaranteed to hold under

$$n^{LD} > A - \frac{e}{\delta}, \tag{81}$$

the same condition for the interval $\left[\max\left\{A - \frac{e}{\delta}, 0\right\}, n^{LD}\right]$ to be non-empty.

Hence, a unique solution $n_1^{LD} \in \left[\max\left\{A - \frac{e}{\delta}, 0\right\}, n^{LD}\right]$ to (80) exists if and only if:

1. The RHS of (80) is not smaller than the LHS at $n_1^{LD} = n^{LD}$; i.e.,

$$\frac{\delta}{1-\delta}\frac{\left(A - n^{LD} - \frac{n^{LD}}{n^D}\frac{e}{\delta}\right)\left(n^{LD}+n^D\right)}{n^{LD}n^{SD}} \leq \mu\left(\frac{n^D - n^{LD}}{\frac{e}{\delta} + n^{LD} - A}\frac{n^D}{n^{LD}+n^D}\right). \tag{82}$$

2. The RHS of (80) is not larger than the LHS at $n_1^{LD} = \max\left\{A - \frac{e}{\delta}, 0\right\}$. One can check that the condition is met.

Notice that (81) and (82) are the two conditions in Proposition 2(a), and that (80) just yields $n_1^{LD} = n^{LD}$ at where (82) holds as an equality.

Balanced Equilibrium In the Balanced Equilibrium, by Table 1, $n_1^{LD} = n^{LD}$, while $n_0^{LD} = n_2^{LD} = 0$. Then, $n_S^D = n_1^{SD} + n^{LD}$ and $n_B^D = n^D - n_1^{SD}$, in which case (75) and (76) specialize to, respectively,

$$\theta = \left(\delta + \eta\left(\theta\right)\frac{n_1^{SD} + n^{LD}}{n^D}\left(1-\delta\right)\right)\frac{n^D - n_1^{SD}}{n^D + n^{LD}}\frac{n^D}{e} \tag{83}$$

$$\frac{\frac{1-\delta}{\delta}\eta\left(\theta\right)\frac{n_1^{SD}+n^{LD}}{n^D} + \frac{n_1^{SD}+n^{LD}}{n^D - n_1^{SD}}}{\delta + \eta\left(\theta\right)\frac{n_1^{SD}+n^{LD}}{n^D}\left(1-\delta\right)}e + n_1^{SD} + n^{LD} = A \tag{84}$$

Solve (84) for

$$\eta\left(\theta\right) = \frac{A\delta - \left(n_1^{SD} + n^{LD}\right)\delta - \frac{n_1^{SD}+n^{LD}}{n^D - n_1^{SD}}e}{\left(\frac{e}{\delta} + n_1^{SD} + n^{LD} - A\right)\left(1-\delta\right)\frac{n_1^{SD}+n^{LD}}{n^D}} \tag{85}$$

Evaluate both sides of (83) by means of η, equate the RHS of the resulting expression to the RHS of (85) and then use $\mu = \eta\left(\theta\right)/\theta$ to obtain

$$\frac{\delta}{1-\delta}\frac{A - n_1^{SD} - n^{LD} - \frac{n_1^{SD}+n^{LD}}{n^D - n_1^{SD}}\frac{e}{\delta}}{\left(n_1^{SD} + n^{LD}\right)\left(n^{SD} - 2n_1^{SD}\right)}\left(n^D + n^{LD}\right) = \mu\left(\frac{n^{SD} - 2n_1^{SD}}{\frac{e}{\delta} + n_1^{SD} + n^{LD} - A}\frac{n^D}{n^D + n^{LD}}\right). \tag{86}$$

Any solution for $n_1^{SD} \in \left[0, n^{SD}\right]$ to (86) is a Balanced Equilibrium.

Consider where

$$A - n^{LD} - \frac{e}{\delta} - \frac{n^{SD}}{2} \leq 0. \tag{87}$$

Then, the RHS of (86) is defined for $n_1^{SD} \in \left[\max\left\{A - n^{LD} - \frac{e}{\delta}, 0\right\}, \frac{n^{SD}}{2}\right]$ over which it is positive and increasing in n_1^{SD}. By (87), the LHS is equal to zero at $n_1^{SD} = \widetilde{n}_1^{SD}$ for some $\widetilde{n}_1^{SD} \leq \frac{n^{SD}}{2}$, satisfying,

$$A - \widetilde{n}_1^{SD} - n^{LD} - \frac{\widetilde{n}_1^{SD} + n^{LD}}{n^D - \widetilde{n}_1^{SD}}\frac{e}{\delta} = 0, \tag{88}$$

and falls below zero for any larger n_1^{SD}. This means that the admissible n_1^{SD} is restricted to the interval $\left[\max\left\{A - n^{LD} - \frac{e}{\delta}, 0\right\}, \widetilde{n}_1^{SD}\right]$. In particular, at $n_1^{SD} = \widetilde{n}_1^{SD}$, the RHS of (86), which is positive, must exceed the LHS, which is equal to 0. Furthermore, we show in the Online Appendix that the LHS is monotone decreasing over where it is positive. Hence, a unique solution $n_1^{SD} \in \left[\max\left\{A - n^{LD} - \frac{e}{\delta}, 0\right\}, \widetilde{n}_1^{SD}\right]$ to (86) exists if and only if the RHS of (86) is not larger than the LHS at $n_1^{SD} = \max\left\{A - n^{LD} - \frac{e}{\delta}, 0\right\}$.

For $A - n^{LD} - \frac{e}{\delta} < 0$ so that the lower bound $n_1^{SD} = 0$, the condition is that of (82) holding in reverse. This proves the first part of Proposition 2(b).

