

# Preliminary and Incomplete Not For Distribution Core-Periphery Trading Networks \*

Chaojun Wang<sup>†</sup>

Stanford University

## Abstract

This paper provides a theory of endogenous network formation in over-the-counter markets based on trade competition and inventory risk balancing. A core-periphery network structure arises as an equilibrium outcome. A small number of agents emerge as core dealers to intermediate among a large number of peripheral agents. The equilibrium level of dealer entry (the size of the core) depends on the combined effect of two countervailing forces: (i) network competition among dealers in their pricing of immediacy to peripheral agents, and (ii) the benefits of a concentrated set of dealers for lowering inventory risk through their ability to quickly offset purchases against sales. The size of the dealer core grows with the total number of agents, and reaches a finite limit size in a market with infinitely many agents. The dealer sector also increases as agents become more risk tolerant. Having more dealers competing to provide liquidity lowers the equilibrium bid-ask spread, individual dealer inventory levels and turnover, but increases market-wide dealer inventory cost. From the viewpoint of market efficiency, surprisingly, dealers tend to under-compete in liquid markets, and over-compete in illiquid markets. This suggests regulation that is differentiated by asset liquidity.

*JEL Classifications:* C73, D43, D85, L13, L14, G14

*Keywords:* Over-the-counter, network, financial intermediation, price competition, inventory risk

---

\*This version is preliminary and incomplete. I am deeply indebted to my advisors Darrell Duffie and Matthew Jackson for their continued guidance and detailed comments and suggestions. I am grateful for early conversations with Arvind Krishnamurthy, Bob Wilson, Shengxing Zhang. Comments are welcome.

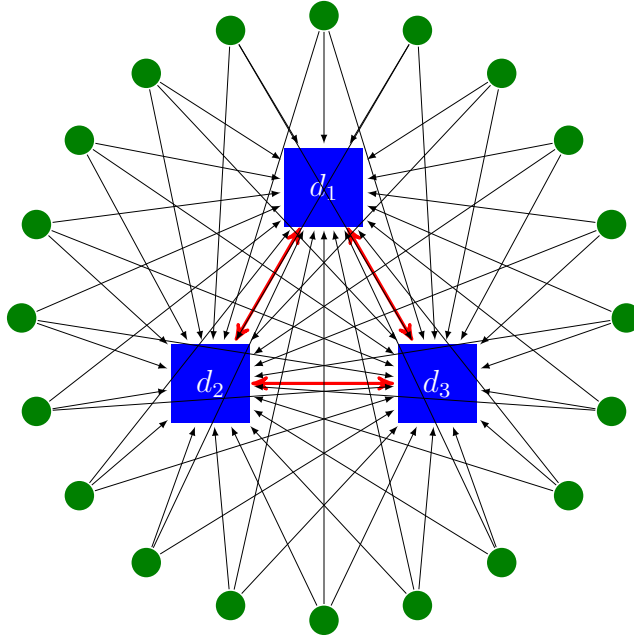
<sup>†</sup>Email address: [chaojunw@stanford.edu](mailto:chaojunw@stanford.edu). Paper URL: <http://ssrn.com/abstract=2747117>

# 1 Introduction

I propose and solve a model of network formation in over-the-counter markets. The equilibrium has an explicitly characterized core-periphery structure. In practice, core-periphery networks dominate conventional OTC markets. Although all agents in the model are ex-ante identical, a small subset of them endogenously arise as core agents, known as “dealers,” to intermediate among a large number of non-dealer (peripheral) agents. The equilibrium number of dealers grows with the total number of agents. However, even for an “infinite” set of agents, one should anticipate only a finite number of dealers. Having more dealers competing to provide liquidity lowers the equilibrium bid-ask spread, individual dealer inventory levels, and turnover. From a welfare viewpoint, the model identifies two sources of externalities. Dealers tend to under-compete in liquid markets and over-compete in illiquid markets.

The model works roughly as follows. A finite number of ex-ante identical agents form trading relationships (links) in a continuous-time bilateral trading game. Dealers arise endogenously to form the core of the market, exploiting their central position to balance inventory risk by netting many purchases against many sales. Peripheral agents establish trading links only with dealers, to benefit from dealers’ ability to engage in greater price competition. The equilibrium level of dealer entry (the size of the core) depends on the combined effect of two countervailing forces: *(i)* network competition among dealers in their pricing of immediacy to peripheral agents, and *(ii)* the benefits of a concentrated set of dealers for lowering inventory risk through their ability to quickly offset purchases against sales. As the number of dealers rises, each dealer receives thinner order flow, and optimally shrinks her target inventory, given her reduced efficiency in netting trades. As the number of dealers falls, peripheral agents obtain insufficient competition from existing dealers (wider bid-ask spreads). More dealers thus enter. In equilibrium, this key trade-off leads to a determinate size for the core. **Figure 1** depicts an example of an equilibrium core-periphery network for a market with 23 agents, 3 of whom have emerged as dealers.

Most over-the-counter markets, such as those for bonds, inter-bank lending, commodities,



**Figure 1** – An example of a core-periphery network with 3 Dealers and 20 Non-Dealers.

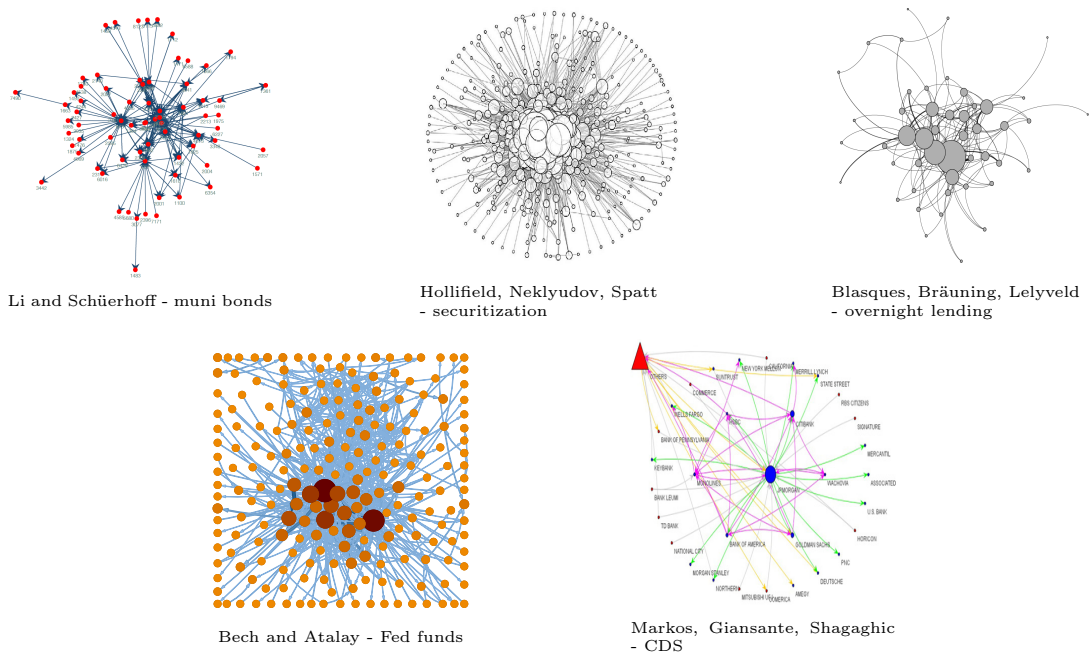
foreign exchange, and swaps, exhibit a clear and stable core-periphery network structure.<sup>1</sup> For many OTC financial markets, roughly the same 10 to 15 dealers, all affiliated with large banks, form the core. The vast majority of trades have one of these dealers on at least one side. For example: The largest fourteen derivatives dealers, known as the “G14,”<sup>2</sup> intermediate 82% of the global total notional amount of outstanding derivatives. Broken down by product, the G14 dealers hold 82% percent of interest rate swaps, 90% of credit default swaps, and 86% of OTC equity derivatives.<sup>3</sup> Figure 2 illustrates some examples of

<sup>1</sup>Bech and Atalay (2010), Allen and Saunders (1986), Afonso, Kovner, and Schoar (2014) provide evidence on the federal funds market, Boss, Elsinger, Summer, and Thurner (2004), Chang, Lima, Guerra, and Tabak (2008), Craig and von Peter (2014), in ’t Veld and van Lelyveld (2014), Blasques, Bräuning, and van Lelyveld (2015) for interbank markets in other countries, Peltonen, Scheicher, and Vuillemy (2014), Siriwardane (2015) for credit default swaps, Di Maggio, Kermani, and Song (2015) for corporate bonds, Li and Schürhoff (2014) for municipal bonds, Hollifield, Neklyudov, and Spatt (2014) for asset-backed securities and James, Marsh, and Sarno (2012) for currencies.

<sup>2</sup>The G14 dealers comprise Bank of America-Merrill Lynch, Barclays Capital, BNP Paribas, Citi, Credit Suisse, Deutsche Bank, Goldman Sachs, HSBC, JP Morgan, Morgan Stanley, Nomura, RBS, Societe Generale, UBS and Wells Fargo Bank.

<sup>3</sup>These statistics are computed by Mengle (2010) using data from the International Swaps and Derivatives Association (ISDA) Market Survey for Mid-Year 2010. Using a different methodology and data, Heller and Vause (2012) estimates that the same G14 dealers account for 96% of trading in CDS markets as of the end of June in 2010. Before 2009, Securities Industry and Financial Markets Association (SIFMA) reports

core-periphery networks in other OTC markets.



**Figure 2** – Core-periphery networks in OTC markets. Source: Li and Schürhoff (2014), Hollifield, Neklyudov, and Spatt (2014), Blasques, Bräuning, and van Lelyveld (2015), Bech and Atalay (2010), Markose, Giansante, and Shaghaghic (2012).

Increasingly, the basic core-periphery network of some OTC markets includes additional structure in the form of trading platforms on which multiple dealers provide quotes. Multilateral trading platforms have appeared in OTC markets for foreign exchange, treasuries, some corporate bonds, and (especially through the force of recent regulation) standardized swaps. Examples of such platforms include MarketAxess and Neptune for bonds, 360T and Hotspot for currencies, and Bloomberg for swaps. This paper restricts its focus, however, to the more “classical” and simple case of purely bilateral OTC trade.

A related issue of concern is whether recent illiquidity in bond markets stems from crisis-induced regulations (such as the Volcker Rule) and higher bank capital requirements. These changes may have caused a reduction in dealers’ effective balance-sheet capacity.<sup>4</sup> It is

trading data for “major firms,” which are sometimes 12 or 13 in number. The identity of major dealers are largely the same across different data sources and asset classes.

<sup>4</sup>Adrian, Fleming, Goldberg, Lewis, Natalucci, and Wu (2013) provide a recent discussion. Prior studies on the relationship between dealer risk tolerance, inventories and market liquidity include Grossman and Miller

also hotly debated whether non-bank firms can substitute for dealer liquidity by taking an effective role as market makers.<sup>5</sup> My model supports a discussion of both issues, as well as some associated welfare implications.

There is rising interest in providing theoretical foundations for the endogenous core-periphery structure of OTC markets. In prior research on this topic, the agents who form the core have some ex-ante special advantages in serving this role. That is, ex-ante heterogeneity of agents has been exploited to explain the separation of dealers and non-dealers. In [Farboodi \(2014\)](#), for example, the core of the federal funds market consists of those banks with risky investment opportunities who, in equilibrium, expose themselves to counterparty risk in order to capture intermediation spreads. As a further distinction, in Farboodi’s model, the endogenous network structure is generated mainly by the effects of counterparty default risk and not (as in my model) by the effects of trade competition and inventory risk management. [Hugonnier, Lester, and Weill \(2015\)](#), [Afonso and Lagos \(2015\)](#), [Shen, Wei, and Yan \(2015\)](#) derive the “coreness” of investors from their preferences for ownership of the asset. Those with average preferences serve as intermediaries between high and low-value investors. The models of [Neklyudov \(2014\)](#) and [Üslu \(2015\)](#), instead, are based on exogenous heterogeneity in investors’ search technologies. These models all have an infinite number of “core” agents, thus missing a key aspect of functioning OTC markets.

My results contribute to this literature in three ways: *(i)* I provide a game-theoretic foundation for the formation of core-periphery networks in OTC markets that is motivated by inventory risk management and competition for trade. Although all agents are ex-ante identical, an ex-post separation of core from peripheral agents is determined solely by endogenous forces that tend to concentrate the provision of immediacy. *(ii)* I show explicitly how the equilibrium number of dealers grows with the total number of agents, and I characterize the limit size of this dealer core in a market with infinitely many agents. *(iii)* The

---

(1988), [Weill \(2007\)](#), [Gromb and Vayanos \(2002, 2010\)](#), [He and Krishnamurthy \(2011\)](#), [Lagos, Rocheteau, and Weill \(2008\)](#), [Rinne and Suominen \(2010, 2011\)](#), [Brunnermeier and Pedersen \(2008\)](#), [Comerton-Forde, Hendershott, Jones, Moulton, and Seasholes \(2010\)](#), [Hendershott and Seasholes \(2007\)](#), [Hendershott and Menkveld \(2013\)](#), [Meli \(2002\)](#), [Adrian, Moench, and Shin \(2010\)](#), [Adrian, Etula, and Shin \(2010\)](#).

<sup>5</sup>Prior studies expressed diverging views, see [Kirilenko, Kyle, Samadi, and Tuzun \(2011\)](#), [Brogaard \(2010\)](#), [Brogaard, Hendershott, and Riordan \(2014\)](#), [Zhang \(2010\)](#).

model relates welfare, market concentration, and asset liquidity, by showing that dealers tend to under-compete in liquid markets and over-compete in illiquid markets.

The paper is organized as follows. [Section 2](#) presents the setup of a symmetric-agent model and defines the equilibrium solution concept. [Section 3](#) shows that a core-periphery network structure emerges in equilibrium, and determines the core size as a function of model parameters. [Section 4](#) provides comparative statics and welfare analysis, and discusses policy implications of the model. [Section 5](#) provides an extension of the symmetric-agent model, in which dealers are allowed to bilaterally negotiate the terms of their trades, rather than merely offer quotes. [Section 6](#) concludes.

## 2 A Symmetric-Agent Model

**Asset and preferences.** I fix the time domain  $[0, \infty)$ , a probability space  $(\Omega, \mathcal{F}, \Pr)$  and a filtration  $\{\mathcal{F}_t : t \geq 0\}$  of  $\sigma$ -algebras satisfying the usual conditions, as defined by [Protter \(2005\)](#). The filtration represents the resolution of information over time.

A finite number  $n$  of ex-ante identical risk-neutral agents consume a single non-storable numéraire good, called “cash.” All costs and monetary payments, to be introduced shortly, are measured in units of cash. All agents are infinitely-lived with time preferences determined by a constant discount rate  $r > 0$ . The agents have access to a risk-free liquid security with interest rate  $r$ .

A common non-divisible asset is traded over the counter. Each unit of the asset generates a sequence  $(D_k)_{k \geq 1}$  of lump-sum payoffs, independent random variables with some finite mean  $v$ , at the event times of an independent Poisson process with intensity  $\alpha$ . The mean  $v$  can be negative, in which case an asset owner pays  $|v|$  in expectation for every unit of the asset she owns at each dividend time. Every agent has 0 initial endowment of the asset, and incurs a quadratic cost  $\beta x^2$  per unit of time when holding an asset inventory of size  $x$ . That is, the agent experiences a quadratic instantaneous disutility when her position deviates from the bliss point, normalized to 0.

This assumption of a quadratic holding cost is common in both static and dynamic trading

models, including those of [Vives \(2011\)](#), [Rostek and Weretka \(2012\)](#), [Du and Zhu \(2014\)](#). One can interpret this flow cost as an inventory cost, which may be related to regulatory capital requirements, collateral requirements, financing costs, as well as the expected cost of being forced to raise liquidity by quickly disposing of inventory into an illiquid market.

Combining the holding cost rate and the expected rate of payment of the asset, it follows that the net mean rate of payoff of an asset inventory of size  $x$  is

$$-\beta \left( x - \frac{\alpha v}{2\beta} \right)^2 + \frac{(\alpha v)^2}{4\beta}. \quad (1)$$

The second term of (1) is constant and thus does not affect an agent's decision making. For simplicity of exposition and without loss of generality for the main purpose of capturing network formation incentives, I assume  $v = 0$ .

**Network formation, search and trade protocols.** Each agent  $i$  supplies a group of risk-neutral “outside investors” that is local to that specific agent. The outside investors of agent  $i$  submit buy and sell orders independently at mean rate  $\lambda$  (that is, at the event times of a Poisson process, independent across agents, with rate parameter  $\lambda$ ). Each order seeks to trade one unit of the asset, and is an “Immediate or Cancel” order. The outside investors pay a premium of  $\pi > 0$  for immediate execution. The premium  $\pi$  may be interpreted as the outside investors' private hedging or liquidity gain. Upon receiving a sell order, agent  $i$  can trade with the outside investors in two distinct ways: she can either (i) take the asset into her own inventory, or (ii) immediately find another agent  $j$  who will buy the asset. These analogous alternatives apply to a buy order from an outside investor. The first choice guarantees a trade, but entails an additional inventory risk for agent  $i$ . The second choice, typically called a “riskless-principal” transaction, allows agent  $i$  to be a pure match-maker. A riskless-principal trade, however, only occurs when agent  $i$  can find another agent  $j$  willing to assume the opposite position. If agent  $i$  fails to find such an agent, no riskless-principal trade occurs. (I will describe the search and trade protocols later.) Due to the quadratic inventory holding cost, an agent may not want to expand her inventory indefinitely. Since the focus of this paper is the network of trading relationships, I simply assume that every

agent conducts only riskless-principal transactions with her outside investors. In this sense, an agent acts as a local agency-based broker-dealer when trading with her outside investors. This assumption is consistent with the empirical observation that peripheral agents tend to “pre-arrange” trades rather than taking assets into inventory. See for example [Li and Schürhoff \(2014\)](#) for evidence.

At any time  $t \geq 0$ , a given agent  $i$  can open a trading account with any other agent  $j$ , giving  $i$  the right to obtain executable price quotes from  $j$ . This trading relationship is represented by a directed link from  $i$  to  $j$  in a “network” with nodes given by the  $n$  agents. In this case, I say that  $j$  is a *quote provider* to  $i$ . Agent  $j$  must separately establish an account with  $i$  in order to obtain quotes from  $i$ . Setting up a trading account is costless, but maintaining an account incurs an ongoing cost of  $c$  per unit time to the account holder. The cost might stem, for example, from operational costs or costly monitoring efforts. An agent is also permitted to terminate some or all of her opened accounts, saving the associated maintenance cost. On the equilibrium path, these trading accounts, once set up, will be maintained forever. But the option to close an account gives a quote seeker the ability to discipline her quote providers from offering aggressively unfavorable prices.

At any time  $t > 0$ , agent  $i$  may search among her current quote providers to request a quote. Search is cost-free, but an agent conducts search only a countable number of times and only when the expected gain from search is strictly positive. This tie-breaking rule can be justified by introducing an arbitrarily small but strictly positive search cost. If agent  $i$  searches among  $m$  quote providers, the search technology is associated with some probability  $\theta_m$  of immediate success. Conditional on a successful search, agent  $i$  reaches any given one of the  $m$  quote providers with equal probability  $1/m$  (independent across searches). The probability  $\theta_m \in [0, 1)$  of a successful search is strictly increasing and strictly concave in  $m$ , and  $\theta_0 = 0$ . For example,  $\theta$  may take the parametric form  $\theta_m = 1 - p^m$  for some  $p \in (0, 1)$ , in which case  $p$  can be interpreted as the probability that the agent fails to reach a given one of the  $m$  quote providers, based on independent matching. The assumption that the probability  $\theta_m$  is strictly increasing in  $m$  implies that agents have incentives to form larger relationship networks in order to mitigate search friction.



Upon reaching a quote provider  $j$ , agent  $i$  submits a Request for Quote (RFQ), which comprises a desired trade direction (sell or buy). Agent  $j$  then posts an executable bid  $b_j$  or offer quote  $a_j$ , a binding offer to buy or sell one unit of the asset at the respective prices. The quote is observed and executable only by agent  $i$ , and is good only when offered. A quote provider is required to make two-way markets by not posting stub quotes.<sup>6</sup> Since the total gain per trade is at most  $\pi$ , the no-stub-quote rule is equivalent to the requirement that  $a_j \leq \pi$  and  $b_j \geq -\pi$ . On many trade platforms, the practice known as “name give up” requires the identity of the quote requester to be revealed. For simplicity, I avoid that requirement, which plays no important role in practice if the trade is centrally cleared, as is increasingly common (and now required in standardized swap markets).<sup>7</sup>

It is more common in OTC markets for the identities of both trading counterparties to be observed by each other. The assumption of quote seeker anonymity does not play a critical role and will be relaxed in [Section 5](#). It is important, however, that the quote provider does not observe the number of other trading accounts maintained by a quote seeker.

As modeled here, quotes are often provided in practice for standard quantities, which could be scaled to one unit without loss of generality. However, in reality the trade size can then be negotiated with price concessions. The restriction here to one unit on trade size is not realistic, especially for inter-dealer trading, and will be relaxed in [Section 5](#).

Due to the inventory holding cost that exhibits increasing returns to scale, an agent may not wish to expand her inventory indefinitely. To ensure the provision of liquidity by every agent to her quote seekers, an abstract mechanism is assumed to be available to every agent as a trading counterparty, in order to prevent an agent from being forced (by the stub-quote rule) to make a trade at a loss. Specifically, whenever an agent receives a Request for Quote from another agent, she can make a riskless-principal transaction with the mechanism, which takes the other side of the same trade. If an agent chooses to trade with the mechanism for this purpose, the price available at the mechanism is the last price, buy or sell, at which

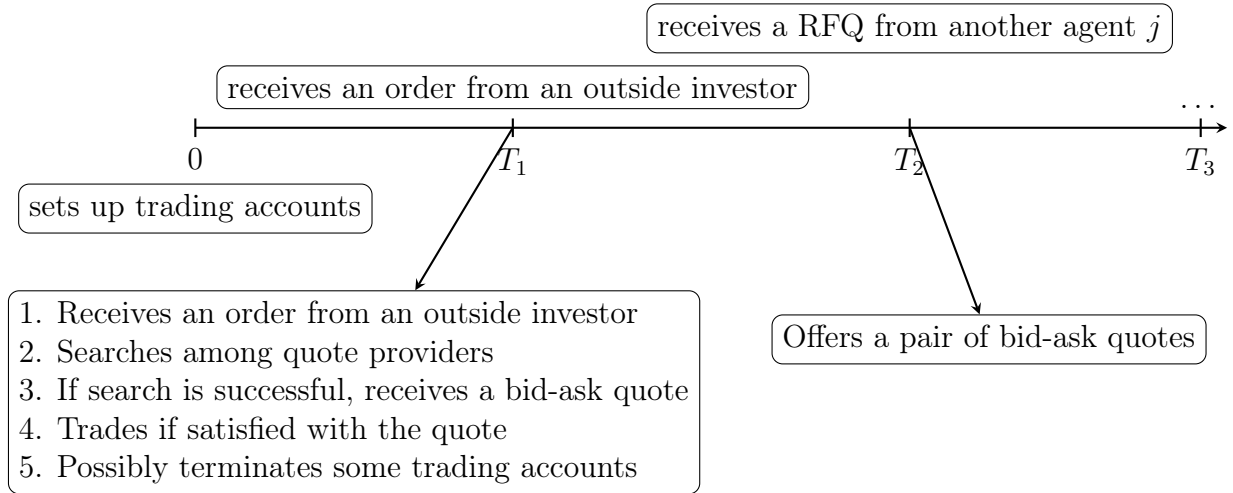
---

<sup>6</sup>A stub quote is a bid or offer price so far away from the prevailing market that it is not intended to be executed. Following the Flash Crash on May 6, 2010, the SEC banned market maker stub quotes.

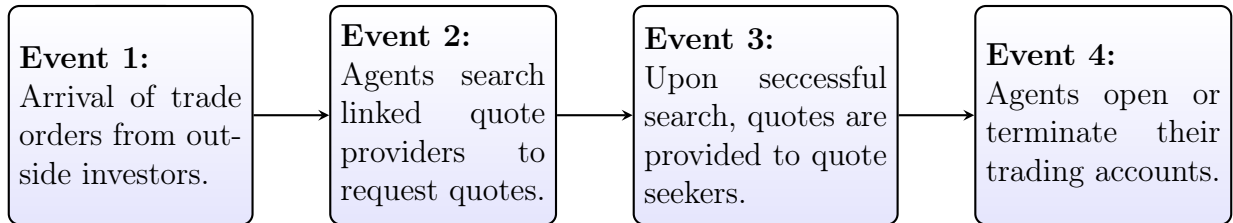
<sup>7</sup>Please see [MFA \(2015\)](#) for the practice of Name Give Up.

the agent bought or sold the asset, respectively. Technically, assuming this abstract trading device ensures the existence of stationary equilibria. In equilibrium, trades with this device only occur when a dealer is "at the boundary" of her inventory position (that is, when the dealer's inventory size is too large for her to profitably provide liquidity). This is an event that occurs with a very small frequency. Only in this case would the dealer route the trade to the mechanism as a final resort. The assumption of the existence of this device does not affect the interpretation of the symmetric-agent model, and will be relaxed in a richer model to be presented in [Section 5](#).

[Figure 3](#) illustrates the order of events from the perspective of a given agent  $i$ , and [Figure 4](#) shows the sequence of events happening at a given time  $t > 0$ .



**Figure 3** – Timeline of trading events for a given agent  $i$ .



**Figure 4** – Sequence of events happening at a given time  $t > 0$ .

**Solution concept.** Given an agent  $i$  and a time  $t \geq 0$ , let  $N_{it}^{\text{out}}$  be the set of quote providers

of agent  $i$  at time  $t$ , and  $N_{it}^{\text{in}}$  be the set of agents who have  $i$  as a quote provider. The sets  $(N_{it})_{t \geq 0} = (N_{it}^{\text{out}}, N_{it}^{\text{in}})_{t \geq 0}$  of potential counterparties of agent  $i$  are taken right continuous with left limits (RCLL) almost surely.<sup>8</sup> Let  $\mathcal{F}_{it}$  represent the information of agent  $i$  up to but excluding time  $t$ . That is,  $\mathcal{F}_{it}$  is the  $\sigma$ -algebra generated by the sets  $(N_{is})_{s < t}$  of all potential counterparties of agent  $i$  up to but excluding time  $t$ , her past inventories  $(x_{is})_{s < t}$ , the quotes  $(p_{is})_{s < t} = (b_{is}, a_{is})_{s < t}$  she offered, the quotes  $(\tilde{p}_{is})_{s < t}$  she was offered, the identities  $(j_{is})_{s < t}$  of the associated quote providers, the monetary transactions  $(\mathcal{P}_{is})_{s < t}$  related to trading, the types  $(O_{is})_{s < t}$  of outside-investor trade orders, and the desired trade directions  $(R_{is})_{s < t}$  of Requests for Quotes from other agents. We can take  $O_{is}, R_{is} \in \{\text{Buy}, \text{Sell}, \text{NA}\}$ , with the obvious meanings for “Buy” and “Sell”. The option “NA” indicates that agent  $i$  does not receive a trade order or a Request for Quote at the given time. The monetary transaction  $\mathcal{P}_{it}$  may be associated with a trade with an outsider, another agent or the abstract mechanism.

- (i) A search strategy  $S_i$  of agent  $i$  specifies, for every time  $t > 0$ , a search decision  $S_{it} \in \{\text{Search}, \text{No Search}\}$ . In addition to the prior information  $\mathcal{F}_{it}$ , agent  $i$  also observes the type of her outside order at time  $t$ , if there is one. Therefore,  $S_{it}$  must be measurable with respect to  $\mathcal{F}_{it}^1$ , where  $\mathcal{F}_{it}^1$  represents the combined information<sup>9</sup> of  $\mathcal{F}_{it}$  and  $O_{it}$ .
- (ii) A quoting strategy  $p_j$  for agent  $j$  specifies, for every  $t > 0$ , a price  $p_{jt}$  that  $j$  would quote upon receiving a Request for Quote. The quote  $p_{jt}$  is measurable with respect to  $\mathcal{F}_{jt}^2$ , where  $\mathcal{F}_{jt}^2$  represents the combined information of  $\mathcal{F}_{jt}$ ,  $R_{jt}$  and  $O_{jt}$ .
- (iii) A response strategy  $\rho_i$  for agent  $i$  specifies, for every  $t > 0$ , a trade decision  $\rho_{it} \in \{\text{Purchase}, \text{Sale}, \text{Rejection}\}$ . The response  $\rho_{it}$  is measurable with respect to  $\mathcal{F}_{it}^3$ , where  $\mathcal{F}_{it}^3$  represents the combined information of  $\mathcal{F}_{it}$ ,  $O_{it}$ ,  $R_{it}$ ,  $p_{it}$ ,  $\tilde{p}_{it}$  and  $j_{it}$ .
- (iv) Let  $N$  be the set of all  $n$  agents. An account maintenance strategy  $N_i^{\text{out}}$  of agent  $i$  specifies, for every  $t \geq 0$ , a set  $N_{it}^{\text{out}} \subseteq N/\{i\}$  of quote providers to  $i$ . The choice  $N_{it}^{\text{out}}$

---

<sup>8</sup>An RCLL function is a function defined on  $\mathbb{R}^+$  that is right-continuous and has left limits everywhere. RCLL functions are standard in the study of jump processes. Please see, for example, [Protter \(2005\)](#). Here, the process  $(N_{it})_{t \geq 0}$  is taken to be RCLL to be consistent with account maintenance behavior in real life. It also ensures that the set  $N_{it-}$  of counterparties available at time  $t$  is well defined for every  $t > 0$ .

