HETEROGENEITY IN DECENTRALIZED ASSET MARKETS

Julien Hugonnier
Benjamin Lester
Pierre-Olivier Weill

Heterogeneity in Decentralized Asset Markets
Julien Hugonnier, Benjamin Lester, and Pierre-Olivier Weill
NBER Working Paper No. 20746
December 2014, Revised February 2016
JEL No. D0,D53,D83,G0,G12

ABSTRACT

We study a search and bargaining model of an asset market, where investors' heterogeneous valuations for the asset are drawn from an arbitrary distribution. Our solution technique makes the model fully tractable and allows us to provide a full characterization of the unique equilibrium, in closed-form, both in and out of steady-state. Using this characterization, we first establish that the model generates aggregate trading patterns that are consistent with those observed in many over-the-counter asset markets. Then, we show that the model can replicate empirical regularities reported from micro-level data sets, including the relationships between the length of the intermediation chains through which assets are reallocated, the network centrality of the dealers involved in these chains, and the markup charged on the asset being passed along the chain. Finally, we show that heterogeneity magnifies the price impact of search frictions, and that this impact is more pronounced on price levels than on price dispersion. Hence, using observed price dispersion to quantify the effect of search frictions on price discounts or premia can be misleading.

Julien Hugonnier
Ecole Polytechnnique Fédérale de Lausanne
Quartier UNIL Dorigny Extranef #212
1015 Lausanne Switzerland
and Swiss Finance Institute
Julien.Hugonnier@epfl.ch

Benjamin Lester
Research Department
Federal Reserve Bank of Philadelphia
10 Independence Mall
Philadelphia, PA 19106
Benjamin.Lester@phil.frb.org

Pierre-Olivier Weill
Department of Economics
University of California, Los Angeles
Bunche Hall 8283
Los Angeles, CA 90095
and NBER
poweill@econ.ucla.edu

1 Introduction

Many assets, including corporate and government bonds, emerging market debt, mortgage-backed securities, and most types of swaps, trade in decentralized or "over-the-counter" (OTC) markets. To study such markets, we construct a search and bargaining model in which investors with heterogeneous valuations are periodically and randomly matched in pairs and given the opportunity to trade. Whereas the existing literature, starting with Duffie, Gârleanu, and Pedersen (2005) (henceforth DGP), has primarily focused on the special case of two valuations, we allow investors' valuations to be drawn from an arbitrary distribution. Despite its greater complexity, this generalized model remains fully tractable: we provide a solution technique that delivers a full characterization of the equilibrium, in closed-form, both in and out of steady state. Equipped with this characterization, we argue that our model provides a structural framework for confronting and understanding empirical regularities that have been documented in OTC markets.

First, we show that our model with an arbitrary distribution of investor valuations generates a number of aggregate trading patterns that are evident in nearly all OTC markets. More specifically, though all investors in our model have the same trading opportunities, in equilibrium assets are reallocated from some investors who are natural sellers to other investors who are natural buyers through a sequence, or *chain*, of investors who act as intermediaries. Moreover, we show that some of these intermediaries trade more frequently than others in equilibrium, so that a *core-periphery* trading network emerges endogenously.

Second, in addition to replicating the aggregate trading patterns mentioned above, we show that our framework can also be used to successfully confront *micro-level* data from OTC markets. In particular, we show analytically that the model can replicate recently documented facts regarding the cross-sectional relationships between the "centrality" of an investor in the trading network, the length of the intermediation chains in which he participates, the duration and volatility of his inventory holdings, and the markup on the assets he trades.

Finally, we use the model to derive novel asset pricing results that highlight the importance of heterogeneity in decentralized asset markets. We show that heterogeneity magnifies the price impact of search frictions, and that this impact is more pronounced on price levels than on price dispersion. Hence, using observed price dispersion to quantify the effect of search frictions on

price discounts or premia can be misleading: price dispersion can essentially vanish while price levels are still far from their frictionless counterpart.

Our model, which we formally describe in Section 2, starts with the basic building blocks of DGP. There is a measure one of investors who can hold either zero or one share of an asset in fixed supply. Investors have stochastic, time-varying utility types that generate heterogenous valuations for the asset. Each investor is periodically and randomly matched with another, and a transaction ensues if there are gains from trade, with prices being determined by Nash bargaining. Our point of departure from DGP is that we allow utility types to be drawn from an arbitrary distribution. Allowing for more than two types changes the nature of the analysis significantly, as it implies that individual investors now face ex ante uncertainty about the utility types of potential trading partners, and hence about the terms of trade. More precisely, the relevant state variable in our model is an infinite-dimensional object: the distributions of the utility types among investors that hold zero and one asset, respectively, over time.

Despite this greater complexity, we show in Section 3 that the model remains fully tractable. In particular, we characterize the equilibrium, in closed form, both in and out of the steady state. This requires deriving explicit solutions for the joint distributions of asset holdings and utility types, and for investors' reservation values; both of these derivations are new to the literature. Moreover, in contrast to the usual guess-and-verify approach, we establish several elementary properties of reservation values directly—without making *a priori* assumptions on the direction of gains from trade—which allows us later to confirm the uniqueness of our equilibrium. Finally, as a by-product of our solution technique, we show that that reservation values can be computed as the present value of utility flows to a hypothetical investor with an appropriately adjusted utility type process that naturally reflects the search and bargaining frictions. This sequential representation of reservation values generalizes the concept of a marginal investor to a decentralized market.

In Section 4, we use our characterization of the equilibrium to highlight a number of the model's implications. Our first set of results are derived by *following investors* with different utility types. More specifically, we analyze how an investor's asset holdings and the frequency with which he trades depend on his utility type, and the implications of these individual trading patterns for aggregate outcomes. Our discussion builds on the simple, yet crucial, observation that

investors who have a lot to gain from trading—typically investors with extreme utility types and the "wrong" asset holdings—tend to trade quickly and then remain inactive for long periods of time. Investors with more moderate utility types, on the other hand, tend to remain active in the market more consistently, buying and selling with equal frequency over time. Hence, these investors with moderate valuations tend to emerge endogenously as intermediaries, even though they are not endowed with a superior search technology. Moreover, such investors tend to trade most often with *each other*. As a result, an asset is typically reallocated from an investor with a low utility type to an investor with a high utility type through a chain of inframarginal trades executed by investors with moderate utility types, so that a core-periphery trading structure emerges endogenously. Existing empirical evidence shows that intermediation chains and core-periphery trading networks are both prevalent in nearly all OTC markets.

Our second set of results are derived by *following assets* as they are reallocated from investors with low utility types to those with high utility types through an intermediation chain. To motivate our analysis, we start by summarizing some key empirical regularities that have recently been documented using micro-level data from a prominent OTC market—specifically, the relationships between the length of the intermediation chain required to transfer an asset from an "initial" seller to a "final" buyer; the network centrality of the dealers involved in this chain; the duration and volatility of the inventory of these dealers; and the total markup realized along the chain. We then analytically derive the distributions of these objects induced by our steady-state equilibrium, and show that the relations they imply are qualitatively consistent with the empirical findings. More broadly, this set of results shows that our generalized model offers a flexible structural framework to confront new facts emerging from micro-level OTC market data regarding the relationship between the structure of the trading network, the nature and efficiency of the process through which assets are reallocated, and the distribution of transaction prices.

Finally, in Section 5, we study equilibrium as trade gets faster and search frictions vanish. This region of the parameter space is important for two reasons. First, it is the empirically relevant case in many financial markets, where trading speeds are indeed becoming faster and faster. Second, studying this region allows us to analytically demonstrate that heterogeneity magnifies the price impact of search frictions: we show that deviations from the Walrasian price are much larger with

a continuum of types than with finitely many types. The reason is that, in the latter case, there is generically an atom of investors at the marginal type. Therefore, the elasticity of demand is infinite at the marginal type and, as a result, a small increase in price drives demand to zero for the entire atom of marginal investors. In contrast, with a continuum of types, we can obtain an arbitrary elasticity of demand by varying the density of investors at the marginal type. A lower elasticity magnifies the bilateral monopoly effects at play in our search-and-matching model by generating much larger price deviations than in previous work. Furthermore, we show that the asymptotic effect of search frictions on price levels and price dispersion are of different magnitudes: prices can be far from their Walrasian counterpart when price dispersion has nearly vanished. Hence, using price dispersion, markups, or the bid-ask spread to quantify frictions may lead one to underestimate the true effect of search frictions on some market outcomes.

1.1 Related Literature

Our paper contributes to the literature that uses search models to study asset prices and allocations in OTC markets. Early papers include Gehrig (1993), Spulber (1996), and Hall and Rust (2003). Most recent papers build on the framework of DGP who assume that investors with one of two valuations for an asset receive infrequent opportunities to trade in either a pure decentralized market—i.e., in bilateral random matches with other investors—or in a pure dealer market—i.e., with an exogenously designated set of marketmakers who have access to a competitive interdealer market. For the most part, this literature has extended DGP in one of two directions.

One strand has dropped the possibility of trading in a pure decentralized market, but has incorporated additional features—including finitely many types —into a model where all trades are executed through a pure dealer market. See, for instance, Weill (2007), Lagos and Rocheteau (2009), Gârleanu (2009), Lagos, Rocheteau, and Weill (2011), Feldhütter (2012), Pagnotta and Philippon (2011), and Lester, Rocheteau, and Weill (2015). Although abstracting from decentralized trade is beneficial for maintaining tractability, it also implies that these models are less helpful in addressing certain markets and issues. For instance, many OTC markets do not have active dealers, so that finding another investor and bargaining over the price is a central feature of these markets.[1] Moreover, even in markets with active intermediaries, the interdealer market itself

---

[1]For example, Ashcraft and Duffie (2007) report that only about one quarter of trades in the federal funds market

is often best characterized by a frictional, bilateral matching market, as documented by Green, Hollifield, and Schürhoff (2006) and Li and Schürhoff (2012) for municipal bonds and Hollifield, Neklyudov, and Spatt (2014) for mortgage-backed securities. Hence, models with only pure dealer markets cannot account for price dispersion and heterogeneous trading times between dealers. Finally, by assuming a priori that some investors intermediate all trades, these models cannot help us understand why reallocation occurs through intermediation chains, and why some investors find themselves in the midst of these chains more than others.

To address such issues, a second strand of the literature has focused exclusively on a pure decentralized market. This approach, however, requires tackling a potentially complex fixed point problem: investors' trading decisions depend on the distributions of asset holdings and utility types—since they determine the option value of search—but these distributions depend on investors' trading decisions. Early models in the literature have dealt with this fixed point problem by limiting heterogeneity to two utility types; see, e.g., Duffie, Gârleanu, and Pedersen (2007), Vayanos and Wang (2007), Vayanos and Weill (2008), Weill (2008), Afonso (2011), Gavazza (2011, 2013), Praz (2013), and Trejos and Wright (2014). While this strand of the literature has revealed a number of important insights related to liquidity and asset prices, the restriction to two types prevents these models from addressing many of the substantive issues analyzed in our paper, such as the reallocation of assets through chains of intermediaries, the structure of the trading network, and the ultimate effect of heterogeneity on price levels and dispersion.

The literature has only recently turned to the analysis of pure decentralized asset markets with more than two types of investors. Perhaps the closest to our work is Afonso and Lagos (2015), who develop a model of purely decentralized exchange to study trading dynamics in the federal funds market. In their model, investors have heterogeneous valuations because they have different levels of asset holdings. Several insights from Afonso and Lagos feature prominently in our analysis. Most importantly, they highlight the fact that investors with moderate asset holdings play the role of "endogenous intermediaries," buying from investors with excess reserves and selling to investors with few. As we discuss at length below, similar investors specializing in intermediation emerge in our environment and have important effects on equilibrium outcomes. However, our work is

---

are brokered. Hall and Rust (2003) highlight a lack of intermediation in certain OTC commodity markets, such as steel coil and plate. The absence of marketmakers is also notable in certain markets for real assets, such as houses.

quite different from that of Afonso and Lagos in a number of important ways, too. For one, since our focus is not exclusively on a market in which payoffs are defined at a predetermined stopping time, we characterize equilibrium both in and out of steady state when the time horizon is infinite. Moreover, while Afonso and Lagos establish many of their results via numerical methods, we can characterize the equilibrium in closed-form for an arbitrary distribution of investor types. This tractability allows us to perform analytical comparative statics and, in particular, derive a number of novel results—for example, our analytical characterization of the relationships between the length of an intermediation chain, the centrality of the intermediaries involved, and the markup are completely new to this literature.[2]

Our paper is also related to the growing literature that studies equilibrium asset pricing and exchange in exogenously specified trading networks. Recent work includes Gofman (2010), Babus and Kondor (2012), Malamud and Rostek (2012), and Alvarez and Barlevy (2014). Atkeson, Eisfeldt, and Weill (2015) and Colliard and Demange (2014) develop hybrid models, blending ingredients from the search and the network literatures. In these models, intermediation chains arise somewhat mechanically; indeed, when investors are exogenously separated by network links, the only feasible way to reallocate assets to those who value them most is to use an intermediation chain. In our dynamic search model, by contrast, intermediation chains arise by choice: though all investors have the option to keep searching until they get an opportunity to trade directly with their best counterparty, they find it optimal to trade indirectly, through intermediation chains. Hence,

---

[2]Several other papers deserve mention here. The present paper merges, replaces, and extends Hugonnier (2012) and Lester and Weill (2013), in which we independently developed the techniques to solve for equilibrium in DGP with a continuum of types. Neklyudov (2012) considers a model with two valuations but introduces heterogeneity in trading speed to study equilibrium prices and allocations in a given core-periphery trading network. In our model, a core-periphery network arises endogenously even though trading speed is constant across investors. In an online Appendix, Gavazza (2011) proposes a model of purely decentralized trade with a continuum of types in which investors have to pay a search cost in order to meet others. He focuses on steady-state equilibria in a region of the parameter space where all investors with the same asset holdings trade at the same frequency, and trade only once between preference shocks; this special case abstracts from most of the interesting dynamics that emerge from our analysis. Shen, Wei, and Yan (2015) incorporate search costs into our framework to endogenize the boundaries of the intermediary sector. Cujean and Praz (2013) study transparency in OTC markets by considering a model with a continuum of types and unrestricted asset holdings, where investors are imperfectly informed about the type of their trading partner. In this environment, Nash bargaining is problematic and, hence, the authors propose a new trading protocol. Üslü (2015) considers a generalized model with heterogenous hedging needs and asset holdings and analyzes, among other things, the determinant of the "speed premium" in OTC markets. Sagi (2015) calibrates a partial equilibrium model with heterogenous types to explain commercial real estate returns.

even though all contacts are random, the endogenous network of actual trades is not, but rather exhibits a core-periphery-like structure that is typical of many OTC markets.[3]

Finally, our paper is related to the literatures that use search-theoretic models to study monetary and labor economics. The former literature, starting with the seminal contribution of Kiyotaki and Wright (1993), has recently incorporated assets into the workhorse model of Lagos and Wright (2005) to study issues related to financial markets, liquidity, and asset pricing.[4] In the latter literature, such as Burdett and Mortensen (1998) and Postel-Vinay and Robin (2002), workers move along a "job ladder" from low- to high-productivity firms much like assets in our model are reallocated from low- to high-valuation investors. Despite this similarity, many of our results are specific to our asset market environment. For example, our analysis of intermediation chains and the trading network is designed to establish contact with micro data from OTC markets, and has no natural analog in labor markets. Likewise, given dramatic increases in trading speed, it is natural for us to study equilibrium outcomes as contact rates tend to infinity, while such analysis has no obvious counterpart in a labor economics.

2 The model

2.1 Preference, endowments, and matching technology

We consider a continuous-time, infinite-horizon model in which time is indexed by $t \geq 0$. The economy is populated by a unit measure of infinitely-lived and risk-neutral investors who discount the future at the same rate $r > 0$. There is one indivisible, durable asset in fixed supply, $s \in (0, 1)$, and one perishable good that we treat as the numéraire.

Investors can hold either zero or one unit of the asset.[5] The utility flow an investor receives at time t from holding a unit of the asset, which we denote by δ_t, differs across investors and, for each investor, changes over time. In particular, each investor receives i.i.d. preference shocks that

---

[3]See Oberfield (2013) for another example of endogenous network formation through search. In a recent paper, Glode and Opp (2014) also examine why intermediation chains are prevalent but their focus is different, as they postulate that these chains help to moderate the inefficiencies induced by asymmetric information.

[4]See, e.g., Lagos (2010), Geromichalos, Licari, and Suárez-Lledó (2007), Lester, Postlewaite, and Wright (2012), and Li, Rocheteau, and Weill (2012).

[5]For the purpose of analyzing steady states, we could equivalently assume that investors can trade any quantity of the asset but are constrained to hold a maximum quantity that is normalized to one share: given linear utility, they would find it optimal to trade and hold either zero or one share.

arrive according to a Poisson process with intensity γ, whereupon the investor draws a new utility flow δ' from some cumulative distribution function $F(\delta')$.[6] We assume that the support of this distribution is a compact interval, and make it sufficiently large so that $F(\delta)$ has no mass points at its boundaries. For simplicity, we normalize this interval to $[0, 1]$. Thus, at this point, we place very few restrictions on the distribution of utility types. In particular, our solution method applies equally well to discrete distributions (such as the two point distribution of Duffie, Gârleanu, and Pedersen, 2005), continuous distributions, and mixtures of the two.

Investors interact in a purely decentralized market in which each investor initiates contact with another randomly selected investor according to a Poisson process with intensity $\lambda/2$.[7] If two investors are matched and there are gains from trade, they bargain over the price of the asset. The outcome of the bargaining game is taken to be the Nash bargaining solution, in which the investor with asset holdings $q \in \{0, 1\}$ has bargaining power $\theta_q \in (0, 1)$, with $\theta_0 + \theta_1 = 1$.

An important object of interest throughout our analysis will be the joint distribution of utility types and asset holdings. The standard approach in the literature, following DGP, is to characterize this distribution by analyzing the density or *measure* of investors across types $(q, \delta) \in \{0, 1\} \times [0, 1]$. Our analysis below reveals that the model becomes much more tractable when we study instead the *cumulative measure*: it is then possible to exhibit a closed-form solution for an arbitrary underlying distribution of types, both in and out of steady state. To this end, let $\Phi_{q,t}(\delta)$ denote the measure of investors at time $t \geq 0$ with asset holdings $q \in \{0, 1\}$ and utility type less than $\delta \in [0, 1]$. Assuming that initial types are randomly drawn from the cumulative distribution $F(\delta)$, the following accounting identities must hold for all $t \geq 0$:[8]

$$\Phi_{0,t}(\delta) + \Phi_{1,t}(\delta) = F(\delta) \tag{1}$$

$$\Phi_{1,t}(1) = s. \tag{2}$$

Equation (1) highlights that the cross-sectional distribution of utility types in the population is

---

[6]The characterization and main properties of the equilibrium remain qualitatively unchanged if we assume that utility types are persistent in the sense that, conditional on experiencing a preference shock, the probability of drawing a utility type in $[0, \delta']$ is given by some function $F(\delta'|\delta)$ that is decreasing in δ for any $\delta' \in [0, 1]$.

[7]We focus on the purely decentralized market for simplicity of exposition. We show in Appendix D that our main results are upheld if we assume, as in DGP, that investors can also periodically trade with a set of market makers who have access to a centralized market.

[8]Most of our results extend to the case in which the initial distribution is not drawn from $F(\delta)$, though the analysis is slightly more complicated; see Appendix C.

constantly equal to $F(\delta)$, which is due to the fact that initial utility types are drawn from $F(\delta)$ and that an investor's new type is independent from his previous type. Equation (2) is a market clearing condition: the total measure of investors who own the asset must equal the total supply of assets available in the economy. Given our previous assumptions, we have that this condition is equivalent to $\Phi_{0,t}(1) = 1 - s$ for all $t \geq 0$.

2.2 The Frictionless Benchmark: Centralized Exchange

Consider a frictionless environment in which there is a competitive, centralized market where investors can buy or sell the asset instantly at some price p_t, which must be constant in equilibrium since the cross-sectional distribution of types in the population is time-invariant.

In such an environment, the objective of an investor is to choose a finite variation asset-holding process $q_t \in \{0, 1\}$ that is progressively measurable with respect to the filtration generated by his utility-type process, and which maximizes

$$\mathbb{E}_{0,\delta}\left[\int_0^\infty e^{-rt}\delta_t q_t dt - \int_0^\infty e^{-rt}p\,dq_t\right] = pq_0 + \mathbb{E}_{0,\delta}\left[\int_0^\infty e^{-rt}(\delta_t - rp)q_t dt\right],$$

where the equality follows from integration by parts. This representation of an investor's objective makes it clear that, at each time t, optimal holdings satisfy

$$q_t^\star = \begin{cases} 0 & \text{if } \delta_t < rp \\ 0 \text{ or } 1 & \text{if } \delta_t = rp \\ 1 & \text{if } \delta_t > rp. \end{cases}$$

This immediately implies that, in equilibrium, the asset is allocated at each time to the investors who value it most. As a result, the distribution of types among investors who own one unit of the asset is time invariant and given by

$$\Phi_1^\star(\delta) = \max\left\{0, F(\delta) - (1-s)\right\} \equiv (F(\delta) - (1-s))^+.$$

It now follows from (1) that the distribution of utility types among investors who do not own the asset is explicitly given by $\Phi_0^\star(\delta) = \min\{F(\delta), 1 - s\}$.

The "marginal" type—i.e., the utility type of the investor who has the lowest valuation among all owners of the asset—is then defined by

$$\delta^\star = \inf\left\{\delta \in [0, 1] : 1 - F(\delta) \leq s\right\},$$

and the equilibrium price of the asset $p^\star$ has to equal $\delta^\star/r$, i.e., the present value of the utility flows enjoyed by a hypothetical investor who holds the asset forever and whose utility type is constantly equal to the marginal type.[9]

3 Equilibrium with search frictions

We now characterize the equilibrium with search frictions in three steps. First, in Section 3.1, we derive the reservation value of an investor with utility type δ, which allows us to characterize optimal trading rules and equilibrium asset prices given the joint distribution of utility types and asset holdings. Then, in Section 3.2, we use the trading rules to derive these joint distributions explicitly. Finally, in Section 3.3, we construct the unique equilibrium and show that it converges to a steady state from any initial allocation.

3.1 Reservation values

Let $V_{q,t}(\delta)$ denote the maximum attainable utility of an investor with $q \in \{0, 1\}$ units of the asset and utility type $\delta \in [0, 1]$ at time $t \geq 0$, and denote this investor's *reservation value* by[10]

$$\Delta V_t(\delta) \equiv V_{1,t}(\delta) - V_{0,t}(\delta).$$

In addition to considering an arbitrary distribution of utility types, our analysis of reservation values improves on the existing literature in several dimensions. First, in Section 3.1.1, we depart from the usual guess-and-verify approach by establishing elementary properties of reservation values directly, without making any *a priori* assumption on the direction of gains from trade. This allows us, down the road in Theorem 1, to claim a general uniqueness result for equilibrium. Second, in Section 3.1.2, we study a differential representation of reservation values which generalizes an earlier closed-form solution for the trading surplus in DGP's two-type model. Third, in Section 3.1.3, we study a sequential representation of reservation values which generalizes the concept of a marginal investor to an asset market with search-and-matching frictions.

---

[9]For simplicity, we will ignore throughout the paper the non-generic case where $F(\delta)$ is flat at the level $1 - s$ because, in such cases, the frictionless equilibrium price is not uniquely defined.

[10]Note that the reservation value function is well defined for all $\delta \in [0, 1]$, and not only for those utility types in the support of the underlying distribution, $F(\cdot)$.

3.1.1 Elementary properties

An application of Bellman's principle of optimality shows that

$$
V_{1,t}(\delta) = \mathbb{E}_t \Bigg[\int_t^\tau e^{-r(u-t)} \delta du + e^{-r(\tau-t)} \bigg(\mathbf{1}_{\{\tau=\tau_1\}} V_{1,\tau}(\delta) \tag{3}
$$

$$
+ \mathbf{1}_{\{\tau=\tau_\gamma\}} \int_0^1 V_{1,\tau}(\delta') dF(\delta')
$$

$$
+ \mathbf{1}_{\{\tau=\tau_0\}} \int_0^1 \max\{V_{1,\tau}(\delta), V_{0,\tau}(\delta) + P_\tau(\delta, \delta')\} \frac{d\Phi_{0,\tau}(\delta')}{1-s} \bigg) \Bigg],
$$

where τ_γ is an exponential random variable with parameter γ that represents the arrival of a preference shock, τ_q is an exponential random variable with parameter λs if $q = 1$ and $\lambda(1-s)$ if $q = 0$ that represents the occurrence of a meeting with a randomly selected investor who owns q units of the asset, the expectation is conditional on $\tau \equiv \min\{\tau_0, \tau_1, \tau_\gamma\} > t$, and

$$
P_\tau(\delta, \delta') \equiv \theta_0 \Delta V_\tau(\delta) + \theta_1 \Delta V_\tau(\delta') \tag{4}
$$

denotes the Nash solution to the bargaining problem at time τ between an asset owner of utility type δ and a non-owner of utility type δ'. Substituting the price (4) into (3) and simplifying shows that the maximum attainable utility of an asset owner satisfies

$$
V_{1,t}(\delta) = \mathbb{E}_t \Bigg[\int_t^\tau e^{-r(u-t)} \delta du + e^{-r(\tau-t)} \bigg(V_{1,\tau}(\delta)
$$

$$
+ \mathbf{1}_{\{\tau=\tau_\gamma\}} \int_0^1 (V_{1,\tau}(\delta') - V_{1,\tau}(\delta)) dF(\delta')
$$

$$
+ \mathbf{1}_{\{\tau=\tau_0\}} \int_0^1 \theta_1 \left(\Delta V_\tau(\delta') - \Delta V_\tau(\delta) \right)^+ \frac{d\Phi_{0,\tau}(\delta')}{1-s} \bigg) \Bigg]. \tag{5}
$$

The first term on the right-hand side of (3) accounts for the fact that an asset owner enjoys a constant flow of utility at rate δ until time τ. The remaining terms capture the three possible events for the asset owner at the stopping time τ: he can receive a preference shock ($\tau = \tau_\gamma$), in which case a new utility type is drawn from the distribution $F(\delta')$; he can meet another asset owner ($\tau = \tau_1$), in which case there are no gains from trade and his continuation payoff is $V_{1,\tau}(\delta)$; or he can meet a non-owner ($\tau = \tau_0$), who is of type δ' with probability $d\Phi_{0,\tau}(\delta')/(1-s)$, in which

case he sells the asset if the payoff from doing so exceeds the payoff from keeping the asset and continuing to search.

Proceeding in a similar way for $q = 0$ shows that the maximum attainable utility of an investor who does not own an asset satisfies

$$V_{0,t}(\delta) = \mathbb{E}_t \left[e^{-r(\tau-t)} \left(V_{0,\tau}(\delta) + \mathbf{1}_{\{\tau=\tau_\gamma\}} \int_0^1 (V_{0,\tau}(\delta') - V_{0,\tau}(\delta)) dF(\delta') \right.\right.$$
$$\left.\left. + \mathbf{1}_{\{\tau=\tau_1\}} \int_0^1 \theta_0 \left(\Delta V_\tau(\delta) - \Delta V_\tau(\delta') \right)^+ \frac{d\Phi_{1,\tau}(\delta')}{s} \right) \right], \qquad (6)$$

and subtracting (6) from (5) shows that the reservation value function satisfies the autonomous dynamic programming equation

$$\Delta V_t(\delta) = \mathbb{E}_t \left[\int_t^\tau e^{-r(u-t)} \delta du + e^{-r(\tau-t)} \left(\Delta V_\tau(\delta) \right.\right.$$
$$+ \mathbf{1}_{\{\tau=\tau_\gamma\}} \int_0^1 (\Delta V_\tau(\delta') - \Delta V_\tau(\delta)) dF(\delta')$$
$$+ \mathbf{1}_{\{\tau=\tau_0\}} \int_0^1 \theta_1 \left(\Delta V_\tau(\delta') - \Delta V_\tau(\delta) \right)^+ \frac{d\Phi_{0,\tau}(\delta')}{1-s}$$
$$\left.\left. - \mathbf{1}_{\{\tau=\tau_1\}} \int_0^1 \theta_0 \left(\Delta V_\tau(\delta) - \Delta V_\tau(\delta') \right)^+ \frac{d\Phi_{1,\tau}(\delta')}{s} \right) \right]. \qquad (7)$$

This equation reveals that an investor's reservation value is influenced by two distinct option values, which have opposing effects. On the one hand, an investor who owns an asset has the option to search and find a non-owner who will pay even more for the asset; as shown on the third line, this option increases her reservation value. On the other hand, an investor who does not own an asset has the option to search and find an owner who will sell at an even lower price; as shown on the fourth line, this option decreases her willingness to pay and, hence, her reservation value.

To guarantee the global optimality of the trading decisions induced by (5) and (6), we further require that the maximum attainable utilities of owners and non-owners, and hence the reservation values, satisfy the transversality conditions

$$\lim_{t\to\infty} e^{-rt} V_{q,t}(\delta) = \lim_{t\to\infty} e^{-rt} \Delta V_t(\delta) = 0, \qquad (q,\delta) \in \{0,1\} \times [0,1]. \qquad (8)$$

The next proposition establishes the existence, uniqueness, and some elementary properties of solutions to (5), (6), and (7) that satisfy (8).