For $A-n^{LD}-\frac{e}{\delta}\geq 0$, which when combined with (87) restricts $A-\frac{e}{\delta}\in\left[n^{LD},n^{LD}+\frac{n^{SD}}{2}\right]$, which constitutes the condition in the second part of Proposition 2(b). At the lower bound $n_1^{SD}=A-n^{LD}-\frac{e}{\delta}$, the RHS is equal to 0. Then, the condition that the LHS of (86) is not smaller than the RHS is met at the lower bound if the LHS is non-negative; i.e.,

$$1+\frac{1}{1-\frac{n^{LD}+n^D}{A-e/\delta}}\geq 0,$$

which holds for

$$n^{LD}+n^D\leq A-e/\delta$$

or

$$\frac{n^{LD}+n^D}{2}\geq A-e/\delta.$$

The first case is ruled out by (87) and so only the second case is relevant, which can be seen to be identical to (87). Notice in case (87) holds as an equality, the two sides of (86) are equal to 0 at $n_1^{SD}=A-n^{LD}-\frac{e}{\delta}=\frac{n^{SD}}{2}$. This completes the proof of Proposition 2(b).

Rewrite (86) as

$$\frac{\delta}{1-\delta}\frac{n_1^{SD}+n^{LD}+\frac{n_1^{SD}+n^{LD}}{n^D-n_1^{SD}}\frac{e}{\delta}-A}{\left(n_1^{SD}+n^{LD}\right)\left(2n_1^{SD}-n^{SD}\right)}\left(n^D+n^{LD}\right)=\mu\left(\frac{2n_1^{SD}-n^{SD}}{A-\frac{e}{\delta}-n_1^{SD}-n^{LD}}\frac{n^D}{n^D+n^{LD}}\right). \quad (89)$$

Consider where

$$A-n^{LD}-\frac{e}{\delta}-\frac{n^{SD}}{2}>0. \quad (90)$$

Then the RHS of (89) is defined for $n_1^{SD}\in\left[\frac{n^{SD}}{2},\min\left\{A-\frac{e}{\delta}-n^{LD},n^{SD}\right\}\right]$ over which it is positive but decreasing in n_1^{SD}. By (90), the LHS is equal to zero at $n_1^{SD}=\widetilde{n}_1^{SD}$ for $\widetilde{n}_1^{SD}\geq\frac{n^{SD}}{2}$ defined in (88) and falls below zero for any smaller n_1^{SD}. This means that the admissible n_1^{SD} is restricted to the interval $\left[\widetilde{n}_1^{SD},\min\left\{A-\frac{e}{\delta}-n^{LD},n^{SD}\right\}\right]$. In particular, at $n_1^{SD}=\widetilde{n}_1^{SD}$, the RHS of (89), which is positive, must exceed the LHS, which is equal to 0. Furthermore, we show in the Online Appendix that the LHS is monotone increasing over where it is positive. Hence, a unique solution $n_1^{SD}\in\left[\widetilde{n}_1^{SD},\min\left\{A-\frac{e}{\delta}-n^{LD},n^{SD}\right\}\right]$ to (89) exists if and only if the RHS of (89) is not larger than the LHS at $n_1^{SD}=\min\left\{A-\frac{e}{\delta}-n^{LD},n^{SD}\right\}$.

For $A-\frac{e}{\delta}-n^{LD}<n^{SD}$, which when combined with (90) restricts $A-\frac{e}{\delta}\in\left(n^{LD}+\frac{n^{SD}}{2},n^D\right)$, which constitutes the condition in the first part of Proposition 2(c). At the upper bound $n_1^{SD}=A-\frac{e}{\delta}-n^{LD}$, the RHS is equal to 0 while the LHS is proportional to $\frac{e}{\delta}\frac{2\left(A-\frac{e}{\delta}-n^{LD}-\frac{n^{SD}}{2}\right)}{n^D+n^{LD}-A+\frac{e}{\delta}}>0$.

For $A-\frac{e}{\delta}-n^{LD}\geq n^{SD}$ so that the upper bound is $n_1^{SD}=n^{SD}$, the condition that RHS of (89) is not larger than the LHS is

$$\frac{\delta}{1-\delta}\frac{n^D+\frac{n^D}{n^{LD}}\frac{e}{\delta}-A}{n^{SD}}\frac{\left(n^D+n^{LD}\right)}{n^D}\geq\mu\left(\frac{n^{SD}}{A-\frac{e}{\delta}-n^{LD}}\frac{n^D}{n^D+n^{LD}}\right), \tag{91}$$

which is the condition in the second part of Proposition 2(c). Notice that at where the above just holds as an equality, (89) yields $n_1^{SD}=n^{SD}$.

Buying Equilibrium In the Buying Equilibrium, by Table 1, $n_1^{SD}=n^{SD}$, $n_0^{SD}=0$, and $n_0^{LD}=0$. Then, $n_S^D=n^D$ and $n_B^D=n_1^{LD}$, in which case (75) and (76) specialize to, respectively,

$$\theta=\left(\delta+\left(1-\delta\right)\eta\left(\theta\right)\right)\frac{n_1^{LD}}{n_1^D+n^D}\frac{n^D}{e}, \tag{92}$$

$$\frac{\frac{1-\delta}{\delta}\eta\left(\theta\right)+\frac{n^D}{n_1^{LD}}}{\delta+\left(1-\delta\right)\eta\left(\theta\right)}e+n^D+n^{LD}-n_1^{LD}=A. \tag{93}$$

Solve (93) for

$$\eta\left(\theta\right)=\frac{\left(A-\left(n^D+n^{LD}-n_1^{LD}\right)\right)\delta-\frac{n^D}{n_1^{LD}}e}{\left(\frac{e}{\delta}+\left(n^D+n^{LD}-n_1^{LD}\right)-A\right)\left(1-\delta\right)} \tag{94}$$

Evaluate both sides of (92) by means of η, equate the RHS of the resulting expression to the RHS of (94) and then use $\mu=\eta\left(\theta\right)/\theta$ to obtain

$$\frac{\delta}{1-\delta}\frac{(-A+n^D+n^{LD}-n_1^{LD}+\frac{n^D}{n_1^{LD}}\frac{e}{\delta})(n_1^{LD}+n^D)}{\left(n^D-n_1^{LD}\right)n^D}= \tag{95}$$
$$\mu\left(\frac{n^D-n_1^{LD}}{A-n^D-n^{LD}+n_1^{LD}-\frac{e}{\delta}}\frac{n^D}{n_1^{LD}+n^D}\right).$$

Any solution for $n_1^{LD}\in\left[0,n^{LD}\right]$ to (95) is a Buying Equilibrium.