<sup>9</sup>That is,  $\mathcal{F}_{it}^1 = \sigma(\mathcal{F}_{it}, O_{it})$ .

is measurable with respect to  $\mathcal{F}_{it}^4$ , where  $\mathcal{F}_{it}^4$  represents the combined information of  $\mathcal{F}_{it}$ ,  $O_{it}$ ,  $R_{it}$ ,  $p_{it}$ ,  $\tilde{p}_{it}$ ,  $j_{it}$ ,  $\mathcal{P}_{it}$  and  $x_{it}$ .

The total net payoff to be achieved by agent  $i$ , beginning at time  $t$ , is

$$U_{it} = \int_t^\infty e^{-r(s-t)} (-\beta x_{is}^2 - |N_{is}^{\text{out}}| c) ds + \mathcal{P}_{is} dM_{is-}, \quad (2)$$

where  $M_{it-}$  is the number of trades that agent  $i$  has made up to but excluding time  $t$ . The continuation utility of agent  $i$  is given by  $E(U_{it} | \mathcal{F}_{it})$ .

In a perfect Bayesian equilibrium (PBE), each agent maximizes her continuation utility at each time  $t \geq 0$ , with the conjecture that other agents will follow their equilibrium strategies. I consider perfect Bayesian equilibria in which each agent's strategy is Markovian and stationary. Formally, let  $Y_{it}$  be the Markov state variable for agent  $i$  at time  $t$ . That is,  $Y_{it} = (x_{it}, N_{it}, O_{it}, R_{it}, p_{it}, \tilde{p}_{it}, j_{it}, \mathcal{P}_{it})$ . A strategy  $\sigma_i$  of agent  $i$  is *Markovian* if  $\sigma_{it}$  is measurable with respect to  $Y_{it}$  for every  $t > 0$ . Thus, a Markovian strategy  $\sigma_i$  of agent  $i$  can be written as  $[f_{it}(Y_{it})]_{t>0}$  for some measurable function  $f_{it}$ . A Markovian strategy  $\sigma_i$  is said to be *stationary* if for every  $t > 0$ ,  $f_{it}(\cdot) = f_i(\cdot)$  for some measurable function  $f_i$ .

Given any strategy profile  $\sigma$ , let  $G_t$  be the associated trading network at time  $t \geq 0$ . That is,  $G_t$  is a directed network with nodes given by the  $n$  agents, with a directed link from node  $i$  to  $j$  if and only if agent  $i$  has a trading account with agent  $j$  at time  $t$ . The network  $G_t$  may depend on the realization of random events before time  $t$ . If there exists a deterministic trading network  $G$  such that for every  $t \geq 0$ ,  $G_t = G$  almost surely, then  $G$  is said to be the *trading network generated by  $\sigma$* . If  $\sigma$  is a PBE in stationary strategies, then  $G$  is simply said to be an *equilibrium trading network* of the model.

### 3 Equilibrium, Core-Periphery Network and Core Size

I show that there exists a family of perfect Bayesian equilibria in stationary strategies that generates trading networks of the form depicted in [Figure 1](#). The family of equilibria is indexed by the number of core dealers, varying from 0 to  $m^*$ , where  $m^*$  is the maximally

sustainable number of dealers. I explicitly characterize the family of equilibria, and determine the maximally sustainable core size. Proofs are given in [Appendix B](#).

Suppose  $I$  is some subset of agents, and  $J$  is the complement of  $I$ , with  $|J| = m$ ,  $|I| = n - m$ . Anticipating the equilibrium, I will call the agents in  $I$  the non-dealers, and agents in  $J$  the dealers. I consider the following strategy profile: At time 0, each non-dealer agent  $i \in I$  opens a trading account with all  $m$  dealers in  $J$ . The dealers in  $J$  set up trading accounts only with each other. In this candidate equilibrium, the  $m$  dealers are thus the common quote providers to all agents.

### Equilibrium spread:

An agent searches among her dealer counterparties whenever she receives an outside order, and does not search otherwise. I denote this search strategy by  $S^*$ . Since quote providers are not allowed to post stub quotes that exceed the total gain  $\pi$  per trade, a quote seeker optimally trades given any quote. I denote this response strategy by  $\rho^*$ . If a given non-dealer  $i \in I$  fails his search, then no trade takes place. The inventory of  $i$  is therefore always 0, because a non-dealer cannot take an outside order into his own inventory.

Dealers' equilibrium bid-ask quotes are anticipated to be of the form  $(-P, P)$ , for some  $0 < P < \pi$ , so that  $P$  is the mid-to-bid spread. Thus, non-dealer  $i$  earns a profit of  $\pi - P$  for every outside order that is followed by a successful search.

Non-dealer  $i$  must also decide when to open or terminate trading accounts. I let  $\Phi_{d,P}$  denote the continuation utility of non-dealer  $i$  if  $i$  maintains  $d$  dealer counterparties, for every  $0 \leq d \leq m$ . Because outside orders arrive at the mean rate  $2\lambda$ , and because the probability of a successful search among the  $d$  dealers is  $\theta_d$ , it must be the case that for every  $0 \leq d \leq m$ ,

$$\Phi_{d,P} = \frac{2\lambda\theta_d(\pi - P) - dc}{r}. \quad (3)$$

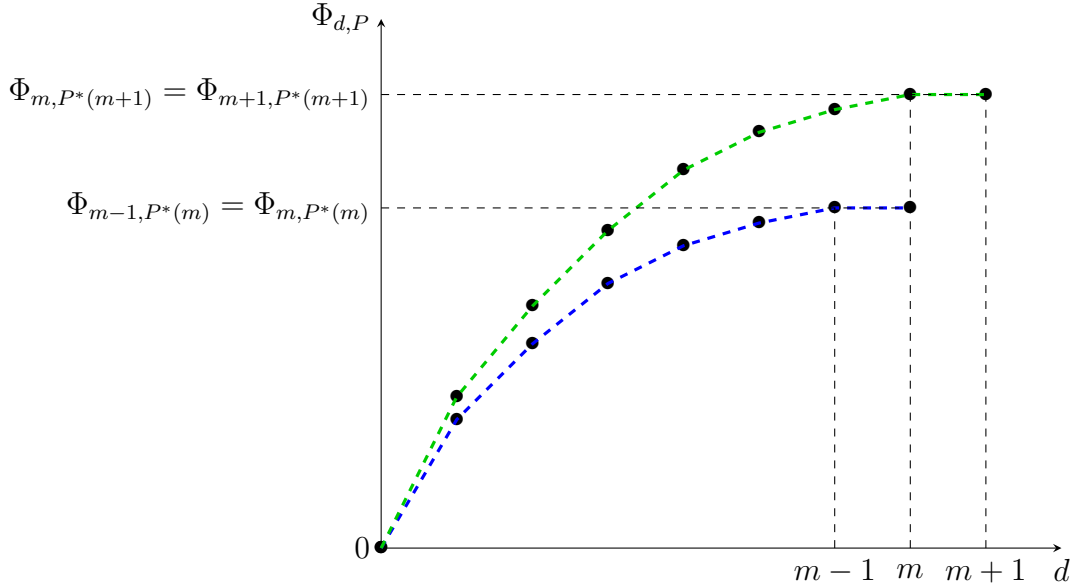
For a given stationary mid-to-bid spread  $P$ , the continuation utility  $\Phi_{d,P}$  of a non-dealer is strictly concave in the number  $d$  of dealer counterparties of the non-dealer. Maintaining and using more dealer trading accounts incurs higher account maintenance costs, which grow

linearly with the number of trading accounts, as captured by the term  $dc$  in expression (3). On the other hand, having more dealer counterparties allows the non-dealer to generate trading profits at a higher mean frequency. The total expected rate of trade profits, as captured by  $2\lambda\theta_d(\pi - P)$ , grow sub-linearly with the number  $d$  of dealer counterparties. The net present value is therefore concave in the number  $d$  of dealer counterparties.

The equilibrium spread  $P$  must satisfy the following non-dealer's indifference condition

$$\Phi_{m,P} = \Phi_{m-1,P}. \quad (4)$$

The indifference condition gives a non-dealer the ability to costlessly discontinue any one of his  $m$  trading accounts, making termination of a trading relationship a credible threat to dealers should they offer aggressively unfavorable prices. Suppose  $P$  is determined by the indifference condition (4). The payoff  $\Phi_{d,P}$  of a non-dealer must be strictly increasing in  $d$  for  $d \leq m-1$ , before plateauing between  $d = m-1$  and  $m$ . Figure 5 provides an illustration of  $\Phi_{d,P}$  as a function  $d$ .



**Figure 5** – A non-dealer's continuation utility  $\Phi_{d,P}$  as a function of the number  $d$  of his dealer counterparties.

The indifference condition (4) uniquely determines the equilibrium mid-to-bid spread

$$P^*(m) = \pi - \frac{c}{2\lambda(\theta_m - \theta_{m-1})}. \quad (5)$$

I will show that, in equilibrium, an agent's (dealer or non-dealer) account maintenance strategy is to always maintain  $m$  or  $m - 1$  dealer counterparties, and to discontinue her trading account with a given dealer whenever the following two conditions are both met:

- (i) The agent has  $m$  dealer counterparties in total.
- (ii) The agent receives an offer price strictly higher than  $P^*(m)$ , or a bid price strictly lower than  $-P^*(m)$ .

Whenever these two conditions are met, the agent, after trading with the dealer at the current price quote, immediately closes her account with the dealer. In particular, any given dealer always maintains trading accounts with all other  $m - 1$  dealers. I denote this account maintenance strategy by  $N_m^*$ . To give an example, suppose non-dealer  $i$  wishes to buy 1 unit of the asset, and dealer  $j$  gouges non-dealer  $i$  by posting an offer price  $a$  that is strictly higher than the equilibrium offer price  $P^*(m)$ . Non-dealer  $i$  buys at the price  $a$  (since otherwise  $i$  would have lost a profitable opportunity to trade), and then immediately terminates his trading account with  $j$ .

No account termination occurs on the equilibrium path. However, the ability of non-dealers to terminate their trading accounts constitutes a credible threat to dealers that would discourage dealers from gouging. I will describe shortly the tradeoff faced by a dealer when providing price quotes to a non-dealer.

From expression (5) of the equilibrium spread  $P^*(m)$ , one obtains the following implication of dealer price competition.

**Proposition 1.** *The equilibrium mid-to-bid spread  $P^*(m)$  is strictly decreasing in the number  $m$  of dealers. When there is only one dealer, the equilibrium spread  $P^*(1)$  is the monopoly price that extracts all rents from non-dealers. That is,  $\Phi_{1,P^*(1)} = 0$ .*

If there are more competing dealers in equilibrium, the dealers compete more fiercely for trades by offering tighter bid-ask quotes. On the other hand, with more competing dealers in the market, the benefits of serving as a dealer decrease, as each dealer receives a thinner order flow from non-dealers. This would limit the equilibrium level of dealer competition. I now demonstrate this intuition, and determine the maximum number  $m^*$  of dealers. I first describe a quote provider's optimization problem and derive her optimal quoting strategy.

### **Tightest dealer-sustainable-spread:**

I fix some candidate spread  $P \in (0, \pi]$ , and some number  $k \geq 1$  of non-dealer customers of a given dealer  $j \in J$ . I consider the following continuous-time stochastic control problem for dealer  $j$ , in which  $j$  is artificially restricted to the given spread  $P$ . (This restriction is relaxed later.)

*Dealer's problem  $\mathcal{P}(k, m, P)$ :*

- The state space is the set  $\mathbb{Z}$  of integers, the inventory space of dealer  $j$ .
- The control space is  $\{-P, -P_{CB}\} \times \{P, P_{CB}\}$ , which is the set of all possible bid-ask quotes that  $j$  may offer. That is,  $j$  charges quote seekers the same price  $P$  for trading (either buying or selling) one unit of the asset, and she can make a riskless principal trade at the price  $P$ , with the abstract mechanism assuming the opposite side of the trade.
- Dealer  $j$  receives Requests for Quote from  $k$  non-dealers and  $m - 1$  dealers at the total mean contact rate  $2\lambda(k\theta_m/m + \theta_{m-1})$ . (The “per capita” mean contact rate is  $2\lambda\theta_m/m$  for non-dealers and  $2\lambda\theta_{m-1}/(m - 1)$  for dealer customers.) Every contacting agent seeks to buy or sell 1 unit of the asset, independently across contacts and with equal probability  $1/2$ .
- The payoff of dealer  $j$  is the expected discounted value of all her monetary transfers and inventory cost, as specified in (2).



I denote this problem by  $\mathcal{P}(k, m, P)$ . In the actual game, a dealer is allowed to post any quote in  $[-\pi, \pi]$  rather than being limited to the prices  $\pm P$  as in the control problem  $\mathcal{P}(k, m, P)$ . However, it will be shown that dealers have no incentive to quote prices other than  $\pm P^*(m)$ , where  $P^*(m)$  is the equilibrium mid-to-bid spread given by (5). Therefore, the auxiliary problem  $\mathcal{P}(k, m, P^*(m))$  characterizes an (unrestricted) optimal quoting strategy of a dealer.

I let  $V_{k,m,P}$  be the value function of the dealer in the control problem  $\mathcal{P}(k, m, P)$ . That is,  $V_{k,m,P}(x)$  is the maximum attainable continuation utility of the dealer in  $\mathcal{P}(k, m, P)$  if her current inventory is  $x$ , for every  $x \in \mathbb{Z}$ . The Bellman principle implies that for every  $x \in \mathbb{Z}$ ,

$$\begin{aligned} rV_{k,m,P}(x) = & -\beta x^2 + \lambda \left( k \frac{\theta_m}{m} + \theta_{m-1} \right) [V_{k,m,P}(x+1) - V_{k,m,P}(x) + P]^+ \\ & + \lambda \left( k \frac{\theta_m}{m} + \theta_{m-1} \right) [V_{k,m,P}(x-1) - V_{k,m,P}(x) + P]^+. \end{aligned} \quad (6)$$

The first term  $-\beta x^2$  is the inventory flow cost for holding  $x$  units of the asset. The second term is the expected rate of change in the dealer's continuation utility associated with requests to sell. The last term is the analogous mean profit rate from requests to buy. Letting

$$\Delta(x) = V_{k,m,P}(x) - V_{k,m,P}(x+1) \quad (7)$$

denote the marginal indirect cost from a purchase, one obtains the following lemma:

**Lemma 1.** *Given an inventory level  $x$ , the optimal bid and offer prices of the dealer are*

$$b^*(x) = \begin{cases} -P, & \text{if } \Delta(x) \leq P, \\ -P_{CB}, & \text{if } \Delta(x) > P, \end{cases} \quad a^*(x) = \begin{cases} P, & \text{if } \Delta(x-1) \geq -P, \\ P_{CB}, & \text{if } \Delta(x-1) < -P. \end{cases}$$

That is, the dealer optimally quotes the offer price  $P$  whenever her marginal indirect cost of selling one unit of the asset is lower than the trading benefit  $P$ . If the marginal inventory cost is strictly higher than the benefit  $P$ , the dealer is not willing to sell from her own inventory, resorting to the abstract mechanism to take the other side of the trade.

From the theory of dynamic programming, one can show that the value function  $V_{k,m,P}$  is even and strictly concave.

**Lemma 2.** *The function  $V_{k,m,P}$  is even and strictly concave, in that for every  $x \in \mathbb{Z}$ ,*

$$V_{k,m,P}(x) = V_{k,m,P}(-x), \quad \Delta(x) < \Delta(x+1).$$

**Lemmas 1** and **2** imply that the dealer's optimal quoting strategy is threshold-type:

**Proposition 2.** *The dealer has a unique optimal quoting strategy  $[b^*(x), a^*(x)]$ , characterized by an inventory threshold level  $\bar{x}_{k,m,P} \in \mathbb{Z}$ :*

- *If  $x \leq -\bar{x}_{k,m,P}$ , then  $[b^*(x), a^*(x)] = [-P, P_{CB}]$ .*
- *If  $x \geq \bar{x}_{k,m,P}$ , then  $[b^*(x), a^*(x)] = [-P_{CB}, P]$ .*
- *If  $-\bar{x}_{k,m,P} < x < \bar{x}_{k,m,P}$ , then  $[b^*(x), a^*(x)] = [-P, P]$ .*

**Proposition 2** indicates that when the dealer is very short (that is, if her inventory is lower than the threshold  $-\bar{x}_{k,m,P}$ ), she is not willing to sell more assets at the price  $P$ , since the trade gain  $P$  no longer covers her marginal inventory cost. Symmetrically, when the dealer's inventory size exceeds  $\bar{x}_{k,m,P}$ , she is not willing to buy more. If her inventory is within the range of  $(-\bar{x}_{k,d,P}, \bar{x}_{k,d,P})$ , she is relatively “flat,” and therefore is willing to warehouse inventory risk.

Now consider whether the dealer has incentive to gouge a non-dealer. If the dealer quotes some offer price strictly lower than  $P$  to an agent requesting to buy, she would have forgone some trade profit. Similarly, the dealer has no incentive to quote a bid price strictly higher than  $-P$ . If the dealer quotes an offer price strictly higher than  $P$ , she increases her trade profit in the current contact at the cost of possibly losing a non-dealer customer for future trading. The highest offer price being  $\pi$ , the one-shot incentive by the dealer to gouge is thus

$$\Pi(P) = \pi - P. \tag{8}$$

The case of a quote seeker requesting to sell is symmetric, and gives the dealer the same one-shot incentive to gouge.

The dealer's future equilibrium profits forgone due to losing one non-dealer customer is given by the loss in her indirect utility:

$$L_{k,m,P}(x) = V_{k,m,P}(x) - V_{k-1,m,P}(x). \quad (9)$$

The dealer has no incentive to deviate from offering the mid-to-bid spread  $P$  if and only if

$$\Pi(P) \leq \mathcal{L}(k, m, P) \equiv \frac{k}{k+m-1} \min_{|x| \leq \bar{x}_{k,m,P}} L_{k,m,P}(x), \quad (10)$$

where  $k/(k+m-1)$  is the probability that the contacting agent is a non-dealer. If inequality (10) is violated, then there exists some  $x = -\bar{x}_{k,m,P}, \dots, \bar{x}_{k,m,P}$  such that if the dealer has a post-trade asset inventory of size  $x$ , then gouging by offering the best deviation quote  $(-\pi, \pi)$  makes the dealer strictly better off. If (10) holds, then the dealer has no incentive to gouge on the equilibrium path by the One-Shot Deviation Principle.

A function  $f : \mathbb{Z} \rightarrow \mathbb{R}$  is said to be *U-shaped* if  $f$  is even and  $f(x+1) > f(x)$ ,  $\forall x \geq 0$ .

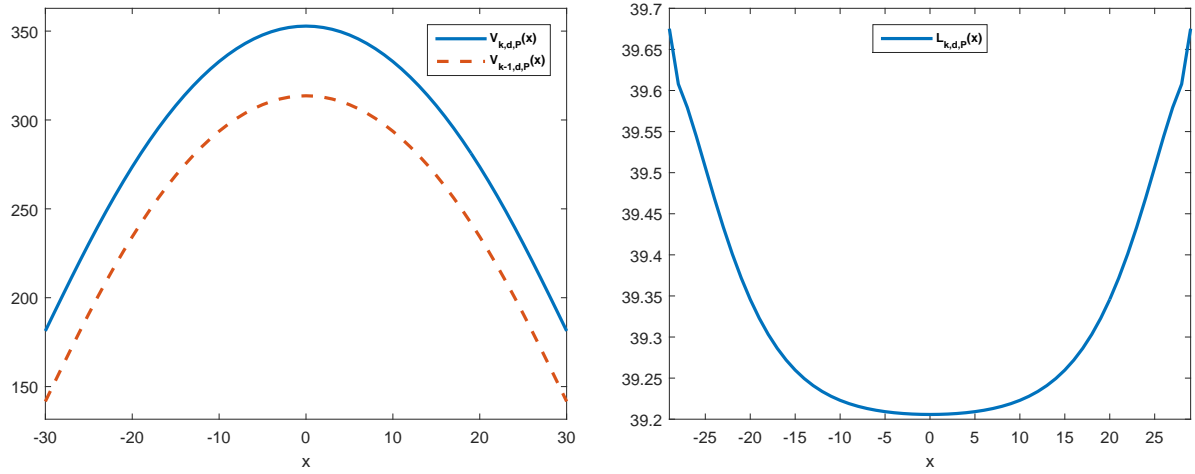
**Lemma 3.** *The loss  $L_{k,m,P}(x)$  due to losing one non-dealer customer is U-shaped in  $x$ .*

**Lemma 3** implies that the minimum on the right hand side of (10) is achieved when  $x = 0$ . That is, the dealer has the strongest incentive to gouge when she has no inventory. When the dealer has a very skewed inventory, she is more reliant on the order flow from her non-dealer customers to balance her inventory quickly, and thus has less incentive to gouge.

One can numerically compute the dealer's value function  $V_{k,m,P}$  through value iteration, and thus calculate her loss  $L_{k,m,P}(x)$  from gouging by (9). **Figure 6** shows the value functions  $V_{k,m,P}$  for  $k = 9, 8$  and the corresponding loss function  $L_{k,m,P}(\cdot)$ . The remaining parameters are:  $m = 3, P = 10, \lambda = 1, \theta_m = 0.9, \beta = 0.1$  and  $r = 0.1$ .

**Lemma 4.** *The equilibrium payoff forgone  $L_{k,m,P}(0)$  is (a) strictly increasing in the total number  $k$  of non-dealer customers, (b) strictly decreasing in the total number  $m$  of dealers, and (c) strictly increasing and continuous in the mid-to-bid spread  $P$ .*

The three properties of  $L_{k,m,P}(0)$  stated in **Lemma 4** can be interpreted as follows:



**Figure 6** – The value functions  $V_{k,m,P}$  and  $V_{k-1,m,P}$ , and the loss function  $L_{k,m,P}$  given by (9).

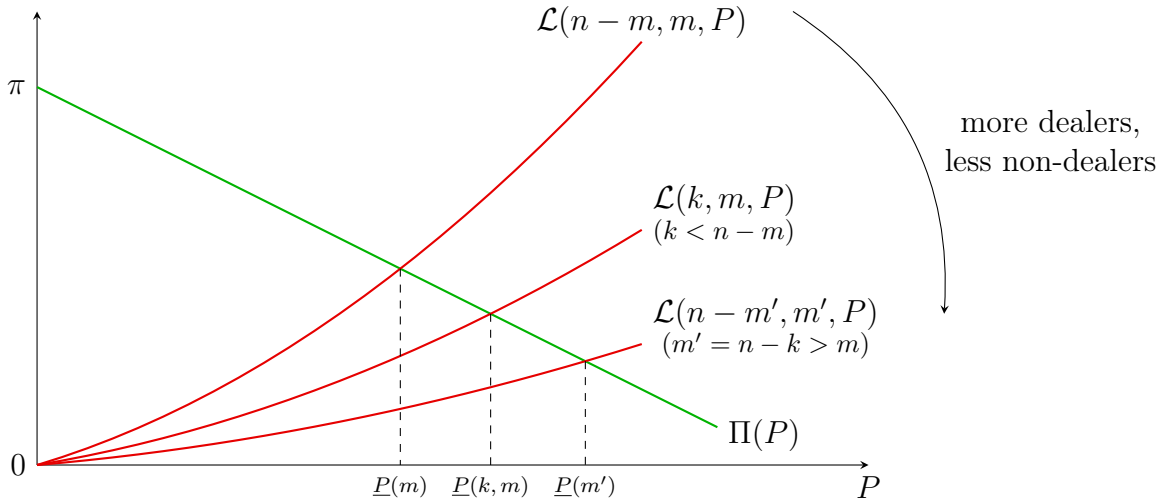
- (a) The marginal cost  $L_{k,m,P}(0)$  to the dealer of losing one non-dealer customer can be broken down into two components: (i) the direct cost of losing the future order flow from the non-dealer customer, which is partially compensated by (ii) the benefit of a reduced balance sheet (that is,  $\bar{x}_{k-1,m,P} < \bar{x}_{k,m,P}$ ) and thus reduced inventory cost. When the dealer has many non-dealer customers, she is efficient in managing her inventory by quickly netting many purchases against sales, and thus her inventory cost is less sensitive to a change in the number of her non-dealer customers (this result about inventory cost is formulated in [Proposition 8](#)). Consequently, the second component (the benefit of reduced inventory cost) becomes negligible. Hence, losing one non-dealer customer is more costly to a dealer who has more non-dealer customers.
- (b) When the total number  $m$  of dealers is larger, each dealer receives a smaller share of the total order flow from non-dealers. The marginal cost  $L_{k,m,P}(0)$  of losing one non-dealer customer is thus lower for any given dealer.
- (c) If the spread is higher, losing one non-dealer customer is more costly to the dealer.

Given that the one-shot incentive  $\Pi(P)$  to gouge is strictly decreasing in  $P$ , property (c) of [Lemma 4](#) leads to the following corollary.

**Corollary 1.** Condition (10) for not gouging is satisfied if and only if  $P \geq \underline{P}(k, m)$ , where  $\underline{P}(k, m)$  is uniquely determined by the equality:

$$\Pi(\underline{P}(k, m)) = \mathcal{L}(k, m, \underline{P}(k, m)). \quad (11)$$

Corollary 1 implies that  $[-\underline{P}(k, m), \underline{P}(k, m)]$  are the tightest quotes the dealer is willing to offer without having incentive to deviate. The mid-to-bid spread  $\underline{P}(k, m)$  is called *the tightest  $(k, m)$ -dealer-sustainable spread*. Any mid-to-bid spread  $P$  is  $(k, m)$ -dealer-sustainable if and only if  $P \geq \underline{P}(k, m)$ . Figure 7 illustrates the tradeoff between the gain  $\Pi(P)$  and the loss  $\mathcal{L}(k, m, P)$  from gouging as  $P$  varies, and shows how the tightest dealer-sustainable spread  $\underline{P}(k, m)$  is determined from the intersection of the two curves.



**Figure 7** – The tradeoff between the one-shot incentive to gouge and the future profits forgone.

In the actual network trading game, a given dealer has  $n - m$  non-dealer customers. Thus, condition (10) needs to be satisfied for  $k = n - m$  for dealers to refrain from gouging. Letting  $\underline{P}(m) \equiv \underline{P}(n - m, m)$ ,  $\underline{P}(m)$  is simply called the *tightest  $m$ -dealer-sustainable spread*. The next corollary regarding  $\underline{P}(k, m)$  and  $\underline{P}(m)$  follows from (a) and (b) of Lemma 4.

**Corollary 2.** The tightest  $(k, m)$ -dealer-sustainable spread  $\underline{P}(k, m)$  is strictly decreasing in the number  $k$  of non-dealer customers of a given dealer, and the tightest  $m$ -dealer-sustainable spread  $\underline{P}(m)$  is strictly increasing in the number  $m$  of dealers.

That is, when a dealer has more non-dealer customers, she is able to offer better quotes, thanks to her ability to efficiently manage inventory risk by more quickly netting purchases against sales. A well connected dealer is in this sense a liquidity hub. When there are more dealers in the market, each dealer has reduced efficiency in managing inventory risk, and is only able to sustain a wider spread. [Figure 7](#) provides an illustration of [Corollary 2](#).

