Heterogeneity in Decentralized Asset Markets[*]

Julien Hugonnier[†]   Benjamin Lester[‡]   Pierre-Olivier Weill[§]

November 10, 2014

Abstract

We study a search and bargaining model of asset markets, where investors' heterogeneous valuations for the asset are drawn from an arbitrary distribution. We use a solution technique that renders the analysis fully tractable: we provide a full characterization of the equilibrium, in closed form, both in and out of steady state. We use this characterization for two purposes. First, we establish that the model can easily and naturally account for a number of stylized facts that have been documented in empirical studies of over-the-counter asset markets. Second, we show that the model generates a number of novel results that underscore the importance of heterogeneity in decentralized markets. We highlight two: first, heterogeneity magnifies the price impact of search frictions; and second, search frictions have larger effects on price levels than on dispersion, so that quantifying the magnitude of search frictions based on observed dispersion can be misleading.

Keywords: search frictions, bargaining, continuum of types, price dispersion

JEL Classification: G11, G12, G21

1

1 Introduction

Many assets trade in decentralized or "over-the-counter" (OTC) markets.[1] To study these markets, we construct a search and bargaining model in which investors with heterogeneous valuations for an asset are periodically and randomly matched in pairs and given the opportunity to trade. Importantly, we allow investors' valuations to be drawn from an arbitrary distribution of types, whereas the existing literature, starting with Duffie, Gârleanu, and Pedersen (2005) (henceforth DGP), has primarily focused on the special case of only two valuations. Despite the potential complexities introduced by this generalization, we use techniques that render the analysis fully tractable: we provide a full characterization of the equilibrium, in closed form, both in and out of steady state.

We use this characterization for two purposes. First, we flesh out many of the model's positive implications and establish that our environment—with both search frictions and many types of investors—can easily and naturally account for a number of stylized facts that have been documented in empirical studies of OTC markets. Since the model is also remarkably tractable, we argue that it constitutes a unified framework to study both theoretical and empirical issues related to these markets.[2] Second, we derive several new results which underscore the importance of heterogeneity in decentralized asset markets. We highlight two. First, we show that heterogeneity magnifies the price impact of search frictions. Second, we show that search frictions have larger effects on price levels than on price dispersion, so that quantifying the price impact of search frictions based on observed dispersion can be misleading.

Our model, which we describe formally in Section 2, starts with the basic building blocks of DGP. We assume that there is a fixed measure of investors in the economy and a fixed measure of indivisible shares of an asset. Investors have heterogeneous valuations for this asset, which can change over time, and they are allowed to hold either zero or one share. Each investor is

---

[1]Examples of assets that trade in decentralized markets include corporate and government bonds, emerging-market debt, mortgage-backed securities, and foreign exchange swaps, to name a few. More generally, our analysis applies to a variety of other decentralized markets where the object being traded is durable, such as capital or real estate.

[2]These issues include, but are not limited to, 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 effect of a large shock (such as a liquidity shock) in an OTC market.

periodically and randomly matched with another investor, and a transaction ensues if there are gains from trade, with prices being determined by Nash bargaining. Our main point of departure from the existing literature is that we allow agents' valuations to be drawn from an arbitrary distribution.

This departure from the canonical model makes the analysis considerably more complex, as the state of the economy now includes two (potentially) infinite-dimensional objects: the distributions of valuations among investors that hold zero and one assets, respectively, over time. However, in Section 3 we show that, using the proper methodology, the solution remains fully tractable, yielding a closed-form characterization of the unique equilibrium.

Then, in Section 4, we exploit our characterization of equilibrium to explore the basic, positive predictions that emerge from our model, and we argue that many of these predictions are consistent with the patterns of trade observed in actual OTC markets. A simple, yet crucial insight that emerges from the analysis is that assets are passed from investors with low valuations to those with high valuations through chains of inframarginal trades. In equilibrium, an investor who owns an asset and has a low valuation tends to sell very quickly and then remain asset-less for a long duration. For similar reasons, an investor who does not own an asset and has a high valuation tends to buy very quickly and then holds on to his asset for a long duration. Finally, investors with moderate, "near-marginal" valuations engage in both purchases and sales with similar frequencies, so that the behavior of these investors resembles that of intermediaries. These patterns of trade have important implications for both individual investors and aggregate outcomes.