The RHS of (95) is defined for $n_1^{LD}\in\left[\max\left\{0,-A+n^D+n^{LD}+\frac{e}{\delta}\right\},n^{LD}\right]$, over which it is increasing in n_1^{LD}. The interval is guaranteed non-empty where

$$-A+\frac{e}{\delta}+n^D\leq 0. \tag{96}$$

The derivative of the LHS has the same sign as

$$\begin{aligned}&-A+n^D+n^{LD}-n_1^{LD}+\frac{e}{\delta}\left(\frac{n^D}{n_1^{LD}}-\frac{1}{2}\left(\frac{n^D}{n_1^{LD}}\right)^2+\frac{1}{2}\right)-\frac{\left(n^D\right)^2-\left(n_1^{LD}\right)^2}{2n^D}\\<&-A+n^D+n^{LD}-n_1^{LD}+\frac{e}{\delta},\end{aligned}$$

where the inequality is due to $\frac{n^D}{n_1^{LD}} - \frac{1}{2}(\frac{n^D}{n_1^{LD}})^2 + \frac{1}{2} \leq 1$. Then, the LHS is decreasing in n_1^{LD} over where the RHS is defined.

Hence, a unique solution $n_1^{LD} \in \left[\max\left\{0, -A + n^D + n^{LD} + \frac{e}{\delta}\right\}, n^{LD}\right]$ to (95) exists if and only if:

1. The RHS of (95) is not larger than the LHS at $n_1^{LD} = \max\left\{0, -A + n^D + n^{LD} + \frac{e}{\delta}\right\}$. One can check that the condition is met.

2. The RHS of (95) is not smaller than the LHS at $n_1^{LD} = n^{LD}$; i.e., where (91) holds in reverse.

Notice that (95) yields $n_1^{LD} = n^{LD}$ (and so $n_2^{LD} = 0$) just at where (91) holds as an equality. This proves Proposition 2(d).

Market clearing We turn next to verify that the solutions for n_1^{LD} to (80) for the Selling Equilibrium, n_1^{SD} to (86) for the Balanced Equilibrium, and n_1^{LD} to (95) for the Buying Equilibrium give rise to allocations that indeed clear the inter-dealer market for the three types of equilibrium, respectively.

In the Selling Equilibrium, the market clearing condition (16), by means of (27), (28), and (30), reduces to

$$n^{LD} - n_1^{LD} \geq 0$$

that obviously must hold. In the Buying Equilibrium, the market clearing condition (19), by means of (26), (28), and (30), reduces to exactly the same condition.

In the Balanced Equilibrium, in case $m_0^{LD} \geq m_2^{LD}$, where small dealers sell in the inter-dealer market, the market clearing condition (17), by means of (27), (28), and (30), reduces to

$$n_1^{SD} \geq 0,$$

whereas in case $m_2^{LD} \geq m_0^{LD}$, where small dealers buy in the inter-dealer market, the market clearing condition (18), by means of (26), (28), and (30), reduces to

$$n^{SD} - n_1^{SD} \geq 0.$$

Both conditions are guaranteed to hold. Furthermore, the condition $m_0^{LD} \geq m_2^{LD}$, by means of (28), (30), (73), (74) and the definitions for n_B^D and n_S^D, can be shown to be equivalent to

$$n_B^I \geq n_S^I \Leftrightarrow n_B^D \geq n_S^D \Leftrightarrow n_1^{SD} \leq \frac{n^{SD}}{2},$$

as presented in Proposition 2(b).

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

Proof of Propositions 3 and 4 By means of (73) and then via (75), we obtain two equations for

$$\eta(\theta)\, n_S^I = \frac{\eta(\theta)\,\frac{n_S^D}{n_B^D} e}{\delta + (1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}}, \tag{97}$$

$$\eta(\theta)\, n_S^I = \mu(\theta)\,\frac{n_S^D n^D}{n_B^D + n_S^D}. \tag{98}$$

By means of (74) and then via (75), we obtain two equations for

$$\eta(\theta)\, n_B^I = \frac{\eta(\theta)\, e}{\delta + (1-\delta)\,\eta(\theta)\,\frac{n_S^D}{n^D}}. \tag{99}$$

$$\eta(\theta)\, n_B^I = \mu(\theta)\,\frac{n_B^D n^D}{n_B^D + n_S^D}. \tag{100}$$

Selling Equilibrium In the Selling Equilibrium, $n_1^{SD} = n_2^{LD} = 0$ and $n_0^{SD} = n^{SD}$ and so $n_B^D = n^D$ and $n_S^D = n_1^{LD}$. In this case, by means of (27) and (30),

$$TV = m_1^{SD} + m_2^{LD} = \eta(\theta)\frac{n_S^I}{n^D}(n_1^{LD} + n^{SD}),$$

whereas, (97) and (98) specialize to, respectively,

$$\eta(\theta) n_S^I = \frac{e\eta(\theta)\frac{n_1^{LD}}{n^D}}{\delta + (1-\delta)\eta(\theta)\frac{n_1^{LD}}{n^D}} \equiv L_S^{SE}(\theta; n_1^{LD}),$$

$$\eta(\theta) n_S^I = \mu(\theta)\frac{n^D n_1^{LD}}{n^D + n_1^{LD}} \equiv R_S^{SE}(\theta; n_1^{LD}).$$

Set n_1^{LD} equal to the equilibrium value. Equilibrium θ satisfies $L_S^{SE}(\theta;n_1^{LD}) = R_S^{SE}(\theta;n_1^{LD})$. Note that $L_1^{SE}(\theta;n_1^{LD})$ is increasing and $R_S^{SE}(\theta;n_1^{LD})$ is decreasing in θ, and that both $L_S^{SE}(\theta;n_1^{LD})$ and $R_1^{SE}(\theta;n_1^{LD})$ are increasing in A, where n_1^{LD} is increasing in A in the Selling Equilibrium as implied by (80). All together then, $\eta(\theta)n_S^I$ is increasing in A and so does TV.