### Maximally sustainable core size:

The equilibrium spread  $P^*(m)$ , as defined in [\(5\)](#), must be  $m$ -dealer-sustainable. Thus,

$$P^*(m) \geq \underline{P}(m). \quad (12)$$

Since  $P^*(m)$  is strictly decreasing in  $m$ , while  $\underline{P}(m)$  is strictly increasing in  $m$  (see [Proposition 1](#) and [Corollary 2](#)), condition [\(12\)](#) is equivalent to

$$m \leq m^*, \quad (13)$$

where  $m^*$  is the largest integer such that

$$P^*(m^*) \geq \underline{P}(m^*). \quad (14)$$

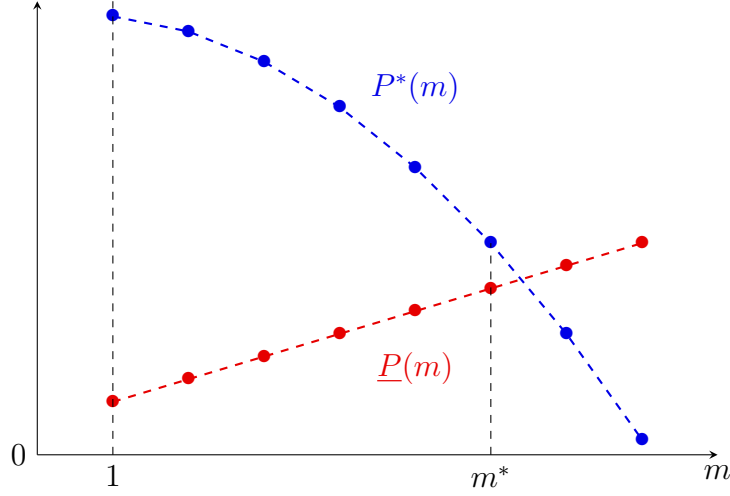
The number  $m^*$  is the *maximally sustainable core size*. [Figure 8](#) illustrates the comparison between the equilibrium spread  $P^*(m)$  and the tightest sustainable spread  $\underline{P}(m)$  as  $m$  varies, and shows how the maximum core size  $m^*$  is determined from the two spread curves.

### Equilibrium existence and uniqueness:

Now I can specify the candidate equilibrium quoting strategy. Suppose a given dealer has  $k$  non-dealer customers (for any  $k \leq n - m$ ) at the time she receives a Request for Quote. I denote the equilibrium bid-ask quote of the dealer by  $[b^*(x, k), a^*(x, k)]$ .

If  $P^*(m) \geq \underline{P}(k, m)$ , then

- If  $-\bar{x}_{k,m,P^*(m)} < x < \bar{x}_{k,m,P^*(m)}$ , then  $b^*(x, k) = -P^*(m)$ ,  $a^*(x, k) = P^*(m)$ .



**Figure 8** – The equilibrium spread  $P^*(m)$ , the tightest dealer-sustainable spread  $\underline{P}(m)$ , and the maximum core size  $m^*$ .

- If  $x \geq \bar{x}_{k,m,P^*(m)}$ , then  $b^*(x, k) = -P_{CB}^*(m)$ ,  $a^*(x, k) = P^*(m)$ .
- If  $x \leq -\bar{x}_{k,m,P^*(m)}$ , then  $b^*(x, k) = -P^*(m)$ ,  $a^*(x, k) = P_{CB}^*(m)$ .

If  $P^*(m) < \underline{P}(k, m)$ , then  $b^*(x, k) = -\pi$ ,  $a^*(x, k) = \pi$ . I denote this quoting strategy by  $p_m^*$ . If a given non-dealer receives a Request for Quote (which is an event off the equilibrium path), his offer price is  $\pi$  and bid price  $-\pi$ . I denote this quoting strategy by  $p_0^*$ .

The candidate equilibrium strategy profile can now be readily defined.

**Definition 1.** The strategy profile  $\sigma^*(m)$  is defined by the following strategies:

- There is a partition  $I \cup J = N$  of agents with  $|J| = m$ . At  $t = 0$ , each agent  $i \in I$ , called a “non-dealer,” opens a trading account with all  $m$  agents in  $J$ , while agents in  $J$ , called “dealers,” set up trading accounts with each other. I denote the associated trading network by  $G(m)$ .
- Each agent in  $N$  (dealer or non-dealer) follows the stationary search strategy  $S^*$ , the response strategy  $\rho^*$  and the account maintenance strategy  $N_m^*$ .
- Each dealer follows the quoting strategy  $p_m^*$ .

(iv) Each non-dealer follows the quoting strategy  $p_0^*$ .

By convention, I denote by  $\sigma^*(0)$  the strategy profile associated with an empty trading network. That is, no agent opens any account, and no search or trade is conducted.

In the remainder of this section, I fix the model parameters  $(n, \beta, \pi, \lambda, \theta, c, r)$ , and let  $m^*$  be defined as in (14). The first main result of the paper concerns equilibrium existence.

**Theorem 1.** *(a) For every  $0 \leq m \leq m^*$ , the strategy profile  $\sigma^*(m)$  is a perfect Bayesian equilibrium in stationary strategies. (b) For every  $m > m^*$ ,  $\sigma^*(m)$  is not a PBE.*

**Theorem 1** shows that there exists a family of perfect Bayesian equilibria in stationary strategies. Each equilibrium generates a trading network of the form depicted in **Figure 1**. The family of equilibria is indexed by the number of core dealers, varying from 0 to  $m^*$ , where  $m^*$  is the maximally sustainable number of dealers. In an equilibrium  $\sigma^*(m)$  for some  $0 \leq m \leq m^*$ , the trading network  $G(m)$  is stable over time. The equilibrium mid-to-bid spread is  $P^*(m)$ , and dealers are deterred from gouging by the fear of losing non-dealer customers. Non-dealers do not terminate any account on the equilibrium path, but their ability to do so constitutes a credible threat that discourages dealers from gouging.

I next provide an equilibrium uniqueness result, stating that  $P^*(m)$  is the unique equilibrium spread in the equilibrium trading network  $G(m)$ .

**Theorem 2.** *For every  $m \geq 0$ , if  $G(m)$  is the trading network generated by some perfect Bayesian equilibrium in stationary strategies, then on the equilibrium path, any given dealer  $j \in J$  always posts some constant bid price  $b_j$  and some constant offer price  $a_j$ , with a spread equal to  $b_j - a_j = 2P^*(m)$ .*

The next theorem follows from **Theorems 1** and **2**:

**Theorem 3.** *Given some integer  $m \geq 0$ , the network  $G(m)$  is the trading network generated by some perfect Bayesian equilibrium in stationary strategies if and only if  $0 \leq m \leq m^*$ .*



### Equilibrium robustness and selection:

The uniqueness of the equilibrium spread  $P^*(m)$  ( $1 \leq m \leq m^*$ ) implies that the equilibrium  $\sigma^*(m)$  is robust against dealer collusion:

**Corollary 3.** *Based on any equilibrium  $\sigma^*(m)$  ( $1 \leq m \leq m^*$ ), no collusive price rigging by any dealer sub-coalition can be sustained by a PBE in stationary strategies.*

Even though the uniqueness result of [Theorem 2](#) is sufficient to guarantee no dealer collusion, it is more intuitive to consider an example of how such a collusion would break down. If all  $m$  dealers were to collude to offer a larger mid-to-bid spread  $P$ , then every non-dealer no longer has incentive to maintain all the  $m$  dealer counterparties given the wider spread. This is because when  $P > P^*(m)$ , the non-dealer indifference condition [\(4\)](#) becomes a strict inequality:

$$\Phi_{m,P} < \Phi_{m-1,P}.$$

Therefore, any collusion by the  $m$  dealers to raise the spread will cause some dealers to lose some non-dealer customers, destabilizing the trading network. Since losing non-dealer customers is costly for any given dealer, the dealers cannot credibly engage in such a collusion.

In an equilibrium  $\sigma^*(m)$  of the model ( $1 \leq m \leq m^*$ ), the equilibrium spread  $P^*(m)$  offered by the  $m$  dealers does not compete away all trading profit, as opposed to the case of Bertrand competition. Imperfect price competition is due to the nature of the sequential search protocol. Specifically, since agents request quotes sequentially, they cannot screen price offers simultaneously to select the best one.

On the other hand, dealers charge an equilibrium spread  $P^*(m)$  that is lower than the monopoly price  $\pi$ , leaving some rent to quote seekers. This is in contrast with the Diamond paradox, in which all price-setting firms post monopoly prices for consumers who search sequentially for price information. In Diamond's setting, the market is static in that every given pair of customer and price setting firm meets only once. In this paper, the continuous-time infinite-horizon trading game features repeated contacts between non-dealers and dealers. Dealer incentive to widen spread is deterred by the implications for repeated business: a

dealer that provides a wider spread increases her trading gain in the current contact, but non-dealers can credibly cut off the dealer in future trading as a punishment for gouging. Moreover, the threat of punishment is not only credible, but also *renegotiation proof*. If a dealer, after gouging, tries to convince the affected non-dealer to refrain from cutting her off, the only way she can possibly achieve this goal is by promising the non-dealer a narrower-than-equilibrium spread in future encounters. However, such a promise is not credible, as it is common knowledge that the dealer never has incentive to provide a spread that is smaller than the equilibrium spread  $P^*(m)$ .

The idea that repeated business may help circumvent the Diamond paradox also appears in Bagwell and Ramey (1992), where firms offer a better-than-monopoly price for the fear of losing customers in future periods. However, Bagwell and Ramey give a continuum of equilibrium prices, even when the number of firms is fixed. The model of this paper makes a unique prediction for the equilibrium spread  $P^*(m)$  for a given number  $m$  of dealers.

However, **Theorem 3** is still not sharp enough, in that it predicts the coexistence of a family of equilibria indexed by the number of core dealers. The next proposition shows that these equilibria are Pareto-ranked for non-dealers. This result helps to offer a criterion for equilibrium selection.

**Proposition 3.** *The equilibrium outcomes induced by the strategy profiles  $\sigma^*(m)$  ( $0 \leq m \leq m^*$ ) are Pareto-ranked for non-dealers, in that every non-dealer enjoys a strictly higher payoff under  $\sigma^*(m')$  than  $\sigma^*(m)$ , for any  $0 \leq m < m' \leq m^*$ .*

**Proposition 3** suggests that all non-dealers prefer an equilibrium with more competing dealers, since the benefit associated with a tighter equilibrium spread and more frequent trading outweighs additional account maintenance costs. A graphical illustration of **Proposition 3** is provided by **Figure 5**. One implication of **Proposition 3** is equilibrium selection: an equilibrium  $\sigma(m)$  with fewer core dealers may be ruled out if agents can actively coordinate the selection of dealers. To see this, suppose  $\sigma^*(m')$  is an equilibrium with more core dealers  $m' > m$ . Any  $m' - m$  non-dealers under the equilibrium  $\sigma^*(m)$  can credibly propose to serve as core dealers, on top of the existing  $m$  dealers, as such a proposal will result in a

new equilibrium  $\sigma^*(m')$ . The  $m' - m$  newly entered dealers improve their payoff by exploiting their central network position to earn intermediation spread. Following [Proposition 3](#), the remaining  $n - m'$  non-dealers also benefit from greater dealer competition. Finally, the  $m$  existing dealers lose some of the market, and are forced to provide a narrower spread  $P^*(m') < P^*(m)$ . However, they have no choice other than adapting to their new equilibrium strategies, as an existing dealer who refuses to lower her spread risks losing future business and ultimately her dealer position.

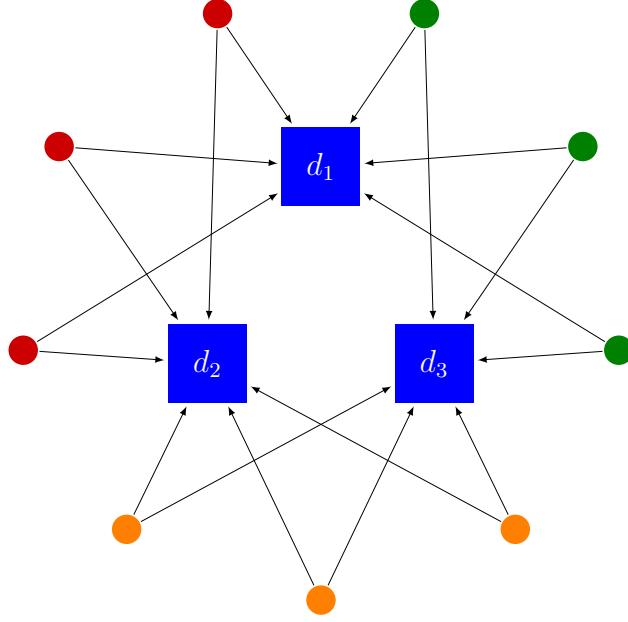
The equilibrium  $\sigma^*(m^*)$  with the maximum number of core dealers cannot be overturned in this manner, since non-dealers cannot credibly propose to enter as additional dealers. Therefore, the equilibrium  $\sigma^*(m^*)$  is uniquely robust against dealer entry. This selection procedure through dealer entry can be regarded as a form of price competition, since non-dealers can enter as dealers to compete for trades. Here, the non-dealers are able to communicate their intention to serve as dealers, even though they are unable to commit to a tighter spread level. The selection procedure closely mimics the logic of *coalition-proof Nash equilibrium* introduced by [Bernheim, Peleg, and Whinston \(1987\)](#). In the family of equilibria  $[\sigma^*(m)]_{0 \leq m \leq m^*}$ , the only coalition-proof equilibrium is  $\sigma^*(m^*)$ .

The discussion above is summarized in the following proposition:

**Proposition 4.** *The equilibrium  $\sigma^*(m^*)$  is the unique equilibrium in the family  $[\sigma^*(m)]_{0 \leq m \leq m^*}$  that is robust against dealer entry.*

### Alternative core-periphery structures:

Thus far, I have considered equilibria that generate a trading network of the form depicted in [Figure 1](#). Alternative network structures have not been examined yet. In practice, observed core-periphery structures in OTC markets are less “concentrated,” in that a typical non-dealer is connected to some dealers, but not to all of them. [Figure 9](#) illustrates such an example of core-periphery structure. I next turn to whether such a trading network could emerge in equilibrium, and if so, how many dealers could reside in the core, and how many trading accounts a non-dealer would have.



**Figure 9** – Example: every Non-Dealer is connected to 2 of the 3 Dealers.

Formally, I consider a class of core-periphery networks with the following property: the network has nodes given by the  $n$  agents, which are partitioned into some  $I \cup J = N$ . Each node  $i \in I$ , representing a “non-dealer,” has a directed link to some nodes in  $J$ , while the nodes in  $J$ , representing “dealers,” have outgoing links to some other nodes in  $J$ . I let  $\mathcal{G}$  denote the class of all such core-periphery networks.

Given a node  $i \in N$  of  $G(I, J)$ , I denote by  $\deg^+(i)$  and  $\deg^-(i)$  be the indegree and outdegree of node  $i$  respectively, and by  $N^+(i)$  and  $N^-(i)$  the set of quote providers and quote seekers of agent  $i$  respectively. Letting

$$m = \max_{i \in N} \deg^+(i)$$

be the maximum outdegree of a node in  $G(I, J)$ , it will be shown that every node must have outdegree  $m$  or  $m - 1$  in equilibrium. For every dealer  $j \in J$ , I let

$$k_j = \sum_{i \in N^{textin}(j)} \mathbb{1}_{\{\deg^+(i)=m\}}, \quad d_j = \sum_{i \in N^{textin}(j)} \mathbb{1}_{\{\deg^+(i)=m-1\}}$$

be the number of quote seekers of  $j$  that have outdegree  $m$  and  $m - 1$  respectively. Only the quote seekers with outdegree  $m$  can costlessly terminate their trading accounts with  $j$ . I let  $V_{k,m,d,P}$  be the value function that solves the following HJB equations: for every  $x \in \mathbb{Z}$ ,

$$\begin{aligned} rV_{k,m,d,P}(x) = & -\beta x^2 + \lambda \left( k \frac{\theta_m}{m} + d \frac{\theta_{m-1}}{m-1} \right) [V_{k,m,d,P}(x+1) - V_{k,m,d,P}(x) + P]^+ \\ & + \lambda \left( k \frac{\theta_m}{m} + d \frac{\theta_{m-1}}{m-1} \right) [V_{k,m,d,P}(x-1) - V_{k,m,d,P}(x) + P]^+. \end{aligned}$$

I let  $L_{k,m,d,P}$  be the loss in the indirect utility function associated with losing one quote seeker:

$$L_{k,m,d,P} = V_{k,m,d,P} - V_{k-1,m,d,P}.$$

Similar to [Lemma 3](#), one can show that  $L_{k,m,d,P}$  is U-shaped. I let

$$\mathcal{L}(k, m, d, P) = \frac{k}{k+d} L_{k,m,d,P}(0).$$

and  $\underline{P}(k, m, d) \in \mathbb{R}^+$  be determined by the equality

$$\Pi(\underline{P}(k, m, d)) = \mathcal{L}(k, m, d, \underline{P}(k, m, d)).$$

Similar to [Corollary 2](#), one can show that  $\underline{P}(k, m, d)$  is strictly decreasing in  $k$ . I let  $k(m, d)$  be the smallest integer such that

$$\underline{P}(k(m, d), m, d) \leq P^*(m),$$

where  $P^*(m)$  is the equilibrium spread as given by expression [\(5\)](#). Any given dealer  $j \in J$  needs at least  $k(m, d_j)$  quote seekers to sustain the equilibrium spread  $P^*(m)$  without having incentive to gouge. The following theorem generalizes [Theorem 3](#) by providing a necessary and sufficient condition for  $G(I, J)$  to be an equilibrium network.

**Theorem 4.** *The core-periphery network  $G(I, J)$  is the equilibrium network generated by some perfect Bayesian equilibrium in stationary strategies if and only if (i) every node has outdegree  $m$  or  $m - 1$ , and (ii) for every dealer  $j \in J$ ,  $k_j \geq k(m, d_j)$ .*

Since  $\underline{P}(k, m, d)$  is strictly increasing in  $m$ , I define  $\tilde{m}^*$  to be the largest integer such that

$$P^*(\tilde{m}^*) \geq \min_{k+d=n-1} \underline{P}(k, \tilde{m}^*, d).$$

Then in order for all dealers to have no incentive to gouge, any agent has at most  $m^*$  dealer counterparties. By the same equilibrium selection procedure through dealer entry following [Proposition 3](#), one can focus on the case of maximum dealer competition  $m = m^*$ . It follows from [Theorem 4](#) that the total number of dealers is bounded above.

**Corollary 4.** *If a core-periphery network  $G(I, J) \in \mathcal{G}$  is an equilibrium network generated by some perfect Bayesian equilibrium in stationary strategies, then the total number of dealers is bounded by*

$$|J| \leq \frac{\tilde{m}^* n}{k(\tilde{m}^*, 0) - 1}.$$

To provide a numerical example, I consider a market with  $n = 1000$  agents, with  $\beta = 0.1$ ,  $\pi = 1$ ,  $\lambda = 3$ ,  $\theta_d = 1 - 0.8^d$ ,  $c = 0.09$ , and  $r = 0.1$ . In equilibrium, every non-dealer has  $\tilde{m}^* = 3$  trading accounts, the equilibrium spread is  $P^*(\tilde{m}^*) \simeq 0.1$ , every dealer needs to have at least  $k(\tilde{m}^*, 0) = 166$  non-dealer customers, and the upper bound on the number of dealers is  $|J| \leq 17$ .

## 4 Comparative Statics, Welfare and Policy Implications

To develop comparative statics on the core size, I focus attention on the equilibrium  $\sigma^*(m^*)$  that generates the core-periphery network  $G(m^*)$  with  $m^*$  dealers, as depicted in [Figure 1](#).

**The core size  $m^*$  and the equilibrium spread  $P^*(m^*)$ :**

The next proposition shows how the core size varies as a function of the model parameters  $(n, \beta, \pi, \lambda, \theta, c, r)$ . I will fix all but one parameter and examine how the core size  $m^*$ , as defined in [\(14\)](#), is affected by the remaining parameter. Proofs are given in [Appendix C](#).

**Proposition 5.** *(i) The core size  $m^*$  is weakly increasing in the total number  $n$  of agents.*

- (ii) For  $n$  sufficiently large, the core size reaches a constant  $m_\infty^*$  that is independent of  $n$ .  
The constant  $m_\infty^*$  is the largest integer  $m$  such that

$$\frac{mr\pi}{2\lambda\theta_m + mr} < P^*(m).$$

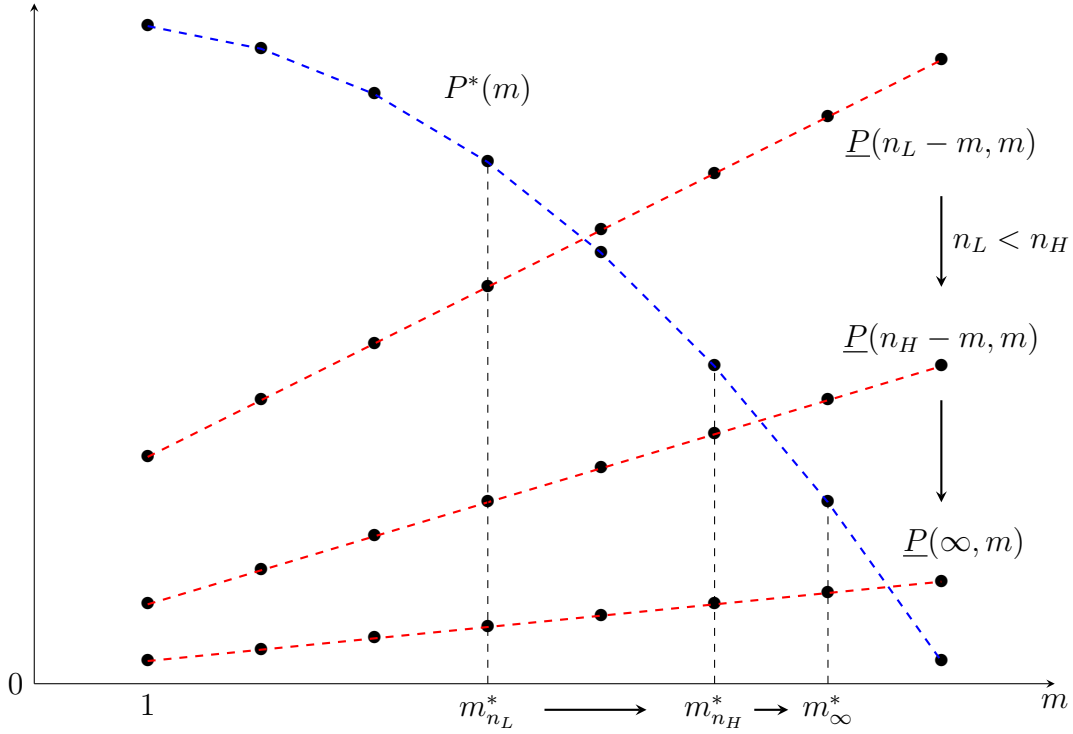
- (iii) The core size  $m^*$  is weakly increasing in the total gain per trade  $\pi$  and the arrival rate  $\lambda$  of trade orders from outside investors, and weakly decreasing in the account maintenance cost  $c$  and the inventory risk coefficient  $\beta$ .
- (iv) The equilibrium spread  $P^*(m^*)$  is weakly increasing in the inventory risk coefficient  $\beta$ , and weakly decreasing in the total number  $n$  of agents.

Part (i) of [Proposition 5](#) has a simple intuitive proof, as follows. Suppose there are  $m$  dealers in the market. As the total number  $n$  of agents increases, each dealer becomes more efficient in inventory management, and thus can support a tighter spread  $\underline{P}(n - m, m)$ . That the dealer-sustainable spread  $\underline{P}(n - m, m)$  is strictly decreasing in  $n$  is a consequence of [Corollary 2](#). In contrast, the equilibrium spread  $P(m)$  does not depend on  $n$ . The core size  $m^*$ , which is the largest integer such that the  $P(m^*) \geq \underline{P}(n - m^*, m^*)$ , is thus weakly increasing in  $n$ . [Figure 10](#) provides a graphical illustration of this result.

Part (ii) says that even for an “infinite” set of investors, one should anticipate only a finite number  $m_\infty^*$  of dealers. A proof is given in [Appendix B](#). To provide a numerical example of the number  $m_\infty^*$  of dealers in a large market, let  $\pi = 1$ ,  $\lambda = 3$ ,  $\theta_d = 1 - 0.8^d$ ,  $c = 0.09$ ,  $s = 0.1$ , and  $r = 0.1$ . Then  $m_\infty^* = 3$ , and the equilibrium spread in the large market is  $P^*(m_\infty^*) \simeq 0.1$ .

As  $\pi$  increases, dealers extract a higher rent per trade (reflected by a wider equilibrium spread  $P^*(m)$ ), but also have stronger incentive to gouge (wider dealer-sustainable spread  $\underline{P}(m)$ ). It will be shown, in [Appendix B](#), that the equilibrium spread  $P^*(m)$  increases more than the dealer-sustainable spread  $\underline{P}(m)$ . The core size  $m^*$  is thus weakly increasing in  $\pi$ .

The parameter  $\lambda$  can be interpreted as a measure of asset liquidity, since it is the rate at which the economy wishes to trade the asset. With a more liquid asset, it is natural that more dealers emerge to facilitate the intermediation of the asset. The negative relationship



**Figure 10** – The core size  $m^*$  is weakly increasing in the total number  $n$  of agents.

between asset liquidity and market concentration predicted by the model is consistent with some empirical evidence in OTC markets. For example, the Herfindahl index is lowest in the interest rate derivatives market, followed by the credit derivatives market and finally the equity derivatives market.<sup>10</sup> Within the foreign exchange derivatives market, the Herfindahl index ranking is, from low to high, USD, EUR, GBP, JPY, CHF, CAD and SEK. The ranking of the outstanding notional amounts of these currencies is almost reversed<sup>11</sup> (with the exception of JPY and GBP, for which the Herfindahl indices are close). As yet another example, [Cetorelli, Hirtle, Morgan, Peristiani, and Santos \(2007\)](#) documented a substantial decline in the market concentration of the credit derivatives market during the 2000-04 period, as “financial institutions have rushed to take part in this exploding market.” However, there is no empirical literature, to my knowledge, that specifically studies the relationship

<sup>10</sup>The statistics are available in the ISDA research notes by [Mengle \(2010\)](#).

<sup>11</sup>The estimates are given in the Semiannual OTC Derivatives Statistics computed by the Bank for International Settlements.



between asset liquidity and market concentration.

If dealers become less risk tolerant (higher  $\beta$ ), they find it more costly to use their own inventories to make markets. As a result, dealers need to be compensated with a wider spread  $\underline{P}(m)$  to continue making a two-way market. On the other hand, the equilibrium spread  $P^*(m)$  is not affected by dealer risk tolerance, as it is determined by an indifference condition for non-dealers. Therefore, the market can only support a smaller core with agents of reduced risk tolerance.

A current hotly debated issue is bond market liquidity. The world's biggest banks are shrinking their bond trading activities to comply with regulations such as the Volcker rule and higher capital requirements imposed after the financial crisis. These restrictions have curbed the ability of banks to build inventory or warehouse risk. On Friday, October 23, 2015, Credit Suisse exited its role as a primary dealer across Europe's bond markets, the latest signal that banks are scaling back bond trading activities. There are also other examples of markets, such as corporate bond markets and currency markets, which are slowly experiencing structural changes due to significant dealer disintermediation. Intermediation in these markets is increasingly agency-based, and many investors report that post-crisis regulatory reforms have reduced market liquidity.

### Dealer inventory levels and turnover, market-wide dealer inventory cost:

Given the model parameters  $(n, \beta, \pi, \lambda, \theta, c, r)$ , the total arrival rate of outside orders is  $n\lambda$ . I fix an equilibrium  $\sigma^*(m)$  for some  $1 \leq m \leq m^*$ , and examine how the dynamic of the dealer inventory process depends on  $n$  and  $\lambda$ . For simplicity of notation, let  $\bar{x}$  denote the dealer inventory level  $\bar{x}_{n-m, m, P^*(m)}$ .

**Proposition 6.** *(i) The dealer inventory threshold level  $\bar{x}$  is weakly decreasing in  $\beta$ . (ii) As  $n\lambda$  increases to infinity, the inventory threshold level  $\bar{x}$  of a given dealer increases to infinity at the speed  $(n\lambda)^{1/3}$ :*

$$\bar{x} = \Theta \left( (n\lambda)^{\frac{1}{3}} \right).$$

Fixing the level of dealer competition, when dealers have reduced ability to warehouse

risk, they optimally shrink their balance sheet to control inventory risk. When the asset is more liquid (either because of a larger rate  $\lambda$  or because of a larger market), dealers expand their inventory size to take advantage of the increased order flow from quote seekers.