Proposition 1 *There exists a unique function $\Delta V : \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ that satisfies (7) subject to (8). This function is uniformly bounded, absolutely continuous in $(t,\delta) \in \mathbb{R}_+ \times [0,1]$, and strictly increasing in $\delta \in [0,1]$ with a uniformly bounded derivative with respect to type. Given $\Delta V_t(\delta)$, there are unique functions $V_{0,t}(\delta)$ and $V_{1,t}(\delta)$ that satisfy (5), (6), and (8).*

The fact that reservation values are strictly increasing in δ implies that, when an asset owner of type δ meets a non-owner of type $\delta' > \delta$, they will always agree to trade. Indeed, these two investors face the same distributions of future trading opportunities and preference shocks. Thus, the only relevant difference between them is the difference in utility flow enjoyed from the asset, which implies that the reservation value of an investor of type δ' is strictly larger than that of an investor of type $\delta < \delta'$. The monotonicity property holds regardless of the distributions $\Phi_{q,t}(\delta)$, which investors take as given when calculating their optimal trading strategy. Moreover, as we establish below, this property greatly simplifies the derivation of closed-form solutions for both reservation values and the equilibrium distribution of asset holdings and utility types.

3.1.2 Differential representation

Integrating both sides of (7) with respect to the distribution of τ, and using the fact that reservation values are strictly increasing in utility type, we obtain that the reservation value function satisfies the integral equation

$$
\Delta V_t(\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)} \bigg(\delta + \lambda \Delta V_u(\delta) + \gamma \int_0^1 \Delta V_u(\delta') dF(\delta') \tag{9}
$$
$$
+ \lambda \int_\delta^1 \theta_1 \left(\Delta V_u(\delta') - \Delta V_u(\delta) \right) d\Phi_{0,u}(\delta')
$$
$$
- \lambda \int_0^\delta \theta_0 \left(\Delta V_u(\delta) - \Delta V_u(\delta') \right) d\Phi_{1,u}(\delta') \bigg) du.
$$

In addition, since Proposition 1 establishes that the reservation value function is absolutely continuous in $(t,\delta) \in \mathbb{R}_+ \times [0,1]$ with a bounded derivative with respect to type, we know that

$$
\Delta V_t(\delta) = \Delta V_t(0) + \int_0^\delta \sigma_t(\delta') d\delta' \tag{10}
$$

for some nonnegative and uniformly bounded function $\sigma_t(\delta)$ that is itself absolutely continuous in time for almost every $\delta \in [0,1]$. We naturally interpret this function as a measure of the local

surplus in the decentralized market, since the gains from trade between a seller of type δ and a buyer of type $\delta + d\delta$ are approximately given by $\sigma_t(\delta)d\delta$.

Substituting the representation (10) into (9), changing the order of integration, and differentiating both sides of the resulting equation with respect to t and δ reveals that the local surplus satisfies the Hamilton-Jacobi-Bellman (HJB) equation

$$(r + \gamma + \lambda\theta_1(1 - s - \Phi_{0,t}(\delta)) + \lambda\theta_0\Phi_{1,t}(\delta))\,\sigma_t(\delta) = 1 + \dot{\sigma}_t(\delta) \tag{11}$$

at almost every point of $\mathbb{R}_+ \times [0, 1]$. To develop some intuition for this equation, consider the steady-state equilibrium characterized by DGP with two utility types, $\delta_\ell \leq \delta_h$. In that equilibrium, the measures $1 - s - \Phi_0(\delta)$ and $\Phi_1(\delta)$ are constant over $[\delta_\ell, \delta_h)$ and correspond to the mass of buyers and sellers, respectively, which DGP denote by μ_{hn} and $\mu_{\ell o}$. Using this property, integrating both sides of (11), and restricting attention to the steady state gives the surplus formula of DGP:

$$(r + \gamma + \lambda\theta_1\mu_{hn} + \lambda\theta_0\mu_{\ell o})\,(\Delta V(\delta_h) - \Delta V(\delta_\ell)) = \delta_h - \delta_\ell.$$

Hence, our local surplus $\sigma_t(\delta)$ is a direct generalization of the trading surplus in DGP to non-stationary environments with arbitrary distributions of utility types.

Given (11) we can now derive a closed-form solution for reservation values. A calculation provided in the Appendix shows that, together with the requirements of boundedness and absolute continuity in time, equation (11) uniquely pins down the local surplus as

$$\sigma_t(\delta) = \int_t^\infty e^{-\int_t^u \left(r + \gamma + \lambda\theta_1(1 - s - \Phi_{0,\xi}(\delta)) + \lambda\theta_0\Phi_{1,\xi}(\delta)\right)d\xi}\,du. \tag{12}$$

Combining this explicit solution for the local surplus with (9) and (10) allows us to derive the reservation value function in closed-form.

Proposition 2 *For any distributions $\Phi_{0,t}(\delta)$ and $\Phi_{1,t}(\delta)$ satisfying (1) and (2), the unique solution to (7) and (8) is explicitly given by*

$$\Delta V_t(\delta) = \int_t^\infty e^{-r(u-t)}\bigg(\delta - \int_0^\delta \sigma_u(\delta')\,(\gamma F(\delta') + \lambda\theta_0\Phi_{1,u}(\delta'))\,d\delta' \tag{13}$$

$$+ \int_\delta^1 \sigma_u(\delta')\,(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_{0,u}(\delta')))\,d\delta'\bigg)du,$$

where the local surplus $\sigma_t(\delta)$ is defined by (12).

We close this sub-section with several intuitive comparative static results for reservation values.

Corollary 1 *For any* $(t, \delta) \in \mathbb{R}_+ \times [0, 1]$, *the reservation value* $\Delta V_t(\delta)$ *increases if an investor can bargain higher selling prices (larger* θ_1*), if he expects to have higher future valuations (a first-order stochastic dominance shift in* $F(\delta')$*), or if he expects to trade with higher-valuation counterparts (a first-order stochastic dominance shift in the path of either* $\Phi_{0,t'}(\delta')$ *or* $\Phi_{1,t'}(\delta')$*).*

To complement these results, note that an increase in the search intensity, λ, can either increase or decrease reservation values. This is because of the two option values discussed above: an increase in λ increases an owner's option value of searching for a buyer who will pay a higher price, which drives the reservation value up, but it also increases a non-owner's option value of searching for a seller who will offer a lower price, which has the opposite effect. As we will see below in Section 5 the net effect is ambiguous and depends on all parameters of the model.

3.1.3 Sequential representation

Differentiating both sides of (9) with respect to time shows that the reservation value function can be characterized as the unique bounded and absolutely continuous solution to the HJB equation

$$r\Delta V_t(\delta) = \delta + \Delta \dot{V}_t(\delta) + \gamma \int_0^1 \left(\Delta V_t(\delta') - \Delta V_t(\delta)\right) dF(\delta') \tag{14}$$

$$+ \lambda \int_\delta^1 \theta_1 \left(\Delta V_t(\delta') - \Delta V_t(\delta)\right) d\Phi_{0,t}(\delta') + \lambda \int_0^\delta \theta_0 \left(\Delta V_t(\delta') - \Delta V_t(\delta)\right) d\Phi_{1,t}(\delta').$$

The following proposition shows that the solution to this equation can be represented as the present value of utility flows from the asset to a hypothetical investor whose utility type process is adjusted to reflect the frictions present in the market.

Proposition 3 *The reservation value function can be represented as*

$$\Delta V_t(\delta) = \mathbb{E}_{t,\delta}\left[\int_t^\infty e^{-r(s-t)} \hat{\delta}_s ds\right], \tag{15}$$

where the market-valuation process, $\hat{\delta}_t$, *is a pure jump Markov process on* $[0, 1]$ *with infinitesimal generator defined by*

$$\mathcal{A}_t[v](\delta) \equiv \int_0^1 \left(v(\delta') - v(\delta)\right) \left(\gamma dF(\delta') + \mathbf{1}_{\{\delta' > \delta\}} \lambda \theta_1 d\Phi_{0,t}(\delta') + \mathbf{1}_{\{\delta' \leq \delta\}} \lambda \theta_0 d\Phi_{1,t}(\delta')\right)$$

for any uniformly bounded function $v : [0, 1] \to \mathbb{R}$.

15

Representations such as (15) are standard in frictionless asset pricing, where private values are obtained as the present value of cash flows under a probability constructed from marginal rates of substitution. The emergence of such a representation in a decentralized market is, to the best of our knowledge, new to this paper and can be viewed as generalizing the concept of the marginal investor. In the frictionless benchmark, the market valuation is constant and equal to the utility flow of the marginal investor, $\delta^\star$, since investors can trade instantly at price $\delta^\star/r$. In a decentralized market, the market valuation differs from $\delta^\star$ for two reasons. First, because meetings are not instantaneous, an owner must enjoy his private utility flow until he finds a trading partner. Second, investors do not always trade with the marginal type. Instead, the terms of trade are random and depend on the distribution of types among trading partners. Importantly, this second channel is only active if there are more than two utility types, because otherwise a single price gets realized in bilateral meetings.

3.2 The joint distribution of asset holdings and types

In this section, we provide a closed-form characterization of the joint equilibrium distribution of asset holdings and utility types, in and out of steady state. To the best of our knowledge, this characterization is new to the literature, even for the special two-type case studied in DGP. We then establish that this distribution converges to the steady-state from any initial conditions satisfying (1) and (2). Finally, we discuss several properties of a steady-state distribution and explain how its shape depends on the arrival rates of preference shocks and trading opportunities.

Since reservation values are increasing in utility type, trade occurs between two investors if and only if one is an owner with utility type δ' and the other is a non-owner with utility type $\delta'' \geq \delta'$. Investors with the same utility type are indifferent between trading or not, but whether they trade is irrelevant since they effectively exchange ownership type. As a result, the rate of change in the measure of owners with utility type less than or equal to a given $\delta \in [0, 1]$ satisfies

$$\dot{\Phi}_{1,t}(\delta) = \gamma \left(s - \Phi_{1,t}(\delta) \right) F(\delta) - \gamma \Phi_{1,t}(\delta) \left(1 - F(\delta) \right) - \lambda \Phi_{1,t}(\delta) \left(1 - s - \Phi_{0,t}(\delta) \right). \quad (16)$$

The first term in equation (16) is the inflow due to type-switching: at each instant, a measure $\gamma \left(s - \Phi_{1,t}(\delta) \right)$ of owners with utility type greater than δ draw a new utility type, which is less than or equal to δ with probability $F(\delta)$. A similar logic can be used to understand the second

term, which is the outflow due to type-switching. The third term is the outflow due to trade. In particular, a measure $(\lambda/2)\Phi_{1,t}(\delta)$ of investors who own the asset and have utility type less than δ initiate contact with another investor, and with probability $1 - s - \Phi_{0,t}(\delta)$ that investor is a non-owner with utility type greater than δ, so that trade ensues. The same measure of trades occur when non-owners with utility type greater than δ initiate trade with owners with utility type less than δ, so that the sum equals the third term in (16).[11]

Using (1), we can rewrite (16) as a first-order ordinary differential equation for the measure of asset owners with utility type less or equal to δ:

$$\dot{\Phi}_{1,t}(\delta) = -\lambda\Phi_{1,t}(\delta)^2 - \Phi_{1,t}(\delta)\left(\gamma + \lambda(1 - s - F(\delta))\right) + \gamma sF(\delta). \tag{17}$$

Importantly, this Riccati equation holds for every $\delta \in [0, 1]$ without imposing any regularity conditions on the distribution of utility types. Proposition 4 below provides an explicit expression for the unique solution to this equation and shows that it converges to a unique, globally stable steady state. To state the result, let

$$\Lambda(\delta) \equiv \sqrt{(1 - s + \gamma/\lambda - F(\delta))^2 + 4s(\gamma/\lambda)F(\delta)},$$

and denote by

$$\Phi_1(\delta) = F(\delta) - \Phi_0(\delta) \equiv -\frac{1}{2}\left(1 - s + \gamma/\lambda - F(\delta)\right) + \frac{1}{2}\Lambda(\delta). \tag{18}$$

the steady-state distribution of owners with utility type less than or equal to δ, i.e., the unique, strictly positive solution to $\dot{\Phi}_{1,t}(\delta) = 0$.

Proposition 4 *At any time $t \geq 0$ the measure of asset owners with utility type less than or equal to $\delta \in [0, 1]$ is explicitly given by*

$$\Phi_{1,t}(\delta) = \Phi_1(\delta) + \frac{(\Phi_{1,0}(\delta) - \Phi_1(\delta))\Lambda(\delta)}{\Lambda(\delta) + (\Phi_{1,0}(\delta) - \Phi_1(\delta) + \Lambda(\delta))(e^{\lambda\Lambda(\delta)t} - 1)} \tag{19}$$

and converges pointwise monotonically to the steady-state measure $\Phi_1(\delta)$ defined in (18) from any initial condition satisfying (1) and (2).

---

[11]Note that trading generates positive *gross* inflow into the set of owners with utility type less than δ, but zero *net* inflow. Indeed, a gross inflow arises when a non-owner with utility type $\delta' \leq \delta$ meets an owner with an even lower type $\delta'' < \delta'$. By trading, the previous owner of utility type δ'' leaves the set, but the new owner of utility type δ' enters the same set, resulting in zero net inflow.

To illustrate the convergence of the equilibrium distributions to the steady state, we introduce a simple numerical example, which we will continue to use throughout the text. In this example, the discount rate is $r = 0.05$; the asset supply is $s = 0.5$; the meeting rate is $\lambda = 12$, so that a given investor meets others on average once a month; the arrival rate of preference shocks is $\gamma = 1$, so that investors change type on average once a year; the initial distribution of utility types among asset owners is given by $\Phi_{1,0}(\delta) = sF(\delta)$; and the underlying distribution of utility types is $F(\delta) = \delta^\alpha$ with $\alpha = 1.5$, so that the marginal type is given by $\delta^\star = 0.6299$.

Using this parameterization, the left panel of Figure 1 plots the equilibrium distributions among owners and non-owners at $t = 0$, after one month, after six months, and in the limiting steady state. As time passes, one can see that the assets are gradually allocated toward investors with higher valuations: the distribution of utility types among owners improves in the sense of first-order stochastic dominance (FOSD). Similarly, the distribution of utility types among non-owners deteriorates, in the FOSD sense, indicating that investors with low valuations are less and less likely to hold the asset over time.

Focusing on the steady-state distributions, (18) offers several natural comparative statics that we summarize in the following corollary.

Corollary 2 *For any $\delta \in [0, 1]$, the steady-state measure $\Phi_1(\delta)$ of asset owners with utility type less than or equal to δ is increasing in γ and decreasing in λ.*

Intuitively, as preference shocks become less frequent (i.e., γ decreases) or trading opportunities become more frequent (i.e., λ increases), the asset is allocated to investors with higher valuations more efficiently, which implies an FOSD shift in the distribution of types among owners. In the limit, where types are permanent ($\gamma \to 0$) or trading opportunities are constantly available ($\lambda \to \infty$), the steady state distributions converge to their frictionless counterparts, as illustrated by the right panel of Figure 1, and the allocation is efficient. We return to this frictionless limit in Section 5, when we study the asymptotic price impact of search frictions.

FIGURE 1: Equilibrium distributions
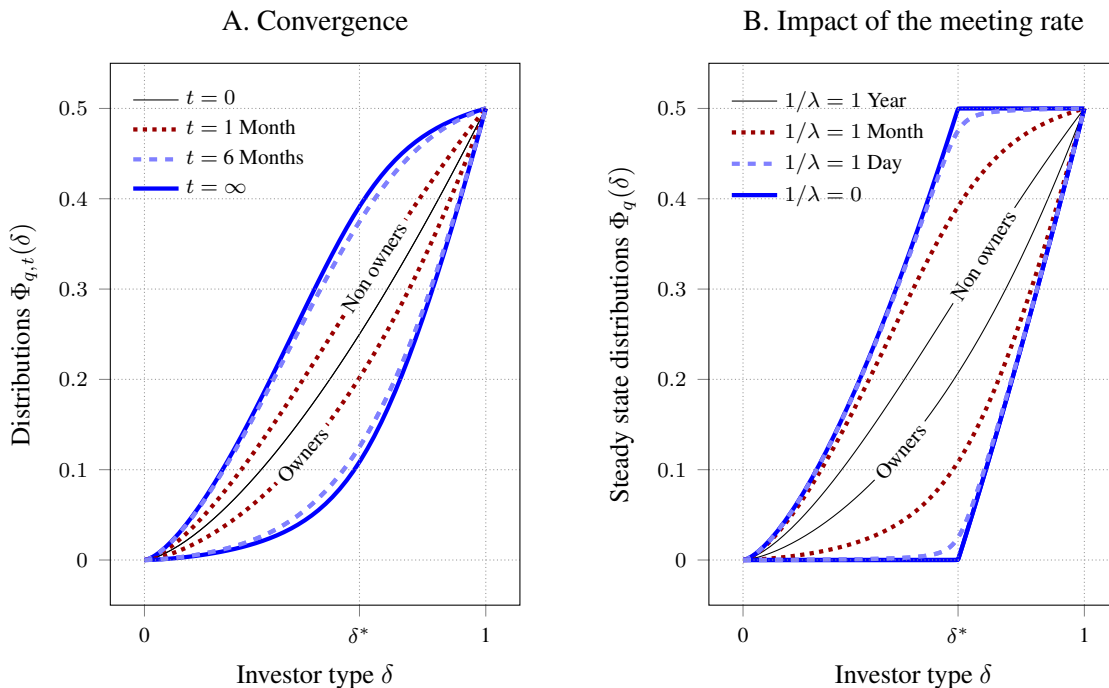
Notes. The left panel plots the cumulative distribution of types among non-owners (upper curves) and owners (lower curves) at different points in time. The right panel plots these distributions in the steady state, for different levels of search frictions, indexed by the average inter-contact time, $1/\lambda$.

3.3 Equilibrium

Definition 1 *An equilibrium is a reservation value function $\Delta V_t(\delta)$ and a pair of distributions $\Phi_{0,t}(\delta)$ and $\Phi_{1,t}(\delta)$ such that the distributions satisfy* (1), (2) *and* (19), *and the reservation value function satisfies* (7) *subject to* (8) *given the distributions.*

Given the analysis above, a full characterization of the unique equilibrium is immediate. Note that uniqueness follows from the fact that we proved reservation values were strictly increasing directly, given arbitrary time paths for the distributions $\Phi_{0,t}(\delta)$ and $\Phi_{1,t}(\delta)$, rather than guessing and verifying that such an equilibrium exists, as was done previously in the literature.

Theorem 1 *There exists a unique equilibrium. Moreover, given any initial conditions satisfying*

19

(1) *and* (2)*, this equilibrium converges to the steady state given by*

$$r\Delta V(\delta) = \delta - \int_0^\delta \sigma(\delta')(\gamma F(\delta') + \lambda\theta_0\Phi_1(\delta'))d\delta' \tag{20}$$

$$+ \int_\delta^1 \sigma(\delta')(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_0(\delta')))d\delta'$$

with the time-invariant local surplus

$$\sigma(\delta) = \frac{1}{r + \gamma + \lambda\theta_1(1 - s - \Phi_0(\delta)) + \lambda\theta_0\Phi_1(\delta)},$$

and the steady-state cumulative distributions defined by (1) *and* (18)*.*

4   Implications of the model

Given our characterization of the steady-state equilibrium above, it is possible to derive many implications of our model analytically. In this section, we will focus on a number of analytical results that are new to the literature, and argue that they are consistent with empirical regularities documented in existing studies of OTC markets. We break this discussion into two parts.

First, in Section 4.1, we follow *investors*: we study how an investor's asset holdings and the frequency with which he trades depend on his utility type, and the implications of these individual trading patterns for aggregate equilibrium outcomes. In particular, we show that investors with utility types near the marginal type $\delta^\star$ account for a disproportionate share of total trading volume, so that a core-periphery trading network emerges endogenously in equilibrium. As we discuss below, core-periphery trading networks are indeed prevalent in many OTC markets.

Second, in Section 4.2, we follow *assets*: we study how an asset is reallocated from investors with low utility types to those with high utility types through a sequence of trades. This sequence corresponds to what the literature often refers to as an "intermediation chain". We first summarize a number of empirical facts about intermediation chains that have been documented using micro-level data from OTC markets—specifically, the relationship between the length of an intermediation chain, the "centrality" of the intermediaries involved in this chain, the speed at which the asset is passed along at each link, and the "markup" between the original purchase price and the final sale price. Then, we derive the theoretical counterparts of these objects and show that our model's predictions are qualitatively consistent with the data.

20

4.1 Following investors: trading intensity and the trading network

In this section, we first establish that, in our model, investors who have the most to gain from trading—i.e., those with extreme utility types and the "wrong" asset holdings—tend to find willing counterparties quickly. An immediate consequence of this seemingly elementary observation is that misallocation clusters around investors with utility types near the marginal type, $\delta^\star$. Since these investors meet relatively frequently with both non-owners with higher utility types than their own and owners with lower utility types than their own, they find themselves intermediating a large fraction of the overall trading volume.

Therefore, even though the network of *meetings* generated by our model is random at any point in time, the network of *trades* is not. In particular, we show that this network has, endogenously, a *core-periphery* structure: over any time interval, if one created a connection between every pair of investors who trade, the network would exhibit what Jackson (2010, p. 67) describes as a "core of highly connected and interconnected nodes and a periphery of less-connected nodes." This type of trading network has been documented in many OTC markets, including the interdealer market for municipal bonds (Green, Hollifield, and Schürhoff, 2006; Li and Schürhoff, 2012), the interdealer market for securitization products (Hollifield, Neklyudov, and Spatt, 2014), the federal funds market (Bech and Atalay, 2010; Afonso and Lagos, 2012), the credit default swap market (Peltonen, Scheicher, and Vuillemey, 2014), several foreign interbank markets (Craig and von Peter, 2014; Boss, Elsinger, Summer, and Thurner, 2004; Chang, Lima, Guerra, and Tabak, 2008), and even interbank flows across Fedwire, the large value transfer system operated by the Federal Reserve (Soramäki, Bech, Arnold, Glass, and Beyeler, 2007).

Trading intensity. The steady-state arrival rate of profitable trading opportunities for an owner with utility type δ, or "selling intensity," is the product of the arrival rate of a meeting and the probability that the investor meets a non-owner with utility type $\delta' \geq \delta$, i.e.,

$$\lambda_1(\delta) \equiv \lambda(1 - s - \Phi_0(\delta)). \tag{21}$$

Similarly, the steady-state arrival rate of profitable trading opportunities for a non-owner with utility type δ, or "buying intensity," is

$$\lambda_0(\delta) \equiv \lambda\Phi_1(\delta). \tag{22}$$

Since $\Phi_0(\delta)$ is non-decreasing, the definition above implies that sellers with a higher utility type trade less often, and thus tend to hold the asset for longer periods. By the same logic, buyers with higher utility types trade more often, and thus tend to remain asset-less for shorter periods.

The left panel of Figure 2 uses the same parameterization of the economic environment as Figure 1 to plot the trading intensities $\lambda_1(\delta)$ and $\lambda_0(\delta)$ as functions of an investor's utility type. In addition to confirming their monotonicity, the figure reveals that the trading intensities fall sharply for owners (non-owners) as their utility type approaches the marginal type $\delta^\star$ from below (above). Intuitively, for sufficiently large λ, the asset allocation becomes close to the frictionless allocation, *especially* at extreme utility types (see Panel B of Figure 1). Hence, owners with utility type $\delta \gg \delta^\star$ and non-owners with utility type $\delta \ll \delta^\star$ have essentially no willing counterparties to trade with, and thus their trading intensities are very low. The figure also reveals that the selling and buying intensities cross at the marginal type. Indeed, $F(\delta^\star) = 1 - s$ when the underlying distribution of utility types is continuous, and it follows that

$$\lambda_1(\delta^\star) = \lambda\left(1 - s - \Phi_0(\delta^\star)\right) = \lambda\left(F(\delta^\star) - \Phi_0(\delta^\star)\right) = \lambda\Phi_1(\delta^\star) = \lambda_0(\delta^\star).$$

Hence, in equilibrium, buyers and sellers whose utility type are close to the marginal type tend to trade at the same speed.

The trading patterns described above illustrate that an investor's utility type endogenously determines his role in the market: those with extreme utility types emerge as natural "customers," trading infrequently and in the same direction, while those with moderate utility types (near $\delta^\star$) emerge as natural "intermediaries," buying and selling more frequently and with approximately equal intensities. As we establish next, these trading patterns have important implications for the tendency of an investor to hold the wrong portfolio, relative to the frictionless benchmark.

FIGURE 2: Trading intensities and equilibrium misallocation
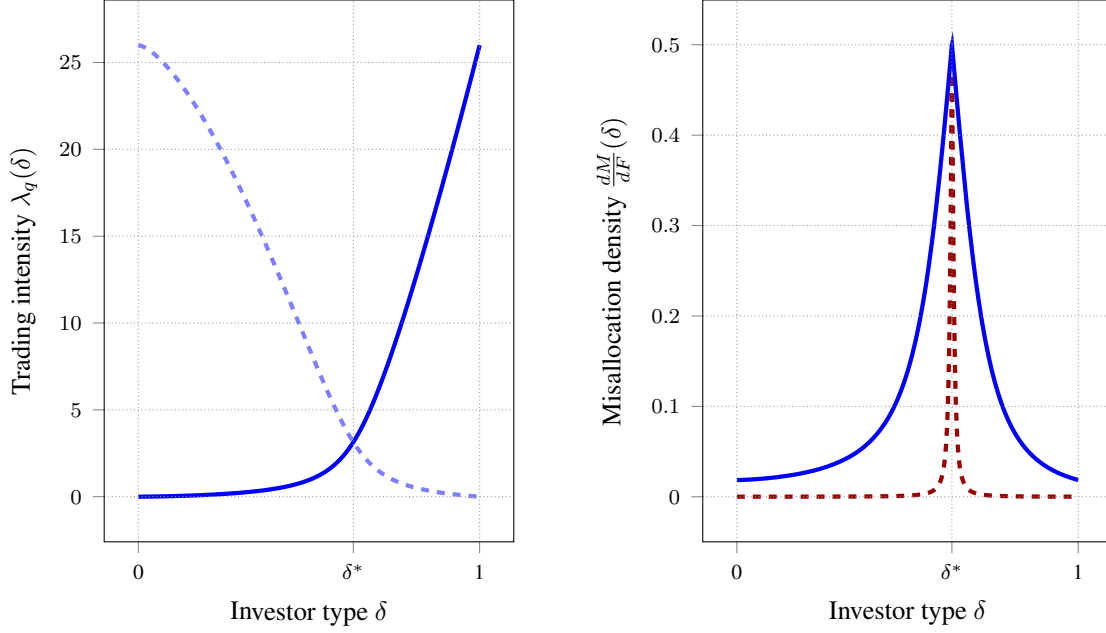
Notes. The left panel plots the trading intensities of owners (dashed) and non owners (solid) when meetings happen on average once every month, while the right panel plots the misallocation density as functions of the investor's utility type when meetings happen on average once every month (solid) and once every hour (dashed). The parameters we use in this figure are otherwise the same as in Figure 1.

Misallocation. We now study *misallocation*, defined as the extent to which the equilibrium asset allocation differs from its frictionless counterpart. To formalize this concept, let

$$M(\delta) = \int_0^\delta \mathbf{1}_{\{\delta' < \delta^\star\}} d\Phi_1(\delta') + \int_0^\delta \mathbf{1}_{\{\delta' \geq \delta^\star\}} d\Phi_0(\delta').$$

This measure is the sum of two types of misallocation: the measure of investors with utility type less than δ who would own the asset in a frictionless environment, but do not own it in the presence of search frictions; and the measure of investors with utility type less than δ who would not own the asset in a frictionless environment, but own it in the presence of search frictions.

To measure the extent of misallocation at a specific utility type, one can simply calculate the Radon-Nikodym density

$$\frac{dM}{dF} = \mathbf{1}_{\{\delta < \delta^\star\}} \frac{d\Phi_1}{dF} + \mathbf{1}_{\{\delta \geq \delta^\star\}} \frac{d\Phi_0}{dF} = \mathbf{1}_{\{\delta < \delta^\star\}} \frac{d\Phi_1}{dF} + \mathbf{1}_{\{\delta \geq \delta^\star\}} \left(1 - \frac{d\Phi_1}{dF}\right) \tag{23}$$

of the misallocation measure with respect to the measure induced by the underlying distribution of utility types; see equation (52) in the Appendix for an explicit expression. The value of the density $\frac{dM}{dF}(\delta)$ represents the *fraction* of investors with utility type δ whose holdings in the environment with search frictions differs from their holdings in the frictionless benchmark.

Lemma 1 *The misallocation density $\frac{dM}{dF}(\delta)$ achieves a global maximum at either $\delta^\star_-$ or $\delta^\star$.*

The misallocation density has two key properties. First, it is non-monotonic and peaks at the marginal type, $\delta^\star$. This arises because the selling intensity is decreasing in utility type, while the buying intensity is increasing. Second, as shown in the right panel of Figure 2, misallocation is highly concentrated near the marginal type. This occurs because there is an equilibrium feedback loop between the trading intensities and the distributions of utility types among owners and non-owners. For example, the non-monotonicity of the misallocation density means that there are relatively more non-owners at low utility types than near the marginal type. This implies that owners with low utility types are able to sell faster than those near the marginal type, which further reduces misallocation away from the marginal type, and increases misallocation near the marginal type. These reinforcing effects ultimately imply that misallocation is not only highest in a neighbourhood of the marginal type but tends to cluster around that point.

We emphasize that these two properties of misallocation arise in a decentralized market because trading intensities differ across utility types. Indeed, when all investors trade with equal intensity—as in frictionless models with centralized markets or in frictional models where all trades are executed by a set of dealers who have access to centralized markets —the measure of misallocation described above would be constant across utility types.