On the other hand, from (99) and (100), respectively, where $n_B^D = n^D$ and $n_S^D = n_1^{LD}$,

$$\eta(\theta)n_B^I = \frac{e\eta(\theta)}{\delta + (1-\delta)\eta(\theta)\frac{n_1^{LD}}{n^D}} \equiv L_B^{SE}(\theta;n_1^{LD}),$$

$$\eta(\theta)n_B^I = \mu(\theta)\frac{(n^D)^2}{n^D + n_1^{LD}} \equiv R_B^{SE}(\theta;n_1^{LD}).$$

Set n_1^{LD} equal to the equilibrium value. Equilibrium θ satisfies $L_B^{SE}(\theta;n_1^{LD}) = R_B^{SE}(\theta;n_1^{LD})$. Note that $L_B^{SE}(\theta;n_1^{LD})$ is increasing and $R_B^{SE}(\theta;n_1^{LD})$ is decreasing in θ, and that both $L_S^{SE}(\theta;n_1^{LD})$ and $R_S^{SE}(\theta;n_1^{LD})$ are decreasing in A, where n_1^{LD} is increasing in A. This implies that $\eta(\theta)n_B^I$ is decreasing in A.

Balanced Equilibrium In the Balanced Equilibrium, $n_0^{LD} = n_2^{LD} = 0$ and $n_1^{LD} = n^{LD}$ so $n_B^D = n^{LD} + n_0^{SD}$ and $n_S^D = n^{LD} + n_1^{SD}$. In this case, by means of (28) and (30),

$$TV = \begin{cases} m_0^{LD} & n_1^{SD} \leq \frac{n_2^{SD}}{2} \\ m_2^{LD} & n_1^{SD} \geq \frac{n_2^{SD}}{2} \end{cases} = \begin{cases} \eta(\theta)\, n_B^I \frac{n^{LD}}{n^D} & n_1^{SD} \leq \frac{n_2^{SD}}{2} \\ \eta(\theta)\, n_S^I \frac{n^{LD}}{n^D} & n_1^{SD} \geq \frac{n_2^{SD}}{2} \end{cases},$$

whereas (99) and (100) specialize to, respectively,

$$\eta(\theta)n_B^I = \frac{e\eta(\theta)}{\delta + (1-\delta)\eta(\theta)\frac{n_1^{SD}+n^D}{n^D}} \equiv L_B^{BaE}(\theta;n_1^{SD}),$$

$$\eta(\theta)n_B^I = \mu(\theta)\frac{(n^D - n_1^{SD})n^D}{n^D + n^{LD}} \equiv R_B^{BaE}(\theta;n_1^{SD}).$$

Set n_1^{SD} equal to the equilibrium value. Equilibrium θ satisfies $L_B^{BaE}(\theta;n_1^{SD}) = R_B^{BaE}(\theta;n_1^{SD})$. Note that $L_B^{BaE}(\theta;n_1^{SD})$ is increasing and $R_B^{BaE}(\theta;n_1^{SD})$ is decreasing in θ. As A increases, both curves shift downward since in the Balanced Equilibrium n_1^{SD} is increasing in A as implied by (86) or (89). This implies that $\eta(\theta)n_B^I$ and TV for $n_1^{SD} \leq \frac{n_2^{SD}}{2}$ are decreasing in A.

On the other hand, from (97) and (98), respectively, where $n_B^D = n^{LD} + n_0^{SD}$ and $n_S^D = n^{LD} + n_1^{SD}$,

$$\eta(\theta)n_S^I = \frac{e\eta(\theta)\frac{n_1^{SD}+n^{LD}}{n^D - n_1^{SD}}}{\delta + (1-\delta)\eta(\theta)\frac{n_1^{SD}+n^{LD}}{n^D}} \equiv L_S^{BaE}(\theta;n_1^{SD}),$$

$$\eta(\theta)n_S^I = \mu(\theta)\frac{n^D(n_1^{SD}+n^{LD})}{n^D+n^{LD}} \equiv R_S^{BaE}(\theta;n_1^{SD}).$$

Set n_1^{SD} equal to the equilibrium value. Equilibrium θ satisfies $L_S^{BaE}(\theta;n_1^{SD}) = R_S^{BaE}(\theta;n_1^{SD})$. Note that $L_S^{BaE}(\theta;n_1^{SD})$ is increasing and $R_S^{BaE}(\theta;n_1^{SD})$ is decreasing in θ. As A increases, both curves shift upward. This implies that $\eta(\theta)n_S^I$ and TV for $n_1^{SD} \geq \frac{n_2^{SD}}{2}$ is increasing in A.

Buying Equilibrium In the Buying Equilibrium, $n_0^{SD} = n_0^{LD} = 0$ and $n_1^{SD} = n^{SD}$ and so $n_B^D = n_1^{LD}$ and $n_S^D = n^D$. In this case, by means of (26) and (28),

$$TV = m_0^{SD} + m_0^{LD} = \eta\left(\theta\right)\frac{n_B^I}{n^D}\left(n^{SD}+n_1^{LD}\right),$$

whereas (97) and (98) specialize to, respectively,

$$\eta(\theta)n_B^I = \frac{e\eta(\theta)}{\delta+(1-\delta)\eta(\theta)} \equiv L_B^{BuE}(\theta),$$

$$\eta(\theta)n_B^I = \mu(\theta)\frac{n^D n_1^{LD}}{n^D+n_1^{LD}} \equiv R_B^{BuE}(\theta;n_1^{LD}).$$

Set n_1^{LD} equal to the equilibrium value. Equilibrium θ satisfies $L_B^{BuE}(\theta) = R_B^{BuE}(\theta;n_1^{LD})$. Note that $L_B^{BuE}(\theta)$ is increasing and $R_B^{BuE}(\theta;n_1^{LD})$ is decreasing in θ. As A increases, $L_B^{BuE}(\theta)$ remains unchanged while $R_B^{BuE}(\theta;n_1^{LD})$ goes down because n_1^{LD} is decreasing in A in the Buying Equilibrium as implied by (95). This implies that TV and $\eta(\theta)n_B^I$ are decreasing in A.