At the individual level, the model implies a tight relationship between an individual's valuation for the asset, his expected duration of search, the expected valuation of his ultimate counterparty, and thus the expected price he pays or receives. In particular, not only does an asset owner with a low valuation tend to trade quickly, but he tends to trade with a counterparty who also has a relatively low valuation, and hence the transaction price tends to be low. As the asset is passed to agents with higher valuations, each agent holds the asset for a longer period of time, on average, before selling at a higher price. These patterns of trade are consistent with observations from certain OTC markets, such as the municipal bonds and the federal funds markets, but they would

be difficult to generate in more traditional models with *either* centralized exchange or only two valuations. More specifically, in a model with a frictionless, centralized market and many types of investors, or in a model with a decentralized market and only two types of investors, all sellers (and buyers) trade at the same frequency and all trades occur at the same price.[3]

At the aggregate level, the trading patterns described above imply that, even though the network of meetings is fully random in our model, the network of trades that emerge in equilibrium are not. In particular, since investors with "near-marginal" valuations tend to specialize in intermediation, a "core-periphery" trading network emerges endogenously, with transactions occurring at different prices at each node. The tendency for OTC markets to exhibit both a core-periphery network structure and substantial price dispersion has been established in a number of empirical studies.

Finally, in Section 5, we ask how search frictions affect prices, paying particular attention to understanding how the answer depends on the degree of heterogeneity in valuations. To obtain closed form comparative statics, we focus on the case where frictions are small, which is the relevant region of the parameter space for many financial markets, where trade occurs very quickly. We highlight two novel results. First, we show that the price impact of frictions is much larger in environments with arbitrarily large amounts of heterogeneity than it is in environments with only a few types of agents. Hence, heterogeneity magnifies the price impact of search frictions. Second, we show that the effect of search frictions on price levels and price dispersion are of different magnitudes: price levels can be far from their Walrasian counterpart when price dispersion has nearly vanished. Hence, using price dispersion to quantify frictions may lead one to underestimate the true effect of search frictions on price levels.

1.1 Related Literature

Our paper builds off of a recent literature that uses search models to study asset prices and allocations in OTC markets. Many of these papers are based on the basic framework developed in Duffie et al. (2005), who study how search frictions in OTC markets affect the bid-ask spread

---

[3]Even in a model of decentralized trade with many types of investors, trading times are constant across investors and prices are independent of the counterparties' type if trade is organized through dealers who have access to a frictionless inter-dealer market, as is the case in Duffie et al. (2005), Weill (2007), and Lagos and Rocheteau (2009).

set by marketmakers who have access to a competitive interdealer market.[4] The current paper is closer in spirit to Duffie et al. (2007), who study a purely decentralized market—i.e., one without any such marketmakers—where investors with one of two valuations meet and trade directly with one another.

This model of purely decentralized trade captures important features of many markets; for example, Li and Schürhoff (2012) show that, in the municipal bond market, even the inter-dealer market is bilateral, with intermediation chains that typically involve more than two transactions. As such, the model of purely decentralized trade has been used to explore a number of important issues related to liquidity and asset prices; see, for example, Vayanos and Wang (2007), Weill (2008), Vayanos and Weill (2008), Afonso (2011), Gavazza (2011a, 2013), and Feldhütter (2012).[5] However, all of these papers have maintained the assumption of only two valuations, and hence cannot be used to address many of the substantive issues that are analyzed in our paper.[6] Neklyudov (2012) considers an environment with two valuations but introduces heterogeneity in trading speed in order to study the terms of trade that emerge in a core-periphery trading network. In our model, a core-periphery network arises endogenously even though trading speed is constant across investors.

To the best of our knowledge, there are very few papers that consider purely decentralized asset markets and allow for agents to have more than two valuations. Perhaps the closest to

---

[4]Other early papers that used search theory to analyze asset markets include include Gehrig (1993), Spulber (1996), Hall and Rust (2003), and Miao (2006).