Trading volume and the trading network. Next, we show that the concentration of misallocation translates into a concentration of trading volume near the marginal type. To see this, let us first define trading volume as the flow rate of trades per unit time:

$$\vartheta = \lambda \int_{[0,1]^2} \mathbf{1}_{\{\delta_0 > \delta_1\}} d\Phi_0(\delta_0) d\Phi_1(\delta_1). \tag{24}$$

When the underlying distribution of utility types is continuous, we can use integration by parts to re-write equation (24) as

$$\vartheta = \lambda \Phi_1(\delta^\star)(1 - s - \Phi_0(\delta^\star))$$
$$+ \lambda \int_0^{\delta^\star} dM(\delta) \left(\Phi_0(\delta^\star) - \Phi_0(\delta)\right) + \lambda \int_{\delta^\star}^1 dM(\delta) \left(\Phi_1(\delta) - \Phi_1(\delta^\star)\right),$$

with the misallocation measure defined in (23). The first term represents the volume generated by trades between owners with utility types in $[0, \delta^\star]$ and non-owners with utility types in $[\delta^\star, 1]$; these would be the only trades taking place in the equilibrium of a model with frictionless exchange. With search frictions, however, there are additional infra-marginal trades, captured by the second and third terms. In particular, the second term accounts for trades between owners with utility types $\delta < \delta^\star$ and non-owners with utility types in $[\delta, \delta^\star]$, while the third term accounts for trades between non-owners with utility types $\delta > \delta^\star$ and owners with utility types in $[\delta^\star, \delta]$.

The formula also highlights the role of misallocation in generating extra volume and suggests that near-marginal investors, who are characterized by greater misallocation, are likely to have a larger contribution to trading volume. This is confirmed in the next proposition.

Proposition 5 *Assume that the distribution of utility types is continuous. Then the steady-state trading volume is explicitly given by*

$$\vartheta \equiv \gamma s (1 - s) \left[(1 + \gamma/\lambda) \log \left(1 + \frac{\lambda}{\gamma} \right) - 1 \right]. \tag{25}$$

In particular, the steady-state trading volume is strictly increasing in the meeting rate λ, with $\lim_{\lambda \to \infty} \vartheta = \infty$ and

$$\lim_{\lambda \to \infty} \frac{\lambda}{\vartheta} \left(\int_{\delta^\star - \varepsilon}^{\delta^\star} \Phi_1(\delta) d\Phi_0(\delta) + \int_{\delta^\star}^{\delta^\star + \varepsilon} (1 - s - \Phi_0(\delta)) d\Phi_1(\delta) \right) = 1$$

for any constant $\varepsilon > 0$ such that $\delta^\star \pm \varepsilon \in [0, 1]$.

Proposition 5 establishes two key results. First, when the underlying distribution of utility types is continuous, the equilibrium trading volume is unbounded as $\lambda \to \infty$. By contrast, the equilibrium trading volume is finite in the frictionless benchmark (see Lemma A.1 in the Appendix).

Therefore, our fully decentralized market can generate arbitrarily large excess volume relative to the frictionless benchmark, as long as search frictions are sufficiently small.[12]

Second, trading volume is, for the most part, generated by investors near the marginal type who assume the role of intermediaries; that is, the trading network has a core-periphery structure. To illustrate this phenomenon, Figure 3 plots the contribution

$$\kappa(\delta_0, \delta_1) = \mathbf{1}_{\{\delta_0 > \delta_1\}} \frac{d\Phi_0}{dF}(\delta_0) \frac{d\Phi_1}{dF}(\delta_1)$$

of each owner-non-owner pair to the equilibrium trading volume. From the figure, one can see that investors with extreme utility types account for a small fraction of total trades and, therefore, lie at the periphery of the trading network. For example, owners with low utility types may trade quickly, but there are very few such owners in equilibrium. Hence, these owners contribute little to the trading volume. Likewise, there are many asset owners with high utility types, but these investors trade very slowly, so they do not account for many trades in equilibrium. Only in the cluster of investors with near-marginal utility types do we find a sufficiently large fraction of individuals who are both holding the "wrong" portfolio and able to meet suitable trading partners at a reasonably high rate—these are the investors that make up the core of the trading network.

4.2 Following assets: intermediation chains, centrality, and markups

In this section, we study the manner in which assets are reallocated through sequences of successive trades, or "intermediation chains." We first summarize a number of stylized facts about these chains that have been documented in the literature. Then, we use our model to explicitly derive the theoretical predictions of our model regarding intermediation chains, and show that these predictions are consistent with these facts.

Stylized Facts. To discuss the empirical regularities about intermediation chains, we focus on evidence from the inter-dealer market for municipal bonds, as reported by Green, Hollifield, and Schürhoff (2006) (henceforth GHS) and Li and Schürhoff (2012) (henceforth LS). This data is particularly appealing for two reasons. First, to date, it is among the best transaction-level data

---

[12]Equation (53) also delivers several additional comparative statics. For example, it shows that trading volume peaks when the asset supply equates the number of potential buyers and sellers—which is well-known from the monetary search literature (Kiyotaki and Wright, 1993)—and that it increases when investors change type more frequently.

FIGURE 3: Contribution to trading volume
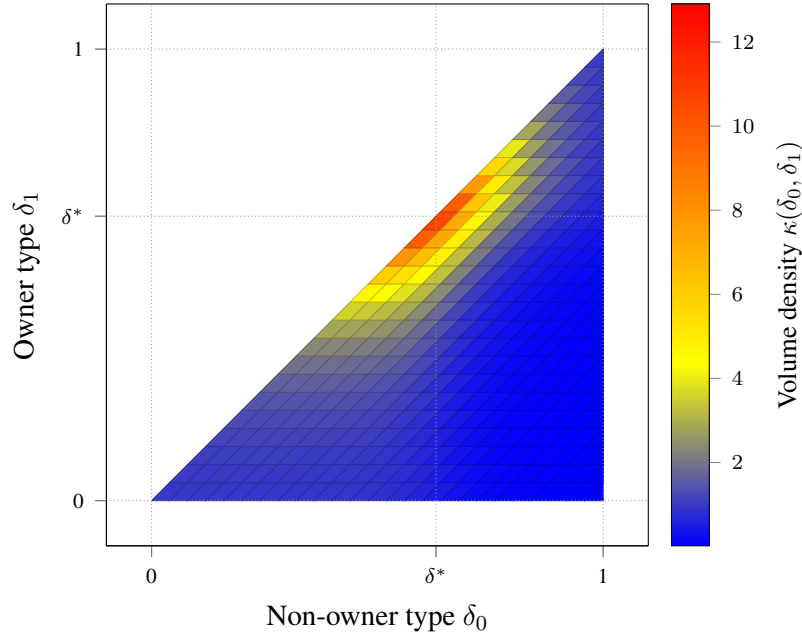
Notes. This figure plots the volume density as a function of the owner's and non-owner's type when meetings occur, on average, once a week. The parameters we use in this figure are otherwise the same as in Figure 1.

from an OTC market. Second, this inter-dealer market is a purely decentralized market, which fits well with the characteristics of our benchmark model.[13]

Using data from the Trade Reporting System of the Municipal Securities Rulemaking Board, GHS and LS observe two types of trades: the trades between dealers and their customers, and the trades among dealers. They first provide evidence of a core-periphery trading network *within the dealer sector* by showing that some dealers are more "central" than others, in the sense that they have more trading links over a given time period. Second, they devise a trade-matching algorithm in order to track blocks of bonds as they are traded among dealers. This algorithm allows them to re-construct the realized intermediation chains. A summary of their findings about dealer centrality and intermediation chains is as follows.

1. More central dealers take longer to sell their inventory (LS, Sec. 5.2).

---

[13]OTC markets with active broker-dealers, on the other hand, might fit better with the extension of our model in Appendix D, where we incorporate explicit marketmakers.

2. More central dealers have larger inventory volatility (LS, Sec. 5.2).

3. Intermediation chains starting with a more central dealer are shorter (LS, Sec. 3.2).

4. Markups increase with the number of dealers in the chain (LS, Sec. 4.1.).

5. Markups increase with the centrality of the first dealer in the chain (LS, Sec. 4.1).

We demonstrate below that our model is consistent with the first three facts for all parameter values. However, the model is only consistent with the last two facts under certain parameter restrictions. This is not necessarily bad news, though, as the relationship between markups, chain length, and the centrality of the first dealer appear to differ across markets. Hence, as we explain below, our environment provides a structural framework to understand *why* these relationships may differ across decentralized asset markets.

Mapping model to data. To confront these stylized facts, we need to formalize the concepts of customers, dealers, intermediation chains, and centrality within the context of our model. It is natural to define a customer as an investor who, over some fixed interval, tends to trade infrequently and in the same direction, while dealers are those investors who trade more frequently and engage in both purchases and sales. This suggests choosing cutoffs $0 < \underline{\delta} < \delta^\star \leq \overline{\delta} < 1$ such that investors with utility types in $[0, \underline{\delta}]$ and $(\overline{\delta}, 1]$ are categorized as customers who sell and buy the asset, respectively, while investors with utility type $\delta \in (\underline{\delta}, \overline{\delta}]$ are categorized as dealers. For the purpose of this discussion, we set $\overline{\delta} = \delta^\star$, so that the customers who buy the asset are the natural holders, i.e., those investors who would hold the asset in a frictionless environment.

There is an obvious caveat to this approach: in the model, investors switch back and forth between assuming the roles of dealer and customer, while in the data this role is more stable. In our analytical and numerical calculations below, we adjust for this instability as follows: when we characterize probability distributions over trade-related random variables, we appropriately condition on the event that the dealers involved in these trades keep stable types (i.e., do not experience a preference shock between buying and selling an asset).[14]

---

[14]This conditioning formalizes the view that, in reality, trading occurs at much higher frequency than type switching. Assuming this view is correct, econometricians of OTC markets are unlikely to observe type switches along intermediation chains. However, even if this view is incorrect, there is an alternative but equivalent interpretation of

Dealer centrality in the model. Proposition 5 showed that investors near the marginal type account for a disproportionate amount of trading volume. Hence, an econometrician who measured centrality using relative trading volume would find that centrality increases as a dealer's utility type approaches $\delta^\star$. However, in LS and other recent empirical studies of decentralized markets, centrality is not measured by relative volume, but instead by the number of trading links established by a dealer over a given time period. This corresponds to the notion of degree centrality in network theory. We now show that our identification of centrality remains the same if we use degree centrality instead of proximity to the marginal type.

To see this, consider the number of trading links established by an investor while keeping a stable utility type, i.e. before he experiences a preference shock. For an investor with utility type δ and asset holdings $q \in \{0, 1\}$, the probability of establishing a trading link before experiencing a preference shock is explicitly given by

$$\pi_q(\delta) \equiv \frac{\lambda_q(\delta)}{\gamma + \lambda_q(\delta)},$$

where $\lambda_q(\delta)$ is the trading intensity, as defined in (21) and (22). Now assume that this investor has established k trading links without ever switching utility type. Conditioning on this event, the investor now has asset holdings $q(k) = q$ if k is even, and $q(k) = 1 - q$ if k is odd, and the probability that he trades once more before switching utility type is explicitly given by $\pi_{q(k)}(\delta)$. Hence, letting $\mathbf{d} \in \{0\} \cup \mathbb{N}$ denote the number of trading links established by the investor before switching utility type, one can easily show that

$$\mathbb{P}\left[\{\mathbf{d} = k\} | \{\delta_0 = \delta, q_0 = q\}\right] = \left(1 - \pi_{q(k)}(\delta)\right) \prod_{n=1}^{k-1} \pi_{q(n)}(\delta). \tag{26}$$

In the left panel of Figure 4, we use this result to plot the probability

$$\mathbb{P}\left[\{\mathbf{d} > k\} | \{\delta_0 = \delta\}\right] = \sum_{q=0}^{1} \mathbb{P}\left[\{\mathbf{d} > k\} | \{\delta_0 = \delta, q_0 = q\}\right] \frac{d\Phi_q}{dF}(\delta)$$

---

the model in which investors have stable types. Namely, one can assume that each investor has a constant utility type but is active in the market for a random period of time that is exponentially distributed with parameter γ. Upon exiting the market, an investor is replaced by another whose utility type is randomly drawn from $F(\delta)$, and who buys the asset at his reservation value if the exiting investor was an asset owner. This alternative model produces the same equilibrium and trading patterns as our original model but has the advantage that the utility type of an investor, and thus his characterization as a customer or a dealer, remains stable over time.

FIGURE 4: Distribution of the number of trading links
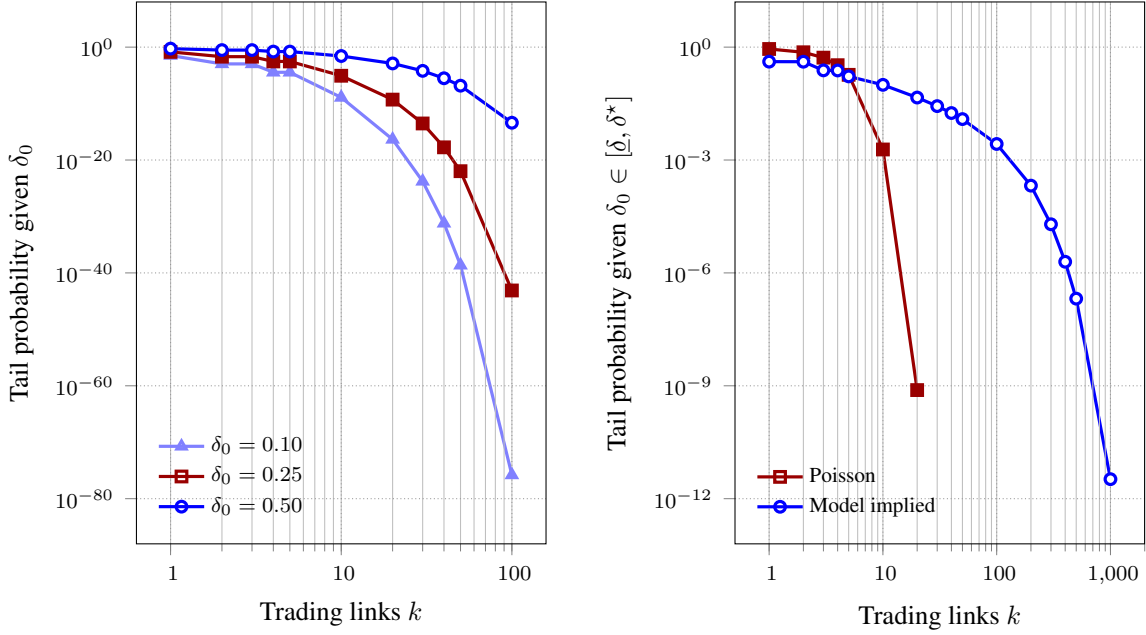
Notes. The left panel plots the probability that a dealer with initial utility type $\delta_0 \in \{0.1, 0.25, 0.5\}$ forms strictly more than a given number of trading links before switching utility type. The right panel plots the same probability but averaged across the dealer space $(\underline{\delta}, \delta^\star]$. To construct this figure we assume that $\underline{\delta} = 0.1$ and that meetings occur on average once every hour. The other parameters are otherwise the same as in Figure 1.

that a dealer with utility type $\delta \in (\underline{\delta}, \delta^\star]$ forms strictly more than $k \in \{0\} \cup \mathbb{N}$ trading links before switching utility type. Consistent with Proposition 5, this figure illustrates that dealers with low valuations tend to establish fewer trading links than dealers with valuations near $\delta^\star$. The following lemma formalizes this result.

Lemma 2 *For any $\delta < \delta' < \delta^\star$ such that $F(\delta) < F(\delta')$, there exists a constant $\bar{\lambda} > 0$ such that the distribution of the random variable $\mathbf{d}$ conditional on $\{\delta_0 = \delta'\}$ dominates the distribution of the random variable $\mathbf{d}$ conditional on $\{\delta_0 = \delta\}$ in the FOSD sense for all $\lambda \geq \bar{\lambda}$.*

Lemma 2 establishes that a dealer with a high utility type is more likely to form more trading links than another dealer whose utility type is lower. This confirms that, within our model, the centrality of a dealer can be adequately measured by his proximity to the marginal type. To conclude, the

30

right panel of Figure 4 plots the tail

$$\mathbb{P}\left[\{\mathbf{d} > k\}|\{\delta_0 \in [\underline{\delta}, \delta^\star]\}\right] = \int_{\underline{\delta}}^{\delta^\star} \mathbb{P}\left[\{\mathbf{d} > k\}|\{\delta_0 = \delta\}\right] \frac{dF(\delta)}{F(\delta^\star) - F(\underline{\delta})}$$

of the degree distribution across the population of dealers and compares it to the tail of a Poisson distribution with the same mean. As in the data of Li and Schürhoff (2012), the figure highlights that the distribution of trading links among dealers has a thick Pareto tail.

Reproducing facts 1 and 2: dealer centrality and inventory. Consider a dealer who owns the asset and has initial utility type $\delta \in (\underline{\delta}, \delta^\star]$. Conditional on type stability, the amount of time that this dealer keeps the asset is exponentially distributed with parameter $\gamma + \lambda_1(\delta)$. Since $\lambda_1(\delta)$ is a decreasing function of the dealer's utility type, it immediately follows that the duration of a dealer's inventory is increasing in the FOSD sense with respect to the dealer's centrality, as measured by the proximity of his utility type to the marginal utility type.[15]

Moreover, since their buying and selling intensities are approximately equal, central dealers also tend to have more volatile inventories. Formally, conditional on keeping a stable utility type, the asset holdings of a dealer follow a continuous-time Markov chain with state space $\{0, 1\}$ that transitions from the state $q = 1$ to the state $q = 0$ with intensity $\lambda_1(\delta)$, and back with intensity $\lambda_0(\delta)$. Hence, the stationary distribution of a dealer's inventory is binomial, with the probability of holding $q = 1$ units of the asset given by

$$\mathbb{P}\left[\{q_t = q\}|\,\mathcal{S}_t \cap \{\delta_0 = \delta\}\right] = \frac{\lambda_0(\delta)}{\lambda_0(\delta) + \lambda_1(\delta)},$$

where $\mathcal{S}_t$ is the event in which the dealer's utility type remains constant over $[0, t]$. It follows that the stationary variance of a dealer's inventory—which is what an empiricist would measure by computing the time-series variance of a dealer's inventory—is

$$\mathrm{Var}\left[q_t = q|\,\mathcal{S}_t \cap \{\delta_0 = \delta\}\right] = \frac{\lambda_0(\delta)\lambda_1(\delta)}{\left(\lambda_0(\delta) + \lambda_1(\delta)\right)^2}.$$

---

[15] At first glance this property might seem inconsistent with the fact that more central dealers tend to form more trading links, but this is not the case. Indeed, conditional on stability, the amount of time that a dealer remains asset-less is exponentially distributed with parameter $\gamma + \lambda_0(\delta)$. Since the buying intensity is an increasing function of the dealer's utility type, it follows that, even though they take longer to sell an asset, more central dealers then buy a new asset more quickly.

As shown above, the buying and selling intensities $\lambda_0(\delta)$ and $\lambda_1(\delta)$ are, respectively, increasing and decreasing with respect to the dealer's utility type, and equal to each other when $\delta = \delta^\star$. As a result, one can easily verify that the stationary variance is an increasing function of the dealer's centrality, as measured by the proximity of his utility type to $\delta^\star$.

The following lemma summarizes the discussion above and confirms that the equilibrium of our search model is consistent with facts 1 and 2.

Lemma 3 *More central dealers have longer inventory duration and larger inventory volatility.*

Reproducing fact 3: intermediation chains and dealer centrality The analysis of the model-implied intermediation chains is more complex because it requires following an asset as it is traded among investors and keeping track of each successive investor's utility type.

Formally, let us fix a given asset and consider the Markov process for the utility type of the asset's owner. This process makes transitions either because the current owner of the asset draws a new utility type, or because a trade has occurred. Let $T^{(k)}$ denote the random time at which the $k-$th transition occurs and denote by $\delta^{(k+1)}$ the utility type of the owner immediately after that transition. Conditional on these two variables, the time until the next transition is

$$T^{(k+1)} - T^{(k)} = \min\left\{\tau_\gamma^{(k+1)}, \tau_0^{(k+1)}\right\},$$

where $\tau_\gamma^{(k+1)}$ is an exponentially distributed random variable with parameter γ that represents the arrival of a preference shock, and $\tau_0^{(k+1)}$ is an exponentially distributed random variable with parameter $\lambda_1(\delta^{(k+1)})$ that represents the occurrence of a sale.

As Figure 5 illustrates, an intermediation chain starts when the asset is sold into the dealer sector by a customer-seller with utility type $[0, \underline{\delta}]$ and ends the first time that the utility type of its holder belongs to the interval $(\delta^\star, 1]$ that identifies the set of customer-buyers. The length of the intermediation chain is then naturally defined as

$$\mathbf{n} = \inf\left\{k \geq 1 : \delta^{(k+1)} > \delta^\star\right\}.$$

As noted above, to ensure that transitions occur through trades rather than preference shocks, we

FIGURE 5: Intermediation chains
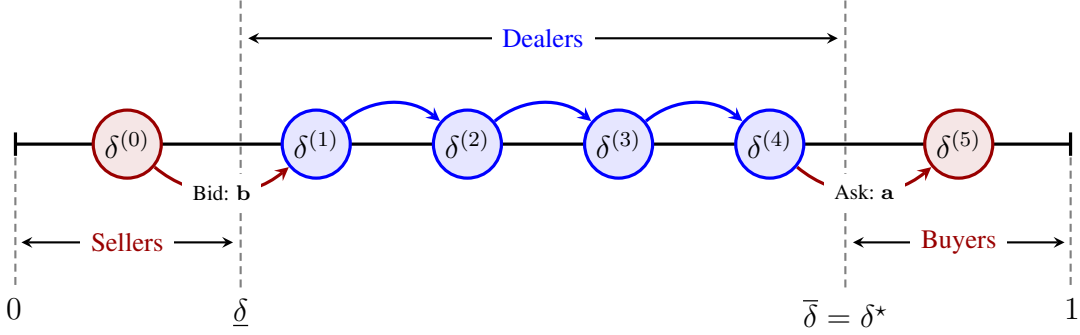
Notes. This figure illustrates an intermediation chain of length $\mathbf{n} = 4$ in which the asset is initially sold by a customer to a dealer at price "Bid" and finally sold by a dealer to a customer at price "Ask".

will condition all our calculations on the stability event defined by

$$\mathcal{S} = \bigcap_{k=1}^{\mathbf{n}} \left\{ \tau_\gamma^{(k)} > \tau_0^{(k)} \right\}.$$

In Appendix A.3, we use Fourier transform techniques to recursively calculate the distribution of the chain length, conditional on the utility type of the first (or "head") dealer in the chain. Though this calculation is quite complex, the distribution itself is surprisingly simple; the next result provides a closed-form characterization of this distribution, and confirms that our model is consistent with the negative relationship between the centrality of the first dealer and the length of the intermediation chain reported by LS.

Proposition 6 *If the distribution of utility types is continuous then*

$$\mathbb{P}\left[\{\mathbf{n} = k\} | \left\{ \delta^{(1)} = \delta \right\} \cap \mathcal{S} \right] = e^{-\Lambda(\delta, \delta^\star)} \frac{\Lambda(\delta, \delta^\star)^{k-1}}{(k-1)!}$$

with the function defined by

$$\Lambda(x, y) \equiv \log\left(\frac{\gamma + \lambda_1(x)}{\gamma + \lambda_1(y)} \right), \qquad x \leq y \leq \delta^\star. \tag{27}$$

In particular, the length of the intermediation chain is decreasing in the FOSD sense with respect to the utility type of the first dealer.

The first part of the proposition shows that, conditional on the utility type of the first dealer being δ, the random variable $\mathbf{n} + 1$ follows a Poisson distribution with parameter $\Lambda(\delta, \delta^\star)$ that reflects the search frictions present in the market and the distance, in the utility type space, between the first dealer and the customer sector. Since the selling intensity $\lambda_1(\delta)$ is a decreasing function it follows from (27) that the Poisson parameter $\Lambda(\delta, \delta^\star)$ is decreasing in the utility type of the first dealer. Clearly, this implies a negative statistical relationship between the chain length, $\mathbf{n}$, and the centrality of the first dealer, as measured by the proximity of his utility type to $\delta^\star$.

Our next result derives the unconditional distribution of the chain length and provides some natural comparative statics.

Corollary 3 *If the distribution of utility types is continuous then*

$$\mathbb{P}\left[\{\mathbf{n} = k\} | \mathcal{S}\right] = \frac{1}{k!} \left(1 - e^{-\Lambda(\underline{\delta}, \delta^\star)}\right)^{-1} \Lambda(\underline{\delta}, \delta^\star)^k.$$

In particular, the chain length is increasing in the FOSD sense with respect to the meeting intensity λ, and decreasing in the FOSD sense with respect to the switching rate γ.

Intuitively, more frequent preference shocks or less frequent trading opportunities coincide with higher levels of misallocation, which make it more likely that an asset owner with a low utility type (a customer-seller) meets an investor with a high utility type who does not already own the asset (a customer-buyer). This diminishes the role of investors with moderate utility types (dealers) in facilitating trade and, ultimately, shortens the intermediation chain.

Reproducing facts 4 and 5: intermediation chain and markups. To analyze markups, we start by defining the bid price

$$\mathbf{b} = \theta_0 \Delta V\left(\delta^{(0)}\right) + \theta_1 \Delta V\left(\delta^{(1)}\right), \tag{28}$$

which is the price at which at which the asset is purchased from a customer by the first dealer in the chain, and the ask price

$$\mathbf{a} = \theta_0 \Delta V\left(\delta^{(\mathbf{n})}\right) + \theta_1 \Delta V\left(\delta^{(\mathbf{n}+1)}\right), \tag{29}$$

which is the price at which the asset is eventually sold to a customer by the last dealer in the chain. The markup along a given intermediation chain is defined as $\mathbf{m} = \mathbf{a}/\mathbf{b} - 1$. Following LS, we study

the statistical relationship between the markup and two key characteristics of the intermediation chain: its length, $\mathbf{n}$, and the centrality of the first dealer, $\delta^{(1)}$.

From (28) and (29), we see that the bid and the ask prices are increasing functions of the utility types of the first and last dealers, $\delta^{(1)}$ and $\delta^{(\mathbf{n})}$, and of the utility types of the customers, $\delta^{(0)}$ and $\delta^{(\mathbf{n+1})}$. However, by virtue of random matching, the utility type of the customers are statistically independent from both the length of the chain and the centrality of the first dealer. Hence it is the type of the first (last) dealer that drives the statistical relationship between the bid (ask) and a given chain characteristic. With this in mind, we first calculate the distribution of the first and last dealers' utility types, conditional on chain length.

Proposition 7 *If the distribution of utility types is continuous then*

$$\mathbb{P}\left[\left\{\delta^{(1)} > x\right\}\middle|\left\{\mathbf{n} = k\right\} \cap \mathcal{S}\right] = \left(\frac{\Lambda(x, \delta^\star)}{\Lambda(\underline{\delta}, \delta^\star)}\right)^k.$$

and

$$\mathbb{P}\left[\left\{\delta^{(\mathbf{n})} > y\right\}\middle|\left\{\mathbf{n} = k\right\} \cap \mathcal{S}\right] = 1 - \left(\frac{\Lambda(\underline{\delta}, y)}{\Lambda(\underline{\delta}, \delta^\star)}\right)^k.$$

In particular, the distributions of the utility type of the first and last dealers are, respectively, decreasing and increasing in the FOSD sense with respect to the length of the chain.

The proposition states that a longer intermediation chain is more likely to start at a dealer with a lower utility type, and to end at a dealer with a higher utility type. By (28) and (29), this implies that longer intermediation chains are characterized by lower bids and higher asks.

Next, we condition on the centrality of the first dealer in the chain as measured by his utility type, $\delta^{(1)} \in [\underline{\delta}, \delta^\star]$. Since equilibrium reservation values are strictly increasing (by Proposition 1), it is immediate to see that the bid price increases with the centrality of the first dealer in the chain: when $\delta^{(1)}$ is larger, the first dealer has a larger reservation value, and hence the initial customer sells the asset into the dealer sector at a higher price. To study the ask price we need to derive the distribution of the utility type of the last dealer conditional on the utility of the first dealer.

Proposition 8 *If the distribution of utility types is continuous then*

$$\mathbb{P}\left[\left\{\delta^{(\mathbf{n})} \in (y, \delta^\star]\right\}\middle|\left\{\delta^{(1)} = \delta\right\} \cap \mathcal{S}\right] = 1 - \mathbf{1}_{\{\delta \le y\}}e^{-\Lambda(y, \delta^\star)}.$$

35

This distribution has an atom at the utility type of the first dealer (because of the possibility that $\mathbf{n} = \mathbf{1}$*, in which case the first dealer is also the last dealer in the chain), and is increasing in the FOSD sense with respect to the utility type of the first dealer.*

The proposition states that the type of the last dealer tends to be increasing in the type of the first dealer. Since the ask price is increasing in the type of the last dealer, this implies that the distribution over ask prices is increasing (in the FOSD sense) in the type of the first dealer.

Propositions 7 and 8 imply that the length of an intermediation chain and the centrality of the first dealer in the chain have unambiguous effects on the level of the bid and the ask prices. However, their overall effect on the markup can still be ambiguous. To see this, consider first the effect of the centrality of the first dealer, as measured by the proximity of his utility type $\delta^{(1)}$ to the marginal type $\delta^{\star}$. By (28) and Proposition 8, an increase in $\delta^{(1)}$ tends to increase both the bid and the ask, so the overall effect on the markup is ambiguous. Consider next the effect of chain length. We have shown in Proposition 7 that, all else equal, a longer chain is associated with a lower bid and a larger ask. Indeed, in a longer chain, the utility type of the first dealer tends to be lower and the utility type of the last dealer tends to be higher. However, we cannot conclude that it is necessarily associated with a larger markup because the utility types of the first and the last dealer are themselves statistically related: if the utility type of the first dealer is larger, then that of the last dealer is also larger, and both move the bid and the ask in the same direction.