On the other hand, from (99) and (100), respectively, where $n_B^D = n_1^{LD}$ and $n_S^D = n^D$,

$$\eta(\theta)n_S^I = \frac{e\eta(\theta)\frac{n^D}{n_1^{LD}}}{\delta+(1-\delta)\eta(\theta)} \equiv L_S^{BuE}(\theta;n_1^{LD}),$$

$$\eta(\theta)n_S^I = \mu(\theta)\frac{(n^D)^2}{n^D+n_1^{LD}} \equiv R_S^{BuE}(\theta;n_1^{LD}).$$

Set n_1^{LD} equal to the equilibrium value. Equilibrium θ satisfies $L_S^{BuE}(\theta;n_1^{LD}) = R_S^{BuE}(\theta;n_1^{LD})$. Note that $L_S^{BuE}(\theta;n_1^{LD})$ is increasing and $R_S^{BuE}(\theta;n_1^{LD})$ is decreasing in θ. As A increases, both $L_S^{BuE}(\theta;n_1^{LD})$ and $R_S^{BuE}(\theta;n_1^{LD})$ go down. This implies that $\eta(\theta)n_S^I$ is increasing in A.

Proof of Proposition 5 We first verify that any investor-dealer match between a seller and a buyer yields a non-negative surplus in any equilibrium in which both small and large dealers are active agents, selling to and buying from investors. To begin, the surpluses of the possible trades are as follows.

$$z_{I_B,S_1} = U_H^{ON} - U_B^I - \left(V_1^{SD} - V_0^{SD}\right),$$

$$z_{I_B,L_i} = U_H^{ON} - U_B^I - \left(V_i^{LD} - V_{i-1}^{LD}\right) \text{ for } i = 1, 2,$$

$$z_{S_0,I_S} = V_1^{SD} - V_0^{SD} - U_S^I,$$

$$z_{L_i,I_S} = V_{i+1}^{LD} - V_i^{LD} - U_S^I \text{ for } i = 0, 1,$$

$$z_{S_0,L_i} = V_1^{SD} - V_0^{SD} - \left(V_i^{LD} - V_{i-1}^{LD}\right) \text{ for } i = 1, 2,$$

$$z_{L_i,S_1} = V_{i+1}^{LD} - V_i^{LD} - \left(V_1^{SD} - V_0^{SD}\right) \text{ for } i = 0, 1,$$

$$z_{L_0,L_2} = V_1^{LD} - V_0^{LD} - \left(V_2^{LD} - V_1^{LD}\right),$$

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

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

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

Expanding the expressions for the zs,

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

Subtracting $U_H^{ON} - U_B^I$ from the two sides of the condition,

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

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

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

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 steady-state equilibrium that involves trading between dealers and investors, indeed at least one type of dealer-seller must do so.

Now, suppose only S_1s sell to I_B where

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

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,

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

must hold given (103). The hypothesis that only S_1s sell to I_Bs 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 I_B 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,

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

must hold given (103). The hypothesis that only S_1s sell to IBs 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

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

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 I_B as well given (106). The hypothesis that only L_1s sell 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 I_B 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.

Given that all investor-dealer trades shall yield a non-negative match surplus, 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.

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

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

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

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

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

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

$$\eta(\theta)\frac{n_B^I}{n^D}\frac{z_{I_B,S_1}}{2} - \eta(\theta)\frac{n_S^I}{n^D}\frac{z_{S_0,I_S}}{2} < \eta(\theta)\frac{n_S^I}{n^D}\frac{z_{L_1,I_S}}{2} + \eta(\theta)\frac{n_B^I}{n^D}\frac{z_{I_B,L_1}}{2} - \eta(\theta)\frac{n_S^I}{n^D}\frac{z_{L_0,I_S}}{2}.$$

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

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

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 $z_{I_B,L_1} = 0$.

References

[1] Adrian, T., M. Fleming, O. Shachar and E. Vogt, 2017, "Market Liquidity after the Financial Crisis," *Annual Review of Financial Economics* 9, 43-83.

[2] Afonso, G., 2011, "Liquidity and Congestion," *Journal of Financial Intermediation* 20 (3), 324-360.

[3] Atkeson, A.G., A.L. Eisfeldt and P.-O. Weill, 2015, "Entry and Exit in OTC Derivatives Markets," *Econometrica* 83 (6), 2231-2292.

[4] 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.

[5] Bethune, Z., B. Sultanum and N. Trachter, 2016, "Private Information in Over-the-Counter Markets," mimeo.

[6] Chang, B. and S. Zhang, 2019, "Endogenous Market Making and Network Formation," mimeo.

[7] Choi, J. and Y. Huh, 2018, "Customer Liquidity Provision: Implications for Corporate Bond Transaction Costs," mimeo.

[8] Cimon, D.A. and C. Garriott, 2018, "Banking Regulation and Market Making," mimeo.

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

[10] Di Maggio, M., A. Kermani and Z. Song, 2017, "The value of trading relations in turbulent times," *Journal of Financial Economics* 124 (2), 266-284.

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

[12] Dunne, P.G., H. Hau and M.J. Moore, 2015, "Dearler Intermediation Between Markets," *Journal of the European Economic Association* 13 (5), 770-804.

[13] Friewald, N. and F. Nagler, 2018, "Over-the-Counter Market Frictions and Yield Spread Changes," *Journal of Finance*, forthcoming.

[14] Farboodi, M., 2017, “Intermediation and Voluntary Exposure to Counterparty Risk,” mimeo.

[15] Farboodi, M., G. Jarosch, G. Menzio and U. Wiriadinata, 2018, “Intermediation as Rent Extraction," mimeo.

[16] Farboodi, M., G. Jarosch and R. Shimer, 2018, “The Emergence of Market Structure," mimeo.

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

[18] Glode, V. and C. Opp, 2016, “Asymmetric Information and Intermediation Chains,” *American Economic Review* 106 (9), 2699-2721.

[19] Henderschott, T., D. Li, D. Livdan and N. Schürhoff, 2017, “ Relationship Trading in OTC Markets,” mimeo.