The inventory process  $(x_{jt})_{t \geq 0}$  of a given dealer  $j$  is a symmetric continuous-time random walk on  $\{-\bar{x}, \dots, \bar{x}\}$ , with jump intensity  $\vartheta$  (the total rate of Requests for Quote received by dealer  $j$ ). The random walk  $x_{jt}$  loops at the boundary points  $\pm\bar{x}$ . Therefore, the inventory process has a unique stationary distribution that is the uniform distribution on its set of states. It is well known that the mixing time of this process is on the order of  $\bar{x}^2/\vartheta$  as  $\vartheta$  goes to infinity.<sup>12</sup> Therefore, one has the following proposition:

**Proposition 7.** *As  $n\lambda$  goes to infinity, the mixing time  $t_{mix}$  of the inventory process  $(x_{jt})_{t \geq 0}$  of a dealer is on the order of  $(n\lambda)^{-1/3}$ :*

$$t_{mix} = \Theta\left((n\lambda)^{-\frac{1}{3}}\right).$$

The proposition above implies that in a liquid asset market, dealer inventory has quick turnover and exhibits fast mixing. The positive relationship between asset liquidity and the speed of dealer inventory rebalancing, as predicted by the model, is consistent with prior empirical studies. Using data on the actual daily U.S.-dollar inventory held by a major dealer, [Duffie \(2012\)](#) estimates that the “expected half-life” of inventory imbalances is approximately 3 days for the common shares of Apple, versus two weeks for a particular investment-grade corporate bond. The data also reveal substantial cross-sectional heterogeneity across individual equities handled by the same market maker, with the expected half-life of inventory imbalances being the highest for (least liquid) stocks with the highest-bid-ask spreads and the lowest trading volume.

The payoff  $V_{n-m,P}(0)$  of a dealer can be decomposed into her inventory cost and her

---

<sup>12</sup>Examples 5.3.1, 12.10 and Theorem 20.3 of [Levin, Peres, and Wilmer \(2009\)](#) provide background on the mixing time of this process.

gain from trading with quote seekers:

$$V_{n-m,m,P}(0) = -C(n, \lambda, m) + \underbrace{2\lambda \left( (n-m) \frac{\theta_m}{m} + \theta_{m-1} \right)}_{:=\vartheta, \text{ total rate of Requests for Quote}} \frac{P}{r}. \quad (15)$$

For a given dealer, the total rate  $\vartheta$  of Requests for Quote and thus the total trade gain grow linearly with  $n$  and  $\lambda$ . The inventory cost  $C(n, \lambda, m)$ , as shown by the next proposition, grows sublinearly with  $n$  and  $\lambda$ .

**Proposition 8.** *(i) The present value  $C(n, \lambda, m)$  of individual dealer inventory cost is strictly increasing and strictly concave in  $n$  and  $\lambda$ . (ii) As  $n\lambda$  increases to infinity, the inventory cost  $C(n, \lambda, m)$  goes to infinity at the speed  $(n\lambda)^{2/3}$ .*

$$C(n, \lambda, m) = \Theta \left( (n\lambda)^{\frac{2}{3}} \right).$$

*(iii) The market-wide total dealer inventory cost  $mC(n, m, \lambda)$  is strictly increasing in  $m$ .*

In the model setup, every agent has a quadratic inventory holding cost  $\beta x^2$ . However, as the total rate of Requests for Quote increases, a given dealer is more efficient in balancing her inventory by quickly netting trades. The netting effect more than offsets the convexity of the inventory cost function. Indeed, property (i) of [Proposition 8](#) probably holds for any convex inventory cost function  $f(x)$ : if the cost  $f(x)$  increases very fast with the inventory size  $|x|$ , the dealer would optimally reduce her inventory level to control her inventory cost.

Property (iii) follows from the decreasing returns to scale of the individual inventory cost function  $C(n, m, \lambda)$  and Jensen's inequality. To minimize the market-wide dealer inventory cost, it is better to concentrate the provision of immediacy at a smaller set of dealers in order to maximize the netting efficiency.

## Welfare analysis and policy discussion:

From the viewpoint of market efficiency, dealers tend to under-compete in liquid markets, and over-compete in illiquid markets. The next proposition establishes this welfare result, which suggests regulation policies on market makers that are differentiated by asset liquidity.

Given a set of model parameters be  $(n, \beta, \pi, \lambda, \theta, c, r)$  and a strategy profile  $\sigma$ , define welfare  $U(\sigma)$  to be the sum of all agents' net payoffs. For any given integer  $m$ , let  $\sigma^{**}(m)$  be an equilibrium with  $m$  dealers in which the dealers are somehow *imposed* to offer the equilibrium spread  $P^*(m)$ . That is, even if dealers have incentive to gouge, they are not allowed to. For  $m \leq m^*$ , since dealers have no incentive to gouge, the strategy profile  $\sigma^{**}(m)$  is identical to  $\sigma^*(m)$ . Let

$$m^{**} = \operatorname{argmax}_{m \geq 0} U(\sigma^{**}(m)).$$

The first result of the following proposition shows that the welfare associated with the strategy profile  $\sigma^{**}(m)$  is increasing in  $m$  as long as the equilibrium spread  $P^*(m)$  is positive. Let  $\bar{m}$  be the largest integer such that

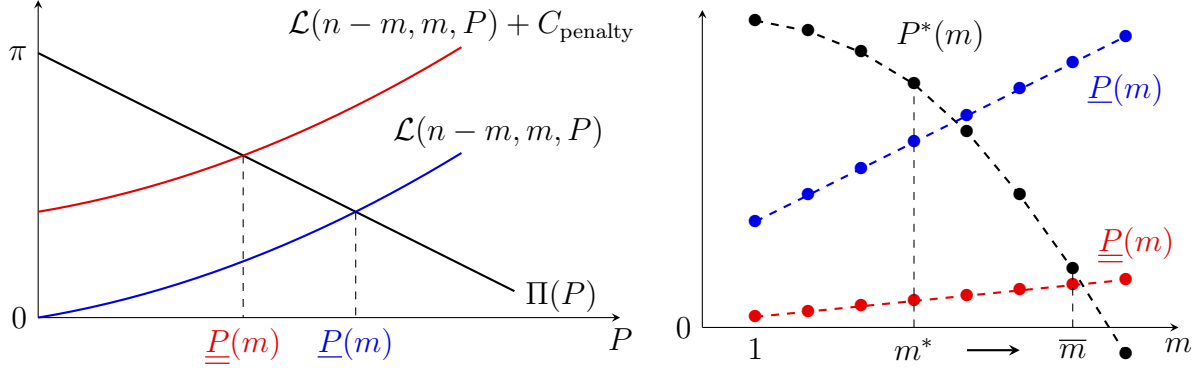
$$P^*(\bar{m}) > 0.$$

**Proposition 9.** *(i) If  $n\lambda$  is sufficiently large, the total welfare  $U(\sigma^{**}(m))$  is strictly increasing in the number  $m$  of core dealers, for  $0 \leq m \leq \bar{m}$ . (ii) Under certain parameter conditions that reduce  $n\lambda$ ,  $m^{**} < m^*$ .*

With a larger core, the increased density of the equilibrium trading network  $G(m)$  results in a higher total cost associated with maintaining trading accounts. However, the first welfare result (i) suggests that the benefit of a market with better asset intermediation outweighs the link cost. This welfare result underlines the importance of efficient asset intermediation and the role of the dealer sector in acting as intermediaries.

Since the dealer-sustainable spread  $\underline{P}(m)$  is always positive, it is clear from [Figure 8](#) that  $m^* \leq \bar{m}$ . In other words, dealers tend to under-compete in a liquid market. This is because dealers' private incentive to gouge limits the scope of price competition, which has a negative externality on all non-dealers. The externality can be eliminated by regulations that aim to discourage dealers from gouging. For example, regulators can impose a penalty on market makers for each trading account that is terminated by one of their quote seekers. In the equilibrium analysis, every dealer would internalize the penalty cost  $C_{\text{penalty}}$  into her

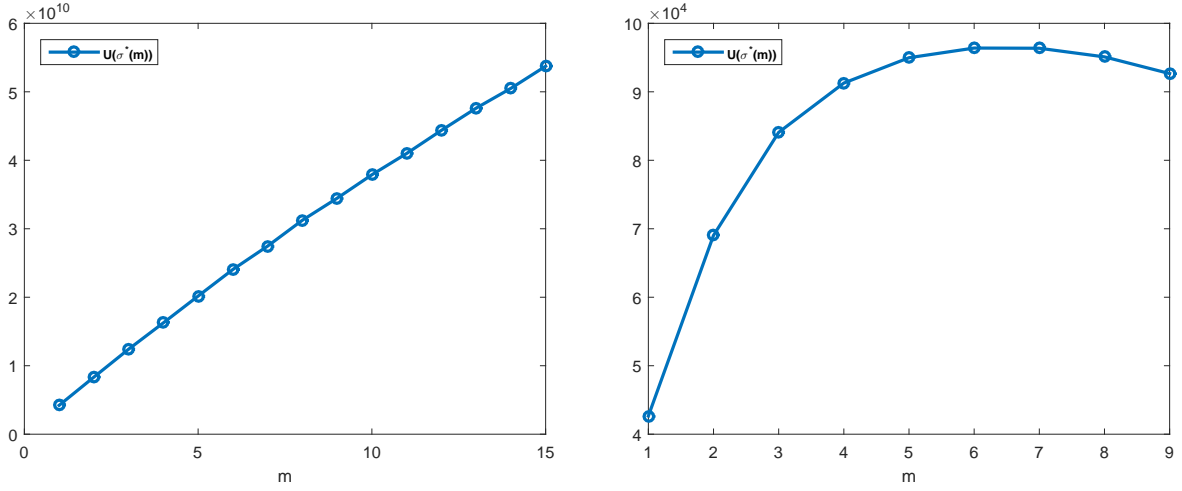
loss function  $\mathcal{L}(n - m, m, P)$  from gouging, and can thus sustain a tighter spread  $\underline{\underline{P}}(m)$ . Such a penalty cost on dealers, surprisingly, encourages dealer entry. This is because the penalty creates room for greater dealer competition by improving the commitment power of a dealer. By choosing an appropriate cost value  $C_{\text{penalty}}$ , the regulator can achieve the socially optimal level of market concentration. **Figure 11** provides a graphical illustration of the effect of such a penalty cost  $C_{\text{penalty}}$ .



**Figure 11** – Introducing a penalty cost  $C_{\text{penalty}}$  on gouging increases dealer competition.

Quite surprisingly, the second welfare result (ii) shows that dealers tend to over-compete in an illiquid market. The reason is as follows: to minimize the market-wide inventory cost  $mC(n - m, m, \lambda)$ , it is better to concentrate the provision of immediacy at a smaller set of dealers in order to maximize the netting efficiency. However, each individual dealer does not internalize her negative externality on the market-wide inventory cost. This source of externality is less pronounced in a liquid market, since the market-wide inventory cost (which is on the order of  $(n\lambda)^{2/3}$ ) is inconsequential to the social welfare (on the order of  $n\lambda$ ) for  $n\lambda$  sufficiently large. In an illiquid market, however, the inventory cost can affect the welfare in a way that favors smaller core size and larger market concentration. **Figure 12** provides a numerical illustration of the welfare  $U(\sigma^*(m))$  as a function of  $m$  in a liquid (left) and an illiquid market (right), respectively.

The welfare results of **Proposition 9** suggest regulation policies that treat market making activities differently according to the liquidity of the underlying asset. It is welfare improving



**Figure 12** – The welfare  $U(\sigma^*(m))$  for  $m = 1, 2, \dots, m^*$ .

to encourage dealer participation in liquid markets, and discourage such in illiquid ones. The current regulatory capital requirement adopted by Basel III uses risk-weighted assets as the denominator of the capital ratio of a bank. This approach can be improved, for example, by adding a liquidity component in the weight calculation.

Of course, excessive risk taking by dealers increases their probability of default. One of the main reasons for regulators to introduce the Volcker rule and more stringent capital requirements is to discipline dealers' risk taking behavior, in order to reduce their default risk and mitigate the impact of a dealer default on the rest of the market. However, these policies may also have a long term impact on market structure and adversely affect market liquidity. These consequences should not be ignored. This paper does not consider dealer default risk or contagion risk in the financial network, and therefore [Propositions 5](#) and [9](#) are not sufficient to provide a complete policy assessment for the Volcker rule or capital requirements. However, it offers an analytic framework from the perspective of intermediation efficiency, and identifies two sources of externalities related to dealer competition for the analysis of these regulation policies on market makers.

Non-bank firms such as fund managers are holding more bonds and starting to act as liquidity providers in daily trading. However, many question whether these buy-side firms

can substitute for dealer liquidity by taking an effective role as market makers. The model outlined in this paper highlights the importance of having a large customer base for an agent to efficiently manage inventory and make markets. Non-bank firms such as fund managers are not naturally liquidity hubs. They do not have the same number of trading lines and global customer base that banks have. Being in a central network position is essential for a financial institution to be able to “lean against the wind” — that is, to provide liquidity during financial disruptions. It is worrisome that the liquidity that non-bank firms are providing may be “illusory”: these firms may be unable or unwilling to absorb external selling pressure in a selloff, and liquidity may vanish. However, this paper does not cover this topic. A different model is needed to illustrate liquidity illusion.

## 5 Inter-Dealer Trading through Nash Bargaining

In practice, dealers form a completely connected core and trade with each other actively to share their inventory risk accumulated when trading with non-dealers. The protocol for bilateral inter-dealer trading usually features Nash bargaining of some sort, rather than a request-for-quote by one counterparty and a take-it-or-leave-it offer by the other. When dealers are allowed to trade larger quantities with each other at a bilaterally negotiated price, they are able to extract gains from trade more efficiently. In this section, I illustrate this intuition through a variant of the symmetric-agent model, in which dealers are allowed to conduct Nash Bargaining when trading with each other. The model no longer preserves agent symmetry, as a subset of agents (dealers) have access to a trading protocol featuring Nash bargaining that is not available to other agents (non-dealers). In practice, dealers resist the participation of non-dealers into the inter-dealer segment, and accuse non-dealers of taking liquidity without exposing themselves to the risks of providing liquidity. Others criticize dealers for trying to prevent competition that would compress bid-ask spreads in the market.

The structure of the trading game to be considered is similar to that of the symmetric-agent model. A common non-divisible asset with 0 expected payments is traded over the

counter by  $n$  agents. Every agent has 0 initial endowment of the asset, and is subject to an instantaneous quadratic inventory cost of  $-\beta x^2$  when holding an asset inventory of size  $x$ . The time discount rate is  $r$ . The total present value of a given agent is given by (2). Every agent serves outside investors, who submit trade orders at the mean arrival rate  $2\lambda$  and pay a fixed premium  $\pi > 0$ , as described in the symmetric-agent model of Section 2. At time  $t = 0$ , agents open trading accounts, which are subject to termination at any time  $t > 0$ . The cost of maintaining an account is  $c$  per unit of time.

All model assumptions made so far are identical to the symmetric-agent model. I next introduce agent asymmetry. There is a partition of agents  $I \cup J = N$  with  $J = m$ . Agents in  $J$ , called “dealers,” give each other Nash bargaining rights. Specifically, if a pair of dealers  $j_1, j_2 \in J$  have trading accounts with each other, they can bilaterally negotiate the terms of trade according to Nash bargaining when they meet. In the current context of spot trading, terms of trade simply consist of price and quantity. Agents in  $I$ , called “non-dealers,” have no access to this inter-dealer trading protocol. That is, even if a non-dealer  $i \in I$  and another agent (dealer or non-dealer) have trading accounts with each other, the pair can only trade following the request-for-quote protocol upon contact.

The search technology also differs between dealer-to-dealer and dealer-to-non-dealer encounters. The probability of a successful search among  $d$  dealers is given by  $\theta_d$ , where  $\theta_d \in [0, 1)$  is strictly increasing and strictly concave in  $d$ , with  $\theta(0) = 0$ . This is the same search technology in the symmetric-agent model. Encounters between pairs of dealers are based on independent random matching, with pair-wise meeting intensity  $\zeta$ .

The information structure is largely identical to that of the symmetric-agent model. In particular, the model maintains quote offer privacy and inventory privacy for trading between dealers and non-dealers. However, there are two differences: (i) the assumption of quote seeker anonymity is relaxed. That is, when a non-dealer requests quotes from another agent (dealer or non-dealer), the identity of the quote seeker is revealed to the quote provider. (ii) When two dealers meet, they observe each others’ inventory in order to conduct Nash bargaining. It would be desirable to maintain inventory privacy for all agents. However, bargaining with incomplete information is a notoriously hard problem in the literature of



economics. I do not explore this direction, and instead assume complete information when it comes to inter-dealer bargaining, so that the paper maintains focus on core-periphery network formation and trading.

The same solution concept of perfect Bayesian equilibrium in stationary strategies, as defined in [Section 2](#), is applied to this network trading game. I focus on large markets when solving this model. This restriction is not necessary when solving the symmetric-agent model.

Formally, consider a core-periphery network defined as follows: The network has nodes given by the  $n$  agents, which are partitioned into some  $I \cup J = N$  with  $|J| = m$  for some  $0 \leq m \leq n$ . Each node  $i \in I$ , representing a “non-dealer,” has a directed link to all the  $m$  nodes in  $J$ , while each node  $j \in J$ , representing a “dealer,” has a bidirectional link with all other  $m - 1$  nodes in  $J$ . I let  $G_n(m)$  denote such a network.

**Theorem 5.** *For a given set of model parameters  $(n, m, \beta, \pi, \lambda, \theta, c, r)$ , suppose  $m^*$  is defined as in [\(14\)](#). There is some constant  $n_0$ , such that if  $n > n_0$ , the following is true: given some integer  $m \geq 0$ , the network  $G_n(m)$  is the equilibrium trading network generated by some symmetric perfect Bayesian equilibrium in stationary strategies if and only if  $0 \leq m \leq m^*$ .*

## 6 Concluding Remarks

Extensive empirical work has shown that core-periphery networks dominate conventional OTC markets. However, few (though an increasing number of) theoretical foundations have been provided. Existing literature exploits some ex-ante heterogeneity of agents to explain the ex-post difference of their network positions. The symmetric-agent model of this paper is original in its ability to provide a separation of core from periphery agents that is determined solely by endogenous forces that tend to concentrate the provision of immediacy. The equilibrium level of dealer entry (the size of the core) depends on the combined effect of two countervailing forces: *(i)* network competition among dealers in their pricing of immediacy to peripheral agents, and *(ii)* the benefits of a concentrated set of dealers for lowering inventory risk through their ability to quickly offset purchases and sales.

Although financial institutions do not have identical characteristics in real OTC markets, the equilibrium results implied by the symmetric-agent model highlight the important role of these two economic forces in the formation of market structure.

One useful direction of future research would be the introduction of agent heterogeneity to study the relationship between dealer centrality and the pricing of immediacy. Recent empirical work suggests that the price-centrality relationship changes across different markets. In the municipal bond market, central intermediaries earn higher markups compared with peripheral intermediaries.<sup>13</sup> The opposite is true in the market for asset-backed securities.<sup>14</sup>

Another feature of the model is that dealer profits only come from earning bid-ask spread through frequent intermediation of the asset, and not from their idiosyncratic taste for the asset. This echoes the finding by [Farboodi, Jarosh, and Shimer \(2015\)](#) that fast intermediation dampens the effect of idiosyncratic taste on a dealer’s net valuation of the asset.

From a welfare viewpoint, the model identifies two sources of externalities: (1) dealers’ private incentive to gouge, which limits the scope of price competition, and (2) the negative externality of each individual dealer on the overall netting efficiency. The first externality dominates in a liquid market, and leads to insufficient dealer competition. Introducing an additional penalty cost on market makers for “each lost customer” deters dealers from gouging, improves their commitment power, and therefore creates room for greater dealer competition. The second externality is more pronounced in an illiquid market, and results in a higher than socially optimal level of dealer competition. The model suggests regulation policies that treat market makers differently according to asset liquidity through, for example, the introduction of a “liquidity weight,” on top of the currently used “risk weight.”

---

<sup>13</sup>Li and Schürhoff (2014) provide evidence for the municipal bond market.

<sup>14</sup>See Hollifield, Neklyudov, and Spatt (2014) provide evidence for the market of asset-backed securities.

# Appendices

## A A Symmetric Tri-Diagonal Matrix

To prove the results stated in the main text, it is useful to first establish some properties about the inverse of a symmetric tri-diagonal matrix.

Let  $n$  be a strictly positive integer. A vector  $\varphi$  of length  $n$  is said to be *U-shaped* if

$$\begin{aligned}\varphi_i &= \varphi_{n+1-i}, & \forall 1 \leq i \leq n, \\ \varphi_1 &> \varphi_2 > \cdots > \varphi_m.\end{aligned}$$

Given two vectors  $\varphi$  and  $\psi$  of the same length, I will write  $\varphi < \psi$  if  $\varphi$  is strictly less than  $\psi$  entry-wise. Given a constant  $\zeta > 1$ , let  $A$  be the following tri-diagonal matrix of size  $n \times n$ :

$$A = \begin{pmatrix} \zeta - \frac{1}{2} & -\frac{1}{2} & & & & \\ -\frac{1}{2} & \zeta & -\frac{1}{2} & & & \\ & -\frac{1}{2} & \zeta & -\frac{1}{2} & & \\ & & \ddots & \ddots & \ddots & \\ & & & -\frac{1}{2} & \zeta & -\frac{1}{2} \\ & & & & -\frac{1}{2} & \zeta - \frac{1}{2} \end{pmatrix} \quad (16)$$

**Lemma 5.** *The matrix  $A$  is invertible, and  $M \equiv A^{-1}$  satisfies the following properties:*

- (i) *The matrix  $M$  is symmetric.*
- (ii) *All entries of the matrix  $M$  are strictly positive.*
- (iii) *For every  $1 \leq i \leq n$ ,*

$$\begin{aligned}M_{i,1} &< M_{i,2} < \cdots < M_{i,i}, \\ M_{i,i} &> M_{i,i+1} < \cdots < M_{i,n}.\end{aligned}$$

(iv) For every  $i \neq j$ ,

$$M_{i,j-1} + M_{i,j+1} > 2M_{i,j}.$$

(v) For every  $i, j$ ,  $M_{i,j} = M_{n+1-i, n+1-j}$ .

(vi) For every vector  $y$  that is U-shaped,  $My$  is U-shaped and  $(My)_n \leq y_n/(\zeta - 1)$ .

(vii) Let  $m = \lfloor (n+1)/2 \rfloor$ . For every  $i \leq m$ ,  $M_{i,1}^2 \geq M_{n+1-i,1}^2$  and

$$M_{m,1}^2 - M_{n+1-m,1}^2 < \frac{2}{\zeta - 1}.$$

*Proof.* Since  $\zeta > 1$ , the matrix  $A$  is diagonally dominant thus invertible. Its inverse  $M$  is symmetric since  $A$  is. The matrix  $A$  can be written as  $A = cI - B/2$ , where

$$B = \begin{pmatrix} 1 & 1 & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & 1 & 1 \\ & & & & & 1 & 1 \end{pmatrix} \quad (17)$$

The sup-norm of the matrix  $B$  is  $\|B\|_\infty = 2$ . Since  $\zeta > 1$ , one can write the following expansion for  $M$ :

$$M = A^{-1} = \zeta^{-1} \left( I - \frac{B}{2\zeta} \right)^{-1} = \zeta^{-1} \underbrace{\left[ I + \frac{B}{2\zeta} + \left( \frac{B}{2\zeta} \right)^2 + \dots \right]}_{\equiv S}. \quad (18)$$

Thus, to show that  $M$  satisfies properties (ii)-(vi) stated in [Lemma 5](#) is equivalent to show that  $S$  satisfies the same properties. One can see that all entries of  $S$  are strictly positive directly from the expansion (18) of  $S$ .

*Properties (iii) and (iv).* One has

$$S \frac{B}{2\zeta} = S - I.$$

Then for every  $i > 1$ ,

$$\frac{S_{i,1} + S_{i,2}}{2\zeta} = S_{i,1} > \frac{S_{i,1}}{\zeta} \implies S_{i,1} < S_{i,2}.$$

Suppose  $S_{i,j-1} < S_{i,j}$  for some  $j \in (1, i)$ , then

$$\frac{S_{i,j-1} + S_{i,j+1}}{2\zeta} = S_{i,j} > \frac{S_{i,j}}{\zeta} \implies S_{i,j} < S_{i,j+1}.$$

Then by induction, one has  $S_{i,j} < S_{i,j+1}$  for every  $i > 1$  and  $j < i$ . Similarly, one can show that  $S_{i,j} < S_{i,j-1}$  for every  $i < n$  and  $j > i$ . Therefore, for every  $i$ ,

$$\begin{aligned} S_{i,1} &< S_{i,2} < \cdots < S_{i,i}, \\ S_{i,i} &> S_{i,i+1} < \cdots < S_{i,n}, \end{aligned} \tag{19}$$

$$S_{i,j-1} + S_{i,j+1} > 2S_{i,j}, \quad \forall i \neq j. \tag{20}$$

*Property (v).* It is clear that  $B_{i,j} = B_{n+1-i,n+1-j}$  for every  $i, j$ . Suppose  $B_{i,j}^\ell = B_{n+1-i,n+1-j}^\ell$  for every  $i, j$ . Then

$$B_{i,j}^{\ell+1} = \sum_k B_{i,k} (B^\ell)_{k,j} = \sum_k B_{n+1-i,n+1-k} (B^\ell)_{n+1-k,n+1-j} = B_{n+1-i,n+1-j}^{\ell+1}.$$

Therefore, for every  $i, j$ ,

$$S_{i,j} = S_{n+1-i,n+1-j}.$$

*Property (vi).* Given a U-shaped vector  $y$ , then for every  $i$ ,

$$(Sy)_i = \sum_j S_{i,j} y_j = \sum_j S_{n+1-i,n+1-j} y_{n+1-j} = (Sy)_{n+1-i}.$$

For every  $0 \leq k \leq m$ , let

$$w(k) = \left( \underbrace{1, \dots, 1}_k, \underbrace{0, \dots, 0}_{(n-2k)}, \underbrace{1, \dots, 1}_k \right)^\top.$$

Any U-shaped vector  $y$  can be written as a linear combination of the vectors  $w(k)$  with strictly positive weights. Thus, to show that  $Sy$  is U-shaped for any U-shaped vectors  $y$ , it is sufficient to show that  $Sw(k)$  is U-shaped for every  $0 \leq k \leq m$ . For every  $k \leq i < m$ ,

$$\begin{aligned} [Sw(k)]_{i+1} &= \sum_{j \leq k} S_{i+1,j} + \sum_{j \geq n-k+1} S_{i+1,j} \\ &= \sum_{j \leq k} (S_{j,i+1} + S_{j,n-i}) < \sum_{j \leq k} (S_{j,i} + S_{j,n-i+1}) = [Sw(k)]_i. \end{aligned}$$

The strict inequality above follows from (20). Let  $e = (1, \dots, 1)^\top$ . Then  $Ae = (\zeta - 1)e$  and thus  $Se = \zeta/(\zeta - 1)e$ . Then for every  $i < k$ ,

$$[Sw(k)]_{i+1} = \frac{\zeta}{\zeta - 1} - \sum_{k+1 \leq j \leq n-k} S_{j,i+1} < \frac{\zeta}{\zeta - 1} - \sum_{k+1 \leq j \leq n-k} S_{j,i} = [Sw(k)]_i.$$

The strict inequality above follows from (19). Thus  $[Sw(k)]_{i+1} < [Sw(k)]_i$  for every  $i < m$ . Therefore,  $Sw(k)$  is U-shaped.