To determine the sign of the overall effect, we return to (28) and (29). These equations reveal that, in our model, the sign of the effect is in part determined by the parameters θ_0 and $\theta_1 = 1 - \theta_0$ that govern the bargaining power of buyers and sellers. Indeed, these two parameters determine the sensitivity of the bid and the ask to the utility type of the first and last dealers in the chain. For example, when buyers have most of the bargaining power, in that $\theta_0 \simeq 1$, the markup

$$\mathbf{m} = \mathbf{a}/\mathbf{b} - 1 \simeq \frac{\Delta V(\delta^{(\mathbf{n})})}{\Delta V(\delta^{(0)})} - 1$$

mostly depends on the utility types $\delta^{(0)}$ and $\delta^{(\mathbf{n})}$. Since the former is independent of the chain characteristics, and the latter is increasing with the respect to the chain length and the centrality of the first dealer (by Propositions 7 and 8), we expect that in this case markups should be increasing with respect to both characteristics of the chain, as documented by LS. On the other hand, when

sellers have most of the bargaining power, in that $\theta_1 \simeq 1$, the markup

$$\mathbf{m} \simeq \frac{\Delta V(\delta^{(\mathbf{n+1})})}{\Delta V(\delta^{(1)})} - 1$$

mostly depends on the utility types $\delta^{(\mathbf{n+1})}$ and $\delta^{(1)}$. Since the former is independent of the chain characteristics and the latter is decreasing with the respect to the chain length and increasing with respect to the centrality of the first dealer (by Propositions 7 and 8), we expect that in this case markups are increasing with respect to the length of the chain and decreasing with respect to the centrality of the first dealer. This intuition is the basis for the following result.

Proposition 9 *If the distribution of utility types is continuous then for any $\delta < \delta'$ and $k < k'$ there are thresholds $0 < \underline{\theta}_0 \leq \overline{\theta}_0 < 1$ such that:*

$$\mathbb{E}\left[\mathbf{m} \,\middle|\, \left\{\delta^{(1)} = \delta'\right\} \cap \mathcal{S}\right] \leq \mathbb{E}\left[\mathbf{m} \,\middle|\, \left\{\delta^{(1)} = \delta\right\} \cap \mathcal{S}\right] \;\; \textit{for all } \theta_0 \in (0, \underline{\theta}_0), \tag{30a}$$

$$\mathbb{E}\left[\mathbf{m} \,\middle|\, \left\{\delta^{(1)} = \delta'\right\} \cap \mathcal{S}\right] \geq \mathbb{E}\left[\mathbf{m} \,\middle|\, \left\{\delta^{(1)} = \delta\right\} \cap \mathcal{S}\right] \;\; \textit{for all } \theta_0 \in (\overline{\theta}_0, 1), \tag{30b}$$

and

$$\mathbb{E}\left[\mathbf{m} \,\middle|\, \{\mathbf{n} = k'\} \cap \mathcal{S}\right] \geq \mathbb{E}\left[\mathbf{m} \,\middle|\, \{\mathbf{n} = k\} \cap \mathcal{S}\right] \tag{31}$$

for all $\theta_0 \in (0, \underline{\theta}_0) \cup (\overline{\theta}_0, 1)$.

Proposition 9 reveals that our model is only consistent with the relationships between markup, chain length, and centrality of the first dealer that LS find in the municipal bond market for some parameter values, e.g., when θ_0 is sufficiently large. However, this ambiguity is potentially helpful, as these relationships appear to differ across markets. For example, while Bech and Atalay (2010) also find a positive relationship between centrality of the first dealer and markup in their study of the federal funds market, Hollifield, Neklyudov, and Spatt (2014) and Di Maggio, Kermani, and Song (2015) document the existence of a centrality discount in their empirical analysis of the markets for asset-backed securities, mortgage-backed securities, collateralized debt obligations, and corporate bonds. Hence, our model provides a structural framework to explore why these relationships are positive in some markets and negative in others.

5 Fast trading and convergence to the frictionless limit

In this section, we study equilibrium allocations and prices as $\lambda \to \infty$. This is an important exercise for two reasons. First, this is the empirically relevant case in many financial markets, where trading speeds are becoming faster and faster. Second, as we establish below, this exercise highlights the effect of heterogeneity in utility types on equilibrium asset prices. In particular, we show that heterogeneity magnifies the price impact of search frictions, and that this impact is more pronounced on price levels than on price dispersion. Hence, using observed price dispersion to quantify the effect of search frictions because price dispersion can essentially vanish while price levels are still far from their frictionless counterpart.

5.1 The frictionless limit

As a first step, we establish two intuitive, but important, results about the economy as $\lambda \to \infty$: first, that the allocation converges to its frictionless counterpart; and second, that the reservation values of all investors converge to the frictionless equilibrium price, $\delta^\star/r$.

Proposition 10 *As search frictions vanish,* $\lim_{\lambda\to\infty} \Phi_0(\delta) = \Phi_0^\star(\delta)$, $\lim_{\lambda\to\infty} \Phi_1(\delta) = \Phi_1^\star(\delta)$, *and* $\lim_{\lambda\to\infty} \Delta V(\delta) = \delta^\star/r = p^\star$ *for every* $\delta \in [0,1]$.

To understand why reservation values converge to the frictionless equilibrium price, consider the market-valuation process of Proposition 3. Since the equilibrium asset allocation becomes approximately efficient as $\lambda \to \infty$, it becomes very easy for an investor with utility type $\delta < \delta^\star$ ($\delta > \delta^\star$) to sell (buy) an asset, but a lot more difficult to buy (sell) one. In particular, we show in Appendix A.4 that the trading intensities satisfy

$$\lim_{\lambda\to\infty} \lambda_0(\delta) \begin{cases} < \infty & \text{if} \quad \delta < \delta^\star \\ = \infty & \text{if} \quad \delta > \delta^\star \end{cases} \quad \text{and} \quad \lim_{\lambda\to\infty} \lambda_1(\delta) \begin{cases} = \infty & \text{if} \quad \delta < \delta^\star \\ < \infty & \text{if} \quad \delta > \delta^\star \end{cases}.$$

Thus, it follows from Proposition 3 that, starting from below (above) the marginal type, the market-valuation process moves up (down) very quickly as the meeting frequency increases. Taken together, these observations imply that the market-valuation process converges to $\delta^\star$ as $\lambda \to \infty$, and it now follows from the sequential representation (15) that all reservation values converge to the frictionless equilibrium price.

38

5.2 Price levels near the frictionless limit

To analyze the behavior of reservation values and prices near the frictionless limit, we study the behavior of the market-valuation process near the marginal type, which yields the following result.

Proposition 11 *Assume that the distribution of utility types is twice continuously differentiable with a derivative that is bounded away from zero. Then,*

$$\Delta V(\delta) = p^\star + \frac{\pi/r}{F'(\delta^\star)} \left(\frac{1}{2} - \theta_0 \right) \left(\frac{\gamma s(1-s)}{\theta_0 \theta_1} \right)^{\frac{1}{2}} \frac{1}{\sqrt{\lambda}} + o\left(\frac{1}{\sqrt{\lambda}} \right), \tag{32}$$

for all utility types $\delta \in [0, 1]$.

The first term in the expansion follows directly from Proposition 10, since all reservation values converge to the frictionless price $p^\star = \delta^\star/r$. The main result of the proposition is the second term in the expansion, which determines the deviation of reservation values from the frictionless price. To calculate this term, we center the market-valuation process defined in Proposition 3 around its frictionless limit and scale it by its convergence rate, which turns out to be $\sqrt{\lambda}$. This delivers an auxiliary process $\hat{x}_t = \sqrt{\lambda}(\hat{\delta}_t - \delta^\star)$ whose limit distribution can be characterized explicitly, and the second term of the expansion is then obtained by calculating the limit of

$$\sqrt{\lambda} \left(\Delta V(\delta) - p^\star \right) = \mathbb{E}_{\sqrt{\lambda}(\delta - \delta^\star)} \left[\int_0^\infty e^{-rt} \hat{x}_t dt \right].$$

We see from the proposition that the deviation from the frictionless price depends on three key features of our decentralized market model.

The first key feature is the average time it takes near-marginal investors to find counterparties, as measured by $1/\sqrt{\lambda}$. The second key feature is the relative bargaining powers of buyers and sellers, which determine whether the asset is traded at a discount or at a premium: if $\theta_0 > 1/2$, the asset is traded at a discount relative to the frictionless equilibrium price in all bilateral meetings, and vice versa if $\theta_0 < 1/2$. When buyers and sellers have equal bargaining powers, the correction term vanishes and all reservation values are well approximated by the frictionless price, irrespective of the other features of the market. The third feature of the market that matters for reservation values is the heterogeneity among investors in a neighborhood of the marginal type, as measured by the derivative $F'(\delta^\star)$ of the distribution at the marginal type. If the derivative is small, then valuations

are dispersed around the marginal type, gains from trade are large, and bilateral bargaining induces significant deviations from the frictionless equilibrium price. On the contrary, if the derivative is large, then valuations are highly concentrated around the marginal type, gains from trade are small, and prices remain closer to their frictionless limit. Interestingly, a direct calculation shows that the derivative is proportional to the elasticity of the Walrasian demand at the frictionless price,

$$\varepsilon(p^\star) = \left. \frac{p^\star}{F(rp^\star) - 1} \frac{d(1 - F(rp))}{dp} \right|_{p=p^\star} = \delta^\star F'(\delta^\star)/s,$$

keeping in mind that $1 - F(\delta^\star) = s$. Hence, holding the marginal investor and the supply the same, if the Walrasian demand is less elastic, price effects in the decentralized market will be larger. It is intuitive that a less elastic demand magnifies the bilateral monopoly effects at play in our search-and-matching market.

To further emphasize the role of heterogeneity, consider what happens when the continuous distribution of utility types approximates a discrete distribution. In such a case, the cumulative distribution function will approach a step function that is vertical at the marginal type, where demand is perfectly elastic. As a result, the derivative $F'(\delta^\star)$ will approach infinity, and it follows from (32) that the corresponding deviation from the frictionless equilibrium price will be very small. This informal argument can be made precise by working out the asymptotic expansion of reservation values with a discrete distribution of utility types.

Proposition 12 *When the distribution of utility types is discrete, the convergence rate of reservation values to the frictionless equilibrium price is generically equal to* $1/\lambda$.

To understand the different convergence rates in Propositions 11 and 12, consider a sequence of discrete distributions converging weakly to some continuous distribution. A simple argument shows that the corresponding allocations and prices converge to their continuous counterparts, but the asymptotic expansions of reservation values do not. Specifically, the proof of Proposition 12 reveals that, in the expansion with a discrete distribution, the coefficient multiplying $1/\lambda$ diverges as the discrete distribution approaches its continuous limit. This means that convergence is slower and slower. Proposition 11 makes this observation mathematically precise by showing that, in the continuous limit, the convergence rate switches from $1/\lambda$ to $1/\sqrt{\lambda}$.

FIGURE 6: Continuous vs. discrete distribution

A. Price deviation   B. Price dispersion
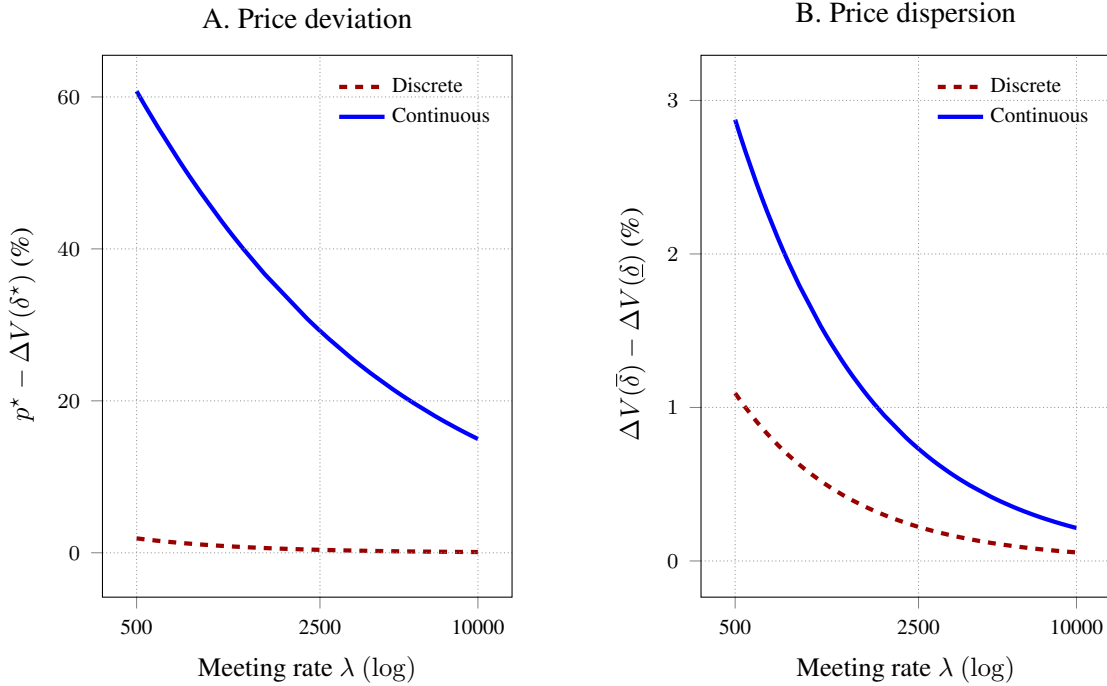
Notes. This figures plots the price deviation relative to the frictionless equilibrium (left panel) and the price dispersion (right panel) as functions of the meeting rate for the base case model of Figure 1 with bargaining power $\theta_0 = 0.75$, and a model with a two point distribution of types constructed to have the same mean and to induce the same marginal investor as the continuous distribution of the base case model.

To see that the difference in convergence rates is economically significant, let us compare the price deviation $p^\star - \Delta V(\delta^\star)$ implied by the continuous distribution of our baseline example with that implied by a two-point distribution, constructed to keep the marginal and average investors the same. The left panel of Figure 6 shows that, when investors meet counterparties twice a day on average (i.e., $\lambda = 500$), the deviation is 60 percent for the continuous distribution, and only about 2 percent for the corresponding discrete distribution. When meetings occur 20 times per day on average (i.e., $\lambda = 10'000$), the deviation is 15 percent for the continuous distribution, but it is now indistinguishable from zero for the discrete distribution. Why is there such a quantitatively large difference in price impact? According to our analysis, the difference is driven by a fundamental economic difference between the two classes of distributions: the elasticity of asset demand is infinite with a discrete distribution, and finite with a continuous one.

5.3 Price dispersion near the frictionless limit

An important implication of Proposition 11 is that, to a first-order approximation, there is no price dispersion. This can be seen by noting that the correction term in (32) does not depend on the investor's utility type. Hence, in order to obtain results about the impact of frictions on price dispersion, it is necessary to work out higher order terms. This is the content of our next result.

Proposition 13 *Assume that the distribution of utility types is twice continuously differentiable with a derivative that is bounded away from zero. Then*

$$\Delta V(1) - \Delta V(0) = \frac{1}{2\theta_0 \theta_1 F'(\delta^\star)} \frac{\log(\lambda)}{\lambda} + O\left(\frac{1}{\lambda}\right),$$

By contrast, with a discrete distribution of utility types, the convergence rate of the price dispersion is generically equal to $1/\lambda$.

Comparing the results of Propositions 11 and 13 shows that, with a continuous distribution of utility types, the price dispersion induced by search frictions vanishes at a rate $\log(\lambda)/\lambda$, which is much faster than the rate $1/\sqrt{\lambda}$ at which reservation values converge to the frictionless equilibrium price. This finding has important implications for empirical analysis of decentralized markets, as it implies that inferring the impact of search frictions based on the observable level of price dispersion can be misleading. In particular, search frictions can have a very small impact on price dispersion and, yet, have a large impact on the equilibrium price level.

This finding is illustrated in Figure 6. Comparing the left and right panels, one sees clearly that the price dispersion induced by search frictions converges to zero much faster than the price deviation. For instance, when investors meet counterparties twice a day on average, the price discount implied by our baseline model is about 60 percent, but the corresponding price dispersion is about 20 times smaller. One can also see from the figure that, in accordance with the result of Proposition 13, price dispersion is larger with a continuous distribution of utility types than with a discrete distribution.[16]

---

[16]In Appendix B, we study the asymptotic welfare cost of misallocation. In line with our results about prices, we show that misallocation has a larger welfare cost when the distribution is continuous than when it is discrete. We also show that the welfare cost of frictions may be accurately measured by the observed amount of price dispersion because these two equilibrium outcomes share the same convergence rate as frictions vanish.

6 Conclusion

In this paper, we develop a search and bargaining model of asset markets that allows investors' utility types to be drawn from an arbitrary distribution. We show that this generalization entails no loss of tractability and has substantial benefits. In particular, the model is able to account for many of the key empirical facts recently reported in studies of OTC markets, which suggests that it could provide a unified structural framework to study a number of important issues such as the effect of trading speed on prices, allocations, and trading volume; the effect of regulation that forces assets trading in an OTC market to trade on a centralized exchange instead; and the propagation of large shocks in a decentralized market. Moreover, the model generates a number of new results, which underscore the importance of heterogeneity in decentralized markets.

References

Milton Abramowitz and Irene A. Stegun. Handbook of Mathematical Functions. Dover, New York, 1964.

Gara Afonso. Liquidity and congestion. Journal of Financial Intermediation, 20(3):324–360, 2011.

Gara Afonso and Ricardo Lagos. An empirical study of trade dynamics in the fed funds market. Technical report, Staff Report, Federal Reserve Bank of New York, 2012.

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

Fernando Alvarez and Gadi Barlevy. Mandatory disclosure and financial contagion. Technical report, Working Paper, Federal Reserve Bank of Chicago, 2014.

Adam Ashcraft and Darrell Duffie. Systemic illiquidity in the federal funds market. American Economic Review, Papers and Proceeding, 97:221–225, 2007.

Andrew Atkeson, Andrea Eisfeldt, and Pierre-Olivier Weill. Entry and exit in otc derivatives markets. Econometrica, 83(6):2231–2292, 2015.

Ana Babus and Peter Kondor. Trading and information diffusion in otc markets. Working Paper FRB Chicago and CEU, 2012.

Morten Bech and Enghin Atalay. The topology of the federal funds market. Physica A, 389:5223–5246, 2010.

Michael Boss, Helmut Elsinger, Martin Summer, and Stefan Thurner. Network topology of the interbank market. Quantitative Finance, 4(6):677–684, 2004.

Kenneth Burdett and Dale T Mortensen. Wage differentials, employer size, and unemployment. International Economic Review, 39(2):257–73, 1998.

Eui Jung Chang, Eduardo José Araujo Lima, Solange Guerra, and Benjamin Tabak. Measures of interbank market structure: An application to brazil. Brazilian Review of Econometrics, 28(2):163–190, 2008.

Jean-Edouard Colliard and Gabrielle Demange. Cash providers: Asset dissemination over intermediation chains. Working Paper, HEC and PSE, 2014.

Ben Craig and Goetz von Peter. Interbank tiering and money center banks. Journal of Financial Intermediation, 23(3):322–347, 2014.

Julienen Cujean and Rémy Praz. Asymmetric information and inventory concerns in over-the-counter markets. Working Paper, University of Maryland, 2013.

Marco Di Maggio, Amir Kermani, and Zhaogang Song. The value of trading relationships in turbulent times. Technical report, Columbia University, 2015.

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

Darrell Duffie, Nicolae Gârleanu, and Lasse H. Pedersen. Valuation in over-the-counter markets. Review of Financial Studies, 20(6):1865, 2007.

Peter Feldhütter. The same bond at different prices: Identifying search frictions and selling pressures. Review of Financial Studies, 25(4):1155–1206, 2012.

Nicolae Gârleanu. Portfolio choice and pricing in illiquid markets. Journal of Economic Theory, 144(2):532–564, 2009.

Alessandro Gavazza. Leasing and secondary markets: Theory and evidence from commercial aircraft. Journal of Political Economy, 119(2):325–377, 2011.

Alessandro Gavazza. An empirical equilibrium model of a decentralized asset market. Working Paper, 2013.

Thomas Gehrig. Intermediation in search markets. Journal of Economics & Management Strategy, 2(1):

97–120, 1993.

Athanasios Geromichalos, Juan Manuel Licari, and José Suárez-Lledó. Monetary policy and asset prices. Review of Economic Dynamics, 10(4):761–779, 2007.

Vincent Glode and Christian Opp. Adverse selection and intermediation chains. 2014. Working Paper, Wharton School, University of Pennsylvania.

Michael Gofman. A network-based analysis of over-the-counter markets. Working Paper, University of Wisconsin-Madison, 2010.

Richard Green, Burton Hollifield, and Norman Schürhoff. Financial intermediation and the costs of trading in an opaque market. Review of Financial Studies, 20:275–314, 2006.

George Hall and John Rust. Middlemen versus market makers: A theory of competitive exchange. Journal of Political Economy, 111(2):353–403, 2003.

Burton Hollifield, Artem Neklyudov, and Chester Spatt. Bid-ask spreads, trading networks and the pricing of securitizations: 144a vs. registered securitization. Working paper, CMU and HEC Lausanne, 2014.

Julien Hugonnier. Speculative behavior in decentralized markets. Working Paper, Swiss Finance Institute, 2012.

Matthew O. Jackson. Social and Economic Networks. Princeton University Press, 2010.

Nobuhiro Kiyotaki and Randall Wright. A search-theoretic approach to monetary economics. American Economic Review, 83(1):63–77, 1993.

Ricardo Lagos. Asset prices and liquidity in an exchange economy. Journal of Monetary Economics, 57(8): 913–930, 2010.

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

Ricardo Lagos and Randall Wright. A unified framework for monetary theory and policy analysis. Journal of political Economy, 113(3):463–484, 2005.

Ricardo Lagos, Guillaume Rocheteau, and Pierre-Olivier Weill. Crises and liquidity in over-the-counter markets. Journal of Economic Theory, 146(6):2169–2205, 2011.

Benjamin Lester and Pierre-Olivier Weill. Over-the-counter markets with continuous valuations. Working Paper, UCLA, 2013.

Benjamin Lester, Andrew Postlewaite, and Randall Wright. Information, liquidity, asset prices, and monetary policy. The Review of Economic Studies, 79(3):1209–1238, 2012.

Benjamin Lester, Guillaume Rocheteau, and Pierre-olivier Weill. Competing for order flow in otc markets. Journal of Money, Credit and Banking, 47(S2):77–126, 2015.

Dan Li and Norman Schürhoff. Dealer networks. Working Paper HEC Lausanne, 2012.

Yiting Li, Guillaume Rocheteau, and Pierre-Olivier Weill. Liquidity and the threat of fraudulent assets. Journal of Political Economy, 120(5):815–846, 2012.

Semyon Malamud and Marzena Rostek. Decentralized exchange. Working Paper, École Polytechnique Fédérale de Lausanne and University of Wisconsin, 2012.

Artem Neklyudov. Bid-ask spreads and the over-the-counter interdealer markets: Core and peripheral dealers. Working Paper HEC Lausanne, 2012.

Ezra Oberfield. Business networks, production chains, and productivity: A theory of input-output architecture. Working Paper, Princeton Universit, 2013.

Emiliano Pagnotta and Thomas Philippon. Competing on speed. Working Paper, NYU Stern School of Business, 2011.

Tuomas A. Peltonen, Martin Scheicher, and Guillaume Vuillemey. The network structure of the cds markets and its determinants. Journal of Financial Stability, 13(118-133), 2014.

Fabien Postel-Vinay and Jean-Marc Robin. Equilibrium wage dispersion with worker and employer heterogeneity. Econometrica, 70(6):2295–2350, 2002.

Rémy Praz. Equilibrium asset pricing with both liquid and illiquid markets. Working Paper, Swiss Finance Institute at EPFL, 2013.

William Reid. Riccati Differential Equations. Mathematics in Science and Engineering. Academic Press, 1972.

Jacob Sagi. Asset-level risk and return in real estate investments. Working paper, UNC Kenan-Flagler Business School, 2015.

Ji Shen, Bin Wei, and Hongjun Yan. Financial intermediation chains in an otc market. 2015.

Kimmo Soramäki, Morten Bech, Jeffrey Arnold, Robert Glass, and Walter Beyeler. The topology of interbank payment flows. Physica A: Statistical Mechanics and its Applications, 379(1):317–333, 2007.

Daniel Spulber. Market making by price-setting firms. The Review of Economic Studies, 63(4):559–580, 1996.

Jiri Sremr. Absolutely continuous functions of two variables in the sense of carathéodory. Electronic Journal of Differential Equations, (154):1–11, 2010.

Alberto Trejos and Randall Wright. Search-based models of money and finance: An integrated approach. Journal of Economic Theory, Forthcoming, 2014.

Semih Üslü. Pricing and liquidity in decentralized asset markets. Technical report, UCLA, 2015.

Dimitri Vayanos and Tan Wang. Search and endogenous concentration of liquidity in asset markets. Journal of Economic Theory, 136(1):66–104, 2007.

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

Pierre-Olivier Weill. Leaning against the wind. The Review of Economic Studies, 74(4):1329–1354, 2007.

Pierre-Olivier Weill. Liquidity premia in dynamic bargaining markets. Journal of Economic Theory, 140 (1):66–96, 2008.

Online Appendix to:

Heterogeneity in Decentralized Asset Markets

J. Hugonnier   B. Lester   P.-O. Weill

A   Proofs

A.1   Volume in the frictionless benchmark

In this section, we briefly study the volume of trade $\vartheta^\star$ that occurs at each instant in the frictionless benchmark equilibrium of Section 2.2. Note that this variable is not uniquely defined. For instance, one can always assume that some investors engage in instantaneous round-trip trades, even if they do not have strict incentives to do so. This leads us to focus on the <u>minimum</u> trading volume necessary to accommodate all investors who have strict incentives to trade.

Lemma A.1 *In the frictionless equilibrium, the minimum volume necessary to accommodate all investors who have strict incentives to trade is given by $\vartheta^\star \equiv \gamma \max\{sF(\delta_-^\star), (1-s)(1-F(\delta^\star))\}$.*

Proof. Consider first the case when there is a point mass at the marginal type, so that $F(\delta^\star) > F(\delta_-^\star)$. In equilibrium, the flow of non-owners who strictly prefer to buy is equal to the set of investors with zero asset holdings who draw a preference shock $\delta' > \delta^\star$. Similarly, the flow investors who own the asset and strictly prefer to sell are those who draw a preference shock $\delta' < \delta^\star$. To implement the equilibrium allocation the volume has to be at least as large as the maximum of these two flows, and the result follows.

    In the continuous case, or more generally when the distribution is continuous at the marginal type, we have $1 - F(\delta^\star) = s$, so that the the minimum volume reduces to $\vartheta^\star = \gamma s(1-s)$.   ■

A.2 Proofs omitted in Section 3

We start by showing that imposing the transversality condition (8) on the reservation value function is equivalent to seemingly stronger requirement of uniform boundedness, and that any such solution to the reservation value equation must be strictly increasing in utility types.

Lemma A.2 *Any solution to* (7) *that satisfies* (8) *is uniformly bounded and strictly increasing in* $\delta \in [0, 1]$.