[20] Ho, T., and H. R. Stoll, 1983, “The Dynamics of Dealer Markets Under Competition,” *Journal of Finance*, 38 (4), 1053–1074.

[21] Hollifield, B., A. Neklyudov and C. S. Spatt, 2015, “Bid-Ask Spreads, Trading Networks and the Pricing of Securitizations: 144a vs Registered Securitizations,” mimeo.

[22] Hollifield, B., A. Neklyudov and C. S. Spatt, 2017, “Bid-Ask Spreads, Trading Networks and the Pricing of Securitizations,” *Review of Financial Studies* 30 (9), 3048-3085

[23] Hugonnier J., B. Lester and P.-O. Weill, 2018, “Frictional Intermediation in Over-the-Counter Markets,” NBER working paper 24956.

[24] Kiyotaki, N. and R. Wright, 1989, “On Money as a Medium of Exchange,” *Journal of Political Economy* 97(4), 927-54.

[25] Li, D and N. Schürhoff, 2019, “Dealer Networks,” *Journal of Finance*, 74 (1), 91-144.

[26] Lagos, R. and G. Rocheteau, 2006, “Search in Asset Markets,” mimeo.

[27] Lagos, R. and G. Rocheteau, 2009, "Liquidity in Asset Markets with Search Frictions," *Econometrica* 77 (2), 403-26.

[28] Lagos, R., G. Rocheteau and P.-O. Weill, 2011, "Crisis and Liquidity in Over-the-Counter Markets," *Journal of Economic Theory* 146 (6), 2169-2205.

[29] Neklyudov, A., 2019, "Bid-Ask Spreads and the Over-the-Counter Interdealer Markets: Core and Peripheral Dealers," *Review of Economic Dynamics*, Forthcoming.

[30] Sambalaibat, B., 2018, "Endogenous Specialization and Dealer Networks," mimeo.

[31] 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 (2), 406-411.

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

[33] Shen, J., B. Wei and H. Yan, 2018, "Financial Intermediation Chains in an OTC Market," mimeo.

[34] Üslü, S., 2019, "Pricing and Liquidity in Decentralized Asset Markets," mimeo.

[35] Vayanos, D. and T. Wang, 2007, "Search and Endogenous Concentration of Liquidity in Asset Markets," *Journal of Economic Theory*, 136, 66-104.

[36] Vayanos, D., and P.-O. Weill, 2008, "A Search-Based Theory of the On-the-Run Phenomenon," *Journal of Finance*, 63, 1361-1398.

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

[38] Weill, P.-O., 2011, "Liquidity Provision in Capacity Constrained Markets," *Macroeconomic Dynamics* 15 (S1), 119-144.

[39] Zhong, Z., 2016, "The Risk Sharing Benefit versus the Collateral Cost: The Formation of the Inter-Dealer Network in Over-the-Counter Trading," mimeo.

8 Online Appendix (Not to be considered for publication)

8.1 Monotonicity of the LHSs of (86) and (89)

We first show that the LHS of (86) is monotone decreasing over $n_1^{SD} \in \left[\max\left\{A - n^{LD} - \frac{e}{\delta}, 0\right\}, \widetilde{n}_1^{SD}\right]$. First suppose $A - n^{LD} - \frac{e}{\delta} \leq 0$, so that the interval to consider is $n_1^{SD} \in \left[0, \widetilde{n}_1^{SD}\right]$. Differentiating with respect to n_1^{SD} yields an expression that has the same sign as

$$\left(-1 - \frac{n^D + n^{LD}}{\left(n^D - n_1^{SD}\right)^2} \frac{e}{\delta}\right)\left(n_1^{SD} + n^{LD}\right)\left(n^{SD} - 2n_1^{SD}\right) + \left(2n^{LD} + 4n_1^{SD} - n^{SD}\right)\left(A - n_1^{SD} - n^{LD} - \frac{n_1^{SD} + n^{LD}}{n^D - n_1^{SD}} \frac{e}{\delta}\right). \tag{115}$$

The above is negative at $n_1^{SD} = \widetilde{n}_1^{SD}$ A priori (115) cannot be signed in general and it seems possible that it can be positive over a range of n_1^{SD} where $2n^{LD} + 4n_1^{SD} - n^{SD} > 0$. In this case, it must change sign, from positive to negative at least once and be equal to zero at some $n_1^{SD} \in \left(0, \widetilde{n}_1^{SD}\right)$, for it to become negative at $n_1^{SD} = \widetilde{n}_1^{SD}$. Differentiating (115) and factoring 2 from the resulting expression,

$$-\frac{n^D + n^{LD}}{\left(n^D - n_1^{SD}\right)^3} \frac{e}{\delta}\left(n_1^{SD} + n^{LD}\right)\left(n^{SD} - 2n_1^{SD}\right) + 2\left(A - n_1^{SD} - n^{LD} - \frac{n_1^{SD} + n^{LD}}{n^D - n_1^{SD}} \frac{e}{\delta}\right). \tag{116}$$

Setting (115) equal to 0,

$$A - n_1^{SD} - n^{LD} - \frac{n_1^{SD} + n^{LD}}{n^D - n_1^{SD}} \frac{e}{\delta} = \left(1 + \frac{n^D + n^{LD}}{\left(n^D - n_1^{SD}\right)^2} \frac{e}{\delta}\right) \frac{\left(n_1^{SD} + n^{LD}\right)\left(n^{SD} - 2n_1^{SD}\right)}{\left(2n^{LD} + 4n_1^{SD} - n^{SD}\right)}$$

and substituting it to (116) and factoring out terms that are guaranteed positive,

$$\frac{2\left(n^D - n_1^{SD}\right)^3 + 3\left(n^{SD} - 2n_1^{SD}\right)\left(n^D + n^{LD}\right)\frac{e}{\delta}}{2n^{LD} + 4n_1^{SD} - n^{SD}} > 0$$

if $2n^{LD} + 4n_1^{SD} - n^{SD} > 0$. We begin the analysis presuming that (115) is positive over a range of $n_1^{SD} \in \left(0, \widetilde{n}_1^{SD}\right)$ from which it follows that there must exist a n_1^{SD} at which (115) is just equal to 0. At this point, we show that (115) must be increasing. In at least one such n_1^{SD}, (115) must be turning from being positive to negative, however. Thus, we have arrived at a contradiction and this confirms that (115) is negative throughout for $n_1^{SD} \in \left(0, \widetilde{n}_1^{SD}\right)$.