Since  $Ae = (\zeta - 1)e$ ,  $\zeta - 1$  is an eigenvalue of  $A$  with the eigenvector  $e$ . The inverse  $M$  thus has the eigenvalue  $1/(\zeta - 1)$  with the same eigenvector  $e$ . Given a U-shaped vector  $y$ ,

$$(My)_n = \sum_j M_{n,j} y_j \leq \sum_j M_{n,j} y_n = y_n (Me)_n = \frac{1}{\zeta - 1} y_n.$$

*Property (vii).* Let  $H = -2A$  and  $W = H^{-1}$ . Then property (vii) is equivalent to  $W_{i,1}^2 \geq W_{n+1-i,1}^2$  for every  $i \leq m$ , and  $W_{m,1}^2 - W_{n+1-m,1}^2 < 1/(2(\zeta - 1))$ . There is nothing to show when  $n = 1$ . If  $n = 2$ , one has

$$\begin{aligned} W &= \frac{1}{4\zeta(1 - \zeta)} \begin{pmatrix} 1 - 2\zeta & -1 \\ -1 & 1 - 2\zeta \end{pmatrix} \\ 0 &< W_{1,1}^2 - W_{2,1}^2 = \frac{1}{4\zeta^2} < \frac{1}{2(\zeta - 1)}. \end{aligned}$$

If  $n > 2$ , define the second-order linear recurrences

$$z_k = -2\zeta z_{k-1} - z_{k-2}, \quad k = 2, 3, \dots, n-1$$

where  $z_0 = 0$ ,  $z_1 = 1 - 2\zeta$ . Let  $\zeta = \cosh \gamma$  where  $\gamma > 0$ . Then it follows by induction that

$$z_k = (-1)^k \frac{\cosh \left( \left( k + \frac{1}{2} \right) \gamma \right)}{\cosh \frac{\gamma}{2}} \quad k = 0, 1, \dots, n-1.$$

It has been shown by [Huang and McColl \(1997\)](#) that the inverse matrix  $W = H^{-1}$  can be expressed as

$$W_{j,j} = \frac{1}{-2\zeta - \frac{z_{j-2}}{z_{j-1}} - \frac{z_{n-j-1}}{z_{n-j}}}$$

where  $j = 2, \dots, n-1$ , and

$$W_{i,j} = \begin{cases} (-1)^{j-i} \frac{z_{i-1}}{z_{j-1}} W_{j,j} & i < j \\ (-1)^{j-i} \frac{z_{n-i}}{z_{n-j}} W_{j,j} & i > j. \end{cases}$$

I first simplify the expression of  $W_{j,j}$  ( $j = 2, \dots, n-1$ ):

$$\begin{aligned} W_{j,j} &= \frac{1}{-2 \cosh \gamma + \frac{\cosh \left( \left( j - \frac{3}{2} \right) \gamma \right)}{\cosh \left( \left( j - \frac{1}{2} \right) \gamma \right)} + \frac{\cosh \left( \left( n - j - \frac{1}{2} \right) \gamma \right)}{\cosh \left( \left( n - j + \frac{1}{2} \right) \gamma \right)}} \\ &= - \frac{\cosh \left( \left( j - \frac{1}{2} \right) \gamma \right) \cosh \left( \left( n - j + \frac{1}{2} \right) \gamma \right)}{\sinh(n\gamma) \sinh \gamma} \end{aligned}$$

I next show that for every  $i, j \leq m$ ,

$$W_{i,j} \geq W_{n+1-i,j}.$$

The inequality above holds if  $i \geq j$  as a consequence of (19). If  $i < j$ , one can verify that

$$\frac{\cosh \left( \left( i - \frac{1}{2} \right) \gamma \right)}{\cosh \left( \left( j - \frac{1}{2} \right) \gamma \right)} - \frac{\cosh \left( \left( i + \frac{1}{2} \right) \gamma \right)}{\cosh \left( \left( n - j + \frac{1}{2} \right) \gamma \right)} > 0.$$

It then follows that

$$W_{i,j} - W_{n+1-i,j} = \left( \frac{\cosh \left( \left( i - \frac{1}{2} \right) \gamma \right)}{\cosh \left( \left( j - \frac{1}{2} \right) \gamma \right)} - \frac{\cosh \left( \left( i + \frac{1}{2} \right) \gamma \right)}{\cosh \left( \left( n - j + \frac{1}{2} \right) \gamma \right)} \right) W_{j,j} > 0.$$

If  $i \leq m$  and  $j > m$ , then

$$W_{i,j} - W_{n+1-i,j} = W_{n+1-i,n+1-j} - W_{i,n+1-j} \leq 0.$$

For every  $i \leq m$ ,

$$\begin{aligned} 2(W_{i,1}^2 - W_{n+1-i,1}^2) &= 2 \sum_j (W_{i,j} - W_{n+1-i,j}) W_{j,1} \\ &= \sum_j (W_{i,j} - W_{n+1-i,j}) W_{j,1} + \sum_j (W_{n+1-i,n+1-j} - W_{i,n+1-j}) W_{n+1-j,n} \\ &= \sum_j (W_{i,j} - W_{n+1-i,j}) (W_{j,1} - W_{j,n}) \geq 0. \end{aligned}$$

Hence,  $W_{i,1}^2 \geq W_{n+1-i,1}^2$ . I next show that  $|W_{i,m} - W_{n+1-i,m}| < 1$  for every  $i \leq m$ :

$$\begin{aligned} |W_{i,m} - W_{n+1-i,m}| &= \left| \frac{z_{i-1}}{z_{m-1}} - \frac{z_{i-1}}{z_m} \right| |W_{m,m}| \\ &= \cosh \left( \left( i - \frac{1}{2} \right) \gamma \right) \left| \frac{1}{\cosh \left( (m - \frac{1}{2}) \gamma \right)} - \frac{1}{\cosh \left( (m + \frac{1}{2}) \gamma \right)} \right| |W_{m,m}| \\ &= \cosh \left( \left( i - \frac{1}{2} \right) \gamma \right) \frac{2 \sinh(m\gamma) \sinh \frac{\gamma}{2}}{\sinh(n\gamma) \sinh \gamma} \\ &< \cosh(m\gamma) \frac{2 \sinh(m\gamma) \sinh \frac{\gamma}{2}}{\sinh(n\gamma) \sinh \gamma} = \frac{\sinh \frac{\gamma}{2}}{\sinh \gamma} < 1. \end{aligned}$$

Since  $2(1 - \zeta)$  is an eigenvalue of  $H$  with the eigenvector  $e$ , then  $1/2(1 - \zeta)$  is an eigenvalue of  $W$  with the same eigenvector  $e$ . Thus,  $\sum_k W_{1,k} = 1/(2(1 - \zeta))$ . It then follows that

$$\begin{aligned} W_{1,m}^2 - W_{1,m+1}^2 &= \sum_k W_{1,k} W_{k,m} - \sum_k W_{1,k} W_{k,m+1} \\ &= \sum_k W_{1,k} (W_{k,m} - W_{n+1-k,m}) < \sum_k W_{1,k} = \frac{1}{2(1 - \zeta)}. \end{aligned} \quad \square$$



## B Proofs from Section 3

### B.1 Proof of Proposition 1

It follows from the expression (5) that the equilibrium spread  $P^*(m)$  is strictly decreasing in the number  $m$  of dealers. When  $m = 1$ , the spread  $P^*(1)$  satisfies the indifference condition

$$\Phi_{1,P^*(1)} = \Phi_{0,P^*(1)}.$$

When a non-dealer is isolated, he does not search or trade at all. The spread  $P$  is thus payoff-irrelevant for an isolated agent, and the total payoff of the agent is  $\Phi_{0,P} = 0$  for every  $P \in \mathbb{R}$ . Therefore,  $\Phi_{1,P^*(1)} = 0$ . That is, when there is only one dealer, the equilibrium spread  $P^*(1)$  is the monopoly price that extracts all rent from non-dealers.

### B.2 Proofs of Lemma 2 and proposition 2

To prove Lemma 2 and proposition 2, I first establish the next lemma.

**Lemma 6.** *Suppose  $T_1$  is an operator that maps a function  $V : \mathbb{Z} \mapsto \mathbb{R}$  to another function in the same space, defined as*

$$T_1(V)(x) = \max\{V(x-1) + a, V(x)\}$$

*for every  $x \in \mathbb{Z}$ , where  $a \in \mathbb{R}$  is a constant. Then  $T_1$  preserves concavity. Likewise, if  $T_2$  is defined as*

$$T_2(V)(x) = \max\{V(x+1) - b, V(x)\}$$

*for every  $x \in \mathbb{Z}$ , then  $T_2$  also preserves concavity.*

*Proof.* Suppose  $V$  is a concave function from  $\mathbb{Z}$  to  $\mathbb{R}$ . Then for every  $x \in \mathbb{Z}$ ,

$$\begin{aligned} T_1(V)(x-1) - T_1(V)(x) &\leq \max\{V(x-2) - V(x-1), V(x-1) - V(x)\} \\ &\leq \min\{V(x-1) - V(x), V(x) - V(x+1)\} \\ &\leq T_1(V)(x) - T_1(V)(x+1). \end{aligned}$$

Therefore,  $T_1$  preserves concavity. The same property holds for  $T_2$ .  $\square$

*Proof of Lemma 2.* Let

$$\vartheta = 2\lambda \left( k \frac{\theta_m}{m} + \theta_{m-1} \right)$$

be the total rate of Requests for Quote, and let  $V_{\vartheta,P}$  denote  $V_{k,m,P}$ . I first write the HJB equations (6) into the following form: for every  $x \in \mathbb{Z}$ ,

$$\begin{aligned} V_{\vartheta,P}(x) &= B_{\vartheta,P,\beta}(V_{\vartheta,P})(x) \\ &\equiv \frac{1}{r + \vartheta} \left( -\beta x^2 + \frac{\vartheta}{2} \max\{V_{\vartheta,P}(x+1) + P, V_{\vartheta,P}(x)\} \right. \\ &\quad \left. + \frac{\vartheta}{2} \max\{V_{\vartheta,P}(x-1) + P, V_{\vartheta,P}(x)\} \right) \end{aligned} \quad (21)$$

With a slight abuse of notation, I sometimes write  $B_{\vartheta}$  or simply  $B$  for the Bellman operator  $B_{\vartheta,P,\beta}$ , and  $V_{\vartheta}$  for  $V_{\vartheta,P}$  whenever there is no ambiguity. Given two functions  $f, g$  from  $\mathbb{Z}$  to  $\mathbb{R}$ , I write  $f \leq g$  if  $f(x) \leq g(x)$  for every  $x \in \mathbb{Z}$ . Let  $\Theta$  be the set of functions from  $\mathbb{Z}$  to  $\mathbb{R}$  bounded above that are also bounded below by  $-\beta x^2/r$ :

$$\Theta \equiv \left\{ f : \mathbb{Z} \rightarrow \mathbb{R} : \forall x, f(x) \geq -\frac{\beta}{r}x^2 \text{ and } f \leq \bar{f} \text{ for some constant } \bar{f} \in \mathbb{R} \right\}.$$

It follows from the HJB equations (6) that the value function  $V_{\vartheta,P}$  is in  $\Theta$ . Let  $\Theta$  be equipped with the metric  $\varrho$  defined as

$$\varrho(f, g) = \sup_{x \in \mathbb{Z}} \frac{|f(x) - g(x)|}{\frac{\beta}{r}x^2 + 1}. \quad (22)$$

The space  $(\Theta, \varrho)$  is a complete metric space.

The Bellman operator  $B$  maps the set  $\Theta$  into itself and satisfies the following properties:

- (monotonicity) Given two functions  $V_1, V_2 \in \Theta$ , if  $V_1 \leq V_2$ , then  $B(V_1) \leq B(V_2)$ .
- (discounting) For every  $V \in \Theta$ ,  $\xi \geq 0$  and  $x \in \mathbb{Z}$ ,  $B(V + \xi) \leq B(V) + \delta \xi$ , where

$$\delta = \frac{\vartheta}{r + \vartheta} < 1. \quad (23)$$

The Bellman operator  $B$  is a contraction on  $\Theta$  as it satisfies *Blackwell's sufficient conditions*. By the Contraction Mapping Theorem, the operator  $B$  admits a unique fixed point in  $\Theta$ , which is the value function  $V_{\vartheta,P}$ .

Let  $T$  be an operator on  $\mathbb{R}^{\mathbb{Z}}$  defined by

$$T(V)(x) = \frac{\vartheta}{2} [\max\{V(x+1) + P, V(x)\} + \max\{V(x-1) + P, V(x)\}]$$

It follows from [Lemma 6](#) that  $T$  preserves concavity. Let  $V^0(x) = -\beta x^2/r$  for every  $x \in \mathbb{Z}$ . Then  $V^0 \in \Theta$ . For every  $h \geq 1$ , let  $V^h = B^h(V^0)$ . As  $h \rightarrow \infty$ , one has

$$\varrho(V^h, V_{\vartheta,P}) \rightarrow 0,$$

which implies that  $V^h$  converges to  $V_{\vartheta,P}$  pointwise.

It follows by induction that for every  $h \geq 0$ ,  $V^h$  is even and strictly concave, with

$$V^h(x+1) + V^h(x-1) - 2V^h(x) \leq -\frac{2\beta}{r+\vartheta}. \quad (24)$$

Letting  $h \rightarrow \infty$ , one obtains that  $V_{\vartheta,P}$  is even and strictly concave.  $\square$

*Proof of [Proposition 2](#).* In [Proposition 2](#), the only statement that needs to be proved is that the inventory threshold level  $\bar{x}_{k,m,P}$  is finite. Let  $\bar{x}_{\vartheta,P}$  denote  $\bar{x}_{k,m,P}$ . Letting  $h \rightarrow \infty$  in (24), one has, for every  $x \geq 0$ ,

$$\Delta(x) = V_{\vartheta,P}(x) - V_{\vartheta,P}(x+1) \geq \frac{(2x+1)\beta}{r+\vartheta}$$

The function  $\Delta$  is strictly increasing (since the value function  $V_{\vartheta,P}$  is strictly concave), and by [Lemma 1](#) the inventory threshold level  $\bar{x}_{\vartheta,P}$  is such that

$$\Delta(\bar{x}_{\vartheta,P} - 1) \leq P, \quad \Delta(\bar{x}_{\vartheta,P}) > P.$$

Thus,  $\bar{x}_{\vartheta,P} < \infty$ .  $\square$

### B.3 Proof of Lemma 3

In the next lemma, I write  $V_{\vartheta,P,\beta}$  for  $V_{\vartheta,P}$  to make clear the dependence on  $\beta$ .

**Lemma 7.** *For every  $x \in \mathbb{Z}$ ,  $V_{\vartheta,P,\beta}(x)$  is jointly continuous in  $(\vartheta, P, \beta) \in \mathbb{R}^3$ .*

*Proof.* First, if  $0 \leq \vartheta_1 \leq \vartheta_2$ , then  $V_{\vartheta_1,P,\beta} \leq V_{\vartheta_2,P,\beta}$ . This is because  $B_{\vartheta_2}^\ell(V_{\vartheta_1,P,\beta})$  converges to  $V_{\vartheta_2,P,\beta}$  pointwise, and  $B_{\vartheta_2}^{\ell+1}(V_{\vartheta_1,P,\beta}) \geq B_{\vartheta_2}^\ell(V_{\vartheta_1,P,\beta})$  for every  $\ell \geq 0$  by induction. Likewise,  $V_{\vartheta,P,\beta}(x)$  is non-decreasing in  $P$  and non-increasing in  $\beta$  for every  $x \in \mathbb{Z}$ .

Let  $(\vartheta_\ell, P_\ell, \beta_\ell)_{\ell \geq 0}$  be a converging sequence of triple of non-negative reals with some limit  $(\vartheta_\infty, P_\infty, \beta_\infty)$ . The sequences  $(\vartheta_\ell, P_\ell, \beta_\ell)_{\ell \geq 0}$  must be bounded. For simplicity of notation, I write  $V_\ell$  for  $V_{\vartheta_\ell, P_\ell, \beta_\ell}$  and  $B_\ell$  for  $B_{\vartheta_\ell, P_\ell, \beta_\ell}$ . For every  $x \in \mathbb{Z}$ , the sequence  $(V_\ell(x))_{\ell \geq 0}$  is bounded. Thus, there exists a subsequence  $(V_{\varphi(\ell)})_{\ell \geq 0}$  that converges pointwise to some  $V$ . Now verify that  $V$  is a solution to the fixed point problem  $B_\infty(V) = V$ : for every  $x \in \mathbb{Z}$ ,

$$\begin{aligned} rV(x) &= \lim_{\ell \rightarrow \infty} rV_{\varphi(\ell)}(x) \\ &= \lim_{\ell \rightarrow \infty} \left( -\beta_{\varphi(\ell)}x^2 + \frac{\vartheta_{\varphi(\ell)}}{2} [V_{\varphi(\ell)}(x+1) - V_{\varphi(\ell)}(x) + P_{\varphi(\ell)}]^+ \right. \\ &\quad \left. + \frac{\vartheta_{\varphi(\ell)}}{2} [V_{\varphi(\ell)}(x-1) - V_{\varphi(\ell)}(x) + P_{\varphi(\ell)}]^+ \right) \\ &= -\beta_\infty x^2 + \frac{\vartheta_\infty}{2} [V(x+1) - V(x) + P_\infty]^+ \\ &\quad + \frac{\vartheta_\infty}{2} [V(x-1) - V(x) + P_\infty]^+. \end{aligned}$$

That is,  $V = B_\infty(V)$ , which implies that  $V = V_\infty$ . The same argument above implies that every subsequence of  $(V_\ell)_{\ell \geq 0}$  admits a sub-subsequence that converges to  $V_\infty$  pointwise. The next lemma shows that the sequence  $(V_\ell)_{\ell \geq 0}$  converges to  $V_\infty$  pointwise. Thus, for every  $x \in \mathbb{Z}$ ,  $V_{\vartheta,P,\beta}(x)$  is jointly continuous in  $(\vartheta, P, \beta) \in \mathbb{R}^3$ .  $\square$

**Lemma 8.** *Suppose a real sequence  $(y_\ell)_{\ell \geq 0}$  satisfies the following property: every subsequence of  $(y_\ell)_{\ell \geq 0}$  admits a sub-subsequence that converges to the same constant  $y_\infty \in \mathbb{R}$ . Then  $y_\ell$  converges to  $y_\infty$  as  $\ell \rightarrow \infty$ .*

*Proof.* Suppose the contrary is true, then there exists some  $\epsilon > 0$  and a subsequence  $(y_{\varphi(\ell)})_{\ell \geq 0}$

that satisfies  $|y_{\varphi(\ell)} - y_{\infty}| > \epsilon$  for every  $\ell \geq 0$ . The subsequence  $(y_{\varphi(\ell)})_{\ell \geq 0}$  does not admit a sub-subsequence that converges to  $y_{\infty}$ .  $\square$

*Proof of Lemma 3.* Let  $\bar{x}$  denote  $\bar{x}_{\vartheta, P, \beta}$ . Formally differentiating (6) with respect to  $\vartheta$ , one obtains

$$\frac{\partial}{\partial \vartheta} V_{\vartheta}(x) = T_{\vartheta} \left( \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) \right) (x) \quad (25)$$

$$\equiv \begin{cases} \delta y(x) + \frac{\delta}{2} \left[ \frac{\partial}{\partial \vartheta} V_{\vartheta}(x+1) + \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) \right], & x \leq -\bar{x}, \\ \delta y(x) + \frac{\delta}{2} \left[ \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) + \frac{\partial}{\partial \vartheta} V_{\vartheta}(x-1) \right], & x \geq \bar{x}, \\ \delta y(x) + \frac{\delta}{2} \left[ \frac{\partial}{\partial \vartheta} V_{\vartheta}(x+1) + \frac{\partial}{\partial \vartheta} V_{\vartheta}(x-1) \right], & |x| < \bar{x}, \end{cases} \quad (26)$$

for every  $x \in \mathbb{Z}$ , where  $\delta$  is defined in (23), and for every  $x \in \mathbb{Z}$ ,

$$y(x) = \frac{rV_{\vartheta}(x) + \beta x^2}{\vartheta^2}.$$

Let  $\tilde{y} = \vartheta^2 y / r$  be a rescaled version of  $y$ . That is, for every  $x \in \mathbb{Z}$ ,  $\tilde{y}(x) = V_{\vartheta}(x) + \beta x^2 / r$ . It follows from the HJB equations (21) that for every  $x \in \mathbb{Z}$ ,

$$\tilde{y}(x) = \frac{\delta}{2} \left( \left[ \tilde{y}(x+1) - \frac{\beta(2x+1)}{r} + P \right] \vee \tilde{y}(x) + \left[ \tilde{y}(x-1) + \frac{\beta(2x-1)}{r} + P \right] \vee \tilde{y}(x) \right).$$

There is a unique solution  $\tilde{y}$  to the fixed point problem above (again from the Blackwell's sufficient conditions and the Contraction Mapping Theorem). By induction, the function  $\tilde{y}$  is even and convex. If  $\tilde{y}$  is not "U-shaped," it must be that the function  $\tilde{y}$  is constant. However, there is no constant function that is a solution to the fixed point problem above. Therefore,  $\tilde{y}$  is U-shaped. The function  $y$  is also U-shaped since  $y$  is a multiple of  $\tilde{y}$ .

There is a unique solution  $\frac{\partial}{\partial \vartheta} V_{\vartheta}$  to the fixed point problem (25) (again from the Blackwell's sufficient conditions and the Contraction Mapping Theorem). One can write the equations

(25), for  $x$  between  $-\bar{x}$  and  $\bar{x}$ , into the following matrix form:

$$A \frac{\partial}{\partial \vartheta} V_\vartheta = y, \quad (27)$$

where  $\zeta = 1/\delta$  and  $A$  is the tri-diagonal matrix of size  $(2\bar{x} + 1) \times (2\bar{x} + 1)$  as in (16). Since  $\zeta > 1$ , Lemma 5 applies. Since  $y$  is U-shaped, the unique solution  $\frac{\partial}{\partial \vartheta} V_\vartheta = A^{-1}y$  to the linear system (27) is U-shaped, and  $\frac{\partial}{\partial \vartheta} V_\vartheta(\bar{x}) < y(\bar{x})/(\zeta - 1)$  (property (vi) of Lemma 5). That is,

$$\begin{aligned} \frac{\partial}{\partial \vartheta} V_\vartheta(x) &= \frac{\partial}{\partial \vartheta} V_\vartheta(-x), \quad \forall |x| \leq \bar{x}, \\ \frac{\partial}{\partial \vartheta} V_\vartheta(0) &< \dots < \frac{\partial}{\partial \vartheta} V_\vartheta(\bar{x}) \leq \frac{1}{\zeta - 1} y(\bar{x}). \end{aligned} \quad (28)$$

For every  $x > \bar{x}$ ,

$$\frac{\partial}{\partial \vartheta} V_\vartheta(x) = \frac{y(x) + \frac{1}{2} \frac{\partial}{\partial \vartheta} V_\vartheta(x-1)}{(\zeta - 1) + \frac{1}{2}}.$$

Since  $y$  is U-shaped, one can show by induction that for every  $x > \bar{x}\vartheta$ ,

$$\frac{\partial}{\partial \vartheta} V_\vartheta(x-1) < \frac{1}{\zeta - 1} y(x).$$

Then for every  $x > \bar{x}_{k,d,P}$ ,

$$\frac{\partial}{\partial \vartheta} V_\vartheta(x) > \frac{(\zeta - 1) \frac{\partial}{\partial \vartheta} V_\vartheta(x-1) + \frac{1}{2} \frac{\partial}{\partial \vartheta} V_\vartheta(x-1)}{(\zeta - 1) + \frac{1}{2}} = \frac{\partial}{\partial \vartheta} V_\vartheta(x-1).$$

Combine the inequalities above with (28), one obtains  $\frac{\partial}{\partial \vartheta} V_\vartheta(x) > \frac{\partial}{\partial \vartheta} V_\vartheta(x-1)$  for every  $x > 1$ . Therefore, the unique solution  $\frac{\partial}{\partial \vartheta} V_\vartheta$  to the fixed problem (25) is U-shaped.

I fix some  $P \in \mathbb{R}^+$ , let  $\vartheta_1, \vartheta_2 \in \mathbb{R}^+$  be such that  $\bar{x}_{\vartheta_1,P} = \bar{x}_{\vartheta_2,P}$ . Integrate equation (25) with respect to  $\vartheta$ , for  $\vartheta$  between  $\vartheta_1$  and  $\vartheta_2$ , and compare to the (6), one obtains,

$$V_{\vartheta_2,P} - V_{\vartheta_1,P} = \int_{\vartheta_1}^{\vartheta_2} \frac{\partial}{\partial \vartheta} V_{\vartheta,P} d\vartheta. \quad (29)$$

Since  $V_{\vartheta,P}(x)$  is continuous with respect to  $\vartheta$  for every  $x \in \mathbb{Z}$  (Lemma 7), it follows that equation (29) holds for every pair  $\vartheta_1, \vartheta_2 \in \mathbb{R}^+$ . Since the function  $\frac{\partial}{\partial \vartheta} V_{\vartheta,P}(\cdot)$  is U-shaped for every  $\vartheta \geq 0$ , the function  $V_{\vartheta_2,P} - V_{\vartheta_1,P}$  is also U-shaped, for every pair  $\vartheta_1, \vartheta_2 \in \mathbb{R}^+$ . In

particular, the loss function  $L_{k,m,P}(x) = V_{k,m,P}(x) - V_{k-1,m,P}(x)$  is U-shaped in  $x$ .  $\square$

## B.4 Proof of Proposition 6

I need the following generalization of Proposition 6 to prove the other results in Section 3.

**Proposition 10.** *Given some  $P \in \mathbb{R}^+$ , the inventory threshold level  $\bar{x}_{\vartheta,P,\beta}$  is weakly increasing in  $\vartheta \in \mathbb{R}^+$  and weakly decreasing in  $\beta \in \mathbb{R}^{++}$ .*

*Proof.* I write  $\Delta_{\vartheta}$  for  $\Delta$  (defined in (7)) to make clear the dependence on  $\vartheta$ . Since  $\frac{\partial}{\partial \vartheta} V_{\vartheta}$  is U-shaped, one has

$$\Delta_{\vartheta_2}(x) - \Delta_{\vartheta_1}(x) = \int_{\vartheta_1}^{\vartheta_2} \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) d\vartheta - \int_{\vartheta_1}^{\vartheta_2} \frac{\partial}{\partial \vartheta} V_{\vartheta}(x+1) d\vartheta < 0,$$

for every  $x \in \mathbb{Z}^+$  and  $0 \leq \vartheta_1 < \vartheta_2$ . Thus

$$\bar{x}_{\vartheta_1,P,\beta} \leq \bar{x}_{\vartheta_2,P,\beta},$$

for every  $\vartheta_1 < \vartheta_2$ . That is,  $\bar{x}_{\vartheta,P,\beta}$  is weakly increasing in  $\vartheta \in \mathbb{R}^+$ . The same technique can be applied to show that  $\bar{x}_{\vartheta,P,\beta}$  is weakly decreasing in  $\beta \in \mathbb{R}^{++}$ . The proof is omitted.  $\square$

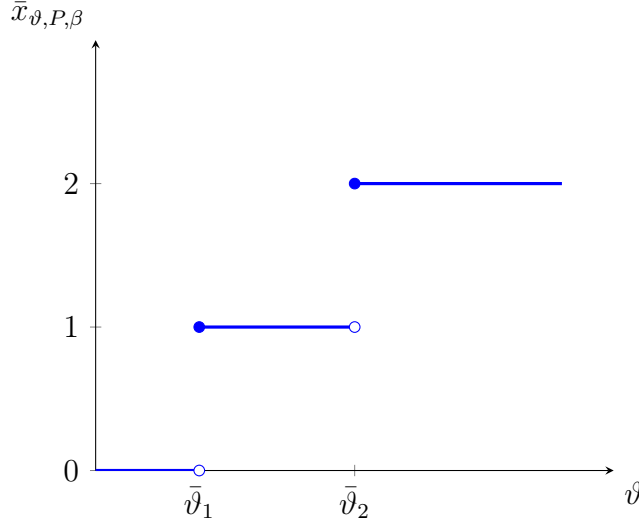
**Corollary 5.** *The inventory threshold level  $\bar{x}_{\vartheta,P,\beta}$  is right-continuous in  $\vartheta \in \mathbb{R}^+$ , left-continuous in  $\beta \in \mathbb{R}^{++}$  and piecewise constant in both parameters.*

*Proof.* Since  $V_{\vartheta,P}$  is continuous in  $\vartheta$  (Lemma 7),  $\Delta_{\vartheta}(x)$  is continuous in  $\vartheta$ . It then follows that  $\bar{x}_{\vartheta,P,\beta}$  is right-continuous in  $\vartheta$ . As  $\bar{x}_{\vartheta,P,\beta} \in \mathbb{Z}^+$  is non-decreasing in  $\vartheta$ , it must be that the set of discontinuity of  $\bar{x}_{\vartheta,P,\beta}$  with respect to  $\vartheta$  is discrete and does not have any accumulation point in  $\mathbb{R}^+$ . Denote the discontinuity points by  $\bar{\vartheta}_1 < \bar{\vartheta}_2 < \dots$ . Figure 13 illustrates  $\bar{x}_{\vartheta,P,\beta}$  as a function of  $\vartheta$ . Likewise,  $\bar{x}_{\vartheta,P,\beta}$  is left-continuous and piecewise constant in  $\beta \in \mathbb{R}^{++}$ .  $\square$

## B.5 Proof of Lemma 4

*Proof of Lemma 4 parts (a) and (b).* It is sufficient to establish that for every  $0 \leq \vartheta_1 < \vartheta_2$ ,

$$\frac{\partial}{\partial \vartheta} V_{\vartheta_1}(0) < \frac{\partial}{\partial \vartheta} V_{\vartheta_2}(0) \tag{30}$$



**Figure 13** – The inventory threshold level  $\bar{x}_{\vartheta, P, \beta}$  as a function of  $\vartheta$ .

Inequality (30) would imply, (a) for every  $1 \leq k_1 < k_2$  and  $1 \leq m_1 < m_2$ ,

$$\begin{aligned} \mathcal{L}(k_1, m, P) &= \int_{\vartheta(k_1-1, m)}^{\vartheta(k_1, m)} \frac{\partial}{\partial \vartheta} V_{\vartheta}(0) d\vartheta < \int_{\vartheta(k_2-1, m)}^{\vartheta(k_2, m)} \frac{\partial}{\partial \vartheta} V_{\vartheta}(0) d\vartheta = \mathcal{L}(k_2, m, P), \\ \mathcal{L}(k, m_1, P) &= \int_{\vartheta(k-1, m_1)}^{\vartheta(k, m_1)} \frac{\partial}{\partial \vartheta} V_{\vartheta}(0) d\vartheta < \int_{\vartheta(k-1, m_2)}^{\vartheta(k, m_2)} \frac{\partial}{\partial \vartheta} V_{\vartheta}(0) d\vartheta = \mathcal{L}(k, m_2, P), \end{aligned}$$

where  $\vartheta(k, m) = 2\lambda(k\theta_m/m + \theta_{m-1})$  is, with a slight abuse of notation, the reparametrization from  $(k, m)$  to  $\vartheta$ .