Proof. To facilitate the presentation we define the operator

$$\mathcal{O}_t[f](\delta) = \int_0^1 \left(f_t(\delta') - f_t(\delta) \right) \left(\gamma dF(\delta') + \lambda \theta_1 \mathbf{1}_{\{f_t(\delta') \geq f_t(\delta)\}} d\Phi_{0,t}(\delta') \right.$$
$$\left. + \lambda \theta_0 \mathbf{1}_{\{f_t(\delta') \leq f_t(\delta)\}} d\Phi_{1,t}(\delta') \right).$$

Integrating with respect to the conditional distribution of the stopping time τ shows that a solution to the reservation value equation (7) is a fixed point of the operator

$$T_t[f](\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)} (\delta + (\gamma + \lambda) f_u(\delta) + \mathcal{O}_u[f](\delta)) du. \tag{33}$$

Assume that $\Delta V_t(\delta) = T_t[\Delta V](\delta)$ is a fixed point that satisfies (8). Since the right-hand side of (33) is absolutely continuous in time, we have that $\Delta V_t(\delta)$ inherits this property, and it thus follows from Lebesgue's differentiation theorem that

$$\dot{\Delta V}_t(\delta) = r \Delta V_t(\delta) - \delta - \mathcal{O}_t[\Delta V](\delta)$$

for every $\delta \in [0, 1]$ and almost every $t \geq 0$. Using this equation together with an integration by parts then shows that the given solution satisfies

$$\Delta V_t(\delta) = e^{-r(H-t)} \Delta V_H(\delta) + \int_t^H e^{-r(u-t)} (\delta + \mathcal{O}_u[\Delta V](\delta)) du \tag{34}$$

$$= \lim_{H \to \infty} \int_t^H e^{-r(u-t)} (\delta + \mathcal{O}_u[\Delta V](\delta)) du \tag{35}$$

for all $(\delta, t) \in \mathcal{S} \equiv \mathbb{R}_+ \times [0, 1]$ and any constant horizon $t \leq H < \infty$ where the second equality follows from the transversality condition. Now assume towards a contradiction that the given solution fails to be nondecreasing in space so that $\Delta V_t(\delta) > \Delta V_t(\delta')$ for some $(t, \delta) \in \mathcal{S}$ and $1 \geq \delta' > \delta$. Because the right-hand side of (33) is absolutely continuous in time, this assumption implies that

$$H^\star \equiv \inf \left\{ u \geq t : \Delta V_u(\delta) \leq \Delta V_u(\delta') \right\} > t.$$

By definition we have that

$$\Delta V_u(\delta) \geq \Delta V_u(\delta'), \qquad t \leq u \leq H^\star \tag{36}$$

and, because the continuous functions $x \mapsto (y - x)^+$ and $x \mapsto -(x - y)^+$ are both non-increasing for every fixed $y \in \mathbb{R}$, it follows that

$$\mathcal{O}_u[\Delta V](\delta) \leq \mathcal{O}_u[\Delta V](\delta'), \qquad t \leq u \leq H^\star. \tag{37}$$

48

To proceed further, we distinguish two cases depending on whether the constant $H^\star$ is finite or not. Assume first that it is finite. In this case it follows from (34) that we have

$$\Delta V_t(\delta) = \int_t^{H^\star} e^{-r(u-t)}(\delta + \mathcal{O}_u[\Delta V](\delta))du + e^{-r(H^\star-t)}\Delta V_{H^\star}(\delta),$$

and combining this identity with (37) then gives

$$\Delta V_t(\delta) \leq \int_t^{H^\star} e^{-r(u-t)}(\delta + \mathcal{O}_u[\Delta V](\delta'))du + e^{-r(H^\star-t)}\Delta V_{H^\star}(\delta)$$
$$= \int_t^{H^\star} e^{-r(u-t)}(\delta + \mathcal{O}_u[\Delta V](\delta'))du + e^{-r(H^\star-t)}\Delta V_{H^\star}(\delta') < \Delta V_t(\delta'), \tag{38}$$

where the equality follows by continuity, and the second inequality follows from the fact that $\delta < \delta'$. Now assume that $H^\star = \infty$ so that (36) and (37) hold for all $u \geq t$. In this case, (35) implies that

$$\Delta V_t(\delta) \leq \lim_{H \to \infty} \int_t^H e^{-r(u-t)}(\delta + \mathcal{O}_u[\Delta V](\delta'))du < \Delta V_t(\delta').$$

Combining this inequality with (38) delivers the required contradiction and establishes that $\Delta V_t(\delta)$ is non-decreasing. To see that it is strictly increasing, rewrite (33) as

$$T_t[f](\delta) = \int_t^\infty e^{-\rho(u-t)}(\delta + \mathcal{M}_u[f](\delta)) \, du. \tag{39}$$

with the operator

$$\mathcal{M}_u[f](\delta) = \lambda\eta f_u(\delta) + \gamma \int_0^1 f_u(\delta')dF(\delta') + \lambda\theta_0 \int_0^1 \min\left\{f_u(\delta'), f_u(\delta)\right\} d\Phi_{1,u}(\delta')$$
$$+ \lambda\theta_1 \int_0^1 \max\left\{f_u(\delta'), f_u(\delta)\right\} d\Phi_{0,u}(\delta'),$$

and the constants $\rho \equiv r + \gamma + \lambda$ and $\eta \equiv 1 - s\theta_0 - (1-s)\theta_1$. Because $\mathcal{M}_u[f](\delta)$ is increasing in $f_u(\delta)$ and the given solution is non-decreasing in space, we have that

$$\Delta V_t(\delta') - \Delta V_t(\delta) = \int_t^\infty e^{-\rho(u-t)}(\delta' - \delta + \mathcal{M}_u[\Delta V](\delta') - \mathcal{M}_u[\Delta V](\delta))du \geq \frac{\delta' - \delta}{\rho}$$

for any $0 \leq \delta \leq \delta' \leq 1$, and the required strict monotonicity follows. To conclude the proof, it remains to establish boundedness. Because the given solution is increasing, we have

$$\sup_{t \geq 0} \mathcal{O}_t[\Delta V](1) \leq 0 \leq \inf_{t \geq 0} \mathcal{O}_t[\Delta V](0)$$

and it now follows from (35) that $0 \leq \Delta V_t(0) \leq \Delta V_t(\delta) \leq \Delta V_t(1) \leq 1/r$ for all $(t, \delta) \in \mathcal{S}$. ∎

Proof of Proposition 1. By Lemma A.2, we have that the existence, uniqueness, and strict (positive) monotonicity of a solution to (7) such that (8) holds is equivalent to the existence and uniqueness of a fixed point of the operator T in the space $\mathcal{X}$ of uniformly bounded, measurable functions from $\mathcal{S}$ to $\mathbb{R}$ equipped with

the sup norm. As is easily seen from (39) we have that T maps $\mathcal{X}$ into itself. On the other hand, using the definition of η together with the fact that the functions $x \mapsto \min\{a; x\}$ and $x \mapsto \max\{a; x\}$ are Lipschitz continuous with constant one for any $a \in \mathbb{R}$ we obtain that

$$\sup_{(t,\delta)\in\mathcal{S}} |\mathcal{M}_t[f](\delta) - \mathcal{M}_t[g](\delta)| \leq (\gamma + \lambda) \sup_{(t,\delta)\in\mathcal{S}} |f_t(\delta) - g_t(\delta)|$$

Combining this bound with (39) then shows that

$$\sup_{(t,\delta)\in\mathcal{S}} |T_t[f](\delta) - T_t[g](\delta)| \leq \left(\frac{\gamma + \lambda}{r + \gamma + \lambda} \right) \sup_{(t,\delta)\in\mathcal{S}} |f_t(\delta) - g_t(\delta)|$$

and the existence of a unique fixed point in the space $\mathcal{X}$ now follows from the contraction mapping theorem because $r > 0$ by assumption. To establish the second part, let $\mathcal{X}_k$ denote the subset of functions $f \in \mathcal{X}$ that are nonnegative and non-decreasing in space with

$$0 \leq f_t(\delta') - f_t(\delta) \leq \frac{\delta' - \delta}{r + \gamma} \equiv k(\delta' - \delta) \tag{40}$$

for all $0 \leq \delta \leq \delta' \leq 1$ and $t \geq 0$. Let further $\mathcal{X}_k^\star$ denote the set of functions $f \in \mathcal{X}_k$ that are strictly increasing in space and absolutely continuous with respect to time and space and observe that, because the set $\mathcal{X}_k$ is closed in $\mathcal{X}$, it suffices to prove that T maps $\mathcal{X}_k$ into $\mathcal{X}_k^\star$. Fix an arbitrary $f \in \mathcal{X}_k$. Since this function is nonnegative, it follows from (39) that $T_t[f](\delta)$ is nonnegative. On the other hand, using the inequalities in (40) in conjunction with the definition of the constant η, the increase of $f_t(\delta)$ and the fact that the functions $x \mapsto \min\{a; x\}$ and $x \mapsto \max\{a; x\}$ are non-decreasing and Lipschitz continuous with constant one, we deduce that

$$0 \leq \mathcal{M}_t[f](\delta'') - \mathcal{M}_t[f](\delta) \leq \lambda k(\delta'' - \delta)$$

for all $0 \leq \delta \leq \delta'' \leq 1$ and $t \geq 0$. Combining these inequalities with (39) and the definition of k then shows that we have

$$\frac{\delta'' - \delta}{r + \gamma + \lambda} \leq T_t[f](\delta'') - T_t[f](\delta) \leq \frac{(1 + \lambda k)(\delta'' - \delta)}{r + \gamma + \lambda} = k(\delta'' - \delta)$$

for all $0 \leq \delta \leq \delta'' \leq 1$ and $t \geq 0$. Taken together, these bounds imply that the function $T_t[f](\delta)$ is strictly increasing in space and belongs to $\mathcal{X}_k$, so it now only remains to establish absolute continuity. By definition of the set $\mathcal{X}_k$ we have that

$$f_t(\delta) = f_t(\delta') + \int_{\delta'}^{\delta} \phi_t(x)dx$$

for all $t \geq 0$, almost every $\delta, \delta' \in [0, 1]^2$, and some $0 \leq \phi_t(x) \leq k$. Substituting this identity into (33) and

changing the order of integration shows that

$$T_t[f](\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)} \bigg(\delta + (\lambda + \gamma)f_u(\delta) - \int_0^\delta \phi_u(\delta')(\gamma F(\delta') + \lambda\theta_0 \Phi_{1,u}(\delta'))d\delta' \tag{41}$$

$$+ \int_\delta^1 \phi_u(\delta')(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_{0,u}(\delta')))d\delta' \bigg) du$$

and the required absolute continuity now follows from Sremr (2010, Theorem 3.1). ∎

Lemma A.3 *Given the reservation value function there exists a unique pair of functions $V_{1,t}(\delta)$ and $V_{0,t}(\delta)$ that satisfy (3) and (6) subject to (8).*

Proof. Assume that $V_{1,t}(\delta)$ and $V_{0,t}(\delta)$ satisfy (3) and (6) subject to (8). Integrating on both sides of (3) and (6) with respect to the conditional distribution of the stopping time τ shows that

$$V_{q,t}(\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)} (\lambda V_{q,u}(\delta) + \mathcal{C}_{q,u}(\delta) + \gamma \int_0^1 V_{q,u}(\delta')dF(\delta'))du. \tag{42}$$

with the uniformly bounded functions defined by

$$\mathcal{C}_{q,t}(\delta) = q\delta + \int_0^1 \lambda\theta_q \left((2q-1)(\Delta V_t(\delta') - \Delta V_t(\delta)) \right)^+ d\Phi_{1-q,t}(\delta'). \tag{43}$$

Because the right-hand side of (42) is absolutely continuous in time, we have that the functions $V_{q,t}(\delta)$ inherit this property, and it thus follows from Lebesgue's differentiation theorem that

$$\dot{V}_{q,t}(\delta) = rV_{q,t}(\delta) - \mathcal{C}_{q,t}(\delta) - \gamma \int_0^1 (V_{q,t}(\delta') - V_{q,t}(\delta))dF(\delta') \tag{44}$$

for all $\delta \in [0,1]$ and almost every $t \geq 0$. Combining this differential equation with the assumed transversality condition then implies that

$$V_{q,t}(\delta) = e^{-r(H-t)}V_{q,H}(\delta) + \int_t^H e^{-r(u-t)}(\mathcal{C}_{q,u}(\delta) + \gamma \int_0^1 (V_{q,u}(\delta') - V_{q,u}(\delta))dF(\delta'))du$$

$$= \lim_{H\to\infty} \int_t^H e^{-r(u-t)}(\mathcal{C}_{q,u}(\delta) + \gamma \int_0^1 (V_{q,u}(\delta') - V_{q,u}(\delta))dF(\delta'))du$$

for any finite horizon and, because the functions $\mathcal{C}_{q,t}(\delta)$ are increasing in space by Lemma A.5 below, the same arguments as in the proof of Lemma A.2 show that the functions $V_{q,t}(\delta)$ are increasing in space and are uniformly bounded. Combining these properties with (44) then shows that the process

$$e^{-rt}V_{q,t}(\delta_t) + \int_0^t e^{-ru}\mathcal{C}_{q,u}(\delta_u)du$$

is a uniformly bounded martingale in the filtration generated by the investor's utility type process, and it follows that we have

$$V_{q,t}(\delta) = \mathbb{E}_{t,\delta} \left[\int_t^\infty e^{-r(u-t)}\mathcal{C}_{q,u}(\delta_u)du \right]. \tag{45}$$

This establishes the uniqueness of the solutions to (3) and (6) subject to (8) and it now only remains to show that these solutions are consistent with the given reservation value function. Applying the law of iterated expectations to (45) at the stopping time τ shows that the function $V_{1,t}(\delta) - V_{0,t}(\delta)$ is a uniformly bounded fixed point of the operator

$$U_t[f](\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)}(\lambda f_u(\delta) + C_{1,u}(\delta) - C_{0,u}(\delta) + \gamma \int_0^1 f_u(\delta')dF(\delta'))du.$$

A direct calculation shows that this operator is a contraction on $\mathcal{X}$ and, therefore, admits a unique fixed point in $\mathcal{X}$. Because the reservation value function is increasing we have

$$C_{1,t}(\delta) - C_{0,t}(\delta) + \gamma \int_0^1 \Delta V_t(\delta')dF(\delta') = \delta + \gamma \Delta V_t(\delta) + \mathcal{O}_t[\Delta V](\delta)$$

and it follows that this fixed point coincides with the reservation value function. ∎

Lemma A.4 *For any fixed $\delta \in [0,1]$, the unique solution to (11) that is both absolutely continuous in time and uniformly bounded is explicitly given by*

$$\sigma_t(\delta) = \int_t^\infty e^{-\int_t^u R_\xi(\delta)d\xi}du, \tag{46}$$

with the effective discount rate $R_t(\delta) = r + \gamma + \lambda\theta_1(1 - s - \Phi_{0,t}(\delta)) + \lambda\theta_0\Phi_{1,t}(\delta)$.

Proof. Fix an arbitrary $\delta \in [0,1]$ and assume that $\sigma_t(\delta)$ is a uniformly bounded solution to (11) that is absolutely continuous in time. Using integration by parts, we easily obtain that

$$\sigma_t(\delta) = e^{-\int_t^T R_\xi(\delta)d\xi}\sigma_T(\delta) + \int_t^T e^{-\int_t^u R_\xi(\delta)d\xi}du, \qquad 0 \le t \le T < \infty.$$

Since $\sigma \in \mathcal{X}$ and $R_t(\delta) \ge \gamma > 0$, we have that

$$\lim_{T\to\infty} e^{-\int_t^T R_\xi(\delta)d\xi}\sigma_T(\delta) = 0$$

and therefore

$$\sigma_t(\delta) = \lim_{T\to\infty}\left(e^{-\int_t^T R_\xi(\delta)d\xi}\sigma_T(\delta) + \int_t^T e^{-\int_t^u R_\xi(\delta)d\xi}du\right) = \int_t^\infty e^{-\int_t^s R_u(\delta)du}ds$$

by monotone convergence. ∎

Lemma A.5 *The functions $C_{q,t}(\delta)$ are increasing in $\delta \in [0,1]$.*

Proof. For $q = 0$ the result follows immediately from (43) and the fact that the reservation value function is increasing in $\delta \in [0,1]$. Assume now that $q = 1$. Using the fact that the reservation value function is increasing and integrating by parts on the right-hand side of equation (43) gives

$$C_{1,t}(\delta) = \delta + \int_\delta^1 \lambda\theta_1\sigma_t(\delta')(1 - s - \Phi_{1,t}(\delta'))d\delta',$$

52

and differentiating this expression shows that

$$\mathcal{C}'_{1,t}(\delta) = 1 - \lambda\sigma_t(\delta)\theta_1(1 - s - \Phi_{1,t}(\delta)) \geq 1 - \frac{\lambda\theta_1(1-s)}{r + \gamma + \lambda(\theta_0 s + \theta_1(1-s))} > 0,$$

where the first inequality follows from (46) and the definition of $R_t(\delta)$, and the last inequality follows from the strict positive of the interest rate. ∎

Proof of Proposition 2. Let the local surplus $\sigma_t(\delta)$ be as above and consider the absolutely continuous function defined by

$$f_t(\delta) = \int_t^\infty e^{-r(u-t)} \left(\delta - \int_0^\delta \sigma_u(\delta') \left(\gamma F(\delta') + \lambda\theta_0\Phi_{1,u}(\delta') \right) d\delta' \right.$$
$$\left. + \int_\delta^1 \sigma_u(\delta') \left(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_{0,u}(\delta')) \right) d\delta' \right) du.$$

Using the uniform boundedness of the functions $\sigma_t(\delta)$, $F(\delta)$, and $\Phi_{q,t}(\delta)$, we deduce that $f \in \mathcal{X}$. On the other hand, Lebesgue's differentiation theorem implies that this function is almost everywhere differentiable in both the time and the space variable with

$$\dot{f}_t(\delta) = r f_t(\delta) - \delta + \int_0^\delta \sigma_t(\delta')(\gamma F(\delta') + \lambda\theta_0\Phi_{1,t}(\delta'))d\delta' \tag{47}$$
$$- \int_\delta^1 \sigma_t(\delta')(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_{0,t}(\delta')))d\delta'$$

for all $\delta \in [0,1]$ and almost every $t \geq 0$, and

$$f'_t(\delta) = \int_t^\infty e^{-r(u-t)} \left(1 - \sigma_u(\delta)(\gamma + \lambda\theta_1(1 - s - \Phi_{0,u}(\delta)) + \lambda\theta_0\Phi_{1,u}(\delta)) \right) du$$
$$= \int_t^\infty e^{-r(u-t)} \left(r\sigma_u(\delta) - \dot{\sigma}_u(\delta) \right) du = \sigma_t(\delta)$$

for all $t \geq 0$ and almost every $\delta \in [0,1]$, where the second equality follows from (11) and the third follows from integration by parts and the boundedness of the local surplus. In particular, the fundamental theorem of calculus implies

$$f_t(\delta') - f_t(\delta) = \int_\delta^{\delta'} \sigma_t(\delta'')d\delta'', \qquad (\delta, \delta') \in [0,1]^2, \tag{48}$$

and it follows that $f_t(\delta)$ is strictly increasing in space. Using this monotonicity in conjunction with (48) and integrating by parts on the right-hand side of (47) shows that

$$\dot{f}_t(\delta) = r f_t(\delta) - \delta - \mathcal{O}_t[f](\delta)$$

for all $\delta \in [0,1]$ and almost every $t \geq 0$. Writing this differential equation as

$$(r + \gamma + \lambda)f_t(\delta) - \dot{f}_t(\delta) = \delta + (\gamma + \lambda)f_t(\delta) + \mathcal{O}_t[f](\delta)$$

53

and integrating by parts then shows that

$$f_t(\delta) = e^{-(r+\gamma+\lambda)(H-t)} f_H(\delta) + \int_t^H e^{-(r+\gamma+\lambda)(u-t)} \left(\delta + (\gamma+\lambda) f_u(\delta) + \mathcal{O}_u[f](\delta) \right) du$$

for any $t \leq H < \infty$, and it now follows from the dominated convergence theorem and the uniform boundedness of the function $f_t(\delta)$ that

$$f_t(\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)} \left(\delta + (\gamma+\lambda) f_u(\delta) + \mathcal{O}_u[f](\delta) \right) du.$$

Comparing this expression with (33), we conclude that $f_t(\delta) = T_t[f](\delta) \in \mathcal{X}$, and the desired result now follows from the uniqueness established in the proof of Proposition 1. ∎

Proof of Corollary 1. As shown in the proof of Proposition 1, we have that $\Delta V_t(\delta)$ is the unique fixed point of the contraction $T : \mathcal{X}_k \to \mathcal{X}_k$ defined by (39) and, by inspection, this mapping is increasing in $f_t(\delta)$ and decreasing in r. Furthermore, it follows from equation (41) that T is increasing θ_1 and decreasing in θ_0, $F(\delta)$ and $\Phi_{q,t}(\delta)$ and the desired monotonicity now follows from Lemma A.6 below. ∎

Lemma A.6 *Let $\mathcal{C} \subseteq \mathcal{X}$ be closed and assume that $A[\cdot;\alpha] : \mathcal{C} \to \mathcal{C}$ is a contraction that is increasing in f and increasing (resp. decreasing) in α. Then its fixed point is increasing (resp. decreasing) in α.*

Proof. Assume that $A_t[f;\alpha](\delta)$ is a contraction on $\mathcal{C} \subset \mathcal{X}$ that is increasing in (α, f) and denote its fixed point by $f_t(\delta;\alpha)$. Combining the assumed monotonicity with the fixed-point property shows that

$$f_t(\delta;\alpha) = A_t[f(\cdot;\alpha);\alpha](\delta) \leq A_t[f(\cdot;\alpha);\beta](\delta), \qquad (t,\delta) \in \mathcal{S}.$$

Iterating this relation gives

$$f_t(\delta;\alpha) \leq A_t^n[f;\beta](\delta), \qquad (t,\delta,n) \in \mathcal{S} \times \{1,2,\ldots\}$$

and the desired result now follows by taking limits on both sides as $n \to \infty$ and using the fact that the mapping $A[\cdot;\beta]$ is a contraction. ∎

Proof of Proposition 3. Using (14) together with the notation of the statement shows that the reservation value function is the unique bounded and absolutely continuous solution to

$$r\Delta V_t(\delta) = \dot{\Delta V}_t(\delta) + \delta + \mathcal{A}_t[\Delta V](\delta).$$

Therefore, it follows from an application of Itô's lemma that the process

$$e^{-rt}\Delta V_t(\hat{\delta}_t) + \int_0^t e^{-ru}\hat{\delta}_u du$$

is a local martingale, and this implies that we have

$$\Delta V_t(\delta) = \mathbb{E}_{t,\delta}\left[e^{-r(\tau_n - t)}\Delta V_{\tau_n}(\hat{\delta}_{\tau_n}) \right] + \mathbb{E}_{t,\delta}\left[\int_0^{\tau_n} e^{-r(u-t)}\hat{\delta}_u du \right]$$

for a non-decreasing sequence of stopping times that converges to infinity. Since the reservation value function is uniformly bounded, we have that the first term on the right-hand side converges to zero as $n \to \infty$, and the desired result now follows by monotone convergence. ∎

Proof of Proposition 4. For a fixed $\delta \in [0,1]$, the differential equation

$$-\dot{\Phi}_{1,t}(\delta) = \lambda \Phi_{1,t}(\delta)^2 + \lambda \Phi_{1,t}(\delta)(1 - s + \gamma/\lambda - F(\delta)) - \gamma s F(\delta)$$

is a Riccati equation with constant coefficients whose unique solution can be found in any textbook on ordinary differential equations; see, for example, Reid (1972). Let us now turn to the convergence part. Using (1) and (2) together with the definition of $\Lambda(\delta)$ and $\Phi_q(\delta)$ shows that the term

$$\Phi_{1,0}(\delta) - \Phi_1(\delta) + \Lambda(\delta) = \Phi_{1,0}(\delta) + \frac{1}{2}(1 - s + \gamma/\lambda - F(\delta) + \Lambda(\delta))$$

$$= \Phi_{1,0}(\delta) + \Phi_1(\delta) + (1 - s + \gamma/\lambda - F(\delta)) = \gamma/\lambda + \Phi_1(\delta) + (1 - s - \Phi_{0,0}(\delta))$$

that appears in the denominator of (19) is nonnegative for all $\delta \in [0,1]$. Since $\lambda \Lambda(\delta) > 0$, this implies that the nonnegative function

$$|\Phi_{1,t}(\delta) - \Phi_1(\delta)| = \frac{|\Phi_{1,0}(\delta) - \Phi_1(\delta)|\Lambda(\delta)}{\Lambda(\delta) + (\Phi_{1,0}(\delta) - \Phi_1(\delta) + \Lambda(\delta))(e^{\lambda \Lambda(\delta)t} - 1)}$$

is monotone decreasing in time and converges to zero as $t \to \infty$. ∎

Lemma A.7 *The steady-state cumulative distribution of types among owners $\Phi_1(\delta)$ is increasing in the asset supply, and increasing and concave in $\phi = \gamma/\lambda$, with*

$$\lim_{\phi \to 0} \Phi_1(\delta) = sF(\delta), \quad \text{and}$$

$$\lim_{\phi \to \infty} \Phi_1(\delta) = (F(\delta) - 1 + s)^+.$$

In particular, the steady-state cumulative distributions functions $\Phi_q(\delta)$ converge to their frictionless counterparts as $\lambda \to \infty$.

Proof of Lemma A.7. A direct calculation shows that

$$\frac{\partial \Phi_1(\delta)}{\partial s} = \frac{\Phi_1(\delta) + \phi F(\delta)}{\Lambda(\delta)}, \tag{49}$$

and the desired monotonicity in s follows. On the other hand, using the definition of the steady-state distribution, it can be shown that

$$\frac{\partial \Phi_1(\delta)}{\partial \phi} = \frac{sF(\delta) - \Phi_1(\delta)}{\Lambda(\delta)} = \frac{s(1-s)F(\delta)(1 - F(\delta))}{(\phi + \Phi_1(\delta) + (1-s)(1 - F(\delta)))\Lambda(\delta)} \tag{50}$$

and the desired monotonicity follows by observing that all the terms on the right-hand side are nonnegative. Knowing that $\Phi_1(\delta)$ is increasing in ϕ, we deduce that

$$\Lambda(\delta) = 2\Phi_1(\delta) + 1 - s + \phi - F(\delta)$$

55

is also increasing in ϕ, and it now follows from the first equality in (50) that

$$\frac{\partial^2 \Phi_1(\delta)}{\partial \phi^2} = -\frac{1}{\Lambda(\delta)}\frac{\partial \Phi_1(\delta)}{\partial \phi}\left(1 + \frac{\partial \Lambda(\delta)}{\partial \phi}\right) \le 0.$$

The expressions for the limiting values follow by sending ϕ to zero and ∞ in the definition of the steady-state distribution. ∎

Proof of Corollary 2. The result follows directly from Lemma A.7. ∎

Proof of Theorem 1. The result follows directly from the definition, Proposition 1, and Proposition 4. We omit the details. ∎

A.3 Proofs omitted in Section 4

To simplify the notation, let $\phi \equiv \gamma/\lambda$. The following lemma follows immediately from the equation defining the steady-state distribution of utility types among asset owners.

Lemma A.8 *The steady-state distributions of types satisfy $\Phi_1(\delta) = F(\delta) - \Phi_0(\delta) = \ell(F(\delta))$, where the bounded function*

$$\ell(x) \equiv -\frac{1}{2}(1 - s + \phi - x) + \frac{1}{2}\sqrt{(1 - s + \phi - x)^2 + 4s\phi x} \tag{51}$$

is the unique positive solution to $\ell^2 + (1 - s + \phi - x)\ell - s\phi x = 0$. Moreover, the function $\ell(x)$ is strictly increasing and convex, and strictly so if $s \in (0, 1)$.

Proof of Lemma A.8. It is obvious that $\ell(x)$ is the unique positive solution of the second-order polynomial shown above. The function $\ell(x)$ is strictly increasing by an application of the implicit function theorem: when $x > 0$, $\ell(x) > 0$, so that the second-order polynomial must be strictly increasing in ℓ and strictly decreasing in x. Convexity follows from the fact that

$$\ell''(x) = \frac{2s(1 - s)\phi\,(1 + \phi)}{\sqrt[3]{4s\phi x + (1 - s + \phi - x)^2}} \ge 0,$$

with a strict inequality if $s \in (0, 1)$. ∎

Proof of Lemma 1. Let $\mathbf{A}$ denote the set of atoms of the distribution $F(\delta)$. With this definition, we have that the Radon-Nikodym density is given by

$$m(\delta) = \frac{dM}{dF}(\delta) \equiv \begin{cases} \mathbf{1}_{\{\delta \notin \mathbf{A}\}}\ell'(F(\delta)) + \mathbf{1}_{\{\delta \in \mathbf{A}\}}\frac{\Delta\ell(F(\delta))}{\Delta F(\delta)}, & \text{if } \delta < \delta^\star, \\ \mathbf{1}_{\{\delta \notin \mathbf{A}\}}(1 - \ell'(F(\delta))) + \mathbf{1}_{\{\delta \in \mathbf{A}\}}\left(1 - \frac{\Delta\ell(F(\delta))}{\Delta F(\delta)}\right), & \text{otherwise.} \end{cases} \tag{52}$$

To establish the result, we need to show that $m(\delta)$ is increasing on $[0, \delta^\star)$ and decreasing on $[\delta^\star, 1]$. As shown in the proof of Lemma A.8, we have that the function $\ell(x)$ is strictly convex on $[0, 1]$. This immediately implies that the functions $\ell'(x)$ and $(\ell(x) - \ell(y))/(x - y)$ are, respectively, increasing in $x \in [0, 1]$ and increasing in $x \in [0, 1]$ and $y \in [0, 1]$, and the desired result now follows from (52). ∎

56

Assuming that a meeting between a buyer and a seller with the same utility results in trade with some constant probability $\pi \in [0, 1]$ we can express the steady state trading volume as

$$\vartheta(\pi) = \lambda \int_{[0,1]^2} \mathbf{1}_{\{\delta_0 > \delta_1\}} d\Phi_0(\delta_0) d\Phi_1(\delta_1) + \pi\lambda \sum_{\delta \in [0,1]} \Delta\Phi_0(\delta)\Delta\Phi_1(\delta),$$

where $\Delta\Phi_q(\delta) = \Phi_q(\delta) - \Phi_q(\delta_-) \geq 0$ denotes the discrete mass of investors who hold $q \in \{0, 1\}$ units of the asset and have a utility type exactly equal to δ.

Lemma 4 *If the distribution of utility types is continuous then*

$$\vartheta(\pi) = \vartheta_c \equiv \gamma s(1-s)\left[(1+\gamma/\lambda)\log\left(1+\frac{\lambda}{\gamma}\right) - 1\right] \tag{53}$$

for all $\pi \in [0, 1]$ and is strictly increasing in both the meeting rate λ and the arrival rate of preference shocks γ. Otherwise, if the distribution of utility types has atoms, then the steady-state trading volume is strictly increasing in $\pi \in [0, 1]$ with $\vartheta(0) < \vartheta_c < \vartheta(1)$.