Because we have not invoked any restriction related to the lower bound $n_1^{SD} = 0$, the same proof above is equally valid where $A - n^{LD} - \frac{e}{\delta} > 0$.

We next show that the LHS of (89) is monotone increasing over $n_1^{SD} \in [\widetilde{n}_1^{SD}, \min\{A - \frac{e}{\delta} - n^{LD}, n^{SD}\}]$. Write $F = L_1 \times L_2$, where

$$L_1 = \frac{-A}{n_1^{SD} + n^{LD}} + \frac{1}{n^D - n_1^{SD}} \frac{e}{\delta} + 1,$$
$$L_2 = \frac{1}{2n_1^{SD} - n^{SD}}.$$

Then the LHS of (89) is equal to

$$\frac{\delta}{1-\delta} \left(n^D + n^{LD}\right) F \tag{120}$$

Suppose for now

$$A - n^{LD} - \frac{e}{\delta} < n^{SD}, \tag{121}$$

so that the upper bound to be considered is $n_1^{SD} = A - \frac{e}{\delta} - n^{LD}$. We first show $F'(A - \frac{e}{\delta} - n^{LD}) \geq 0$.

$$\begin{aligned}
&F'(A - \frac{e}{\delta} - n^{LD}) \\
=&L_1'(A - \frac{e}{\delta} - n^{LD})L_2(A - \frac{e}{\delta} - n^{LD}) + L_1(A - \frac{e}{\delta} - n^{LD})L_2'(A - \frac{e}{\delta} - n^{LD}) \\
=&L_1'(A - \frac{e}{\delta} - n^{LD})L_2(A - \frac{e}{\delta} - n^{LD}) - 2L_1(A - \frac{e}{\delta} - n^{LD})L_2^2(A - \frac{e}{\delta} - n^{LD}) \quad \text{because } L_2' = -2L_2^2 \\
\propto&L_1'(A - \frac{e}{\delta} - n^{LD}) - 2L_1(A - \frac{e}{\delta} - n^{LD})L_2(A - \frac{e}{\delta} - n^{LD}) \\
\propto&\frac{A}{(A - \frac{e}{\delta})^2}[4(A - \frac{e}{\delta}) - (2n^{LD} + n^{SD})] - \frac{e}{\delta}\frac{1}{[(2n^{LD} + n^{SD}) - (A - \frac{e}{\delta})]^2}[3(2n^{LD} + n^{SD}) - 4(A - \frac{e}{\delta})] - 2.
\end{aligned} \tag{122}$$

Denote $2n^{LD} + n^{SD}$ as x. Differentiate (122) with respect to x yields,

$$-\frac{A}{(A - \frac{e}{\delta})^2} - \frac{e}{\delta}\frac{5(A - \frac{e}{\delta}) - 3x}{[x - (A - \frac{e}{\delta})]^3},$$

which is maximized at $x = 2(A - \frac{e}{\delta})$ and equal to

$$-\frac{1}{A - \frac{e}{\delta}} < 0.$$

Therefore, (122) is decreasing in x when $x \in [A - \frac{e}{\delta}, 2(A - \frac{e}{\delta})]$. By (90), x is indeed smaller than $2(A - \frac{e}{\delta})$. Then,

$$\begin{aligned}
F'(A - \frac{e}{\delta} - n^{LD}) \geq& \frac{A}{(A - \frac{e}{\delta})^2}2(A - \frac{e}{\delta}) - \frac{e}{\delta}\frac{1}{(A - \frac{e}{\delta})^2}2(A - \frac{e}{\delta}) - 2 \\
\propto& A - \frac{e}{\delta} - A + \frac{e}{\delta} = 0.
\end{aligned}$$

This proves that $F'(A - \frac{e}{\delta} - n^{LD}) \geq 0$.

Next, given that $L_2' = -2L_2^2$ and $L_2'' = -4L_2^2L_2'$,

$$\begin{aligned} F'' &= L_1''L_2 + L_1L_2'' + 2L_1'L_2' \\ &= L_1''L_2 - 4L_1L_2^2L_2' - 4L_1'L_2^2 \\ &\propto L_1'' - 4L_1L_2' - 4L_1'L_2 \\ &= L_1'' - 4F', \end{aligned}$$

where

$$L_1'' = \frac{-A}{(n_1^{SD} + n^{LD})^3} + \frac{1}{(n^D - n_1^{SD})^3}\frac{e}{\delta},$$

which is strictly increasing in n_1^{SD}. Therefore, there exist a unique $\check{n}_1^{SD}$ such that $L_1''(\check{n}_1^{SD}) = 0$.

If F is not monotone increasing in $n_1^{SD} \in \left(\widetilde{n}_1^{SD}, A - \frac{e}{\delta} - n^{LD}\right)$ and given $F'(A - \frac{e}{\delta} - n^{LD}) \geq 0$, there must exist some n_{1h}^{SD} and n_{1l}^{SD} such that $n_{1h}^{SD} > n_{1l}^{SD} \in (\widetilde{n}_1^{SD}, A - n^{LD} - \frac{e}{\delta}]$, $F'(n_{1l}^{SD}) = F'(n_{1l}^{SD}) = 0$, $F''(n_{1l}^{SD}) < 0$, $F''(n_{1h}^{SD}) > 0$ and $F'(n_1^{SD}) < 0$ for any $n_1^{SD} \in (n_{1l}^{SD}, n_{1h}^{SD})$. If $\check{n}_1^{SD} \leq \widetilde{n}_1^{SD}$, then $L_1'' > 0$ for any relevant n_1^{SD}. In this case, $F' = 0$ implies $F'' > 0$. This is a contradiction. If $\check{n}_1^{SD} > \widetilde{n}_1^{SD}$, then $\check{n}_1^{SD}$ must lie in the interval of $(n_{1l}^{SD}, n_{1h}^{SD})$. This means that $F'(\check{n}_1^{SD}) < 0$. On the other hand, from $F''(\check{n}_1^{SD}) = 0$ we can calculate

$$\check{n}_1^{SD} = \frac{n^D - (\frac{e}{A\delta})^{1/3} n^{LD}}{1 + (\frac{e}{A\delta})^{1/3}}, \tag{124}$$

and we know that

$$\frac{-A}{(\check{n}_1^{SD} + n^{LD})^3} = \frac{e}{\delta}\frac{1}{(n^D - \check{n}_1^{SD})^3}. \tag{125}$$