To show (30), I formally differentiate equation (25) with respect to  $\vartheta$  to obtain

$$\zeta \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(x) = \begin{cases} z(x) + \frac{1}{2} \left[ \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(x+1) + \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(x) \right], & x = -\bar{x}_{\vartheta}, \\ z(x) + \frac{1}{2} \left[ \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(x) + \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(x-1) \right], & x = \bar{x}_{\vartheta}, \\ z(x) + \frac{1}{2} \left[ \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(x+1) + \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(x-1) \right], & |x| < \bar{x}_{\vartheta}, \end{cases} \quad (31)$$



where  $\zeta = 1/\delta > 1$ , and

$$z(x) = \begin{cases} \frac{1}{\vartheta} \left[ \frac{\partial}{\partial \vartheta} V_{\vartheta}(x+1) - \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) \right], & x = -\bar{x}_{\vartheta}, \\ \frac{1}{\vartheta} \left[ \frac{\partial}{\partial \vartheta} V_{\vartheta}(x-1) - \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) \right], & x = \bar{x}_{\vartheta}, \\ \frac{1}{\vartheta} \left[ \frac{\partial}{\partial \vartheta} V_{\vartheta}(x+1) + \frac{\partial}{\partial \vartheta} V_{\vartheta}(x-1) - 2 \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) \right], & |x| \leq \bar{x}_{\vartheta}, \end{cases}$$

Since the function  $\frac{\partial}{\partial \vartheta} V_{\vartheta}$  is U-shaped, it follows that the function  $z$  is even and

$$\sum_{\tilde{x}=-x}^x z(\tilde{x}) = \begin{cases} \frac{2}{\vartheta} \left[ \frac{\partial}{\partial \vartheta} V_{\vartheta}(x+1) - \frac{\partial}{\partial \vartheta} V_{\vartheta}(x) \right] > 0, & \text{for } 0 \leq x < \bar{x}_{\vartheta}, \\ 0, & \text{for } x = \bar{x}_{\vartheta}. \end{cases} \quad (32)$$

The linear system (31) can be written into the matrix form

$$A \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta} = z, \quad (33)$$

Solving the linear system (33), one obtains  $\frac{\partial^2}{\partial \vartheta^2} V_{\vartheta} = A^{-1}z$ . In particular,

$$\frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(0) = \sum_{x=-\bar{x}_{\vartheta}}^{\bar{x}_{\vartheta}} A_{0,x}^{-1} \cdot z(x) = \sum_{x=0}^{\bar{x}_{\vartheta}-1} (A_{0,x}^{-1} - A_{0,x+1}^{-1}) \sum_{\tilde{x}=-x}^x z(\tilde{x}) > 0.$$

The last inequality follows from (32) and property (iii) of Lemma 5.

Let  $\vartheta_1, \vartheta_2 \in R^+$  be such that  $\bar{x}_{\vartheta_1} = \bar{x}_{\vartheta_2}$ . Integrate equation (31) with respect to  $\vartheta$ , for  $\vartheta$  from  $\vartheta_1$  to  $\vartheta_2$ , and compare to the equations (25), one obtains,

$$\frac{\partial}{\partial \vartheta} V_{\vartheta_2} - \frac{\partial}{\partial \vartheta} V_{\vartheta_1} = \int_{\vartheta_1}^{\vartheta_2} \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta} d\vartheta.$$

In particular,

$$\frac{\partial}{\partial k} V_{k_2,d,P}(0) - \frac{\partial}{\partial k} V_{k_1,d,P}(0) > 0.$$

Let  $\vartheta_0$  be a discontinuity point of  $\bar{x}_{\vartheta_0}$ . That is,  $\vartheta_0 \in \mathbb{R}^+$  is such that

$$\lim_{\vartheta \uparrow \vartheta_0} \bar{x}_\vartheta \equiv \bar{x}^- < \bar{x}_{\vartheta_0}. \quad (34)$$

The same argument used in the proof of [Lemma 7](#) implies that for every  $x \in \mathbb{Z}$ ,

$$\lim_{\vartheta \uparrow \vartheta_0} \frac{\partial}{\partial \vartheta} V_\vartheta(x) = \tilde{V}(x),$$

where the function  $\tilde{V}$  is the unique solution to the fixed point problem

$$\tilde{V}(x) = \begin{cases} \delta y(x) + \frac{\delta}{2} [\tilde{V}(x+1) + \tilde{V}(x)], & x \leq -\bar{x}^-, \\ \delta y(x) + \frac{\delta}{2} [\tilde{V}(x) + \tilde{V}(x-1)], & x \geq \bar{x}^-, \\ \delta y(x) + \frac{\delta}{2} [\tilde{V}(x+1) + \tilde{V}(x-1)], & |x| < \bar{x}^-, \end{cases}$$

On the other hand, the function  $\frac{\partial}{\partial \vartheta} V_\vartheta$  is the unique solution to the fixed point problem

$$\frac{\partial}{\partial \vartheta} V_{\vartheta_0} = T_{\vartheta_0} \left( \frac{\partial}{\partial \vartheta} V_{\vartheta_0} \right),$$

where the operator  $T_{\vartheta_0}$  is defined in [\(26\)](#). Starting from the function  $\tilde{V}^0 = \tilde{V}$ , let the sequence  $(\tilde{V}^\ell)_{\ell \geq 0}$  be generated through iterations of the operator  $T_{\vartheta_0}$ . Because  $\bar{x}^- < \bar{x}_{\vartheta_0}$  (as in [\(34\)](#)) and the function  $\tilde{V}^0$  is U-shaped, one has

$$\tilde{V}^1(x) \begin{cases} = \tilde{V}^0(x) & |x| < \bar{x}^-, \\ > \tilde{V}^0(x) & \bar{x}^- \leq |x| < \bar{x}_{\vartheta_0}, \\ = \tilde{V}^0(x) & |x| \geq \bar{x}_{\vartheta_0}. \end{cases}$$

Thus,  $\tilde{V}^0 \leq \tilde{V}^1$ . It then follows from the monotonicity of  $T_{\vartheta_0}$  that  $\tilde{V}^h \leq \tilde{V}^{h+1}$  for every  $h \geq 0$ . Letting  $h \rightarrow \infty$ , one has

$$\lim_{\vartheta \uparrow \vartheta_0} \frac{\partial}{\partial \vartheta} V_\vartheta = \tilde{V} \leq \frac{\partial}{\partial \vartheta} V_{\vartheta_0}.$$

For every pair  $0 \leq \vartheta_1 < \vartheta_2$ ,

$$\frac{\partial}{\partial \vartheta} V_{\vartheta_2}(0) - \frac{\partial}{\partial \vartheta} V_{\vartheta_1}(0) \geq \int_{\vartheta_1}^{\vartheta_2} \frac{\partial^2}{\partial \vartheta^2} V_{\vartheta}(0) d\vartheta > 0.$$

This establishes (30) and thus completes the proof of part (a) of [Lemma 4](#).  $\square$

I write  $\mathcal{L}(k, d, P, \beta)$  for  $\mathcal{L}(k, d, P)$  to make clear the dependence on  $\beta$ .

**Proposition 11.** *The loss  $\mathcal{L}(k, d, P, \beta)$  is strictly decreasing in  $\beta \in \mathbb{R}^{++}$ .*

The same technique used in the proof of parts (a) and (b) of [Lemma 4](#) can be applied to show [Proposition 11](#). One should apply the continuity property of  $V_{k,d,P,\beta}$  in  $\beta$  (in [Lemma 7](#)) and the monotonicity of  $\bar{x}_{k,d,P,\beta}$  in  $\beta$  (in [Figure 13](#)). The proof is omitted.

*Proof of [Lemma 4](#) part (c).* It is easy to see that for every constant  $\eta > 0$ ,  $V_{k,d,\eta P,\eta\beta} = \eta V_{k,d,P,\beta}$ . For every  $P_2 > P_1$ , let  $\eta = P_2/P_1 > 1$ , then

$$\mathcal{L}(k, d, P_2, \beta) > \mathcal{L}(k, d, P_2, \eta\beta) = \eta \mathcal{L}(k, d, P_1, \beta) > \mathcal{L}(k, d, P_1, \beta).$$

That is,  $\mathcal{L}(k, d, P, \beta)$  is strictly increasing in  $P \in \mathbb{R}^+$ . The continuity of  $\mathcal{L}(k, d, P, \beta)$  in  $P$  is implied by that of  $V_{k,d,P}(0)$  as per [Lemma 7](#).  $\square$

## B.6 Proof of [Theorem 1](#)

*Part (a):* I fix some integer  $m$  such that  $0 \leq m \leq m^*$ . Given an agent  $i \in N$ , suppose all other agents follow their respective equilibrium strategies in  $\sigma^*(m)$ . When  $i$  receives a trade order from her outside investors, her expected gain from search is  $\theta_m[\pi - P^*(m)] > 0$ . Otherwise, her expected gain from search is 0. Thus, agent  $i$  optimally searches among her dealer counterparties upon receiving an outside order. Conditional on successfully reaching a quote provider, agent  $i$  optimally trades at a given quote offer if and only if her intermediation profit is non-negative. Since quote providers are not allowed to post stub quotes that exceed the total gain  $\pi$ , it is optimal for  $i$  to trade at any given quote price.

If  $i$  has trading accounts with  $d$  dealers, for some  $0 \leq d \leq m$ , the total net payoff of  $i$  is given by  $\Phi_{d,P^*(m)}$  defined in (3). Since  $\Phi_{d,P^*(m)} < \Phi_{d+1,P^*(m)}$  for  $0 \leq d \leq m-2$  and

$\Phi_{m-1, P^*(m)} = \Phi_{m, P^*(m)}$ , it is strictly optimal for  $i$  to have trading accounts with  $m - 1$  or  $m$  dealers. Therefore, the account maintenance strategy  $N_m^*$  is optimal for  $i$ .

Given a dealer  $j \in J$  and let all other agents follow their respective equilibrium strategies in  $\sigma^*(m)$ . Suppose  $j$  is the quote provider to  $k$  non-dealers (for any  $k \leq n - m$ ) at the time she receives a Request for Quote.

- If the equilibrium spread  $P^*(m)$  is weakly greater than the  $(k, m)$ -tightest dealer sustainable spread  $\underline{P}(k, m)$ , then dealer  $j$  cannot be better off by posting an offer (or bid) price that is not  $P^*(m)$  ( $-P^*(m)$  respectively). Then the optimization problem faced by dealer  $j$  can be equivalently reduced to the problem  $\mathcal{P}_{k, m, P^*(m)}$ . That is, the optimal quoting strategy of  $j$  is characterized by the inventory threshold level  $\bar{x}_{k, m, P^*(m)}$ .
- If  $P^*(m) < \underline{P}(k, m)$ , then it is optimal for  $j$  to gouge by offering the bid-ask price  $[-\pi, \pi]$ .

Finally, if a non-dealer  $i \in I$  receives a Request for Quote at time  $t$ , he conjectures that the associated quote seeker will follow her equilibrium strategy by discontinuing her trading account with  $i$  immediately after time  $t$ . Then it is optimal for  $i$  to gouge by offering the bid-ask price  $[-\pi, \pi]$ . Therefore, the strategy profile  $\sigma^*(m)$  is a PBE in stationary strategy.

*Part (b):* Given an integer  $m > m^*$ , the equilibrium spread  $P^*(m)$  is strictly less than the tightest dealer-sustainable spread  $\underline{P}(m)$ . Consequently, any given dealer is strictly better off by gouging quote seekers. Thus,  $\sigma^*(m)$  is not a PBE.

## B.7 Proof of Theorem 2

I fix an integer  $m \geq 0$ , and suppose  $G(m)$  is the trading network generated by some perfect Bayesian equilibrium  $\sigma = (S, p, \rho, N^{\text{out}})$  in stationary strategies.

*Step 1:* I first show that on the equilibrium path, the offer price of any given dealer  $j \in J$  is some constant  $a_j^*$ , and her bid price is some constant  $b_j^*$ . Formally, for every non-dealer  $i \in I$ , every dealer  $j \in J$  and every  $\ell = 1, 2, \dots$ , let  $\tau_{i\ell}$  denote the time of the  $\ell$ 'th Request

for Quote of non-dealer  $i$ , and  $\tau_{j\ell}$  denote the time of dealer  $j$  providing the  $\ell$ 'th quote. I use the subscripts  $i\ell$ ,  $j\ell$  and  $i\ell^-$  to denote “at time  $\tau_{i\ell}$ ,” “at time  $\tau_{j\ell}$ ,” and “right before time  $\tau_{i\ell}$ ” respectively. It will be shown that almost surely,

$$a_{j\ell} = a_j^*, \quad b_{j\ell} = b_j^*$$

for every  $\ell = 1, 2, \dots$ .

Since the search strategy of every given agent  $i \in N$  is stationary, and  $i$  searches at most a countable number of times on the equilibrium path almost surely, it must be that  $i$  only searches upon receiving an outsider order. Conditional on successfully reaching a quote provider, agent  $i$  optimally trades at any given quote offer (as stub quotes are banned).

On the equilibrium path, a given non-dealer  $i$  does not have asset inventory, and has trading accounts with all the  $m$  dealers. For every dealer  $j \in J$ ,  $a \in \mathbb{R}$  and  $\ell = 1, 2, \dots$ , let

$$A_{i\ell}(a, j) = \{O_{i\ell} = \text{Buy}, \tilde{a}_{i\ell} = a, j_{i\ell} = j, x_{i\ell^-} = 0, N_{i\ell^-}^{\text{out}} = J, N_{i\ell^-}^{\text{in}} = \emptyset\}$$

denote the event that non-dealer  $i$  receives an offer price  $a$  from dealer  $j$  at time  $\tau_{i\ell}$ . Since the account maintenance strategy  $N_i^{\text{out}}$  of non-dealer  $i$  is stationary, the set  $N_{i\ell}^{\text{out}}$  is deterministic and independent of  $\ell$  on the event  $A_{i\ell}(a, j)$ . For every dealer  $j \in J$ , let

$$a_j^* = \sup\{a \in [-\pi, \pi] : \forall i \in I, N_{i\ell}^{\text{out}} = J \text{ almost surely on the event } A_{i\ell}(a, j)\} \quad (35)$$

be the highest ask price that  $j$  may offer without triggering account termination by any non-dealer on the equilibrium path. Likewise, for every  $b \in \mathbb{R}$ , let

$$B_{i\ell}(b, j) = \{O_{i\ell} = \text{Sell}, \tilde{b}_{i\ell} = b, j_{i\ell} = j, x_{i\ell^-} = 0, N_{i\ell^-}^{\text{out}} = J, N_{i\ell^-}^{\text{in}} = \emptyset\},$$

and

$$b_j^* = \sup\{b \in [-\pi, \pi] : \forall i \in I, N_{i\ell}^{\text{out}} = J \text{ almost surely on the event } B_{i\ell}(a, j)\}. \quad (36)$$

be the lowest bid price that  $j$  may post.

Suppose dealer  $j$  offers some ask price  $a_{j\ell} = a$  ( $\ell = 1, 2, \dots$ ) on the equilibrium path. If  $a < a_j^*$ , then there exists some  $\epsilon > 0$  such that  $a + \epsilon \leq \pi$ , and offering an ask price of  $a + \epsilon$  would not trigger any account termination. Dealer  $j$  is thus strictly better off if she raises her offer price by  $\epsilon$ , contradicting the optimality of her quoting strategy  $p_j^*$ . If  $a_j^* < a \leq \pi$ , then by the definition (35) of  $a_j^*$ , the offer price  $a$  triggers at least one account termination by some non-dealer on the equilibrium path. This contradicts that  $G(m)$  is the equilibrium trading network. Therefore, the offer price of dealer  $j$  is always  $a_j^*$  on the equilibrium path. Likewise, the bid price of dealer  $j$  is always  $b_j^*$  on the equilibrium path.

Consequently, the offer price  $a_j^*$  and the bid price  $b_j^*$  do not trigger account termination by any non-dealer. That is, for every non-dealer  $i \in I$  and  $\ell = 1, 2, \dots$ ,

$$N_{i\ell}^{\text{out}} = J \text{ almost surely on the event } A_{i\ell}(a_j^*, j) \cup B_{i\ell}(b_j^*, j).$$

*Step 2:* I next show that for every dealer  $j \in J$ , the bid-ask spread  $a_j^* - b_j^*$  is equal to  $P^*(m)$ .

The equilibrium continuation utility of any non-dealer  $i \in I$  at any given time  $t \geq 0$  is

$$\Phi = \sum_{j \in J} \lambda \frac{\theta_m}{m} (\pi - a_j^* + \pi + b_j^*) - \frac{mc}{r}.$$

The first term is the total expected gain from trading, the second term is the total account maintenance cost. If non-dealer  $i$  terminates his trading account with a given dealer  $j \in J$ , his continuation utility is

$$\Phi_{-j} = \sum_{j' \in J/\{j\}} \lambda \frac{\theta_{m-1}}{m-1} (\pi - a_{j'}^* + \pi + b_{j'}^*) - \frac{(m-1)c}{r}.$$

It must be that  $i$  is indifferent towards maintaining all his trading accounts or terminating one of them. That is, for every given dealer  $j \in J$ ,

$$\Phi = \Phi_{-j}. \tag{37}$$

The indifference condition holds because:

- (i) There is no account termination on the equilibrium path, it is thus optimal for  $i$  to maintain all his trading accounts. That is,  $\Phi \geq \Phi_{-j}$ .
- (ii) If  $i$  has *strict* preference over maintaining all his trading accounts (that is, if  $\Phi > \Phi_{-j}$ ), then  $i$  would never terminate his account with dealer  $j$  given any quote offer. Likewise, all other quote seekers would never terminate their accounts with  $j$ . It is then optimal for dealer  $j$  to post the highest offer price  $\pi$  and the lowest bid price  $-\pi$ . Non-dealer  $i$  would then be strictly better off terminating his account with  $j$ , a contradiction.

It then follows from the indifference conditions (37) that  $P_j^*$  does not depend on  $j \in J$ . To show this, let  $j_1 = \operatorname{argmax}_j P_j^*$  and  $j_2 = \operatorname{argmax}_j P_j^*$  be the dealer offering the largest and the smallest spread respectively. If  $P_{j_1}^* > P_{j_2}^*$ , then terminating the trading account with  $j_1$  is strictly better for a non-dealer than terminating the trading account with  $j_2$ . That is,  $\Phi_{-j_1} > \Phi_{-j_2}$ , which contradicts the indifference conditions (37). Thus,  $P_{j_1}^* = P_{j_2}^*$ . Therefore, there exists some constant  $P^*$  such that  $P_j^* = P^*$  for every  $j \in J$ .

The total net equilibrium payoff of a non-dealer is then  $\Phi = \Phi_{m,P^*}$ , where  $\Phi_{d,P}$  is given by (3). The total continuation utility of a non-dealer after terminating one trading account is given by  $\Phi_{m-1,P^*}$ . The non-dealer indifference condition  $\Phi_{m,P^*} = \Phi_{m-1,P^*}$  implies  $P^* = P^*(m)$ , where  $P^*(m)$  is the equilibrium spread given in (5).

## B.8 Proof of Theorem 3

For any given integer  $m \geq 0$ , suppose  $G(m)$  is the trading network generated by some perfect Bayesian equilibrium in stationary strategies. For a given dealer  $j \in J$ , let  $a^* = P^*(m) + h$  and  $b^* = -P^*(m) + h$  be her equilibrium offer and bid prices respectively, for some  $h \in \mathbb{R}$ . The optimal quoting strategy of dealer  $j$  should solve the following stochastic control problem:

*Dealer's problem:*

- The state space is the set  $\mathbb{Z}$  of integers, the inventory space of dealer  $j$ .
- The control space is  $\{b^*, b_{CB}^*\} \times \{a^*, a_{CB}^*\}$ , which is the set of all possible bid-ask quotes that dealer  $j$  may offer.

- Dealer  $j$  receives Requests for Quote from  $k$  non-dealers and  $m - 1$  dealers at the total mean contact rate  $2\lambda(k\theta_m/m + \theta_{m-1})$ . Every contacting non-dealer seeks to buy or sell 1 unit of the asset, independently across contacts and with equal probability  $1/2$ .
- The payoff of dealer  $j$  is the expected discounted value of all her monetary transfers and inventory cost, as specified in (2).

Let  $\mathcal{P}(k, h)$  denote this problem. The number  $k$  of non-dealer customers will be set to be  $k = n - m$ . Let  $V_{k,h}$  be the value function of the dealer in the control problem  $\mathcal{P}(k, h)$ . Then  $V_{k,h}$  satisfies the HJB equations

$$\begin{aligned} rV_{k,h}(x) = & -\beta x^2 + \lambda \left( k \frac{\theta_m}{m} + \theta_{m-1} \right) [V_{k,h}(x+1) - V_{k,h}(x) + P^*(m) - h]^+ \\ & + \lambda \left( k \frac{\theta_m}{m} + \theta_{m-1} \right) [V_{k,h}(x-1) - V_{k,h}(x) + P^*(m) + h]^+. \end{aligned} \quad (38)$$

I extend the domain of the value function  $V_{k,h}$  from  $\mathbb{Z}$  to  $\mathbb{R}$ . That is, the HJB equation (38) holds for every  $x \in \mathbb{R}$ . It follows again from the Blackwell's sufficient conditions and the Contraction Mapping Theorem that there is a unique function  $V_{k,h} : \mathbb{R} \rightarrow \mathbb{R}$  such that  $V_{k,h}$  is bounded above by some constant and below by the function  $\mathbb{R} \ni x \mapsto -\beta x^2/r$ , and  $V_{k,h}$  satisfies the HJB equations (38). It follows from Lemma 6 and value iteration that the function  $V_{k,h}$  is strictly concave.

Let  $y : \mathbb{R} \rightarrow \mathbb{R}$  be defined by

$$y_h(x) = V_{k,h}(x) + \frac{\beta}{r}x^2$$

for every  $x \in \mathbb{R}$ . Then it follows from the HJB equations (38) that for every  $x \in \mathbb{R}$ ,

$$\begin{aligned} y_h(x) = & \frac{\delta}{2} \left( \left[ y_h(x+1) - \frac{\beta(2x+1)}{r} + P - h \right] \vee y_h(x) \right. \\ & \left. + \left[ y_h(x-1) + \frac{\beta(2x-1)}{r} + P + h \right] \vee y_h(x) \right), \end{aligned} \quad (39)$$

where  $\delta$  is given by equation (23) with  $d = m$ . Another application of the Contraction



Mapping Theorem implies that the function  $y_h$  is uniquely determined by the equations (39). It follows from value iteration that the function  $y_0$  is even and strictly convex. Thus, it must be that  $y_0$  is a continuous function on  $\mathbb{R}$ . It can be verified that

$$y_h(x) = y_0\left(x + \frac{rh}{2\beta}\right)$$

for every  $x \in \mathbb{R}$ . That is,  $y_h$  is obtained by simply shifting the function  $y_0$  to the left by  $rh/(2\beta)$ . The functions  $y_h$  and thus  $V_{h,k}$  are continuous. Since  $V_{h,k}$  is strictly concave, there exists a unique  $\bar{x}_{k,h} \in \mathbb{R}$  and  $\tilde{x}_{k,h} \in \mathbb{R}$  such that

$$\begin{aligned} V_{k,h}(\bar{x}_{k,h} - 1) - V_{k,h}(\bar{x}_{k,h}) &= P^*(m) - h, \\ V_{k,h}(-\tilde{x}_{k,h} + 1) - V_{k,h}(-\tilde{x}_{k,h}) &= P^*(m) + h. \end{aligned}$$

The threshold inventory levels  $\bar{x}_{k,h}$  and  $-\tilde{x}_{k,h}$  can be obtained by shifting  $\bar{x}_{k,0}$  and  $-\bar{x}_{k,0}$  to the left by  $rh/(2\beta)$  units:

$$\bar{x}_{k,h} = \bar{x}_{k,0} - \frac{rh}{2\beta}, \quad -\tilde{x}_{k,h} = -\bar{x}_{k,0} - \frac{rh}{2\beta}.$$

In the problem  $\mathcal{P}(k, h)$ , Dealer  $j$  optimally controls her inventory size within the range  $[-\tilde{x}_{k,h}, \bar{x}_{k,h}]$ . I next consider whether Dealer  $j$  has incentive to gouge. The static incentive by the dealer to gouge is

$$\Pi(h) = \max\{\pi - a^*, \pi + b^*\} = \pi - P^*(m) + |h|.$$

If dealer  $j$  gouges, her continuation value is at least  $V_{k-1,h}$  (she can achieve at least this continuation utility by simply following the quoting strategy characterized by the problem  $\mathcal{P}(k-1, h)$ ). Thus, the dealer's future equilibrium profits forgone due to losing one non-dealer customer is at most

$$L_{k,h}(x) \equiv V_{k,h}(x) - V_{k-1,h}(x).$$

Therefore, a necessary condition for dealer  $j$  to have no incentive to gouge is given by

$$\Pi(h) \leq \mathcal{L}(k, h) \equiv \frac{k}{k+m-1} \min_{\substack{x \in \mathbb{Z} \\ -\tilde{x}_{k,h} \leq x \leq \bar{x}_{k,h}}} L_{k,h}(x). \quad (40)$$

I will show that, for every  $h \in \mathbb{R}$ ,

$$\mathcal{L}(k, h) \leq \mathcal{L}(k, 0) + |h|. \quad (41)$$

Then condition (40) would imply

$$\Pi(0) \leq \mathcal{L}(k, 0).$$

It follows from [Corollary 1](#) that the condition above is equivalent to

$$P^*(m) \geq \underline{P}(k, m), \quad (42)$$

where  $\underline{P}(k, m)$  is the tightest  $(k, m)$ -dealer-sustainable spread. When the number of non-dealer customers of dealer  $j$  is  $k = n - m$ , condition (42) is equivalent to  $m \leq m^*$ , where  $m^*$  is maximum core size defined in (14). This would complete the proof of [Theorem 3](#). Hence, it remains to show inequality (41).