Proof of Proposition 4. Consider the continuous functions defined by

$$G_1(x) = \frac{\ell(x)}{s} \quad \text{and} \quad G_0(x) = \frac{x - \ell(x)}{1-s}.$$

Rearranging the quadratic equation for $\ell(x)$ given in Lemma A.8, it can be shown that these functions satisfy the identity

$$G_1(x) = \frac{\phi G_0(x)}{1 + \phi - G_0(x)}, \tag{54}$$

where $\phi = \gamma/\lambda$. Since the functions $G_q(x)$ are continuous, strictly increasing, and map $[0, 1]$ onto itself, we have that they each admit a continuous and strictly increasing inverse $G_q^{-1}(y)$, and it follows that identity (54) can be written equivalently as

$$G_1(G_0^{-1}(y)) = \frac{\phi y}{1 + \phi - y}. \tag{55}$$

Consider the class of tie-breaking rules whereby a fraction $\pi \in [0, 1]$ of the meetings between an owner and a non-owner of the same utility type lead to a trade. By definition, the trading volume associated with such a tie breaking rule can be computed as

$$\vartheta(\pi) = \lambda s(1-s)\left(\mathbb{P}[\delta_0 > \delta_1] + \pi\mathbb{P}[\delta_0 = \delta_1]\right),$$

where the random variables $(\delta_0, \delta_1) \in [0, 1]^2$ are distributed according to $\Phi_0(\delta)/(1-s) = G_0(F(\delta))$ and $\Phi_1(\delta)/s = G_1(F(\delta))$ independently of each other. A direct calculation shows that the quantile functions of these random variables are given by

$$\inf\{x \in [0, 1] : G_q(F(x)) \geq u\} = \inf\{x \in [0, 1] : F(x) \geq G_q^{-1}(u)\} = \Delta(G_q^{-1}(u))$$

where $\Delta(y)$ denotes the quantile function of the underlying distribution of utility types, and it thus follows

57

from Lemma A.9 below that the trading volume satisfies

$$\frac{\vartheta(\pi)}{\lambda s(1-s)} = \mathbb{P}\left[\Delta(G_0^{-1}(u_0)) > \Delta(G_1^{-1}(u_1))\right] + \pi \mathbb{P}\left[\Delta(G_0^{-1}(u_0)) = \Delta(G_1^{-1}(u_1))\right],$$

where u_0 and u_1 denote a pair of iid uniform random variables. If the distribution is continuous, then its quantile function is strictly increasing, and the above identity simplifies to

$$\frac{\vartheta(\pi)}{\lambda s(1-s)} = \mathbb{P}\left[G_0^{-1}(u_0) > G_1^{-1}(u_1)\right] = \mathbb{P}\left[u_1 < G_1(G_0^{-1}(u_0))\right]$$

$$= \mathbb{E}\left[G_1(G_0^{-1}(u_0))\right] = \int_0^1 G_1(G_0^{-1}(x))dx = \int_0^1 \frac{\phi x}{1+\phi-x}dx = \frac{\vartheta^\star}{\lambda s(1-s)},$$

where we used formula (55) for $G_1(G_0^{-1}(y))$, and the last equality follows from the calculation of the integral. If the distribution fails to be continuous, then its quantile function will have flat spots that correspond to the levels across which the distribution jumps, but it will nonetheless be weakly increasing. As a result, we have the strict inclusions

$$\left\{\Delta(G_0^{-1}(u_0)) > \Delta(G_1^{-1}(u_1))\right\} \subset \left\{G_0^{-1}(u_0) > G_1^{-1}(u_1)\right\} \subset \left\{\Delta(G_0^{-1}(u_0)) \geq \Delta(G_1^{-1}(u_1))\right\},$$

and it follows that

$$\frac{\vartheta(0)}{\lambda s(1-s)} = \mathbb{P}\left[\Delta(G_0^{-1}(u_0)) > \Delta(G_1^{-1}(u_1))\right]$$

$$< \mathbb{P}\left[G_0^{-1}(u_0) > G_1^{-1}(u_1)\right] = \frac{\vartheta^\star}{\lambda s(1-s)}$$

$$= \mathbb{P}\left[G_0^{-1}(u_0) \geq G_1^{-1}(u_1)\right] < \mathbb{P}\left[\Delta(G_0^{-1}(u_0)) \geq \Delta(G_1^{-1}(u_1))\right] = \frac{\vartheta(1)}{\lambda s(1-s)}.$$

Since the function $\vartheta(\pi)$ is continuous and strictly increasing in π, this further implies that there exists a unique tie-breaking probability $\pi^\star$ such that $\vartheta^\star = \vartheta(\pi^\star)$ and the proof is complete. ∎

Lemma A.9 *Let $H(x)$ be a cumulative probability distribution function on $[0,1]$. If the random variable U is uniformly distributed on $[0,1]$, then the random variable $\inf\{x \in [0,1] : H(x) \geq U\}$ is distributed according to $H(x)$.*

Proof. Let $\mathcal{X}(q) \equiv \{x' \in [0,1] : H(x') \geq q\}$ and $X(q) \equiv \inf \mathcal{X}(q)$. We show that $X(q) \leq x$ if and only if $H(x) \geq q$. For the "if" part, suppose that $H(x) \geq q$, then x belongs to $\mathcal{X}(q)$ and is therefore larger than its infimum, which is $X(q) \leq x$. For the "only if" part, let $(x_n)_{n=1}^\infty \subseteq \mathcal{X}(q)$ be a decreasing sequence converging toward $X(q)$. For each n, we have that $H(x_n) \geq q$. Going to the limit and using the fact that $H(x)$ is right continuous, we obtain that $H(X(q)) \geq q$, which implies $H(x) \geq q$ since $H(x)$ is increasing and $X(q) \leq x$ for all $x \in \mathcal{X}(q)$. ∎

Proof of Proposition 5. The first part of the result follows directly from Lemma 4. To establish the second part Let ε be as in the statement and assume that the distribution of utility types is continuous. In this case

the equilibrium trading volume can be decomposed as

$$\vartheta_c = \lambda\Phi_1(\delta^\star)(1 - s - \Phi_0(\delta^\star)) + \lambda\int_0^{\delta^\star-\varepsilon}\Phi_1(\delta)d\Phi_0(\delta) + \lambda\int_{\delta^\star+\varepsilon}^1 (1 - s - \Phi_0(\delta))d\Phi_1(\delta)$$

$$+ \lambda\int_{\delta^\star-\varepsilon}^{\delta^\star}\Phi_1(\delta)d\Phi_0(\delta) + \lambda\int_{\delta^\star}^{\delta^\star+\varepsilon}(1 - s - \Phi_0(\delta))\, d\Phi_1(\delta). \tag{56}$$

We show that all the terms on the first line remain bounded as $\lambda \to \infty$. Since $F(\delta^\star) = 1 - s$ when the distribution of type is continuous we have that the first term is equal to

$$\lambda\Phi_1(\delta^\star)\left(1 - s - F(\delta^\star) + \Phi_1(\delta^\star)\right) = \lambda\Phi_1(\delta^\star)^2.$$

and we know from Lemma A.8 that the measure $\Phi_1(\delta^\star)$ of owners below the marginal type solves

$$\lambda\Phi_1(\delta^\star)^2 + \gamma\Phi_1(\delta^\star) - \gamma s(1 - s) = 0.$$

This immediately implies that

$$\lambda\Phi_1(\delta^\star)^2 \le \gamma s(1 - s)$$

and it follows that the first term on the first line of (56) remains bounded as $\lambda \to \infty$. Turning to the second term, we note that

$$\lambda\int_0^{\delta^\star-\varepsilon}\Phi_1(\delta)d\Phi_0(\delta) \le \lambda\Phi_1(\delta^\star - \varepsilon)F(\delta^\star - \varepsilon), \tag{57}$$

where the inequality follows (1) and the increases of $\Phi_1(\delta)$. From Lemma A.8, we have that the steady-state measure of owners with valuations below $\delta^\star - \varepsilon$ solves

$$\lambda\Phi_1(\delta^\star - \varepsilon)^2 + \left(1 - s - F(\delta^\star - \varepsilon) + \frac{\gamma}{\lambda}\right)\lambda\Phi_1(\delta^\star - \varepsilon) - \gamma s F(\delta^\star - \varepsilon) = 0.$$

This immediately implies that

$$\lambda\Phi_1(\delta^\star - \varepsilon) \le \frac{\gamma s F(\delta^\star - \varepsilon)}{1 - s - F(\delta^\star - \varepsilon)}.$$

and combining this inequality with (57) shows that the second term on the first line of (56) remains bounded as $\lambda \to \infty$. Proceeding similarly, one can show that the third term also remains bounded as frictions vanish, and the desired result now follows by observing that $\lim_{\lambda\to\infty}\vartheta_c = \infty$. ∎

Proof of Lemma 2. Relying on (26) we have that

$$\mathbb{P}\left[\{\mathbf{d} > k\}|\{\delta_0 = \delta\}\right] = \begin{cases} (\pi_0(\delta)\pi_1(\delta))^{\frac{1+k}{2}}, & k \text{ is odd} \\ (\pi_0(\delta)\pi_1(\delta))^{\frac{k}{2}}\left[\pi_0(\delta)\frac{d\Phi_0(\delta)}{dF(\delta)} + \pi_1(\delta)\frac{d\Phi_1(\delta)}{dF(\delta)}\right], & \text{otherwise.} \end{cases} \tag{58}$$

On the other hand, combining (70) with the result of Proposition 10 we deduce that for all utility types below

the marginal type we have

$$\lim_{\lambda \to \infty} \pi_1(\delta) = 1,$$

$$\lim_{\lambda \to \infty} \pi_0(\delta) = \frac{sF(\delta)}{(1-s)(1-F(\delta))},$$

and

$$\lim_{\lambda \to \infty} \frac{d\Phi_q(\delta)}{dF(\delta)} = 1 - q.$$

Therefore, the limit distribution is given by

$$\lim_{\lambda \to \infty} \mathbb{P}\left[\{\mathbf{d} > k\} \middle| \{\delta_0 = \delta\}\right] = a(\delta, k) \equiv \left(\frac{sF(\delta)}{(1-s)(1-F(\delta))}\right)^{1+\left[\!\!\left[\frac{k}{2}\right]\!\!\right]}, \qquad \forall \delta < \delta^\star, \tag{59}$$

where $[\![x]\!]$ denotes the integer part of a real number x. Now fix two arbitrary utility types (δ, δ') as in the statement. Since $F(\delta) < F(\delta')$ we have that

$$a(\delta, k) < a(\delta', k), \qquad k \in \{0, 1\},$$

and, because all the terms on the right hand side of (58) are continuous in λ, it follows from (59) that there exists a threshold $\overline{\lambda}$ such that

$$\mathbb{P}\left[\{\mathbf{d} > k\} \middle| \{\delta_0 = \delta\}\right] < \mathbb{P}\left[\{\mathbf{d} > k\} \middle| \{\delta_0 = \delta'\}\right], \qquad (\lambda, k) \in (\overline{\lambda}, \infty) \times \{0, 1\}. \tag{60}$$

Using this inequality with $k = 1$ we deduce that

$$\pi_0(\delta)\pi_1(\delta) < \pi_0(\delta')\pi_1(\delta'), \qquad \lambda \in (\overline{\lambda}, \infty), \tag{61}$$

and it now follows from (58) that (60) holds not only for $k = 1$ but for all odd values of $k \in \mathbb{N}$. On the other hand, using (60) with $k = 0$ in conjunction with (61) we deduce that

$$\left(\frac{\pi_0(\delta)\pi_1(\delta)}{\pi_0(\delta')\pi_1(\delta')}\right)^{\frac{k}{2}} < 1 < \frac{\pi_0(\delta')\frac{d\Phi_0(\delta')}{dF(\delta')} + \pi_1(\delta')\frac{d\Phi_1(\delta')}{dF(\delta')}}{\pi_0(\delta)\frac{d\Phi_0(\delta)}{dF(\delta)} + \pi_1(\delta)\frac{d\Phi_1(\delta)}{dF(\delta)}}, \qquad (\lambda, k) \in (\overline{\lambda}, \infty) \times \mathbb{N},$$

and rearranging this inequality shows that (60) holds for all even values of $k \in \mathbb{N}$ ∎

Proof of Lemma 3. The result follows directly from the properties listed above the statement. ∎

Lemma A.10 *If the distribution of utility types is continuous then the distribution of $\left(\mathbf{n}, \delta^{(\mathbf{n})}\right)$ conditional on the event $\mathcal{S}_\delta = \mathcal{S} \cap \{\delta^{(1)} = \delta\}$ is given by*

$$e^{\Lambda(\delta, \delta^\star)} \mathbb{P}\left[\{\mathbf{n} = k\} \cap \{\delta^{(\mathbf{n})} \in dx\} \middle| \mathcal{S}_\delta\right] = \mathbf{1}_{\{k=1\}}\varepsilon_{\{\delta\}}(dx) + \mathbf{1}_{\{k>1\}}d_x\left(\frac{\Lambda(\delta, x)^{k-1}}{(k-1)!}\right)$$

where $\varepsilon_{\{\delta\}}(dx)$ denotes the Dirac measure at the point $\delta \in (\underline{\delta}, \delta^\star]$ and $d_x(\cdot)$ indicates a differential with respect to the variable x.

Proof of Lemma A.10. By Bayes' rule we have that the Fourier transform of the desired distribution can be computed as

$$
\mathbb{E}\left[e^{-\beta \mathbf{n} - \alpha \delta^{(\mathbf{n})}} \,\Big|\, \mathcal{S}_\delta \right] = \frac{h_{\alpha,\beta}(\delta)}{h_{0,0}(\delta)}, \qquad (\delta, \alpha, \beta) \in (\underline{\delta}, \delta^\star] \times \mathbb{C}^2
\tag{62}
$$

with the function

$$
h_{\alpha,\beta}(\delta) = \mathbb{E}\left[e^{-\beta \mathbf{n} - \alpha \delta^{(\mathbf{n})}} \mathbf{1}_{\{\mathcal{S}\}} \,\Big|\, \{\delta^{(1)} = \delta\} \right].
$$

By application of the law of iterated expectations we have that this function satisfies the recursive integral equation given by

$$
\begin{aligned}
h_{\alpha,\beta}(\delta) = \frac{\lambda \Phi_0(1)}{\gamma + \lambda \Phi_0(1)} \Bigg[& h_{\alpha,\beta}(\delta) \frac{\Phi_0(\delta)}{\Phi_0(1)} \\
& + \int_\delta^{\delta^\star} e^{-\beta} h_{\alpha,\beta}(x) \frac{d\Phi_0(x)}{\Phi_0(1)} + e^{-\beta - \alpha \delta} \left(1 - \frac{\Phi_0(\delta^\star)}{\Phi_0(1)} \right) \Bigg]
\end{aligned}
\tag{63}
$$

and evaluating both sides of this equation at the marginal type shows that

$$
h_{\alpha,\beta}(\delta^\star) = e^{-\beta - \alpha \delta^\star} h_{0,0}(\delta^\star) = e^{-\beta - \alpha \delta^\star} \pi_1(\delta^\star).
\tag{64}
$$

When the distribution of utility types is continuous, (63) and (64) can be solved as an ordinary differential equation. Indeed, differentiating (63) and solving the resulting equation for the differential of the unknown functions gives

$$
dh_{\alpha,\beta}(\delta) = h_{\alpha,\beta}(\delta) \left(e^{-\beta} - 1 \right) d\log(\gamma + \lambda_1(\delta)) - \alpha e^{\alpha(\delta^\star - \delta) - \Lambda(\delta, \delta^\star)} h_{\alpha,\beta}(\delta^\star) d\delta
$$

which is a first order linear differential equation. Integrating this differential equation subject to the boundary condition (64) we obtain

$$
h_{\alpha,\beta}(\delta) = e^{-\Lambda(\delta, \delta^\star)} h_{\alpha,\beta}(\delta^\star) \left(e^{\alpha(\delta^\star - \delta)} + \int_\delta^{\delta^\star} e^{\alpha(\delta^\star - x)} d_x \left(e^{a_\beta \Lambda(\delta, x)} \right) \right)
\tag{65}
$$

with the constant defined by $a_\beta = e^{-\beta}$. Evaluating this expression at the point $(\alpha, \beta) = (0,0)$ and simplifying the result then shows that

$$
h_{0,0}(\delta) = e^{-\Lambda(\delta, \delta^\star)} h_{0,0}(\delta^\star) \left(1 + e^{\Lambda(\delta, x)} \Big|_{x=\delta}^{\delta^\star} \right) = h_{0,0}(\delta^\star)
$$

does not depend on the utility type of the first dealer, and it now follows from (62), (64) and (65) that the Fourier transform of the desired distribution is explicitly given by

$$
\mathbb{E}\left[e^{-\beta \mathbf{n} - \alpha \delta^{(\mathbf{n})}} \,\Big|\, \mathcal{S}_\delta \right] = e^{-\Lambda(\delta, \delta^\star)} \left(e^{-\alpha \delta} a_\beta + \int_\delta^{\delta^\star} a_\beta^2 e^{-\alpha x + a_\beta \Lambda(\delta, x)} d_x \left(\Lambda(\delta, x) \right) \right).
$$

Using the chain rule and the fact that

$$a_\beta^2 e^{a_\beta \Lambda(\delta, x)} = \sum_{k=2}^{\infty} e^{-\beta k} \frac{\Lambda(\delta, x)^{k-2}}{(k-2)!}$$

to invert the transform with respect to β shows that we have

$$e^{\Lambda(\delta, \delta^\star)} \mathbb{E}\left[e^{-\alpha \delta^{(\mathbf{n})}} \mathbf{1}_{\{\mathbf{n}=k\}} \middle| \mathcal{S}_\delta \right] = \mathbf{1}_{\{k=1\}} e^{-\alpha\delta} + \mathbf{1}_{\{k>1\}} \int_\delta^{\delta^\star} e^{-\alpha x} d_x \left(\frac{\Lambda(\delta, x)^{k-1}}{(k-1)!} \right)$$

and the desired result now follows by inspection. ∎

Proof of Proposition 6. The first part of the statement follows by integrating the joint distribution of Lemma A.10 with respect to $x \in [\delta, \delta^\star]$. Since $\lambda_1(\delta)$ is decreasing in δ we have that $\Lambda(\delta, \delta^\star)$ is decreasing in δ and the second part of the statement follows by noting that the tail probability

$$\mathbb{P}\left[\{\mathbf{n} > k\} | \mathcal{S}_\delta \right] = \sum_{n=k+1}^{\infty} \frac{e^{-\Lambda(\delta, \delta^\star)} \Lambda(\delta, \delta^\star)^{n-1}}{(n-1)!} = 1 - \frac{1}{\Gamma(k)} \int_{\Lambda(\delta, \delta^\star)}^{\infty} t^{k-1} e^{-t} dt$$

is an increasing function of $\Lambda(\delta, \delta^\star)$. ∎

Proof of Corollary 3. By virtue of random matching we have that the distribution of the utility type of the first dealer in a chain is given by

$$\mathbb{P}\left[\{\delta^{(1)} \in d\delta\} \middle| \mathcal{S} \right] = \frac{d\Phi_0(\delta)}{\Phi_0(\delta^\star) - \Phi_0(\underline{\delta})}, \tag{66}$$

and integrating this expression against the conditional distribution of the chain length derived in Proposition 6 shows that we have

$$\mathbb{P}\left[\{\mathbf{n} = k\} | \mathcal{S} \right] = \int_{\underline{\delta}}^{\delta^\star} \mathbb{P}\left[\{\mathbf{n} = k\} \middle| \mathcal{S} \cap \{\delta^{(1)} = \delta\} \right] \mathbb{P}\left[\{\delta^{(1)} \in d\delta\} \middle| \mathcal{S} \right]$$

$$= \int_{\underline{\delta}}^{\delta^\star} \frac{\Lambda(\delta, \delta^\star)^{k-1} e^{-\Lambda(\delta, \delta^\star)}}{(k-1)!} \frac{d\Phi_0(\delta)}{\Phi_0(\delta^\star) - \Phi_0(\underline{\delta})} = \frac{1}{k!} \left(e^{\Lambda(\underline{\delta}, \delta^\star)} - 1 \right)^{-1} \Lambda(\underline{\delta}, \delta^\star)^k$$

where the last two equalities follow from (27). The tail probability function associated with this zero-truncated Poisson distribution is given by

$$\mathbb{P}\left[\{\mathbf{n} > k\} | \mathcal{S} \right] = \frac{1}{\left(1 - e^{-\Lambda(\underline{\delta}, \delta^\star)} \right) \Gamma(1+k)} \int_0^{\Lambda(\underline{\delta}, \delta^\star)} t^k e^{-t} dt \equiv H(\Lambda(\underline{\delta}, \delta^\star))$$

and, since

$$H'(x) = \frac{1}{(1 - e^{-x})^2 \Gamma(1+k)} \int_0^x e^{x-t} \left(x^k - t^k \right) dt > 0,$$

the proof will be complete once we show that $\Lambda(\underline{\delta}, \delta^\star)$ is decreasing with respect to the parameter γ and

increasing with respect to the parameter λ. To this end we start by observing that

$$\Lambda(\underline{\delta}, \delta^\star) = \log\left(\frac{\phi + 1 - s - F(\underline{\delta}) + \ell(F(\underline{\delta}); \phi)}{\phi + \ell(1 - s; \phi)}\right)$$

depends on the parameters γ and λ only through their ratio $\phi = \gamma/\lambda$. Differentiating both sides with respect to ϕ and simplifying the result shows that

$$\text{sign}\left(\frac{\partial \Lambda(\underline{\delta}, \delta^\star)}{\partial \phi}\right) = \text{sign}\left(G(F(\underline{\delta}); \phi) - G(F(\delta^\star); \phi)\right)$$

where we have set

$$G(x; \phi) = \frac{1 + \frac{\partial \ell}{\partial \phi}(x; \phi)}{1 - s - x + \phi + \ell(x; \phi)}.$$

Now, a direct calculation using the definition of the function $\ell(x, \phi)$ shows that the derivative of this function with respect to $x \in [0, 1]$ is explicitly given by

$$\frac{\partial G}{\partial x}(x; \phi) = \frac{(1 - s)(1 + s - x + \phi)}{((1 - s + \phi - x)^2 + 4s\phi x)^{3/2}} > 0$$

and the desired result follows. ∎

Proof of Proposition 7. Consider the utility type of the first dealer in the chain. By application of Bayes' rule we have that the distribution that underlies the desired tail probability can be computed as

$$\mathbb{P}\left[\{\delta^{(1)} \in d\delta\} \Big| \mathcal{S} \cap \{\mathbf{n} = k\}\right] = \frac{\mathbb{P}\left[\{\delta^{(1)} \in d\delta\} \cap \{\mathbf{n} = k\} \big| \mathcal{S}\right]}{\mathbb{P}\left[\{\mathbf{n} = k\} \big| \mathcal{S}\right]}$$

$$= \frac{\mathbb{P}\left[\{\mathbf{n} = k\} \big| \mathcal{S} \cap \{\delta^{(1)} = \delta\}\right] \mathbb{P}\left[\{\delta^{(1)} \in d\delta\} \big| \mathcal{S}\right]}{\mathbb{P}\left[\{\mathbf{n} = k\} \big| \mathcal{S}\right]}.$$

Substituting (66) and the results of Proposition 6 and Corollary 3 into the right hand side of this expression and simplifying the result shows that

$$\mathbb{P}\left[\{\delta^{(1)} \in d\delta\} \Big| \mathcal{S} \cap \{\mathbf{n} = k\}\right] = \frac{e^{-\Lambda(\delta, \delta^\star)}}{\Lambda(\delta, \delta^\star)} \left(e^{\Lambda(\underline{\delta}, \delta^\star)} - 1\right) \left(\frac{\Lambda(\delta, \delta^\star)}{\Lambda(\underline{\delta}, \delta^\star)}\right)^k \frac{k \, d\Phi_0(\delta)}{\Phi_0(\delta^\star) - \Phi_0(\underline{\delta})}$$

$$= -d_\delta\left[\left(\frac{\Lambda(\delta, \delta^\star)}{\Lambda(\underline{\delta}, \delta^\star)}\right)^k\right], \tag{67}$$

where the second equality follows from the chain rule and the fact that

$$e^{-\Lambda(\delta, \delta^\star)}\left(e^{\Lambda(\underline{\delta}, \delta^\star)} - 1\right) = \frac{\gamma + \lambda_1(\delta^\star)}{\gamma + \lambda_1(\delta)}\left(\frac{\gamma + \lambda_1(\underline{\delta})}{\gamma + \lambda_1(\delta^\star)} - 1\right)$$

$$= \frac{\gamma + \lambda_1(\delta^\star)}{\gamma + \lambda_1(\delta)}\left(\frac{\lambda_1(\underline{\delta}) - \lambda_1(\delta^\star)}{\gamma + \lambda_1(\delta^\star)}\right) = \frac{\lambda(\Phi_0(\delta^\star) - \Phi_0(\underline{\delta}))}{\gamma + \lambda_1(\delta)}$$

by definition of the functions $\lambda_1(\delta)$ and $\Lambda(x, y)$. Integrating both sides of (67) with respect to δ over the

interval $(x, \delta^\star]$ then gives

$$\mathbb{P}\Big[\{\delta^{(1)} > x\}\Big| \mathcal{S} \cap \{\mathbf{n} = k\}\Big] = \left(\frac{\Lambda(x, \delta^\star)}{\Lambda(\underline{\delta}, \delta^\star)}\right)^k$$

and the required monotonicity follows by observing that, because the function $\Lambda(x, y)$ is decreasing in its first argument, the term inside the bracket on the right hand side is smaller than unity.

Let us now turn to the utility type of the last dealer in the chain. Proceeding in the same way we have that the distribution that underlies the desired tail probability can be computed as

$$\mathbb{P}\Big[\{\delta^{(\mathbf{n})} \in dx\}\Big| \mathcal{S} \cap \{\mathbf{n} = k\}\Big] = \frac{\mathbb{P}\big[\{\delta^{(\mathbf{n})} \in dx\} \cap \{\mathbf{n} = k\}\big| \mathcal{S}\big]}{\mathbb{P}\big[\{\mathbf{n} = k\}\big| \mathcal{S}\big]}$$

$$= \frac{1}{\mathbb{P}\left[\{\mathbf{n} = k\}|\,\mathcal{S}\right]} \int_{\underline{\delta}}^x \mathbb{P}\Big[\{\delta^{(\mathbf{n})} \in dx\} \cap \{\mathbf{n} = k\}\Big| \mathcal{S}_\delta\Big] \mathbb{P}\Big[\{\delta^{(1)} \in d\delta\}\Big| \mathcal{S}\Big].$$

Substituting (66) and the results of Lemma A.10 and Corollary 3 into the right hand side of this expression shows that

$$\mathbb{P}\Big[\{\delta^{(\mathbf{n})} \in dx\}\Big| \mathcal{S} \cap \{\mathbf{n} = k\}\Big] = d_x\left[\left(\frac{\Lambda(\underline{\delta}, x)}{\Lambda(\underline{\delta}, \delta^\star)}\right)^k\right],$$

and the desired results now follow by integrating this expression with respect to x on the interval $(y, \delta^\star]$ and noting that the function $\Lambda(x, y)$ is increasing in its second argument. ∎

Proof of Proposition 8. Using the result of Lemma A.10, we have that the distribution that underlies the required tail probability is given by

$$\mathbb{P}\Big[\{\delta^{(\mathbf{n})} \in dx\}\Big| \mathcal{S}_\delta\Big] = \sum_{k=1}^\infty \mathbb{P}\Big[\{\delta^{(\mathbf{n})} \in dx\} \cap \{\mathbf{n} = k\}\Big| \mathcal{S}_\delta\Big]$$

$$= e^{-\Lambda(\delta, \delta^\star)} \varepsilon_{\{\delta\}}(dx) + \sum_{k=2}^\infty e^{-\Lambda(\delta, \delta^\star)} \frac{\Lambda(\delta, x)^{k-2}}{(k-2)!} d_x\left(\Lambda(\delta, x)\right)$$

$$= e^{-\Lambda(\delta, \delta^\star)} \left(\varepsilon_{\{\delta\}}(dx) + e^{\Lambda(\delta, x)} d_x\left(\Lambda(\delta, x)\right)\right) = e^{-\Lambda(\delta, \delta^\star)} \varepsilon_{\{\delta\}}(dx) + d_x\left(e^{-\Lambda(x, \delta^\star)}\right)$$

for all $x \in [\delta, \delta^\star]$ where the third equality follows from the calculation of the infinite sum and the last equality follows from the chain rule and the definition of $\Lambda(x, y)$. Integrating this expression with respect to x on the interval $(y, \delta^\star] \cap [\delta, \delta^\star]$ then gives

$$\mathbb{P}\Big[\{\delta^{(\mathbf{n})} > y\}\Big| \mathcal{S}_\delta\Big] = 1 - \mathbf{1}_{\{\delta \le y\}} e^{-\Lambda(y, \delta^\star)}$$

and the required monotonicity follows by observing that the function $\Lambda(x, y)$ is monotone decreasing in its first argument. ∎

Proof of Proposition 9. From (20) we deduce that the reservation value function $\Delta V(\delta; \theta_0)$ is continuous in $\theta_0 \in (0, 1)$ and can be extended by continuity to a function that is strictly increasing with respect to

$\delta \in [0, 1]$ for all bargaining powers $\theta_0 \in [0, 1]$ and such that

$$0 < V_0 \leq \inf_{(\delta, \theta_0) \in [0,1]^2} \Delta V(\delta; \theta_0) \leq \sup_{(\delta, \theta_0) \in [0,1]^2} \Delta V(\delta; \theta_0) \leq V_1 < \infty \tag{68}$$

for some constants V_0 and V_1. Using these properties in conjunction with Proposition 8, the dominated convergence theorem, the integration by parts formula and the fact that $\delta^{(0)}$ and $\delta^{(\mathbf{n})+1}$ are both independent from $\delta^{(1)}$, $\delta^{(\mathbf{n})}$ and $\mathcal{S}$ shows that the function

$$h(\delta; \theta_0) = 1 + \mathbb{E}\left[\mathbf{m} \,\middle|\, \{\delta^{(1)} = \delta\} \cap \mathcal{S}\right]$$
$$= \mathbb{E}\left[\frac{\theta_0 \Delta V(\delta^{(\mathbf{n})}; \theta_0) + (1 - \theta_0)\Delta V(\delta^{(\mathbf{n}+1)}; \theta_0)}{\theta_0 \Delta V(\delta^{(\mathbf{0})}; \theta_0) + (1 - \theta_0)\Delta V(\delta; \theta_0)} \,\middle|\, \{\delta^{(1)} = \delta\} \cap \mathcal{S}\right]$$

is continuous in $\theta_0 \in [0, 1]$ and satisfies

$$\lim_{\theta_0 \to 0} h(\delta; \theta_0) = h_0(\delta) = \frac{1}{\Delta V(\delta; 0)} \mathbb{E}\left[\Delta V(\delta^{(\mathbf{n}+1)}; 0)\right],$$
$$\lim_{\theta_0 \to 1} h(\delta; \theta_0) = h_1(\delta) = \mathbb{E}\left[\frac{1}{\Delta V(\delta^{(0)}; 1)}\right]\left(\Delta V(\delta^{\star}; 1) - \int_{\delta}^{\delta^{\star}} \sigma(x; 1)e^{-\Lambda(x, \delta^{\star})} dx\right).$$

Since reservation values are strictly increasing in δ we deduce that the functions $h_0(\delta)$ and $-h_1(\delta)$ are both strictly decreasing and the existence of thresholds such that (30) holds now follows by continuity.