From (124), $\check{n}_1^{SD} < A - n^{LD} - \frac{e}{\delta}$ if and only if

$$2n^{LD} + n^{SD} < (A - \frac{e}{\delta})(1 + (\frac{e}{A\delta})^{1/3}). \tag{126}$$

Then,

$$
\begin{aligned}
F'(\check{n}_1^{SD}) \propto & L_1'(\check{n}_1^{SD}) - 2L_1(\check{n}_1^{SD})L_2(\check{n}_1^{SD}) \\
= & \frac{A}{(\check{n}_1^{SD} + n^{LD})^3}[\check{n}_1^{SD} + n^{LD} + 2\frac{(\check{n}_1^{SD} + n^{LD})^2}{2\check{n}_1^{SD} - n^{SD}}] \\
& - \frac{e}{\delta}\frac{1}{(n^D - \check{n}_1^{SD})^3}[2\frac{(n^D - \check{n}_1^{SD})^2}{2\check{n}_1^{SD} - n^{SD}} - n^D + \check{n}_1^{SD}] - \frac{2}{2\check{n}_1^{SD} - n^{SD}} \\
= & \frac{3A}{(\check{n}_1^{SD} + n^{LD})^3}(2n^{LD} + n^{SD}) - \frac{2}{2\check{n}_1^{SD} - n^{SD}} \text{ (after plugging in (125))} \\
= & 3A\frac{(1 + (\frac{e}{A\delta})^{1/3})^3}{(2n^{LD} + n^{SD})^2} - 2\frac{1 + (\frac{e}{A\delta})^{1/3}}{(1 - (\frac{e}{A\delta})^{1/3})(2n^{LD} + n^{SD})} \text{ (after plugging in (124))} \\
\propto & 3A\frac{(1 + (\frac{e}{A\delta})^{1/3})^2}{2n^{LD} + n^{SD}} - \frac{2}{1 - (\frac{e}{A\delta})^{1/3}} \\
> & 3\frac{1 + (\frac{e}{A\delta})^{1/3}}{1 - \frac{e}{A\delta}} - \frac{2}{1 - (\frac{e}{A\delta})^{1/3}} \text{ (after plugging in (126))} \\
\propto & 3\frac{1 + (\frac{e}{A\delta})^{1/3}}{1 + (\frac{e}{A\delta})^{1/3} + (\frac{e}{A\delta})^{2/3}} - 2 \\
> & 0 \text{ (because } \frac{e}{A\delta} < 1).
\end{aligned}
$$

This is a contradiction and so n_1^{SD} is monotone increasing over $(\widetilde{n}_1^{SD}, A - n^{LD} - \frac{e}{\delta})$.

We can construct a proof similar to the above for the case of $A - n^{LD} - \frac{e}{\delta} \geq n^{SD}$.

8.2 Inter-dealer market price

In the main text, 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* $p \in (\beta(V_1^{SD} - V_0^{SD}), \beta(V_1^{LD} - V_0^{LD}))$ *only for* $A = A_S(e)$. *Equilibrium obtains for a given* $p \in (\beta(V_2^{LD} - V_1^{LD}), \beta(V_1^{SD} - V_0^{SD}))$ *only for* $A = A_B(e)$.

Proof. Consider an inter-dealer market price such that

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

The buyers are L_0s, whereas the sellers are S_1s and L_2s. This means that $n_1^{LD} = n^{LD}$, $n_0^{SD} = n^{SD}$, and $n_0^{LD} = n_2^{LD} = n_1^{SD} = 0$, whereby $n_S^D = n^{LD}$ and $n_B^D = n^D$. Substituting into (77) and (78) and manipulating yields (82) or $\Omega_S(A) = 0$.

Consider a price such that

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

The buyers are L_0s and S_0, whereas the sellers are L_2s. This means that $n_1^{LD} = n^{LD}$, $n_1^{SD} = n^{SD}$, and $n_0^{LD} = n_2^{LD} = n_0^{SD} = 0$, whereby $n_S^D = n^D$ and $n_B^D = n^{LD}$. Substituting into (83) and (84) and manipulating yields (91) or $\Omega_B(A) = 0$. ■

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.

8.3 Investors' Bid and Ask Prices

Given the surpluses of trade in (6) and (10), the prices that investors sell to dealers and buy from dealers are equal to, respectively,

$$p_{I_S} = p - \frac{z_{I_S}}{2}, \tag{127}$$

$$p_{I_B} = p + \frac{z_{I_B}}{2}. \tag{128}$$

Combining (127) and (61) yields,

$$p_{I_S} = \frac{1}{2}\frac{1 - \beta + \frac{\eta(\theta)n_B^D}{n^D}\beta}{1 - \beta + \frac{\eta(\theta)n_B^D}{2n^D}\beta}p.$$

Combining (128) and (64) yields,

$$p_{I_B} = \frac{1}{2}\frac{(1-\delta)\beta\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\upsilon+\left(\beta\frac{\eta(\theta)}{2n^D}\delta n_B^D+\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)\left(1-\left(1-\frac{\eta(\theta)n_S^D}{n^D}\right)(1-\delta)\beta\right)\right)p}{\left(1-\left(1-\frac{\eta(\theta)n_S^D}{2n^D}\right)(1-\delta)\beta\right)\left(1-\beta+\frac{\eta(\theta)n_B^D}{2n^D}\beta\right)}.$$