Let  $\vartheta = 2\lambda(k\theta_m/m + \theta_{m-1})$ . I write  $V_{\vartheta,h}$  for  $V_{k,h}$ ,  $\tilde{x}_{\vartheta,h}$  for  $\tilde{x}_{k,h}$  and  $\bar{x}_{\vartheta,h}$  for  $\bar{x}_{k,h}$ . I formally differentiate  $V_{\vartheta,h}(x)$  with respect to  $h$ , and then with respect to  $\vartheta$  in (38), to obtain

$$\zeta \frac{\partial}{\partial h} V_{\vartheta,h}(x) = \begin{cases} \frac{1}{2} \left[ \frac{\partial}{\partial h} V_{\vartheta,h}(x+1) + \frac{\partial}{\partial h} V_{\vartheta,h}(x-1) \right], & -\tilde{x}_{\vartheta,h} + 1 \leq x \leq \bar{x}_{\vartheta,h} - 1, \\ \frac{1}{2} \left[ \frac{\partial}{\partial h} V_{\vartheta,h}(x+1) + \frac{\partial}{\partial h} V_{\vartheta,h}(x) - 1 \right], & x < -\tilde{x}_{\vartheta,h} + 1, \\ \frac{1}{2} \left[ \frac{\partial}{\partial h} V_{\vartheta,h}(x-1) + \frac{\partial}{\partial h} V_{\vartheta,h}(x) + 1 \right], & x > \bar{x}_{\vartheta,h} - 1. \end{cases} \quad (43)$$

$$\begin{aligned}
& \zeta \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x) \\
&= \begin{cases} \varphi(x) + \frac{1}{2} \left[ \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x+1) + \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x-1) \right], & -\tilde{x}_{\vartheta, h} + 1 \leq x \leq \bar{x}_{\vartheta, h} - 1, \\ \varphi(x) + \frac{1}{2} \left[ \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x+1) + \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x) \right], & x < -\tilde{x}_{\vartheta, h} + 1, \\ \varphi(x) + \frac{1}{2} \left[ \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x) + \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x-1) \right], & x > \bar{x}_{\vartheta, h} - 1, \end{cases} \quad (44)
\end{aligned}$$

where  $\zeta = 1/\delta$  and  $\varphi(x) = (\zeta - 1) \frac{\partial}{\partial h} V_{\vartheta, h}(x)/\vartheta$ . For any given  $x \in [-\tilde{x}_{\vartheta, h}, \bar{x}_{\vartheta, h}]$ , let  $\ell = \lfloor \bar{x}_{\vartheta, h} - x \rfloor + \lfloor x + \tilde{x}_{\vartheta, h} \rfloor + 1$ ,  $s = \lfloor (\ell + 1)/2 \rfloor$ , and  $\varpi$  be the following vector of length  $\ell$ ,

$$\varpi = (-1/2, 0, \dots, 0, 1/2)^\top,$$

where  $A$  is the matrix (16) of size  $\ell \times \ell$ . The two systems (43) and (44) of linear equations can be written into following matrix forms:

$$A \frac{\partial}{\partial h} V = \varpi, \quad A \frac{\partial^2}{\partial \vartheta \partial h} V = \frac{(\zeta - 1) \frac{\partial}{\partial h} V}{\vartheta} \implies \frac{\partial^2}{\partial \vartheta \partial h} V = \frac{\zeta - 1}{k} A^{-2} \varpi. \quad (45)$$

For every  $x, h \geq 0$ , let  $i = \lfloor \bar{x}_{\vartheta, h} - x \rfloor + 1$ . Then  $i \leq s$ . It follows from (vii) of Lemma 5 that

$$\frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x) = \frac{\zeta - 1}{2\vartheta} (-A_{\ell+1-i, 1}^{-2} + A_{\ell+1-i, \ell}^{-2}) = \frac{\zeta - 1}{2\vartheta} (-A_{\ell+1-i, 1}^{-2} + A_{i, 1}^{-2}) \geq 0. \quad (46)$$

For every  $h \in \mathbb{R}$  such that  $|rh/(2\beta)| \leq 1/2$ ,

$$\left| \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(0) \right| = \frac{\zeta - 1}{\vartheta} (-A_{\ell+1-s, 1}^{-2} + A_{s, 1}^{-2}) < \frac{1}{\vartheta} \leq \frac{m}{2\lambda\theta_m}. \quad (47)$$

For every  $x \in \mathbb{R}$ , one can show that  $V_{\vartheta, h}(x)$  is jointly continuous in  $(\vartheta, h) \in \mathbb{R}^+ \times \mathbb{R}$  (a continuity proof can be obtained by adapting that of Lemma 7). It then follows that for every  $(\vartheta_1, h_1)$  and  $(\vartheta_2, h_2)$ ,

$$[V_{\vartheta_2, h_2}(x) - V_{\vartheta_1, h_2}(x)] - [V_{\vartheta_2, h_1}(x) - V_{\vartheta_1, h_1}(x)] = \int_{\vartheta_1}^{\vartheta_2} \int_{h_1}^{h_2} \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta, h}(x) dh d\vartheta.$$

The equality above follows from a similar argument that implies [equation \(29\)](#). In particular,

$$L_{k,h}(x) - L_{k,0}(x) = \int_{\vartheta(k-1,m)}^{\vartheta(k,m)} \int_0^h \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta,h}(x) dh d\vartheta,$$

where  $\vartheta(k, m) = 2\lambda(k\theta_m/m + \theta_{m-1})$ . It then follows from [\(46\)](#) that for every  $x, h \in \mathbb{R}^+$ ,

$$L_{k,0}\left(x + \frac{rh}{2\beta}\right) - L_{k,0}(x) = L_{k,h}(x) - L_{k,0}(x) \geq 0.$$

That is, the function  $L_{k,0}(\cdot)$  is weakly increasing on  $\mathbb{R}^+$ . Since  $L_{k,0}(\cdot)$  is even, it is weakly decreasing on  $\mathbb{R}^-$ . Hence, for every  $h \in \mathbb{R}$ ,

$$\begin{aligned} \frac{k+m-1}{k} \mathcal{L}(k, h) &= \min_{\substack{x \in \mathbb{Z} \\ -\bar{x}_{k,h} \leq x \leq \bar{x}_{k,h}}} L_{k,h}(x) = \min_{\substack{x \in \mathbb{Z} \\ -\bar{x}_{k,0} \leq x \leq \bar{x}_{k,0}}} L_{k,0}\left(x + \frac{rh}{2\beta}\right) \\ &= \min \left\{ L_{k,0}\left(\frac{rh}{2\beta} - \left\lfloor \frac{rh}{2\beta} \right\rfloor\right), L_{k,0}\left(\frac{rh}{2\beta} - \left\lceil \frac{rh}{2\beta} \right\rceil\right) \right\} \\ &= \min \left\{ L_{k,0}\left(\frac{rh}{2\beta} - \left\lfloor \frac{rh}{2\beta} \right\rfloor\right), L_{k,0}\left(\left\lceil \frac{rh}{2\beta} \right\rceil - \frac{rh}{2\beta}\right) \right\} \end{aligned}$$

If  $rh/(2\beta) - \lfloor rh/(2\beta) \rfloor \leq 1/2$ , then  $rh/(2\beta) - \lfloor rh/(2\beta) \rfloor \leq \lceil rh/(2\beta) \rceil - rh/(2\beta)$ . Hence,

$$\frac{k+m-1}{k} \mathcal{L}(k, h) = L_{k,0}\left(\frac{rh}{2\beta} - \left\lfloor \frac{rh}{2\beta} \right\rfloor\right) = L_{k,0}(0) + \int_{\vartheta(k-1,m)}^{\vartheta(k,m)} \int_0^{h^*} \frac{\partial^2}{\partial \vartheta \partial h} V_{\vartheta,h}(0) dh d\vartheta.$$

where  $h^*$  is such that  $rh^*/(2\beta) = rh/(2\beta) - \lfloor rh/(2\beta) \rfloor \leq 1/2$ . It then follows from [\(47\)](#) that

$$\frac{k+m-1}{k} \mathcal{L}(k, h) < L_{k,0}(0) + |h^*| \leq L_{k,0}(0) + |h| = \frac{k+m-1}{k} \mathcal{L}(k, 0) + |h|.$$

Thus,  $\mathcal{L}(k, h) < \mathcal{L}(k, 0) + |h|$ . Likewise, the same inequality  $\mathcal{L}(k, h) < \mathcal{L}(k, 0) + |h|$  holds if  $rh/(2\beta) - \lfloor rh/(2\beta) \rfloor > 1/2$ . Therefore, [\(41\)](#) holds for every  $h \in \mathbb{R}$ .

## B.9 Proof of [Proposition 3](#)

Given some integers  $m < m'$ , one has  $\Phi_{m,P^*(m)} < \Phi_{m',P^*(m')}$ . Therefore, every non-dealer enjoys a higher payoff under  $\sigma^*(m')$  than  $\sigma^*(m)$ .

## B.10 Proof of Theorem 4 and Corollary 4

The proof of Theorem 4 closely parallels those of Theorems 1 to 3, and is thus omitted. In the directed network  $G(I, J)$ , the sum of outdegrees and the sum of indegrees must be equal:

$$m^*n - \frac{\sum_{j \in J} d_j}{m-1} = \sum_{j \in J} (k_j + d_j).$$

Since  $k_j \geq k(m^*, d_j)$  for every  $j \in J$  (Theorem 4), it follows from the next lemma that

$$m^*n \geq \sum_{j \in J} \left( k(m^*, d_j) + \frac{m}{m-1} d_j \right) \geq |J| [k(m^*, 0) - 1],$$

which implies

$$|J| \leq \frac{m^*n}{k(m^*, 0) - 1}.$$

**Lemma 9.** *For every integers  $m > 0$  and  $d \geq 0$ ,*

$$k(m, d) + \frac{m}{m-1} d \geq k(m, 0) - 1.$$

*Proof.* It follows from the definition of  $k(m, d)$  that

$$\Pi(P^*(m)) > \mathcal{L}(k(m, 0) - 1, m, 0, P^*(m)). \quad (48)$$

Let  $\vartheta = 2\lambda(k\theta_m/m + d\theta_{m-1}/(m-1))$ . With reparametrization, I write  $V_{\vartheta, P}$  for  $V_{k, m, d, P}$ , and  $L_{\vartheta, P}$  for  $L_{k, m, d, P}$ . It follows from (29) and (30) in the proof of Lemma 4 that  $L_{\vartheta, P}(0)$  is strictly increasing in  $\vartheta \in \mathbb{R}^+$ . It then follows that

$$\mathcal{L}(k(m, 0) - 1, m, 0, P^*(m)) > \mathcal{L}\left(\left\lfloor k(m, 0) - 1 - \frac{m}{m-1} d \right\rfloor, m, d, P^*(m)\right).$$

Combining with (48), one has

$$\Pi(P^*(m)) > \mathcal{L}\left(\left\lfloor k(m, 0) - 1 - \frac{m}{m-1} d \right\rfloor, m, d, P^*(m)\right).$$

Therefore,

$$k(m, d) \geq \left\lfloor k(m, 0) - \frac{m}{m-1} d \right\rfloor \geq k(m, 0) - 1 - \frac{m}{m-1} d. \quad \square$$

## C Proofs from Section 4

### C.1 Proof of Proposition 5

Part (i) is proved in the main text. Proofs for the remaining results are provided here.

*Part (ii):* I fix  $m \geq 1$  and  $P > 0$ , and suppress  $m$  and  $P$  from the subscripts to simplify notations. For example,  $V_n$  denotes  $V_{n-m,m,P}$ . Since the function  $\frac{\partial}{\partial \vartheta} V_\vartheta$  is U-shaped, one has, for every integer  $x \geq 0$ ,

$$\begin{aligned} & [V_{n+1}(x) - V_{n+1}(x+1)] - [V_n(x) - V_n(x+1)] \\ &= \int_{\vartheta(n-m,m)}^{\vartheta(n-m+1,m)} \left[ \frac{\partial}{\partial \vartheta} V_\vartheta(x) - \frac{\partial}{\partial \vartheta} V_\vartheta(x+1) \right] d\vartheta < 0. \end{aligned}$$

That is, the sequence  $[V_n(x) - V_n(x+1)]_{n \geq m}$  is strictly decreasing. Thus, the sequence admits some limit  $\Delta_\infty(x)$ .

It will be shown in the proof of Proposition 6 that  $\bar{x}_n \rightarrow \infty$  as  $n \rightarrow \infty$ . Then for every  $x \in \mathbb{Z}^+$ ,  $x < \bar{x}_n$  for  $n$  sufficiently large, and it follows from (21) that

$$\begin{aligned} & rV_n(x) \\ &= -\beta x^2 + \lambda \left( (n-m) \frac{\theta_m}{m} + \theta_{m-1} \right) [V_n(x+1) + V_n(x-1) - 2V_n(x) + 2P] \\ &\sim n\lambda \frac{\theta_m}{m} [\Delta_\infty(x-1) - \Delta_\infty(x) + 2P] \end{aligned} \quad (49)$$

Where the symbol  $\sim$  indicates asymptotic equivalence as  $n \rightarrow \infty$ . Letting  $x = 0$ , one has

$$rV_n(0) \sim n\lambda \frac{\theta_m}{m} [-2\Delta_\infty(0) + 2P]. \quad (50)$$

For every  $x \geq 0$ ,

$$r[V_n(0) - xP] \leq rV_n(x) \leq rV_n(0) \implies rV_n(x) \sim n\lambda \frac{\theta_m}{m} [-2\Delta_\infty(0) + 2P]. \quad (51)$$

By comparing the asymptotic equivalences in (49) and (51), one obtains

$$\Delta_\infty(x) - \Delta_\infty(x-1) = 2\Delta_\infty(0),$$

for every  $x \in \mathbb{Z}^+$ . Thus

$$\Delta_\infty(x) = (2x + 1)\Delta_\infty(0).$$

If  $\Delta_\infty > 0$ , then  $\Delta_\infty(x) > P$  for  $x > [P/\Delta_\infty(0) - 1]/2$ , which implies  $\bar{x}_n < [P/\Delta_\infty(0) - 1]/2$ . The last inequality contradicts with the fact that  $\bar{x}_n$  goes to infinity as  $n \rightarrow \infty$ . Therefore,  $\Delta_\infty(0) = 0$ . It then follows from (50) that

$$rV_n(0) \sim 2n\lambda \frac{\theta_m}{m} P. \quad (52)$$

Since  $L_n(0)$  is strictly increasing in  $n \geq m$  (see Lemma 4), it has a limit (possibly) as  $n \rightarrow \infty$ . It then follows Cesàro's Theorem that

$$\frac{V_n(0)}{n - m} = \frac{\sum_{k=m+1}^n L_k(0)}{n - m} \xrightarrow{n \rightarrow \infty} \lim_{n \rightarrow \infty} L_n(0).$$

The equivalence in (52) then implies that

$$\lim_{n \rightarrow \infty} L_n(0) = \frac{2\lambda\theta_m}{rm} P$$

One can then solve  $\Pi(P) = \mathcal{L}(n - m, m, P)$  to obtain

$$\lim_{n \rightarrow \infty} \underline{P}(n - m, m) = \frac{mr\pi}{2\lambda\theta_m + mr}.$$

Finally, since  $\mathcal{L}(n - m, m, P)$  is strictly increasing in  $n$ , the tightest dealer sustainable spread  $\underline{P}(n - m, m)$  is strictly decreasing in  $n$ . Therefore, the limiting number  $m_\infty^*$  of dealer is the largest integer  $m$  such that

$$\frac{mr\pi}{2\lambda\theta_m + mr} = \lim_{n \rightarrow \infty} \underline{P}(n - m, m) < P^*(m).$$

with a strict inequality.

*Part (iii) (dependence of  $m^*$  on  $\pi$ ):* To indicate the dependence of endogenous variables on the parameter  $\pi$ , I will write  $P^*(m, \pi)$  for the equilibrium spread, and  $\underline{P}(m, \pi)$  for the sustainable dealer spread. The loss function  $\mathcal{L}(k, m, P)$  does not depend on  $\pi$ .

Given some  $\pi_1 < \pi_2$ , one has, for every  $m \geq 1$ ,  $P > 0$ , and  $\ell = 1, 2$ ,

$$\pi_\ell - \underline{P}(m, \pi_\ell) = \mathcal{L}(n - m, m, \underline{P}(m, \pi_\ell)).$$

Since the loss function  $\mathcal{L}(n - m, m, P)$  is strictly increasing in  $P$  ([Lemma 4](#)), it must be that

$$\pi_1 - \underline{P}(m, \pi_1) < \pi_2 - \underline{P}(m, \pi_2), \quad \underline{P}(m, \pi_2) > \underline{P}(m, \pi_1).$$

That is, when the total gain per trade increases from  $\pi_1$  to  $\pi_2$ , the increase in the dealer sustainable spread  $\underline{P}(m, \pi)$  is strictly less than  $\pi_2 - \pi_1$ . On the other hand, one has

$$P^*(m, \pi_2) - P^*(m, \pi_1) = \pi_2 - \pi_1.$$

That is, the equilibrium spread increases more than the dealer sustainable spread. Therefore, the core size  $m^*$  is weakly increasing in  $\pi$ .

*Part (iii) (dependence of  $m^*$  on  $\lambda$ ):* The same technique used in the proof of [Lemma 4](#) can be applied to show that the loss  $\mathcal{L}(k, m, P)$  from gouging is strictly increasing in  $\lambda$ , for every  $k \geq 1$  and  $P > 0$ . Hence, the dealer-sustainable spread  $\underline{P}(m)$  is strictly decreasing in  $\lambda$ . On the other hand, the equilibrium spread  $P^*(m)$  is strictly increasing in  $\lambda$  (see expression (5)). Therefore, the core size  $m^*$  is weakly increasing in  $\lambda$ .

*Part (iii) (dependence of  $m^*$  on  $c$ ):* As  $c$  decreases,  $P^*(m)$  increases, while  $\underline{P}(m)$  is not affected. The core size  $m^*$  thus weakly increases.

*Part (iii) (dependence of  $m^*$  on  $\beta$ ):* For every  $m \geq 1$ ,  $k \geq 1$  and  $P > 0$ , the loss function  $\mathcal{L}(k, m, P, \beta)$  is strictly decreasing in  $\beta \in \mathbb{R}^{++}$  ([Proposition 11](#)). Thus, the dealer sustainable spread  $\underline{P}(m)$  is strictly increasing in  $\beta$ . However, the equilibrium spread  $P^*(m)$  does not depend on  $\beta$ . Therefore, the core size  $m^*$  is weakly decreasing in  $\beta$ .

*Part (iv):* When  $\beta$  increases or  $n$  decreases, the equilibrium number  $m^*$  of dealers weakly decreases. With less competition, dealers widen their equilibrium spread offer. This can be seen directly from expression (5) of the equilibrium spread  $P^*(m^*)$ .



## C.2 Proof of **Proposition 6**

**Proposition 10** shows that  $\bar{x}_{n-m,m,P^*(m)}$  is weakly decreasing in  $\beta$  and weakly increasing in  $n\lambda$ . To drive the desired asymptotic, I fix some  $m \geq 1$  and  $P > 0$ , and let  $\vartheta = 2\lambda((n-m)\theta_m/m + \theta_{m-1})$ . With reparametrization, I write  $V_\vartheta$  for  $V_{n-m,m,P}$  and  $\bar{x}_\vartheta$  for  $\bar{x}_{n-m,m,P}$ . It is sufficient to show that  $\bar{x}_\vartheta = \Theta(\vartheta^{1/3})$  as  $\vartheta$  goes to infinity. It follows from (6) that

$$V_\vartheta(x) = T_1(V_\vartheta)(x), \quad -\bar{x}_\vartheta < x < \bar{x}_\vartheta. \quad (53)$$

$$V_\vartheta(x) = T_2(V_\vartheta)(x), \quad x \geq \bar{x}_\vartheta. \quad (54)$$

where for every function  $V : \mathbb{Z} \rightarrow \mathbb{R}$ ,

$$\begin{aligned} T_1(V)(x) &= \frac{1}{\vartheta + r} \left( -\beta x^2 + \frac{\vartheta}{2} [V(x-1) + V(x+1) + 2P] \right) \\ T_2(V)(x) &= \frac{1}{\vartheta + r} \left( -\beta x^2 + \frac{\vartheta}{2} [V(x-1) + V(x) + P] \right). \end{aligned}$$

Each of the two equations above is a difference equation. The strategy of the proof is to solve for solutions of the two difference equations in their respective regions, then pin down the undetermined coefficients of the solutions by some conditions on the boundaries  $\bar{x}_\vartheta$ .

One can solve for a quadratic solution  $U_\vartheta^0$  of the first difference equation ((53)):

$$U_\vartheta^0(x) = -\frac{\beta}{r}x^2 + \frac{\vartheta}{r} \left( P - \frac{\beta}{r} \right).$$

To solve for all solutions of **equation (53)**, I consider its homogeneous version:

$$rV(x) = \frac{\vartheta}{2} [V(x-1) + V(x+1) - 2V(x)]. \quad (55)$$

The set of solutions to the difference equation above forms a 2-dimensional vector space

$$\{ae^{d_\vartheta x} + \tilde{a}e^{d_\vartheta x} : a, \tilde{a} \in \mathbb{R}\}.$$

where

$$d_\vartheta = \sqrt{\frac{2r}{\vartheta}} + O\left(\vartheta^{-\frac{3}{2}}\right). \quad (56)$$

Therefore, the solutions of [equation \(53\)](#) are

$$\mathbb{Z} \ni x \mapsto -\frac{\beta}{r}x^2 + \frac{\vartheta}{r} \left( P - \frac{\beta}{r} \right) + ae^{d_\vartheta x} + \tilde{a}e^{-d_\vartheta x},$$

where  $a, \tilde{a} \in \mathbb{R}$ . The value function  $V_\vartheta$  must be equal to one of the solutions  $U_\vartheta$  in the region  $-\bar{x}_n \leq x \leq \bar{x}_n$ , for some  $a = a_\vartheta$  and  $\tilde{a} = \tilde{a}_\vartheta$ . Since the function  $V_\vartheta$  is even, one must have  $a_\vartheta = \tilde{a}_\vartheta$ . Hence, for every integer  $x \in [-\bar{x}_\vartheta, \bar{x}_\vartheta]$ ,

$$V_\vartheta(x) = U_\vartheta(x) \equiv -\frac{\beta}{r}x^2 + \frac{\vartheta}{r} \left( P - \frac{\beta}{r} \right) + a_\vartheta \cosh(d_\vartheta x), \quad (57)$$

By solving the second difference equation ([\(54\)](#)), one obtains, for every integer  $x \geq \bar{x}_\vartheta - 1$ ,

$$\begin{aligned} V_\vartheta(x) = W_\vartheta(x) &\equiv W_n^0(x) + b_\vartheta e^{c_\vartheta x} \\ &\equiv -\frac{\beta}{r}x^2 + \frac{\vartheta}{r} \frac{\beta}{r} x - \left( \frac{\vartheta}{r} \right)^2 \frac{\beta}{2r} + \frac{\vartheta}{2r} \left( P - \frac{\beta}{r} \right) + b_\vartheta e^{c_\vartheta x}, \end{aligned} \quad (58)$$

for some  $b_\vartheta \in \mathbb{R}$ , where

$$c_n = -\frac{2r}{\theta} + 2 \left( \frac{r}{\theta} \right)^2 + O(\vartheta^{-3}). \quad (59)$$

I show that the undetermined coefficients  $a_\vartheta$  and  $b_\vartheta$  are non-negative. For this purpose, I define  $V_\vartheta^0$  as an even function from  $\mathbb{Z}$  to  $\mathbb{R}$  such that for every  $x \in \mathbb{Z}^+$ ,

$$V_n^0(x) = \max \{ U_n^0(x), W_n^0(x) \}.$$

Let  $B_\vartheta$  be the Bellman operator defined in [\(21\)](#). Then one has

$$\begin{cases} B_\vartheta(V_\vartheta^0) \geq T_1(V_\vartheta^0) \geq T_1(U_\vartheta^0) = U_\vartheta^0, \\ B_\vartheta(V_\vartheta^0) \geq T_2(V_\vartheta^0) \geq T_2(W_\vartheta^0) = W_\vartheta^0, \end{cases} \implies B_\vartheta(V_\vartheta^0) \geq \max \{ U_\vartheta^0, W_\vartheta^0 \} = V_\vartheta^0.$$

By iterating the Bellman operator  $B_\vartheta$ , one obtains  $V_\vartheta \geq V_\vartheta^0$ , which implies  $a_\vartheta \geq 0, b_\vartheta \geq 0$ .

It follows from [equations \(57\) and \(58\)](#) that the functions  $U_\vartheta$  and  $W_\vartheta$  must have same values at  $x = \bar{x}_\vartheta - 1, \bar{x}_\vartheta$ . Therefore, one has the following boundary conditions:

$$U_\vartheta(\bar{x}_\vartheta - 1) = W_\vartheta(\bar{x}_\vartheta - 1), \quad U_\vartheta(\bar{x}_\vartheta) = W_\vartheta(\bar{x}_\vartheta). \quad (60)$$

One also has

$$T_1(U_\vartheta)(\bar{x}_\vartheta) = U_\vartheta(\bar{x}_\vartheta) = W_\vartheta(\bar{x}_\vartheta) = T_2(W_\vartheta)(\bar{x}_\vartheta) = T_2(U_\vartheta)(\bar{x}_\vartheta),$$

where the last equality uses the boundary conditions (60). It then follows that

$$U_\vartheta(\bar{x}_\vartheta) - U_\vartheta(\bar{x}_\vartheta + 1) = P.$$

By an abuse of notation, I use  $U_\vartheta$  and  $W_\vartheta$  to denote the functions on the entire real line  $\mathbb{R}$  with the same expression given by (57) and (58) respectively. Given that  $U_\vartheta$  is infinitely differentiable, there exists some  $\tilde{x}_\vartheta \in (\bar{x}_\vartheta, \bar{x}_\vartheta + 1)$  such that

$$U'_\vartheta(\tilde{x}_\vartheta) = -P. \quad (61)$$

Similarly, there exists some  $\hat{x}_\vartheta \in (\bar{x}_\vartheta - 1, \bar{x}_\vartheta + 1)$  such that

$$W'_\vartheta(\hat{x}_\vartheta) = -P, \quad (62)$$

The boundary conditions (60) to (62) are sufficient to pin down the asymptotics of the undetermined coefficients  $a_\vartheta, b_\vartheta$  as well as  $\bar{x}_\vartheta$ , as  $\vartheta$  goes to infinity. Plugging the expressions of  $U_\vartheta$  and  $W_\vartheta$  into (60) to (62), one obtains

$$-\frac{2\beta}{r}\tilde{x}_\vartheta + a_\vartheta d_\vartheta \sinh(d_\vartheta \tilde{x}_\vartheta) = -P, \quad (63)$$

$$-\frac{2\beta}{r}\hat{x}_\vartheta + \frac{\vartheta}{r} \frac{\beta}{r} + b_\vartheta c_\vartheta e^{c_\vartheta \hat{x}_\vartheta} = -P. \quad (64)$$

$$\frac{\vartheta}{2r} \left( P - \frac{\beta}{r} \right) + a_\vartheta \cosh(d_\vartheta \bar{x}_\vartheta) = \frac{\vartheta}{r} \frac{\beta}{r} \bar{x}_\vartheta - \left[ \frac{\vartheta}{r} \right]^2 \frac{\beta}{2r} + b_\vartheta e^{c_\vartheta \bar{x}_\vartheta}, \quad (65)$$

Equation (64) and  $b_n \geq 0$  imply that

$$0 \leq -\frac{r}{\vartheta} b_\vartheta c_\vartheta e^{c_\vartheta \hat{x}_\vartheta} = \frac{\beta}{r} \left( 1 - \frac{2r}{\vartheta} \hat{x}_\vartheta \right) + \frac{r}{\vartheta} P.$$

Thus,  $b_\vartheta e^{c_\vartheta \bar{x}_\vartheta} = O(\vartheta^2)$  and  $\hat{x}_\vartheta \leq \frac{\vartheta}{2r}$  for  $\vartheta$  sufficiently large. By multiplying equation (64) by

$(n - m)\vartheta/(2r)$  and subtracting by [equation \(65\)](#), one obtains

$$a_\vartheta \cosh(d_\vartheta \bar{x}_\vartheta) = b_\vartheta O(\vartheta^{-1}) e^{c_\vartheta \bar{x}_\vartheta} + O(\vartheta) = O(\vartheta). \quad (66)$$

I show that  $\hat{x}_\vartheta = o(\vartheta)$ . Suppose this is not the case, then there exists a sequence  $(\vartheta_\ell)_{\ell \geq 0}$  going to infinity and  $\hat{x}_{\vartheta_\ell} = \Theta(\vartheta_\ell)$  as  $\ell$  goes to infinity. It then follows from [\(63\)](#) that

$$a_{\vartheta_\ell} \sinh(d_{\vartheta_\ell} \tilde{x}_{\vartheta_\ell}) = \Theta\left(\vartheta_\ell^{3/2}\right), \quad \text{thus} \quad a_{\vartheta_\ell} \cosh(d_{\vartheta_\ell} \tilde{x}_{\vartheta_\ell}) = \Theta\left(\vartheta_\ell^{3/2}\right).$$

This contradicts [equation \(66\)](#). Therefore,  $\hat{x} = o(\vartheta)$ , and thus  $\bar{x}_\vartheta = o(\vartheta)$ .

I multiply [\(64\)](#) by  $\vartheta/(2r)$  and subtract by [\(65\)](#), to derive a higher order Taylor expansion

$$\begin{aligned} a_\vartheta \cosh(d_\vartheta \bar{x}_\vartheta) &= \frac{\beta}{r} \frac{\vartheta}{r} + O(1). \\ \implies a_\vartheta d_\vartheta \sinh(d_\vartheta \bar{x}_\vartheta) &\sim \frac{2\beta}{r} \sqrt{\frac{\vartheta}{r}} \tanh(d_\vartheta \bar{x}_\vartheta). \end{aligned} \quad (67)$$

It then follows from [equation \(63\)](#) that

$$d_\vartheta \bar{x}_\vartheta \sim \tanh(d_\vartheta \bar{x}_\vartheta).$$

However, the equation  $y = \tanh y$  does not have any non-zero solution. Thus,  $\lim_\vartheta d_\vartheta \bar{x}_\vartheta = 0$ .