Similarly, using (68) and the dominated convergence theorem in conjunction with Proposition 7, the integration by parts formula and the fact that $\delta^{(0)}$ and $\delta^{(\mathbf{n})+1}$ are both independent from $\delta^{(1)}$, $\delta^{(\mathbf{n})}$ and $\mathcal{S}$ we obtain that the function

$$f(k; \theta_0) = 1 + \mathbb{E}\left[\mathbf{m} \,\middle|\, \{\mathbf{n} = k\} \cap \mathcal{S}\right]$$

is continuous in θ_0 and satisfies

$$\lim_{\theta_0 \to 1} f(k; \theta_0) = f_1(k) = \mathbb{E}\left[\frac{\Delta V(\delta^{(\mathbf{n})}; 1)}{\Delta V(\delta^{(0)}; 1)} \,\middle|\, \{\mathbf{n} = k\} \cap \mathcal{S}\right]$$
$$= \mathbb{E}\left[\frac{1}{\Delta V(\delta^{(0)}; 1)}\right]\left(\Delta V(\underline{\delta}; 1) - \int_{\underline{\delta}}^{\delta^{\star}} \sigma(x; 1)\left(\frac{\Lambda(\underline{\delta}, x)}{\Lambda(\underline{\delta}, \delta^{\star})}\right)^k dx\right),$$
$$\lim_{\theta_0 \to 0} f(k; \theta_0) = f_0(k) = \mathbb{E}\left[\frac{\Delta V(\delta^{(\mathbf{n}+1)}; 0)}{\Delta V(\delta^{(1)}; 0)} \,\middle|\, \{\mathbf{n} = k\} \cap \mathcal{S}\right]$$
$$= \mathbb{E}\left[\Delta V(\delta^{(\mathbf{n}+1)}; 1)\right]\left(\frac{1}{\Delta V(\underline{\delta}; 0)} - \int_{\underline{\delta}}^{\delta^{\star}} \frac{\sigma(x; 0)}{\Delta V(x; 0)^2}\left(\frac{\Lambda(x, \delta^{\star})}{\Lambda(\underline{\delta}, \delta^{\star})}\right)^k dx\right).$$

Since reservation values are strictly increasing in δ and $\Lambda(x, y)$ is increasing in its first argument we deduce that the functions $f_0(k)$ and $f_1(k)$ are both increasing and the existence of thresholds such that (31) holds now follows by continuity. ∎

A.4 Proofs omitted in Section 5

Proof of Proposition 10. From equation (18), it follows that we have

$$\lim_{\lambda \to \infty} \Phi_1(\delta) = \frac{|1 - s - F(\delta)|}{2} - \frac{1 - s - F(\delta)}{2} = (1 - s - F(\delta))^- = \Phi_1^\star(\delta)$$

and therefore $\lim_{\lambda \to \infty} \Phi_0(\delta) = \Phi_0^\star(\delta)$ for all $\delta \in [0, 1]$. By Theorem 1, we have that the steady state reservation value function is explicitly given by

$$r \Delta V(\delta) = \delta - \int_0^\delta k_0(\delta') d\delta' + \int_\delta^1 k_1(\delta') d\delta'$$

with the uniformly bounded functions defined by

$$k_0(\delta') = \frac{\gamma F(\delta') + \lambda \theta_0 \Phi_1(\delta')}{r + \gamma + \lambda \theta_1 (1 - s - \Phi_0(\delta')) + \lambda \theta_0 \Phi_1(\delta')}$$

$$k_1(\delta') = \frac{\gamma(1 - F(\delta')) + \lambda \theta_1 (1 - s - \Phi_0(\delta'))}{r + \gamma + \lambda \theta_1 (1 - s - \Phi_0(\delta')) + \lambda \theta_0 \Phi_1(\delta')}.$$

Using the first part of the proof and the assumption that $\theta_q > 0$, we obtain

$$\lim_{\lambda \to \infty} k_q(\delta') = \frac{\theta_q \Phi_{1-q}^\star(\delta')}{\theta_0 \Phi_1^\star(\delta') + \theta_1 \Phi_0^\star(\delta')} = \mathbf{1}_{\{q=0\}} \mathbf{1}_{\{\delta \geq \delta^\star\}} + \mathbf{1}_{\{q=1\}} \mathbf{1}_{\{\delta < \delta^\star\}},$$

and the required result now follows from an application of the dominated convergence theorem because the functions $k_q(\delta')$ take values in $[0, 1]$. ∎

Convergence rates of the distributions. To derive the rates at which the equilibrium distributions converge to their frictionless counterparts, recall the inflow-outflow equation that characterizes the steady-state equilibrium distributions:

$$\gamma F(\delta) (s - \Phi_1(\delta)) = \gamma \Phi_1(\delta) (1 - F(\delta)) + \lambda \Phi_1(\delta) (1 - s - \Phi_0(\delta)). \tag{69}$$

By Proposition 10 we have that $\Phi_1(\delta) \to 0$ and $\Phi_0(\delta) \to F(\delta) < 1 - s$ for all utility types $\delta < \delta^\star$ as the meeting frequency becomes infinite, and it thus follows from (69) that for $\delta < \delta^\star$ the distribution of utility types among asset owners admits the approximation

$$\Phi_1(\delta) = \frac{\gamma F(\delta) s}{1 - s - F(\delta)} \left(\frac{1}{\lambda}\right) + o\left(\frac{1}{\lambda}\right). \tag{70}$$

Similarly, by Proposition 10 we have that $\Phi_1(\delta) \to F(\delta) - 1 + s > 0$ and $\Phi_0(\delta) \to 1 - s$ for all utility types $\delta > \delta^\star$ as the meeting frequency becomes infinite, and it thus follows from (69) that for $\delta > \delta^\star$ the distribution of utility types among non-owners admits the approximation

$$1 - s - \Phi_0(\delta) = \frac{\gamma(1 - s)(1 - F(\delta))}{F(\delta) - (1 - s)} \left(\frac{1}{\lambda}\right) + o\left(\frac{1}{\lambda}\right). \tag{71}$$

To derive the convergence rate at the point $\delta = \delta^\star$, assume first that the distribution of utility types crosses

66

the level $1-s$ continuously and observe that in this case we have

$$1 - s - \Phi_0(\delta^\star) = 1 - s - F(\delta^\star) + \Phi_1(\delta^\star) = \Phi_1(\delta^\star).$$

Substituting these identities into (69) evaluated at the marginal type and letting $\lambda \to \infty$ on both sides shows that the equilibrium distributions admit the approximation given by

$$\Phi_1(\delta^\star) = 1 - s - \Phi_0(\delta^\star) = \sqrt{\gamma s(1-s)}\left(\frac{1}{\sqrt{\lambda}}\right) + o\left(\frac{1}{\sqrt{\lambda}}\right). \tag{72}$$

If the distribution of utility types crosses $1 - s$ by a jump, we have $F(\delta^\star) > 1 - s$, and it follows that the approximation (71) also holds at the marginal type. ∎

Proof of Proposition 11. Assume without loss of generality that the support of the distribution of utility types is the interval $[0,1]$. Evaluating (20) at $\delta^\star$ and making the change of variable $x = \sqrt{\lambda}(\delta' - \delta^\star)$ in the two integrals shows that

$$r\sqrt{\lambda}\left(\Delta V(\delta^\star) - p^\star\right) = P(\lambda) - D(\lambda), \tag{73}$$

where the functions on the right-hand side are defined by

$$D(\lambda) \equiv \int_{-\infty}^{0} \mathbf{1}_{\{x + \delta^\star \sqrt{\lambda} \geq 0\}} \frac{\gamma F(\delta^\star + x/\sqrt{\lambda}) + \theta_0 \sqrt{\lambda} g_1(x)}{r + \gamma + \theta_0 \sqrt{\lambda} g_1(x) + \theta_1 \sqrt{\lambda} g_0(x)} dx$$

and

$$P(\lambda) \equiv \int_{0}^{\infty} \mathbf{1}_{\{x + \delta^\star \sqrt{\lambda} \leq 1\}} \frac{\gamma(1 - F(\delta^\star + x/\sqrt{\lambda})) + \theta_1 \sqrt{\lambda} g_0(x)}{r + \gamma + \theta_0 \sqrt{\lambda} g_1(x) + \theta_1 \sqrt{\lambda} g_0(x)} dx$$

with the functions

$$g_q(x) \equiv \frac{\lambda_{1-q}(\delta^\star + x/\sqrt{\lambda})}{\sqrt{\lambda}} = \sqrt{\lambda}(1-q)(1 - s - F(\delta^\star + x/\sqrt{\lambda})) + \sqrt{\lambda}\Phi_1(\delta^\star + x/\sqrt{\lambda}).$$

Letting the meeting rate $\lambda \to \infty$ on both sides of equation (73) and using the convergence result established by Lemma A.13 below we obtain that

$$\lim_{\lambda \to \infty} r\sqrt{\lambda}(\Delta V(\delta^\star) - p^\star) = \int_0^\infty \frac{\theta_1 g(-x) dx}{\theta_0 g(x) + \theta_1 g(-x)} - \int_{-\infty}^0 \frac{\theta_0 g(z) dz}{\theta_0 g(z) + \theta_1 g(-z)}$$

$$= \int_0^\infty \frac{(1 - 2\theta_0) g(x) g(-x) dx}{(\theta_0 g(x) + \theta_1 g(-x))(\theta_0 g(-x) + \theta_1 g(x))} dx$$

$$= \int_0^\infty \frac{\gamma s(1-s)(1 - 2\theta_0) dx}{\gamma s(1-s) + \theta_0 \theta_1 (x F'(\delta^\star))^2} = \frac{\pi}{F'(\delta^\star)}\left(\frac{1}{2} - \theta_0\right)\left(\frac{\gamma s(1-s)}{\theta_0 \theta_1}\right)^{\frac{1}{2}},$$

where the function

$$g(x) = \frac{1}{2} x F'(\delta^\star) + \frac{1}{2}\sqrt{(x F'(\delta^\star))^2 + 4\gamma s(1-s)}$$

is the unique positive solution to (74), the second equality follows by making the change of variable $-z = x$ in the second integral, the third equality follows from the definition of the function $g(x)$, and the last equality follows from the fact that

$$\int_0^\infty \frac{dx}{a + x^2} = \left[\frac{\arctan(x/\sqrt{a})}{\sqrt{a}} \right]_0^\infty = \frac{\pi}{2\sqrt{a}}, \qquad a > 0.$$

This shows that the asymptotic expansion of the statement holds at the marginal type and the desired result now follows from the fact that $\Delta V(\delta) = \Delta V(\delta^\star) + o(1/\sqrt{\lambda})$ by Proposition 13. ∎

Lemma A.11 *Assume that the conditions of Proposition 11 hold and denote by $g(x)$ the positive solution to the quadratic equation*

$$g^2 - gF'(\delta^\star)x - \gamma s(1 - s) = 0. \tag{74}$$

Then we have that $g_1(x) \to g(x)$ and $g_0(x) \to g(-x)$ for all $x \in \mathbb{R}$ as $\lambda \to \infty$.

Proof. Evaluating (17) at the steady-state shows that the function $g_1(x)$ is the unique positive solution to the quadratic equation given by

$$g^2 + \left[\frac{\gamma}{\sqrt{\lambda}} + \sqrt{\lambda} \left(F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right) \right] g - \gamma s F(\delta^\star + x/\sqrt{\lambda}) = 0. \tag{75}$$

Because the left hand side of this quadratic equation is negative at the origin and positive at $g = 1$ we have that $g_1(x) \in [0, 1]$. This implies that $g_1(x)$ has a well-defined limit as $\lambda \to \infty$, and it now follows from (75) that this limit is given by the positive solution to (74). Next, we note that

$$g_0(x) = g_1(x) + \sqrt{\lambda} \left(F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right).$$

Substituting this expression into equation (75) then shows that the function $g_0(x)$ is the unique positive solution to the quadratic equation given by

$$g^2 + \left[\frac{\gamma}{\sqrt{\lambda}} - \sqrt{\lambda} \left(F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right) \right] g - \gamma(1 - s) \left(1 - F(\delta^\star + x/\sqrt{\lambda}) \right),$$

and the desired result follows from the same arguments as above. ∎

Lemma A.12 *Assume that the conditions of Proposition 11 hold. Then*

(a) *There exists a finite $K \geq 0$ such that*

$$\begin{aligned}
g_1(x) &\leq K/|x|, & x \in I_-^\lambda \equiv [-\delta^\star \sqrt{\lambda}, 0], \\
g_0(x) &\leq K/|x|, & x \in I_+^\lambda \equiv [0, (1 - \delta^\star)\sqrt{\lambda}].
\end{aligned} \tag{76}$$

(b) *For any given $\bar{x} \in I_+^\lambda \cap (-I_-^\lambda)$, there exists a strictly positive k such that*

$$\begin{aligned}
g_1(x) &\geq k|x|, & x \in I_+^\lambda \cap [\bar{x}, \infty), \\
g_0(x) &\geq k|x|, & x \in I_-^\lambda \cap (-\infty, -\bar{x}]
\end{aligned} \tag{77}$$

for all sufficiently large λ.

Proof. Because $g_1(x)$ is the positive root of (75) we have that (76) holds if and only if

$$\min_{x \in I_-^\lambda} \left\{ \frac{K^2}{x^2} + \frac{K}{|x|} \left(\frac{\gamma}{\sqrt{\lambda}} + \sqrt{\lambda} \left[F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right] \right) - \gamma s F(\delta^\star + x/\sqrt{\lambda}) \right\} \geq 0,$$

and a sufficient condition for this to be the case is that

$$\min_{x \in I_-^\lambda} \left\{ \frac{K}{|x|} \sqrt{\lambda} \left[F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right] - \gamma s(1 - s) \right\} \geq 0. \tag{78}$$

By the mean value theorem, we have that for any $x \in I_-^\lambda \cup I_+^\lambda$ there exists $\hat{\delta}(x) \in [0, 1]$ such that

$$F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) = -\frac{x F'(\hat{\delta}(x))}{\sqrt{\lambda}}, \tag{79}$$

and substituting this expression into (78) shows that a sufficient condition for the validity of equation (76) is that we have

$$K \geq K^\star \equiv \max_{\delta \in [0, 1]} \frac{\gamma s(1 - s)}{F'(\delta)}.$$

Because the derivative of the distribution of utility types is assumed to be bounded away from zero on the whole interval $[0, 1]$, we have that $K^\star$ is finite and equation (76) follows. One obtains the same constant when applying the same calculations to the function $g_0(x)$ over the interval I_+^λ.

Now let us turn to the second part of the statement and fix an arbitrary $\bar{x} \in I_+^\lambda \cap (-I_-^\lambda)$. Because the function $g_1(x)$ is the positive root of (75) we have that (77) holds if and only if

$$\max_{x \in I_+^\lambda \cap [\bar{x}, \infty)} \left\{ k^2 x^2 + kx \left(\frac{\gamma}{\sqrt{\lambda}} + \sqrt{\lambda} \left[F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right] \right) - \gamma s F(\delta^\star + x/\sqrt{\lambda}) \right\} \leq 0.$$

Combining this inequality with equation (79) then shows that a sufficient condition for the validity of equation (77) is given by

$$k \leq k^\star \equiv \inf_{\delta \in [0, 1]} \left(F'(\delta) - \frac{\gamma}{\bar{x}\sqrt{\lambda}} \right),$$

and the desired result now follows by noting that, because the derivative of the distribution of utility types is assumed to be strictly positive on the whole interval $[0, 1]$, we can pick the meeting rate λ large enough for the constant $k^\star$ to be strictly positive. One obtains the same constant when applying the same calculations to the function $g_0(x)$ over the interval $I_-^\lambda \cap (-\infty, -\bar{x}]$. ∎

Lemma A.13 *Assume that the conditions of Proposition 11 hold. Then*

$$\lim_{\lambda \to \infty} D(\lambda) = \int_{-\infty}^0 \frac{\theta_0 g(x) \, dx}{\theta_0 g(x) + \theta_1 g(-x)} \quad \text{and} \quad \lim_{\lambda \to \infty} P(\lambda) = \int_0^\infty \frac{\theta_1 g(-x) \, dx}{\theta_0 g(x) + \theta_1 g(-x)},$$

where the function $g(x)$ is defined as in Lemma A.11.

Proof. By Lemma A.11 we have that the integrand

$$H(x; \lambda) \equiv \mathbf{1}_{\{x \in I_-^\lambda\}} \left(\frac{\gamma F(\delta^\star + x/\sqrt{\lambda}) + \theta_0 \sqrt{\lambda} g_1(x)}{r + \gamma + \theta_0 \sqrt{\lambda} g_1(x) + \theta_1 \sqrt{\lambda} g_0(x)} \right)$$

in the definition of $D(\lambda)$ satisfies

$$\lim_{\lambda \to \infty} H(x; \lambda) = \frac{\theta_0 g(x)}{\theta_0 g(x) + \theta_1 g(-x)}. \tag{80}$$

Now fix an arbitrary $\bar{x} \in I_+^\lambda \cap (-I_-^\lambda)$ and let the meeting rate λ be large enough. On the interval $[-\bar{x}, 0]$, we can bound the integrand above by 1 and below by zero, while on the interval $I_-^\lambda \setminus [-\bar{x}, 0]$ we can use the bounds provided by Lemma A.12 to show that

$$0 \leq H(x; \lambda) \leq \frac{\gamma |x| + \theta_0 \sqrt{\lambda} K}{\sqrt{\lambda}(\theta_0 K + \theta_1 k |x|^2)} \leq \frac{\gamma \delta^\star + \theta_0 K}{\theta_0 K + \theta_1 k |x|^2},$$

where the inequality follows from the definition of I_-^λ. Combining these bounds shows that the integrand is bounded by a function that is integrable on $\mathbb{R}_-$ and does not depend on λ. This allows us to apply the dominated convergence theorem, and the result for $D(\lambda)$ now follows from (80). The result for the other integral follows from identical calculations. We omit the details. ∎

Proof of Proposition 12. Assume that there are $I \geq 2$ utility types $\delta_1 < \delta_2 < \ldots < \delta_I$, identify the marginal type with the index $m \in \{1, \ldots, I\}$ such that:

$$1 - F(\delta_m) \leq s < 1 - F(\delta_{m-1})$$

and set $\delta_0 \equiv 0$ and $\delta_{I+1} \equiv 1$. Assume further that $1 - F(\delta_m) < s$, which occurs generically when the distribution of utility types is restricted to be discrete. Under these assumptions, the same algebraic manipulations that we used to establish (70) and (71) show that we have

$$\Phi_1(\delta) = \Phi_1(\delta_i) = \begin{cases} \frac{1}{\lambda} \frac{\gamma s F(\delta_i)}{1 - s - F(\delta_i)} + o\left(\frac{1}{\lambda}\right) & \text{if } i < m \\ F(\delta_i) - (1 - s) + \frac{1}{\lambda} \frac{\gamma (1 - F(\delta_i))(1-s)}{F(\delta_i) - (1-s)} + o\left(\frac{1}{\lambda}\right) & \text{if } i \geq m, \end{cases} \tag{81}$$

for all $\delta \in [\delta_i, \delta_{i+1})$ and $i \in \{1, \ldots, I\}$. Likewise, we have that the local surplus satisfies

$$\sigma(\delta) = \sigma(\delta_i) = \begin{cases} \frac{1}{\lambda \theta_1 (1 - s - F(\delta_i))} + o\left(\frac{1}{\lambda}\right) & \text{if } i < m \\ \frac{1}{\lambda \theta_0 (F(\delta_i) - (1-s))} + o\left(\frac{1}{\lambda}\right) & \text{if } i \geq m \end{cases}$$

for all $\delta \in [\delta_i, \delta_{i+1})$ and $i \in \{1, \ldots, I\}$, and it follows that the steady-state reservation values satisfy

$$\Delta V(\delta_m) - \Delta V(\delta_i) = \sum_{j=i}^{m-1} (\delta_{j+1} - \delta_j) \sigma(\delta_j) = \sum_{j=i}^{m-1} \frac{\delta_{j+1} - \delta_j}{\theta_1 (1 - s - F(\delta_j))} + o\left(\frac{1}{\lambda}\right).$$

for every $i < m$, and

$$\Delta V(\delta_i) - \Delta V(\delta_m) = \sum_{j=m}^{i-1} (\delta_{j+1} - \delta_j)\, \sigma(\delta_j) = \sum_{j=m}^{i-1} \frac{\delta_{j+1} - \delta_j}{\theta_0 \left(F(\delta_j) - (1-s) \right)} + o\left(\frac{1}{\lambda}\right).$$

for every $i > m$. To complete the proof we calculate the steady-state reservation value $\Delta V(\delta_m)$ of the marginal investor using formula (13). This gives

$$
\begin{aligned}
r\Delta V(\delta_m) = \delta_m &+ \sum_{i=m}^{I} (\delta_{i+1} - \delta_i) \frac{\gamma\left(1 - F(\delta_i)\right) + \lambda\theta_1\left(1 - s - \Phi_0(\delta_i)\right)}{r + \gamma + \lambda\theta_0\Phi_1(\delta_i) + \lambda\theta_1\left(1 - s - \Phi_0(\delta_i)\right)} \\
&- \sum_{i=0}^{m-1} (\delta_{i+1} - \delta_i) \frac{\gamma F(\delta_i) + \lambda\theta_0\Phi_1(\delta_i)}{r + \gamma + \lambda\theta_0\Phi_1(\delta_i) + \lambda\theta_1\left(1 - s - \Phi_0(\delta_i)\right)} \\
= \delta_m &+ \frac{1}{\lambda} \sum_{i=m}^{I} (\delta_{i+1} - \delta_i) \frac{\gamma\left(1 - F(\delta_i)\right)\left(F(\delta_i) - (1-s)(1-\theta_1)\right)}{\left(F(\delta_i) - (1-s)\right)^2} \\
&- \frac{1}{\lambda} \sum_{i=0}^{m-1} (\delta_{i+1} - \delta_i) \frac{\gamma F(\delta_i)\left(1 - F(\delta_i) - s(1 - \theta_0)\right)}{\left(F(\delta_i) - (1-s)\right)^2} + o\left(\frac{1}{\lambda}\right)
\end{aligned}
$$

where the second equality follows from condition (1) and the asymptotic expansion of $\Phi_1(\delta)$ given in equation (81). ∎

Proof of Proposition 13. The result follows from Lemmas A.14, A.15, and A.16. To simplify the presentation we assume without loss of generality in these lemmas that the endpoints of the support of the distribution of utility types are given by $\underline{\delta} = 0$ and $\overline{\delta} = 1$. ∎

Lemma A.14 *Assume that the conditions of Proposition 13 hold true. Then*

$$A(\lambda) \equiv \lambda \int_0^{\delta^\star} \sigma(\delta)\, d\delta - \int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} = O(1) \tag{82}$$

$$B(\lambda) \equiv \lambda \int_{\delta^\star}^1 \sigma(\delta)\, d\delta - \int_{\delta^\star}^1 \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta^\star - \delta) + 1 - s - \Phi_0(\delta)} = O(1). \tag{83}$$

as the meeting rate $\lambda \to \infty$.

Proof. To establish (82) we start by noting that

$$\lambda\sigma(\delta) = \frac{\lambda}{r + \gamma + \lambda\theta_1\left(1 - s - \Phi_0(\delta)\right) + \lambda\theta_0\Phi_1(\delta)} = \frac{1}{\frac{r+\gamma}{\lambda} + \theta_1\left(F(\delta^\star) - F(\delta)\right) + \Phi_1(\delta)}, \tag{84}$$

where we used the facts that $\Phi_0(\delta) = F(\delta) - \Phi_1(\delta)$, and $F(\delta^\star) = 1 - s$ due to the assumed continuity of the distribution. Substituting this identity into (82), we obtain:

$$|A(\lambda)| \leq \int_0^{\delta^\star} \frac{\left| F'(\delta^\star)(\delta^\star - \delta) - (F(\delta^\star) - F(\delta)) \right|}{\theta_1 F'(\delta^\star)\left(\delta^\star - \delta\right)\left(F(\delta^\star) - F(\delta)\right)}\, d\delta.$$

Under our assumption that the distribution of utility types is is twice continuously differentiable, we can use Taylor's Theorem to extend the integrand by continuity at $\delta^\star$, with value

$$\lim_{\delta \to \delta^\star} \frac{|F'(\delta^\star)(\delta^\star - \delta) - (F(\delta^\star) - F(\delta))|}{\theta_1 F'(\delta^\star)(\delta^\star - \delta)(F(\delta^\star) - F(\delta))} = \frac{|F''(\delta^\star)|}{2\theta_1 (F'(\delta^\star))^2}.$$

Since the derivative is bounded away from zero this shows that the integrand is bounded and (82) follows. Turning to (83) we start by observing that because of (1) and the assumed continuity of the distribution of utility types we have

$$\Phi_1(\delta) = F(\delta) - F(\delta^\star) + F(\delta^\star) - \Phi_0(\delta) = F(\delta) - F(\delta^\star) + 1 - s - \Phi_0(\delta).$$

Substituting this identity into (84) shows that

$$\lambda \sigma(\delta) = \frac{1}{\frac{r+\gamma}{\lambda} + \theta_0 (F(\delta) - F(\delta^\star)) + 1 - s - \Phi_0(\delta)},$$

and the desired result now follows from the same argument as above. ∎

Lemma A.15 *Assume that the conditions of Proposition 13 hold true. Then*

$$A_0(\lambda) \equiv \int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} = \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1).$$

as the meeting rate $\lambda \to \infty$.

Proof. To establish a lower bound we start by noting that $\Phi_1(\delta) \leq \Phi_1(\delta^\star)$ for all $\delta \leq \delta^\star$. Substituting this into the definition of $A_0(\lambda)$ and integrating we find that

$$\begin{aligned} A_0(\lambda) &\geq \int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta^\star)} \\ &= \left[\frac{-1}{\theta_1 F'(\delta^\star)} \log \left(\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta^\star) \right) \right]_0^{\delta^\star} \\ &= \frac{-1}{\theta_1 F'(\delta^\star)} \log \left(\sqrt{(\gamma/\lambda)s(1-s)} + o\left(\frac{1}{\sqrt{\lambda}}\right) \right) + O(1) = \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1), \end{aligned}$$

where the second equality follows from the asymptotic expansion of $\Phi_1(\delta^\star)$ given in equation (72) above. To establish the reverse inequality let us break down the integral into an integral over the interval $[0, \delta^\star - 1/\sqrt{\lambda}]$, and an integral over the interval $[\delta^\star - 1/\sqrt{\lambda}, \delta^\star]$. A direct calculation shows that the first integral can be bounded above by:

$$\int_0^{\delta^\star - 1/\sqrt{\lambda}} \frac{d\delta}{\theta_1 F'(\delta^\star)(\delta^\star - \delta)} = \frac{1}{\theta_1 F'(\delta^\star)} \log \left(\delta^\star \sqrt{\lambda} \right) = \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1).$$

On the other hand, noting that

$$\inf_{\delta \in [\delta^\star - 1/\sqrt{\lambda}, \delta^\star]} \Phi_1(\delta) \geq \Phi_1 \left(\delta^\star - \frac{1}{\sqrt{\lambda}} \right) = \frac{g_1(-1)}{\sqrt{\lambda}}$$

and integrating we find that the second integral can be bounded from above by

$$\int_{\delta^\star-1/\sqrt{\lambda}}^{\delta^\star} \frac{d\delta}{\theta_1 F'(\delta^\star)(\delta^\star - \delta) + g_1(-1)/\sqrt{\lambda}} = \frac{1}{\theta_1 F'(\delta^\star)} \log\left(1 + \frac{\theta_1 F'(\delta^\star)}{g_1(-1)}\right) = O(1)$$

where the last equality follows Lemma A.11. ∎

Lemma A.16 *Assume that the conditions of Proposition 13 hold true. Then*

$$B_0(\lambda) \equiv \int_{\delta^\star}^{1} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta - \delta^\star) + 1 - s - \Phi_0(\delta)} = \frac{\log(\lambda)}{2\theta_0 F'(\delta^\star)} + O(1)$$

as the meeting rate $\lambda \to \infty$.

Proof. The proof is similar to that of Lemma A.15. We omit the details. ∎

A.5   Additional results

Lemma A.17 *The steady-state expected time to trade is given by*

$$\eta_q(\delta) = \left(1 - \int_0^1 \frac{\gamma dF(\delta')}{\gamma + \lambda_q(\delta')}\right)^{-1} \frac{1}{\gamma + \lambda_q(\delta)}.$$

The steady-state expected time to trade is decreasing in δ for non-owners, increasing in δ for owners, and decreasing in λ for both owners and non-owners.