A Taylor expansion applied to [equation \(67\)](#) leads to

$$a_\vartheta = \frac{\beta \vartheta}{r^2} + O\left(\frac{\beta \vartheta}{r^2} d_\vartheta^2 \bar{x}_\vartheta^2\right). \quad (68)$$

Using [equation \(68\)](#), another Taylor expansion applied to [equation \(63\)](#) leads to

$$a_\vartheta d_\vartheta^4 \tilde{x}_\vartheta^3 + O\left(\frac{\beta \vartheta}{r^2} d_\vartheta^4 \bar{x}_\vartheta^3\right) = -P,$$

which implies

$$\tilde{x}_\vartheta = \Theta\left(\vartheta^{1/3}\right), \quad \bar{x}_\vartheta = \Theta\left(\vartheta^{1/3}\right).$$

### C.3 Proof of Proposition 8

*Part (i):* For some integer  $m \geq 1$  and  $P > 0$ , let  $\vartheta = 2\lambda((n - m)\theta_m/m + \theta_{m-1})$ . With reparametrization, I write  $C(\vartheta)$  for  $C(n, \lambda, m)$ . The dealer payoff  $V_\vartheta(0)$  increases superlinearly with  $\vartheta$ , as shown by the proof of Lemma 4. Thus, the individual dealer inventory cost  $C(\vartheta)$  is strictly concave in  $\vartheta$ . On the other hand,

$$V_\vartheta(0) \leq \frac{\vartheta P}{r}.$$

Thus,  $C(\vartheta) \geq 0$ . Hence, it must be that  $C(\vartheta)$  is strictly increasing in  $\vartheta \in \mathbb{R}^+$ .

*Part (ii):* Equation (68) implies that  $a_\vartheta = \beta\vartheta/r^2 + O(\vartheta^{2/3})$ . Letting  $x = 0$  in (57), one has

$$V_\vartheta(0) = \frac{\vartheta P}{r} + O\left(\vartheta^{\frac{2}{3}}\right),$$

which implies  $C(\vartheta) = O(\vartheta^{2/3})$ .

*Part (iii):* Let  $\vartheta_m = 2\lambda((n - m)\theta_m/m + \theta_{m-1})$ , to make clear the dependence of the total rate  $\vartheta$  of Requests for Quote on  $m$ . Then

$$m\vartheta_m < (m + 1)\vartheta_{m+1}.$$

Since the individual dealer inventory cost  $C(\vartheta)$  is strictly increasing and strictly concave in  $\vartheta$ , it follows from Jensen's inequality that

$$mC(\vartheta_m) = mC(\vartheta_m) + C(0) < (m + 1)C\left(\frac{m\vartheta_m}{m + 1}\right) < (m + 1)C(\vartheta_{m+1}).$$

### C.4 Proof of Proposition 9

In the outcome induced by  $\sigma^{**}(m)$  ( $0 \leq m \leq \bar{m}$ ), the payoff  $\Phi_{m, P^*(m)}$  of a non-dealer is given by.

$$\Phi_{m, P^*(m)} = \frac{2\lambda\theta_m(\pi - P^*(m)) - mc}{r},$$

As  $n$  goes to infinity, the net payoff  $V_{n-m,m,P^*(m)}(0)$  of a given dealer has the following asymptotic equivalence (see (52)):

$$V_{n-m,m,P^*(m)}(0) \sim 2n\lambda \frac{\theta_m}{m} P^*(m).$$

Thus, the welfare  $U(\sigma^*(m))$  satisfies

$$\begin{aligned} U(\sigma^*(m)) &= (n-m)\Phi_{m,P^*(m)} + mV_{n-m,m,P^*(m)}(0) \\ &\sim \left( \frac{2\lambda\theta_m\pi - mc}{r} \right) n = \left[ \sum_{1 \leq m' \leq m} (\theta_{m'} - \theta_{m'-1}) P^*(m') \right] \frac{2\lambda n}{r} \equiv g(m) \frac{2\lambda n}{r}. \end{aligned} \quad (69)$$

Since  $P^*(m') > 0$  for every  $1 \leq m' \leq \bar{m}$ , then  $g(m)$  is strictly increasing in  $m$  for  $0 \leq m \leq \bar{m}$ . The asymptotic equivalence (69) implies that there exists some integer  $n_0 > 0$ , if the total number of agent  $n > n_0$ , the welfare  $U(\sigma^*(m))$  is strictly increasing in  $m$  for  $0 \leq m \leq \bar{m}$ .

## D Proof of Theorem 5

To prove Theorem 5, I will need the next result about Markov Chain mixing:

**Lemma 10.** *Let  $(\bar{x}_n)_{n \geq 0}$  be a sequence of strictly positive integers that goes to infinity, and  $(\xi_n)_{n \geq 0}$  be a sequence of strictly positive reals that goes to infinity, such that  $\bar{x}_n^2/\xi_n$  goes to 0 as  $n$  goes to infinity. For every  $n \geq 0$ , let  $f_n, g_n$  be some functions from  $\{-\bar{x}_n, \dots, \bar{x}_n\}^2$  to  $\{-\bar{x}_n, \dots, \bar{x}_n\}$  such that  $f_n(x, y) + g_n(x, y) = x + y$  for every couple  $(x, y)$ . Fix some constant  $\zeta > 0$ .*

*For every  $n$ , let  $X_t^n = (X_{jt}^n)_{j \in J}$  be a continuous-time Markov process on  $\mathbb{Z}^{|J|}$  with  $|J|$  components. Suppose  $(X_t^n)_{t \geq 0}$  evolves according to the following rule:*

- (i) *At  $t = 0$ ,  $X_0^n \sim \nu_n$  for some probability distribution  $\nu_n$  on  $\{-\bar{x}_n, \dots, \bar{x}_n\}^{|J|}$ .*
- (ii) *Every  $X_{jt}^n$  independently jumps at Poisson arrival times with intensity  $\xi_n$ . If  $-\bar{x}_n < X_{jt}^n < \bar{x}_n$ , then conditional on a jump,  $X_{jt}^n$  moves up by 1 unit with probability 1/2 and moves down by 1 unit with probability 1/2. If  $X_{jt}^n = \bar{x}_n$ , then conditional on a jump,*

$X_{jt}^n$  moves down by 1 unit with probability  $1/2$  and stays at  $\bar{x}$  with probability  $1/2$ . If  $X_{jt}^n = \bar{x}_n$ , then conditional on a jump,  $X_{jt}^n$  moves up by 1 unit with probability  $1/2$  and stays at  $-\bar{x}_n$  with probability  $1/2$ . These jumps are called “individual jumps.”

(iii) Independently of (i), every couple  $(X_{jt}^n, X_{j't}^n)$  “meet” at Poisson arrival times with intensity  $\zeta$ . The Poisson “meeting” times are mutually independent across different meeting pairs  $(j, j')$ . Conditional on a “meeting,”  $X_{jt}^n$  jumps to  $f_n(X_{jt}^n, X_{j't}^n)$  and  $X_{j't}^n$  jumps to  $g(X_{jt}^n, X_{j't}^n)$ . As opposed to (ii), these jumps are called “joint jumps.”

Let  $\mu_n$  denote the uniform distribution on the set  $\{-\bar{x}_n, \dots, \bar{x}_n\}^{|J|}$ . Let  $d_n(t) = \|X_t^n, \mu_n\|_{TV}$  for every  $t \geq 0$ , where  $\|\cdot, \cdot\|_{TV}$  denotes the total variation distance between two probability distributions. Then for every  $t > 0$ ,

$$d_n(t) \xrightarrow{n \rightarrow \infty} 0.$$

*Proof.* For every  $n$  and  $j \in J$ , let  $(\tilde{X}_{jt}^n)_{t \geq 0}$  be the Markov process on  $\{-\bar{x}_n, \dots, \bar{x}_n\}$  which loops at the endpoints, and otherwise jumps in the same direction as  $X_{jt}^n$  every time  $X_{jt}^n$  experiences an individual jump. That is, if  $\tilde{X}_{jt}^n = \bar{x}_n$  (or  $-\bar{x}_n$ ), then  $\tilde{X}_{jt}^n$  jumps down (up) by 1 unit when  $X_{jt}^n$  individually jumps down by 1 unit; If  $-\bar{x}_n \tilde{X}_{jt}^n < \bar{x}_n$ , then  $\tilde{X}_{jt}^n$  jumps whenever  $X_{jt}^n$  jumps individually, in the same direction. At  $t = 0$ , every process  $\tilde{X}_{j0}^n = X_{j0}^n$ .

Then every process  $\tilde{X}_{jt}^n$  is a symmetric continuous-time random walk on  $\{-\bar{x}_n, \dots, \bar{x}_n\}$ , with jump intensity  $\xi_n$ . The random walk  $\tilde{X}_{jt}^n$  loops at the end points  $\pm \bar{x}$ . The processes  $\tilde{X}_{jt}^n$  are mutually independent. Therefore, the joint inventory process  $\tilde{X}_t^n \equiv (\tilde{X}_{jt}^n)_{j \in J}$  has a stationary distribution which is the uniform distribution  $\mu_n$  on the set  $\{-\bar{x}_n, \dots, \bar{x}_n\}$ . Thus the mixing time of this process is on the order of  $\bar{x}_n^2/\xi_n$  as  $n$  goes to infinity. Since  $\bar{x}_n^2/\xi_n$  converges to 0, then one has for every  $t > 0$ ,

$$\|\tilde{X}_t^n, \mu_n\|_{TV} \xrightarrow{n \rightarrow \infty} 0.$$

It should be noted that the convergence above is uniform in the initial distribution  $X_0^n \sim \nu_n$ . By Proposition 4.7 of [Levin, Peres, and Wilmer \(2009\)](#), there exists a random variable

$Y_t^n \sim \mu_n$  for every  $n$  and  $t \geq 0$  such that

$$\mathbb{P}(\tilde{X}_t^n \neq Y_t^n) = \|\tilde{X}_t^n, \mu_n\|_{TV}.$$

Then for every  $t > 0$  and  $\epsilon > 0$ , there exists  $n_\epsilon > 0$  such that for every  $n > n_\epsilon$ ,

$$\mathbb{P}(\tilde{X}_t^n \neq Y_t^n) < \epsilon.$$

For every  $n$  and  $t > 0$ , let  $D_t^n$  be the event that no pairs of  $X_{jt}^n$  has met by time  $t$ . Then on the event  $D_t^n$ ,  $X_t^n = \tilde{X}_t^n$  almost surely. Let  $(t_n)_{n \geq 0}$  be a sequence of times converging to 0. Then one has

$$\mathbb{P}(D_{t_n}^n) = e^{-\zeta t_n} \rightarrow 1.$$

It then follows that for every  $t > 0$  and  $n > n_\epsilon$ ,

$$\begin{aligned} \|X_{t_n}^n, \mu_n\|_{TV} &\leq \mathbb{P}(X_{t_n}^n \neq Y_{t_n}^n) = 1 - \mathbb{P}(X_{t_n}^n = Y_{t_n}^n) \\ &\leq 1 - \mathbb{P}(X_{t_n}^n = Y_{t_n}^n \mid D_{t_n}^n) \mathbb{P}(D_{t_n}^n) \leq 1 - \mathbb{P}(\tilde{X}_{t_n}^n = Y_{t_n}^n) \mathbb{P}(D_{t_n}^n) \\ &\leq 1 - (1 - \epsilon) \mathbb{P}(D_{t_n}^n) \rightarrow \epsilon. \end{aligned}$$

Therefore, as  $n \rightarrow \infty$ ,

$$\|X_{t_n}^n, \mu_n\|_{TV} \rightarrow 0 \tag{70}$$

The convergence above is uniform in the initial distribution  $X_0^n \sim \nu_n$ . Thus for every  $t > 0$ ,

$$\begin{aligned} \|X_t^n, \mu_n\|_{TV} &= \sup_{A \in \mathbb{Z}^{|J|}} |\mathbb{P}(X_t^n \in A) - \mu_n(A)| \\ &= \sup_{A \in \mathbb{Z}^{|J|}} |\mathbb{P}(X_t^n \in A \mid X_{t-t_n}^n) - \mu_n(A)| \\ &= \sup_{A \in \mathbb{Z}^{|J|}} \left| \mathbb{P}(\hat{X}_{t_n}^n \in A \mid \hat{X}_0^n \stackrel{\mathcal{L}}{=} X_{t-t_n}^n) - \mu_n(A) \right| \rightarrow 0, \end{aligned}$$

where  $(\hat{X}_t^n)$  is an independent copy of  $(X_t^n)$ , with the initial distribution  $\hat{X}_0^n \stackrel{\mathcal{L}}{=} X_{t-t_n}^n$ .  $\square$

*Proof of Theorem 5.* The joint inventory process  $(x_{jt})$  of dealers is a Markov Chain that



satisfies the conditions stated in [Lemma 10](#), with  $\xi_n = 2(n - m)\lambda\theta_m/m$  and  $\zeta = (m - 1)\varsigma$ . Thus the dealer's problem is modified to include inter-dealer trading as follows:

*Dealer's problem  $\widehat{\mathcal{P}}(k, m, P)$ :*

- The state space is  $\mathbb{Z}$ , which is the inventory space of dealer  $j$ .
- The control space is  $\{-P, -P_{CB}\} \times \{P, P_{CB}\}$ , which is the set of all possible bid-ask quotes that dealer  $j$  may offer.
- Dealer  $j$  receives Requests for Quote from  $k$  non-dealers and  $m - 1$  dealers at the total mean contact rate  $2\lambda(\theta_m/m + \theta_{m-1})$ . Every contacting non-dealer seeks to buy or sell 1 unit of the asset, independently and with equal probability  $1/2$ .
- The payoff of dealer  $j$  is the expected discounted value of all her payments and inventory flow cost, as specified in [\(2\)](#).
- Dealer  $j$  connects with another dealer  $j'$  at the total intensity rate of  $(m - 1)\varsigma$ .
- If dealer  $j$  has inventory  $x$ , and dealer  $j'$  has inventory  $y$  at the time of connection, then dealer  $j$  sells  $q(x, y)$  units of the asset to her counterpart at the price  $p(x, y)$ .
- The inventory size of dealer  $j'$  follows the uniform distribution  $\nu_{\bar{x}}$  on  $\{-\bar{x}, \dots, \bar{x}\}$ .

Let  $\widehat{V}_{k,d,P}$  be the value function of the dealer in  $\widehat{\mathcal{P}}(k, d, P)$ . The terms of trade,  $q(x, y)$  and  $p(x, y)$  are determined by maximizing the Nash Bargaining product:

$$[q(x, y), p(x, y)] = \operatorname{argmax}_{q \in \mathbb{Z}, p \in \mathbb{R}} \left[ V_{k,d,P}(x - q) - \widehat{V}_{k,d,P}(x) + p \right] \left[ \widehat{V}_{k,d,P}(y + q) - \widehat{V}_{k,d,P}(y) - p \right], \quad (71)$$

such that

$$\begin{aligned} \widehat{V}_{k,d,P}(x - q) - \widehat{V}_{k,d,P}(x) + p &\geq 0, \\ \widehat{V}_{k,d,P}(y + q) - \widehat{V}_{k,d,P}(y) - p &\geq 0, \end{aligned}$$

That is, dealers have equal bargaining power in inter-dealer trading.

I conjecture, and verify later, that the value function  $\widehat{V}_{k,d,P}$  is strictly concave on  $\mathbb{Z}$ . With this conjecture, the Nash product (71) is maximum when

$$\begin{aligned} & \widehat{V}_{k,d,P}(x - q) - \widehat{V}_{k,d,P}(x) + p \\ &= \widehat{V}_{k,d,P}(y + q) - \widehat{V}_{k,d,P}(y) - p \\ &= \frac{\widehat{V}_{k,d,P}(\lfloor \frac{x+y}{2} \rfloor) + \widehat{V}_{k,d,P}(\lceil \frac{x+y}{2} \rceil) - \widehat{V}_{k,d,P}(x) - \widehat{V}_{k,d,P}(y)}{2} \end{aligned}$$

That is, the two dealers trade  $q(x, y)$  units of the asset to equalize their inventories, and the monetary transfer  $p(x, y)$  is such that the total gain from trade is equally shared by the two dealers.

Therefore, the value function  $\widehat{V}_{k,d,P}$  satisfies the Hamilton–Jacobi–Bellman equation: for every  $x \in \mathbb{Z}$ ,

$$\begin{aligned} r\widehat{V}_{k,d,P}(x) &= -\beta x^2 + k\lambda \frac{\theta_d}{d} \left[ \widehat{V}_{k,d,P}(x+1) - \widehat{V}_{k,d,P}(x) + P \right]^+ \\ &\quad + k\lambda \frac{\theta_d}{d} \left[ \widehat{V}_{k,d,P}(x-1) - \widehat{V}_{k,d,P}(x) + P \right]^+ \\ &\quad + \frac{1}{2}(m-1)\varsigma \sum_{y \in \mathbb{Z}} \nu_{\bar{x}}(y) \left[ \widehat{V}_{k,d,P} \left( \left\lfloor \frac{x+y}{2} \right\rfloor \right) + \widehat{V}_{k,d,P} \left( \left\lceil \frac{x+y}{2} \right\rceil \right) \right. \\ &\quad \left. - \widehat{V}_{k,d,P}(x) - \widehat{V}_{k,d,P}(y) \right]. \end{aligned} \tag{72}$$

The last term on the right hand side of the equation above corresponds to the expected change in the dealer’s continuation utility due to inter-dealer trading. This term differentiates the HJB equations above from the HJB equations (21) in Section 3.

The same technique can be applied to analyze the the HJB equations (72). The details are omitted.  $\square$

## References

- Adrian, T., E. Etula, and H. S. Shin, 2010, “Risk Appetite and Exchange Rates,” *2010 Meeting Papers*.
- Adrian, T., M. Fleming, J. Goldberg, M. Lewis, F. Natalucci, and J. Wu, 2013, “Dealer Balance Sheet Capacity and Market Liquidity during the 2013 Selloff in Fixed Income Markets,” *FEDS Notes*.
- Adrian, T., E. Moench, and H. S. Shin, 2010, “Macro Risk Premium and Intermediary Balance Sheet Quantities,” *IMF Economic Review*, 58(1), 179–207.
- Afonso, G., A. Kovner, and A. Schoar, 2014, “Trading partners in the interbank lending market,” working paper.
- Afonso, G., and R. Lagos, 2015, “Trade Dynamics in the Market for Federal Funds,” *Econometrica*, 83(1), 263–313.
- Allen, L., and A. Saunders, 1986, “The large-small bank dichotomy in the federal funds market,” *Journal of Banking & Finance*, 10(2), 219–230.
- Babus, A., 2007, “The Formation of Financial Networks,” *SSRN Electronic Journal*.
- Babus, A., and T.-W. Hu, 2015, “Endogenous Intermediation in Over-the-Counter Markets,” *SSRN Electronic Journal*.
- Bagwell, K., and G. Ramey, 1992, “The Diamond Paradox: A Dynamic Resolution,” working paper.
- Bech, M. L., and E. Atalay, 2010, “The topology of the federal funds market,” *Physica A: Statistical Mechanics and its Applications*, 389(22), 5223–5246.
- Bernheim, B., B. Peleg, and M. D. Whinston, 1987, “Coalition-Proof Nash Equilibria I. Concepts,” *Journal of Economic Theory*, 42(1), 1–12.

- Bertsekas, D. P., 2000, *Dynamic Programming and Optimal Control*. Athena Scientific, 2nd edn.
- Blasques, F., F. Bräuning, and I. van Lelyveld, 2015, “A Dynamic Network Model of the Unsecured Interbank Lending Market,” *Social Science Research Network Working Paper Series*.
- Boss, M., H. Elsinger, M. Summer, and S. Thurner, 2004, “Network topology of the interbank market,” *Quantitative Finance*, 4(6), 677–684.
- Brogaard, J., 2010, “High frequency trading and its impact on market quality,” . . . *University Kellogg School of Management Working . . .*
- Brogaard, J., T. Hendershott, and R. Riordan, 2014, “High-Frequency Trading and Price Discovery,” *Review of Financial Studies*, 27(8), 2267–2306.
- Brunnermeier, M. K., and L. H. Pedersen, 2008, “Market Liquidity and Funding Liquidity,” *Review of Financial Studies*, 22(6), 2201–2238.
- Cetorelli, N., B. Hirtle, D. P. Morgan, S. Peristiani, and J. Santos, 2007, “Trends in financial market concentration and their implications for market stability,” *Economic Policy Review*, (Mar), 33–51.
- Chang, B., and S. Zhang, 2015, “Endogenous Market Making and Network Formation,” *SSRN Electronic Journal*.
- Chang, E. J., E. J. A. Lima, S. M. Guerra, and B. M. Tabak, 2008, “Measures of Interbank Market Structure: An Application to Brazil,” .
- Comerton-Forde, C., T. Hendershott, C. Jones, P. Moulton, and M. Seasholes, 2010, “Time Variation in Liquidity: The Role of Market-Maker Inventories and Revenues,” *The Journal of Finance*, 65(1), 295–331.
- Craig, B., and G. von Peter, 2014, “Interbank tiering and money center banks,” *Journal of Financial Intermediation*, 23(3), 322–347.

- Di Maggio, M., A. Kermani, and Z. Song, 2015, “The Value of Trading Relationships in Turbulent Times,” *SSRN Electronic Journal*.
- Du, S., and H. Zhu, 2014, “Ex Post Equilibria in Double Auctions of Divisible Assets,” *SSRN Electronic Journal*.
- Duffie, D., 2012, “Market Making Under the Proposed Volcker Rule,” *SSRN Electronic Journal*.
- Farboodi, M., 2014, “Intermediation and Voluntary Exposure to Counterparty Risk,” *SSRN Electronic Journal*.
- Farboodi, M., G. Jarosh, and R. Shimer, 2015, “Who is at the Core of Financial Networks? The Role of Meeting Technologies,” *Working Paper*.
- Fricke, D., and T. Lux, 2014, “Core–Periphery Structure in the Overnight Money Market: Evidence from the e-MID Trading Platform,” *Computational Economics*, 45(3), 359–395.
- Gârleanu, N., 2009, “Portfolio choice and pricing in illiquid markets,” *Journal of Economic Theory*, 144(2), 532–564.
- Gromb, D., and D. Vayanos, 2002, “Equilibrium and welfare in markets with financially constrained arbitrageurs,” *Journal of Financial Economics*, 66(2-3), 361–407.
- , 2010, “Limits of Arbitrage,” *Annual Review of Financial Economics*, 2(1), 251–275.
- Grossman, S. J., and M. H. Miller, 1988, “Liquidity and Market Structure,” *Journal of Finance*, 43(3), 617–37.
- Guo, X., and O. Hernández-Lerma, 2009, “Continuous-Time Markov Decision Processes,” in *Continuous-Time Markov Decision Processes*. Springer Berlin Heidelberg, vol. 62 of *Stochastic Modelling and Applied Probability*, pp. 9–18.
- He, Z., and A. Krishnamurthy, 2011, “A Model of Capital and Crises,” *The Review of Economic Studies*, 79(2), 735–777.

- Heller, D., and N. Vause, 2012, “Collateral Requirements for Mandatory Central Clearing of Over-the-Counter Derivatives,” *BIS Working Papers*, pp. 1–51.
- Hendershott, T., D. Li, D. Livdan, and N. Schürhoff, 2015, “Relationship Trading in OTC Markets,” *Working Paper*.
- Hendershott, T., and A. J. Menkveld, 2013, “Price Pressures,” *Working Paper*, June(3), 1–40.
- Hendershott, T., and M. S. Seasholes, 2007, “Market Maker Inventories and Stock Prices,” *American Economic Review*, 97(2), 210–214.
- Hollifield, B., A. Neklyudov, and C. S. Spatt, 2014, “Bid-Ask Spreads, Trading Networks and the Pricing of Securitizations: 144a vs. Registered Securitizations,” *SSRN Electronic Journal*.
- Huang, Y., and W. F. McColl, 1997, “Analytical inversion of general tridiagonal matrices,” *Journal of Physics A: Mathematical and General*, 30(22), 7919–7933.
- Hugonnier, J., B. Lester, and P.-O. Weill, 2015, “Heterogeneity in decentralized asset markets,” *Working Papers*.
- in ’t Veld, D., and I. van Lelyveld, 2014, “Finding the core: Network structure in interbank markets,” *Journal of Banking & Finance*, 49, 27–40.
- in ’t Veld, D. L., M. van der Leij, and C. H. Hommes, 2014, “The Formation of a Core Periphery Structure in Heterogeneous Financial Networks,” *SSRN Electronic Journal*.
- James, J., I. W. Marsh, and L. Sarno (eds.), 2012, *Handbook of Exchange Rates*. John Wiley & Sons, Inc., Hoboken, NJ, USA.
- Kirilenko, A. A., A. S. Kyle, M. Samadi, and T. Tuzun, 2011, “The Flash Crash: The Impact of High Frequency Trading on an Electronic Market,” *SSRN Electronic Journal*.
- Kyle, A. S., A. A. Obizhaeva, and Y. Wang, 2013, “Smooth Trading with Overconfidence and Market Power,” *SSRN Electronic Journal*.

- Lagos, R., G. Rocheteau, and P.-O. Weill, 2008, “Crashes and Recoveries in Illiquid Markets,” *NBER working paper*.
- Levin, D. A., Y. Peres, and E. L. Wilmer, 2009, *Markov Chains and Mixing Times*. American Mathematical Soc.
- Li, D., and N. Schürhoff, 2014, “Dealer Networks,” *SSRN Electronic Journal*.
- Markose, S., S. Giansante, and A. R. Shaghaghi, 2012, “‘Too interconnected to fail’ financial network of US CDS market: Topological fragility and systemic risk,” *Journal of Economic Behavior & Organization*, 83(3), 627–646.
- Markose, S. M., 2012, “Systemic Risk from Global Financial Derivatives: A Network Analysis of Contagion and Its Mitigation with Super-Spreader Tax,” *IMF Working Papers*, 12(282).
- Meli, J. A., 2002, “Do capital constraints on market makers matter? : Evidence from the United States treasury market,” working paper.
- Mengle, D., 2010, “Concentration of OTC Derivatives among Major Dealers,” working paper, International Swaps and Derivatives Association.
- MFA, 2015, “Why Eliminating Post-Trade Name Disclosure will Improve the Swaps Market,” .
- Neklyudov, A. V., 2014, “Bid-Ask Spreads and the Over-the-Counter Interdealer Markets: Core and Peripheral Dealers,” *Working Paper*.
- Peltonen, T. A., M. Scheicher, and G. Vuillemeys, 2014, “The network structure of the CDS market and its determinants,” *Journal of Financial Stability*, 13, 118–133.
- Protter, P. E., 2005, *Stochastic Integration and Differential Equations*, vol. 21 of *Stochastic Modelling and Applied Probability*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Rinne, K., and M. Suominen, 2010, “Short Term Reversals, Returns to Liquidity Provision and the Costs of Immediacy,” *SSRN Electronic Journal*.

- , 2011, “A Structural Model of Short-Term Reversals,” *SSRN Electronic Journal*.
- Rostek, M., and M. Weretka, 2012, “Price Inference in Small Markets,” *Econometrica*, 80(2), 687–711.
- Shen, J., B. Wei, and H. Yan, 2015, “Financial Intermediation Chains in a Search Market,” *SSRN Electronic Journal*.
- Siriwardane, E., 2015, “Concentrated Capital Losses and the Pricing of Corporate Credit Risk,” *SSRN Electronic Journal*.
- Stahl, Dale O. I., 1989, “Oligopolistic Pricing with Sequential Consumer Search,” *American Economic Review*, 79(4), 700–712.
- Üslu, S., 2015, “Pricing and Liquidity in Decentralized Asset Markets,” *Working Paper*.
- Vives, X., 2011, “Strategic Supply Function Competition With Private Information,” *Econometrica*, 79(6), 1919–1966.
- Weill, P.-O., 2007, “Leaning Against the Wind,” *Review of Economic Studies*, 74(4), 1329–1354.
- Zhang, F., 2010, “High-Frequency Trading, Stock Volatility, and Price Discovery,” *SSRN Electronic Journal*.
- Zhong, Z., 2013, “The Network of Inter-Dealer Trading in an Over-the-Counter Market,” *SSRN Electronic Journal*.