Proof of Lemma A.17. Consider an agent of ownership type q and denote his utility type process by δ_t. The next time this agent trades is the first time ϱ_q at which he meets an agent of ownership type $1 - q$ whose utility type is such that

$$(2q - 1)(\delta' - \delta_\tau) \geq 0.$$

In the steady state, the arrival rate of this event is

$$\lambda_q(\delta_t) = \lambda q(1 - s - \Phi_0(\delta_t)) + \lambda(1 - q)\Phi_1(\delta_t),$$

and it follows that

$$\eta_q(\delta) = \mathbb{E}[\varrho_q] = \mathbb{E}\left[\int_0^\infty t\, d\left(1 - e^{-\int_0^t \lambda_q(\delta_s)ds}\right)\right] = \mathbb{E}\left[\int_0^\infty e^{-\int_0^t \lambda_q(\delta_s)ds}dt\right].$$

Let σ denote the first time that the agent's utility type changes. Combining the above expression with the law of iterated expectations gives

$$\eta_q(\delta) = \mathbb{E}\left[\int_0^\sigma e^{-\int_0^t \lambda_q(\delta_s)ds}dt + e^{-\int_0^\sigma \lambda_q(\delta_s)ds}\eta_q(\delta_\sigma)\right] \tag{85}$$

$$= \mathbb{E}\left[\int_0^\sigma e^{-\lambda_q(\delta)t}dt + e^{-\lambda_q(\delta)\sigma}\eta_q(\delta_\sigma)\right] = \frac{1}{\gamma + \lambda_q(\delta)}\left(1 + \gamma \int_0^1 \eta_q(\delta')dF(\delta')\right),$$

where the second equality follows from the fact that the agent's utility type rate is constant over $[\![0, \sigma]\!]$, and the third equality follows from the fact that

$$\mathbb{P}\left(\{\sigma \in dt\} \cup \{\delta_\sigma \leq \delta'\}\right) = \gamma e^{-\gamma t} F(\delta') dt.$$

Integrating both sides of (85) against the cumulative distribution function $F(\delta)$ and solving the resulting equation gives

$$1 + \gamma \int_0^1 \eta_q(\delta') dF(\delta') = \left(1 - \gamma \int_0^1 \frac{dF(\delta')}{\gamma + \lambda_q(\delta')}\right)^{-1},$$

and substituting back into (85) gives

$$\eta_q(\delta) = \frac{1}{\gamma + \lambda_q(\delta)} \left(1 - \gamma \int_0^1 \frac{dF(\delta')}{\gamma + \lambda_q(\delta')}\right)^{-1}. \tag{86}$$

Now assume that the cumulative distribution function $F(\delta)$ is continuous. Combining Proposition 4 with Lemma A.8 and the change of variable formula for Stieltjes integrals shows that the integral on the right-hand side can be calculated as

$$\gamma \int_0^1 \frac{dF(\delta')}{\gamma + \lambda_q(\delta')} = \int_0^1 \frac{\gamma dx}{\gamma + \lambda q(1 - s - x) + \lambda \ell(x)} = \kappa(\gamma/\lambda, \Phi_q(1)),$$

where the function $\ell(x)$ is in (51) and we have set

$$\kappa(a, x) = 1 + a \log\left(\frac{1+a}{a}\right) + \left(1 - \frac{1+a}{x}\right) \log\left(\frac{1+a}{1+a-x}\right).$$

Substituting this expression back into (86) and simplifying the result gives the explicit formula for the waiting time reported in the statement.

The comparative statics with respect to δ follow from (86) and the fact that $\lambda_q(\delta)$ is increasing in δ for owners and decreasing for non-owners. On the other hand, a direct calculation shows that

$$\lambda \frac{\partial^2 \lambda_q(\delta)}{\partial \lambda^2} = \phi^2 \frac{\partial^2 \Phi_1(\delta)}{\partial \phi^2},$$

where we have set $\phi = \gamma/\lambda$. Since $\Phi_1(\delta)$ is concave in ϕ (see the proof of Lemma A.7 below), this shows that $\lambda_q(\delta)$ is concave in the meeting rate, and it follows that

$$\frac{\partial \lambda_q(\delta)}{\partial \lambda} \geq \lim_{\lambda \to \infty} \frac{\partial \lambda_q(\delta)}{\partial \lambda} = \lim_{\lambda \to \infty} q(1 - s - \Phi_0(\delta)) + \lim_{\lambda \to \infty} (1 - q) \Phi_1(\delta) + \lim_{\lambda \to \infty} \lambda \frac{\partial \Phi_1(\delta)}{\partial \lambda}$$
$$= q(1 - s - F(m))^+ + (1 - q)(1 - s - F(m))^- \geq 0,$$

where the second equality results from Lemma A.7 and the fact that

$$\lim_{\lambda \to \infty} \lambda \frac{\partial \Phi_1(\delta)}{\partial \lambda} = -\lim_{\phi \to 0} \phi \frac{\partial \Phi_1(\delta)}{\partial \phi}$$

$$= -\lim_{\phi \to 0} \frac{\phi s(1-s)F(\delta)(1-F(\delta))}{(\phi + \Phi_1(\delta) + (1-s)(1-F(\delta)))\Lambda(\delta)} = 0$$

due to (50). This shows that $\lambda_q(\delta)$ is an increasing function of λ, and the desired result now follows from (86) by noting that the distribution function $F(\delta)$ does not depend on the meeting intensity.

To complete the proof, it remains to establish the comparative statics with respect to the asset supply. An immediate calculation shows that

$$\frac{\partial \lambda_q(\delta)}{\partial s} = \lambda \left(\frac{\partial \Phi_1(\delta)}{\partial s} - q \right)$$

and the result for non-owners follows from Lemma A.7 below. Now consider asset owners. Since

$$\frac{\partial^2 \lambda_1(\delta)}{\partial s^2} = \frac{\partial^2 \Phi_1(\delta)}{\partial s^2} = \frac{2\gamma(1+\phi)F(\delta)(1-F(\delta))}{\Lambda(m)^3} \geq 0,$$

we have that $\lambda_1(\delta)$ is convex in s, and it now follows from (49) that

$$\frac{\partial \lambda_1(\delta)}{\partial s} \leq \left. \frac{\partial \lambda_1(\delta)}{\partial s} \right|_{s=1} = \lambda \left(\frac{(1+\phi)F(\delta)}{\phi + F(\delta)} - 1 \right) \leq 0.$$

This shows that $\lambda_1(\delta)$ is decreasing in s, and the desired result now follows from (86) by noting that the function $F(\delta)$ does not depend on s. ∎

Because a non-owner with utility type $\delta \in [0,1]$ only buys from owners with utility types $\delta' \leq \delta$, we have that the expected buying price is given by

$$p_0(\delta) = \theta_1 \Delta V(\delta) + \theta_0 \int_0^\delta \Delta V(\delta') \frac{d\Phi_1(\delta')}{\Phi_1(\delta)}.$$

Similarly, because an owner with utility type δ only sells to non-owners with utility types $\delta' \geq \delta$, we have that the expected selling price is given by

$$p_1(\delta) = \theta_0 \Delta V(\delta) + \theta_1 \int_{\delta_-}^1 \Delta V(\delta') \frac{d\Phi_0(\delta')}{1 - s - \Phi_0(\delta_-)}.$$

Our next result establishes formally that the expected buying and selling prices are both increasing in utility type, while the effect of increasing the meeting rate is ambiguous.

Lemma A.18 *The expected trading price $p_q(\delta)$ is increasing in $\delta \in [0,1]$ for $q \in \{0,1\}$, but can be non-monotonic in λ.*

Proof of Lemma A.18. Because the reservation value function is absolutely continuous in $\delta \in [0,1]$, it

follows from an integration by parts that the expected buyer price can be written as

$$p_0(\delta) = \theta_1 \Delta V(\delta) + \theta_0 \Delta V(0) + \theta_0 \int_0^\delta \sigma(\delta') \left(1 - \frac{\Phi_1(\delta')}{\Phi_1(\delta)} \right) d\delta'$$

$$= \Delta V(\delta) - \theta_0 \int_0^\delta \sigma(\delta') \frac{\Phi_1(\delta')}{\Phi_1(\delta)} d\delta' = \Delta V(0) + \int_0^\delta \sigma(\delta') \left(1 - \theta_0 \frac{\Phi_1(\delta')}{\Phi_1(\delta)} \right) d\delta'.$$

The required monotonicity now follows by observing that $\sigma(\delta')$ is nonnegative and that the function in the bracket under the integral sign is increasing in δ. Similarly, the expected seller price can be written as

$$p_1(\delta) = \Delta V(1) - \int_\delta^1 \sigma(\delta') \left(1 - \theta_1 \frac{1 - s - \Phi_0(\delta')}{1 - s - \Phi_0(\delta-)} \right) d\delta',$$

and the required monotonicity follows by observing that the function in the bracket under the integral sign is decreasing in δ. ∎

Lemma A.19 *The spread between the highest and lowest realized price is*

$$\Delta V(\overline{\delta}) - \Delta V(\underline{\delta}) = \int_{\underline{\delta}}^{\overline{\delta}} \frac{d\delta}{r + \gamma + \lambda \theta_1 (1 - s - \Phi_0(\delta)) + \lambda \theta_0 \Phi_1(\delta)}.$$

where the constants $0 < \underline{\delta} \leq \overline{\delta} < 1$ denote the endpoints of the support of the distribution of utility types. The spread between the highest and lowest realized price is decreasing in both the meeting rate, λ, and the arrival rate of preference shocks, γ.

Proof of Lemma A.19. The expression for the spread follows directly from the explicit expression for the steady state reservation values in (20). To establish the second part we write

$$\frac{1}{\sigma(\delta)} = r + \gamma + \lambda(1 - s - F(\delta)) + \lambda \Phi_1(\delta) = r + \gamma + \theta_1 \lambda_0(\delta) + \theta_0 \lambda_1(\delta)$$

and observe that the steady state distribution $\Phi_1(\delta)$ is increasing in γ by Corollary 2, and that $\lambda_q(\delta)$ is increasing in λ as shown in the proof of Lemma A.17. ∎

B The welfare cost of frictions

In this appendix, we briefly study the asymptotic impact of frictions on welfare. To this end we start by noting that in the context of our model the welfare cost of misallocation can be defined as

$$C(\lambda) \equiv - \int_0^{\delta^\star} \delta d\Phi_1(\delta) + \int_{\delta^\star}^1 \delta d\Phi_0(\delta) = \int_0^{\delta^\star} \Phi_1(\delta) d\delta + \int_{\delta^\star}^1 (1 - s - \Phi_0(\delta)) d\delta$$

where the second equality follows from an integration by parts. The two terms on the right-hand side of this definition capture the two types of misallocation arising in our model. The first term accounts for the utility derived by investors who hold an asset when they should not, and the second term account for the utility not derived by investors who should hold an asset.

Proposition B.1 *Assume that the distribution of utility types is twice continuously differentiable with a derivative that is bounded away from zero. Then*

$$C(\lambda) = \frac{\gamma s(1-s)}{F'(\delta^\star)} \frac{\log(\lambda)}{\lambda} + O\left(\frac{1}{\lambda}\right).$$

By contrast, with a discrete distribution of utility types, the convergence rate of the welfare cost to zero is generically equal to $1/\lambda$.

Proof. The quadratic equation for the equilibrium distribution and the assumed continuity of the distribution of utility types jointly imply that

$$\lambda\Phi_1(\delta) = \frac{\gamma s F(\delta)}{\gamma/\lambda + \Phi_1(\delta) + F(\delta^\star) - F(\delta)},$$

and combining this identity with arguments similar to those we used in the proof of Lemma A.14 shows that the first integral in the definition of the welfare cost satisfies

$$\left| \int_0^{\delta^\star} \left(\lambda\Phi_1(\delta) - \frac{\gamma s F(\delta^\star)}{\gamma/\lambda + \Phi_1(\delta) + F'(\delta^\star)(\delta^\star - \delta)} \right) d\delta \right| = O(1). \tag{87}$$

On the other hand, the same arguments as in the proof of Lemma A.15 imply that

$$\int_{\delta^\star - 1/\sqrt{\lambda}}^{\delta^\star} \frac{\gamma s F(\delta^\star) d\delta}{\gamma/\lambda + \Phi_1(\delta) + F'(\delta^\star)(\delta^\star - \delta)} \leq \frac{\gamma s F(\delta^\star)}{F'(\delta^\star)} \log\left(1 + \frac{F'(\delta^\star)}{g_1(-1)}\right) = O(1)$$

and combining this inequality with (87) gives

$$\int_0^{\delta^\star} \lambda\Phi_1(\delta) d\delta = \int_0^{\delta^\star - 1/\sqrt{\lambda}} \frac{\gamma s F(\delta^\star)}{\gamma/\lambda + \Phi_1(\delta) + F'(\delta^\star)(\delta^\star - \delta)} d\delta + O(1).$$

To obtain a lower bound for the integral, we can bound $\Phi_1(\delta)$ above by $\Phi_1(\delta^\star - 1/\sqrt{\lambda})$, and to obtain an upper bound, we can bound $\Phi_1(\delta)$ below by zero. In both cases, we can compute the resulting integral explicitly and we find that the upper and the lower bound can both be written as

$$\frac{\gamma s F(\delta^\star)}{2F'(\delta^\star)} \log(\lambda) + O(1) = \frac{\gamma s(1-s)}{2F'(\delta^\star)} \log(\lambda) + O(1).$$

Going through the same steps shows that the second integral satisfies

$$\int_{\delta^\star}^1 \lambda\left(1 - s - \Phi_0(\delta)\right) d\delta = \frac{\gamma s(1-s)}{2F'(\delta^\star)} \log(\lambda) + O(1)$$

and the desired result now follows by adding up the asymptotic expansions of the two integrals. In order to complete the proof assume that the distribution of utility types is discrete. Using the same notation as in the

proof of Proposition 12 we find that

$$C(\lambda) = \sum_{i=0}^{m-1} (\delta_{i+1} - \delta_i)\Phi_1(\delta_i) + \sum_{i=m}^{I} (\delta_{i+1} - \delta_i)(1 - s - F(\delta_i) + \Phi_1(\delta_i))$$

and the desired conclusion follows from the expansion of $\Phi_1(\delta_i)$ given in equation (81). ∎

The proposition shows that, as was the case for price levels and price dispersion, search frictions have a larger welfare impact when the distribution of utility types is continuous, than when it is discrete. It also shows that the welfare cost of frictions may be accurately measured by the observed amount of price dispersion because the two quantities converge to their frictionless counterparts at the same speed.

C Non-stationary initial conditions

Assume that the initial distribution of utility types in the population is given by an arbitrary cumulative distribution function $F_0(\delta)$, which need not even be absolutely continuous with respect to $F(\delta)$. Since the reservation values of Proposition 2 are valid for any joint distribution of types and asset holdings, we need only to determine the evolution of the equilibrium distributions in order to derive the unique equilibrium.

Consider first the distribution of utility types in the whole population. Since upon a preference shock each agent draws a new utility type from $F(\delta)$ with intensity γ, we have that

$$\dot{F}_t(\delta) = \gamma(F(\delta) - F_t(\delta)).$$

Solving this ordinary differential equation shows that the cumulative distribution of utility types in the whole population is explicitly given by

$$F_t(\delta) = F(\delta) + e^{-\gamma t}\left(F_0(\delta) - F(\delta)\right)$$

and converges to the long-run distribution $F(\delta)$ in infinite time. On the other hand, the same arguments as in Section 3.2 show that in equilibrium the distributions of perceived growth rate among the population of asset owners solves the differential equation

$$\dot{\Phi}_{1,t}(\delta) = -\lambda\Phi_{1,t}(\delta)^2 - \lambda(1 - s - F_t(m) + \Phi_{1,t}(\delta)) + \gamma(sF(\delta) - \Phi_{1,t}(\delta)).$$

Given an initial condition satisfying the accounting identity

$$\Phi_{0,0}(\delta) + \Phi_{1,0}(\delta) = F_0(\delta)$$

this Riccati equation admits a unique solution that can be expressed in terms of the confluent hypergeometric function of the first kind $M_1(a, b; x)$ (see Abramowitz and Stegun (1964)) as

$$\lambda\Phi_{1,t}(\delta) = \lambda(F_t(m) - \Phi_{0,t}(\delta)) = \frac{\dot{Y}_{+,t}(\delta) - A(\delta)\dot{Y}_{-,t}(\delta)}{Y_{+,t}(\delta) - A(\delta)Y_{-,t}(\delta)} \tag{88}$$

with

$$Y_{\pm,t}(\delta) = e^{-\lambda Z_{\pm}(\delta)t}W_{\pm,t}(\delta)$$

$$Z_{\pm}(\delta) = \frac{1}{2}\left(1 - s + \gamma/\lambda - F(\delta)\right) \pm \frac{1}{2}\Lambda(\delta) \tag{89}$$

$$W_{\pm,t}(\delta) = M_1\left(\frac{\lambda}{\gamma}Z_{\pm}(\delta), 1 \pm \frac{\lambda}{\gamma}\Lambda(m); e^{-\gamma t}\frac{\lambda}{\gamma}(F(\delta) - F_0(\delta))\right)$$

and

$$A(\delta) = \frac{\dot{Y}_{+,0}(\delta) - \lambda\Phi_{1,0}(\delta)Y_{+,0}(\delta)}{\dot{Y}_{-,0}(\delta) - \lambda\Phi_{1,0}(\delta)Y_{-,0}(\delta)}.$$

The following lemma relies on standard properties of confluent hypergeometric functions to show that the above cumulative distribution function converges to the same steady-state distribution as in the case with stationary initial condition.

Lemma C.1 *The equilibrium distributions defined by* (88) *satisfy* $\lim_{t\to\infty}\Phi_{q,t}(\delta) = \Phi_q(\delta)$ *for any initial distributions* $F_0(\delta)$ *and* $F_{1,0}(\delta)$.

Proof. Straightforward algebra shows that (88) can be rewritten as

$$\lambda\Phi_{1,t}(\delta) = \frac{\lambda Z_+(\delta)W_{+,t}(\delta) - \dot{W}_{-,t}(\delta) + e^{\lambda\Lambda(\delta)t}A(\delta)(\dot{W}_{+,t}(\delta) - \lambda Z_-(\delta)W_{-,t}(\delta))}{e^{\lambda\Lambda(\delta)t}A(\delta)W_{-,t}(\delta) - W_{+,t}(\delta)}.$$

On the other hand, using standard properties of the confluent hypergeometric function of the first kind it can be shown that we have

$$\lim_{\delta\to\infty}\dot{W}_{\pm,t}(\delta) = \lim_{\delta\to\infty}(1 - W_{\pm,t}(\delta)) = 0$$

and combining these identities we deduce that

$$\lim_{\delta\to\infty}\lambda\Phi_{1,t}(\delta) = -\lambda Z_-(\delta) + \lim_{\delta\to\infty}\frac{\dot{W}_{+,t}(\delta)}{W_{-,t}(\delta)} = -\lambda Z_-(\delta) = \lambda\Phi_1(\delta),$$

where the last equality follows from (89) and the definition of the steady-state distribution $\Phi_1(\delta)$. ∎

Given the joint distribution of types and asset holdings the equilibrium can be computed by substituting the equilibrium distributions into (12) and (13), and the same arguments as in the stationary case show that this equilibrium converges to the same steady-state equilibrium as in Theorem 1.

D Marketmakers

Assume that in addition to a continuum of agents, the market also includes a unit mass of competitive marketmakers who have access to a frictionless interdealer market and keep no inventory. An agent contacts marketmakers with intensity $\alpha \geq 0$. When an agent meets a market maker, they bargain over the terms of a potential trade, and we assume that the result of this negotiation is given by the Nash bargaining solution with bargaining power $1 - z \in [0, 1]$ for the market maker.

D.1 Pricing in the interdealer market

Let Π_t denote the asset price on the interdealer market and consider a meeting between a market maker and an investor of type $\delta \in [0, 1]$ who owns $q \in \{0, 1\}$ units of the asset. Such a meeting results in a trade if and only if the trade surplus

$$S_{q,t}(\delta_t) = (2q - 1)(\Pi_t - \Delta V_t(\delta)) = (2q - 1)(\Pi_t - V_{1,t}(\delta) + V_{0,t}(\delta))$$

is nonnegative, in which case the assumption of Nash bargaining implies that the realized price is

$$\hat{P}_t(\delta) = (1 - z)\Delta V_t(\delta) + z\Pi_t.$$

If reservation values are increasing in type, which we show is the case below, then there must be a cutoff $w_t \in [0, 1]$ such that only owners of type $\delta \le w_t$ are willing to sell, while only those non-owners of type $\delta \ge w_t$ are willing to buy. Since marketmakers must be indifferent to trading with marginal agents this, in turn, implies that the price on the interdealer market is $\Pi_t = \Delta V_t(w_t)$

To determine the cutoff we use the fact that the positions of marketmakers must net out to zero because they keep no inventory. The total mass of owners who contact marketmakers to sell is $\alpha \Phi_{1,t}(w_t)$. On the other hand, the total mass of non-owners who contact marketmakers to buy is

$$\alpha(1 - s - \Phi_{0,t}(w_t) + \Delta\Phi_{0,t}(w_t)).$$

Because the distribution of utility types can have atoms, some randomization may be required at the margin. Taking this into account shows that the interdealer market clearing condition is

$$\Phi_{1,t}(w_t) - (1 - \pi_{1,t})\Delta\Phi_{1,t}(w_t) = 1 - s - \Phi_{0,t}(w_t) + \pi_{0,t}\Delta\Phi_{0,t}(w_t) \tag{90}$$

where $\pi_{q,t} \in [0, 1]$ denotes the probability with which marketmakers execute orders from marginal agents. Since the distribution is by assumption not flat at the supply level, this condition implies that the cutoff is uniquely given by $w_t \equiv \delta^\star$ for all $t \ge 0$, and it now remains to determine the execution probabilities. Two cases may arise depending on the properties of the distribution. If $F(\delta^\star) = 1 - s$, then the execution probabilities are uniquely defined by $\pi_{q,t} = q$, and only marginal buyers get rationed in equilibrium. On the contrary, if the cutoff is an atom, then the execution probabilities are not uniquely defined. In this case, one may, for example, take

$$\pi_{0,t} = 1 - \pi_{1,t} = \frac{F(\delta^\star) - (1 - s)}{\Delta F(\delta^\star)}$$

so that a fraction of both marginal buyers and marginal sellers get rationed in equilibrium, but many other choices are also compatible with market clearing. By construction, this choice has no influence on the welfare of agents, and we verify below but it also does not have any impact on the evolution of the equilibrium distribution of utility types.

D.2 Equilibrium value functions

Taking as given the evolution of the joint distribution of types and asset holdings, and proceeding as in Section 3.1 shows that the reservation value function solves

$$\Delta V_t(\delta) = \mathbb{E}_t \Bigg[\int_t^\sigma e^{-r(u-t)} \delta du + e^{-r(\sigma-t)} \bigg(\Delta V_\sigma(\delta) + \mathbf{1}_{\{\sigma=\tau_\alpha\}} z \left(\Delta V_\sigma(\delta^\star) - \Delta V_\sigma(\delta) \right) \tag{91}$$

$$+ \mathbf{1}_{\{\sigma=\tau_\gamma\}} \int_0^1 (\Delta V_\sigma(\delta') - \Delta V_\sigma(\delta)) dF(\delta')$$

$$+ \mathbf{1}_{\{\sigma=\tau_0\}} \theta_1 \int_0^1 \left(\Delta V_\sigma(\delta') - \Delta V_\sigma(\delta) \right)^+ \frac{d\Phi_{0,\sigma}(\delta')}{1-s}$$

$$- \mathbf{1}_{\{\sigma=\tau_1\}} \theta_0 \int_0^1 \left(\Delta V_\sigma(\delta) - \Delta V_\sigma(\delta') \right)^+ \frac{d\Phi_{1,\sigma}(\delta')}{s} \bigg) \Bigg]$$

subject to (8), where τ_α is an exponentially distributed random variable with parameter α that represents a meeting with a market maker, and $\sigma = \min\{\tau_0, \tau_1, \tau_\gamma, \tau_\alpha\}$. Comparing this equation with (7) shows that the reservation value function in an environment with marketmakers is isomorphic to that which would prevail in an environment where there are no marketmakers, the distribution of types is

$$\hat{F}(\delta) \equiv \frac{\gamma}{\gamma + \alpha z} F(\delta) + \left(1 - \frac{\gamma}{\gamma + \alpha z} \right) \mathbf{1}_{\{\delta \geq \delta^\star\}}, \qquad \delta \in [0,1],$$

and the arrival rate of type changes is $\hat{\gamma} = \gamma + \alpha z$. Combining this observation with Proposition 1 delivers the following characterization of reservation values in the model with marketmakers.

Lemma D.1 *There exists a unique function that satisfies* (91) *subject to* (8). *This function is uniformly bounded, strictly increasing in space, and given by*

$$\Delta V_t(\delta) = \int_t^\infty e^{-r(u-t)} \bigg(\delta - \int_0^\delta \hat{\sigma}_u(\delta') \left(\hat{\gamma} \hat{F}(\delta') + \lambda \theta_0 \Phi_{1,u}(\delta') \right) d\delta'$$

$$+ \int_\delta^1 \hat{\sigma}_u(\delta') \left(\hat{\gamma}(1 - \hat{F}(\delta')) + \lambda \theta_1 (1 - s - \Phi_{0,u}(\delta')) \right) d\delta' \bigg) du,$$

where the local surplus $\hat{\sigma}_t(\delta)$ is defined as in (12) *albeit with $\hat{\gamma}$ in place of γ.*

D.3 Equilibrium distribution of types

Because agents can trade both among themselves and with marketmakers, the evolution of the equilibrium distributions must include additional terms to reflect the new trading opportunities.

Consider the group of asset owners with utility type $\delta' \leq \delta$. In addition to the entry channels of the benchmark model, an agent may enter this group because he is a non-owner with $\delta'' \leq \delta$ who buys one unit of the asset from a market maker. The contribution of such entries is

$$\mathcal{E}_t(\delta) = \alpha \left((\Phi_{0,t}(\delta) - \Phi_{0,t}(\delta^\star))^+ + \mathbf{1}_{\{\delta^\star \leq \delta\}} \pi_{0,t} \Delta\Phi_{0,t}(\delta^\star) \right)$$

where the last term takes into account the fact that not all meetings with marginal buyers result in a trade. On the other hand, an agent may exit this group because he is an asset owner with $\delta'' \leq \delta$ who sells to a

market maker. The contribution of such exits is

$$\mathcal{X}_t(\delta) = \alpha \left(\Phi_{1,t}(\delta \wedge \delta^\star) - \mathbf{1}_{\{\delta^\star \leq \delta\}}(1 - \pi_{1,t})\Delta\Phi_{1,t}(\delta^\star) \right).$$

Gathering these contributions and using (90) shows that the total contribution of intermediated trades is explicitly given by

$$\mathcal{E}_t(\delta) - \mathcal{X}_t(\delta) = -\alpha\Phi_{1,t}(\delta) + \alpha(1 - s - F(\delta))^-.$$

Finally, combining this with (17) shows that the rate of change is given by

$$\dot{\Phi}_{1,t}(\delta) = -\lambda\Phi_{1,t}(\delta)^2 - \Phi_{1,t}(\delta)\left(\gamma + \lambda(1 - s - F(\delta))\right)$$
$$- \alpha\Phi_{1,t}(m) + \gamma s F(\delta) + \alpha(1 - s - F(m))^-$$

and does not depends on the choice of the probabilities $\pi_{0,t}$ and $\pi_{1,t}$ with which marketmakers execute orders from marginal agents. To solve this Riccati differential equation, let

$$\Phi_1(\delta) = -\frac{1}{2}\left(1 - s + \gamma/\lambda + \alpha/\lambda - F(\delta)\right) + \frac{1}{2}\Psi(\delta) \tag{92}$$

with

$$\Psi(\delta) = \sqrt{(1 - s + \gamma/\lambda + \alpha/\lambda - F(\delta))^2 + 4s(\gamma/\lambda)F(\delta) + 4(\alpha/\lambda)(1 - s - F(\delta))^-}$$

denote the steady-state distribution of owners with utility type less than δ, i.e., the unique strictly positive solution to $\dot{\Phi}_{1,t}(\delta) = 0$. The following results are the direct counterparts of Proposition 4 and Corollary 2 for the model with marketmakers.

Proposition D.2 *At any time $t \geq 0$ the measure of the set asset owners with utility type less than or equal to $\delta \in [0, 1]$ is explicitly given by*

$$\Phi_{1,t}(\delta) = \Phi_1(\delta) + \frac{(\Phi_{1,0}(\delta) - \Phi_1(\delta))\,\Psi(\delta)}{\Psi(\delta) + (\Phi_{1,0}(\delta) - \Phi_1(\delta) + \Psi(\delta))\left(e^{\lambda\Psi(\delta)t} - 1\right)} \tag{93}$$

and converges pointwise monotonically to $\Phi_1(\delta)$ from any initial condition satisfying (1) and (2).

Proof. The proof is analogous to that of Proposition 4. ∎

Corollary D.3 *The steady-state measure $\Phi_1(\delta)$ is increasing in γ and decreasing in λ, and it converges to the frictionless measure $\Phi_1^\star(\delta)$ as $\gamma/\lambda \to 0$.*

Proof. The proof is analogous to that of Corollary 2 and Proposition 10. ∎

D.4 Equilibrium

Definition 2 *An equilibrium is a reservation value function $\Delta V_t(\delta)$ and a pair of distributions $\Phi_{0,t}(\delta)$ and $\Phi_{1,t}(\delta)$) such that: the distributions satisfy (1), (2) and (93), and the reservation value function satisfies (7) subject to (8) given the distributions.*

As in the benchmark model without marketmakers, a full characterization of the unique equilibrium follows immediately from our explicit characterization of the reservation value function and the joint distribution of types and asset holdings.

Theorem D.4 *There exists a unique equilibrium with marketmakers. Moreover, given any initial conditions satisfying (1) and (2), this equilibrium converges to the steady-state distributions of equation (92) and the reservation value function of equation (20) albeit with $(\hat{\gamma}, \hat{F}(\delta))$ in place of $(\gamma, F(\delta))$.*

Relying on Theorem D.4, it is possible to derive the counterparts of our results regarding the expected time to trade, the equilibrium trading volume, the equilibrium misallocation, and the asymptotic price impact of search frictions for the model with marketmakers, and verify that the corresponding predictions are qualitatively similar to those of the benchmark model.