[5]For example, Vayanos and Wang (2007) and Vayanos and Weill (2008) show that these models can generate different prices for identical assets, which can help explain the "on-the-run" phenomenom in the Treasury market; Weill (2008) explores how liquidity differentials, which emerge endogenously in these models, can help explain the cross-sectional returns of assets with different quantities of tradeable shares; Gavazza (2011a, 2013) uses these models to explain price differences in the market for commercial aircraft; and Feldhütter (2012) uses these models to generate a measure of selling pressure in the OTC market for corporate bonds.

[6]The framework of Duffie et al. (2005), with a predetermined set of marketmakers, has also been extended in a number of directions. Lagos and Rocheteau (2009), Gârleanu (2009) show how to accommodate additional heterogeneity in this framework, as they allow agents to choose arbitrary asset holdings Lagos, Rocheteau, and Weill (2011) extend this framework even further to study market crashes, as was done previously by Weill (2007) with restricted asset holdings and discrete types. Praz (2013) extend this framework to study asset pricing with correlated assets trading in centralized and decentralized markets. As we discuss below, allowing for heterogeneity in an environment in which investors trade with intermediaries is a simpler and conceptually different exercise since the expected time to trade is constant across all agents on one side of the market, and the price they pay does not depend on the distribution of valuations of agents on the other side of the market. Lester, Rocheteau, and Weill (2014) consider a model in the spirit of Lagos and Rocheteau (2009) with *directed*, instead of random search, and show how this framework can generate a relationship between an investor's valuation, the price he pays, and the time it takes to trade.

our work is Afonso and Lagos (2012b), who develop a model of purely decentralized exchange to study trading dynamics in the Fed Funds market. In their model, agents have heterogenous 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 agents with moderate asset holdings play the role of "endogenous intermediaries," buying from agents with excess reserves and selling to agents with few. As we discuss at length below, similar agents specializing in intermediation emerge in our environment, and have important effects on equilibrium outcomes. However, our work is quite different from 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. This tractability allows us to perform analytical comparative statics and, in particular, derive a number of new results about the patterns of trade, volume, prices, and allocations.[7]

Our paper is also related to an important, 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 (2012) and Colliard and Demange (2014) develop hybrid frameworks, which blend influences from both 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 re-allocate assets towards investors who value them

---

[7]Several other papers also deserve mention here. First, in an online Appendix, Gavazza (2011a) proposes a model of purely decentralized trade with a continuum of types in which agents have to pay a search cost, c, in order to meet others. The optimal strategy is then for an agent with zero (one) asset to search only if his valuation is above (below) a certain threshold R_b (R_s). He focuses on a steady state equilibrium when c is large, so that $R_b > R_s$. Focusing on this case simplifies the analysis considerably, since all investors with the same asset holdings trade at the same frequency, and they trade only once between preference shocks. However, this special case also abstracts from many of the interesting dynamics that emerge from our analysis about trading patterns, the network structure, misallocation, and prices (both in and out of steady state). Second, many of the results presented below were derived independently in two working papers—Hugonnier (2012) and Lester and Weill (2013)—which were later combined to form the current paper. Finally, in a recent working paper, Shen and Yan (2014) exploit a methodology related to ours in an environment with two assets to study the relationship between liquid assets (that trade in frictionless markets) and less liquid assets (that trade in OTC markets).

most is to use an intermediation chain. In our dynamic search model, by contrast, intermediation chains arise by choice. In particular, all investors have the option to search long enough in order to trade directly with their best counterparty. In equilibrium, however, they find it optimal to trade indirectly, through intermediation chains. Hence, even though all contacts are random, the network of actual trades is not random, but rather exhibits a core-periphery-like structure that is typical of many OTC markets in practice.[8]

Finally, our paper is also related to the literatures that use search-theoretic models to study monetary theory and labor economics. For example, a focal point in the former literature is understanding how the price of one particular asset—fiat money—depends on its value or "liquidity" in future transactions; for a seminal contribution, see Kiyotaki and Wright (1993).[9] Naturally, understanding such liquidity premia is central to our analysis as well. 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, which is intuitively similar to the process by which assets in our model are reallocated from low to high valuation investors. The dynamics differ, however, because firms in on-the-job search models can grow without bound, subject to a flow constraint: they hire or lose only one worker at a time.[10] In models of decentralized financial markets, by contrast, investors are typically assumed to hold a bounded number of assets, but they face no flow constraints: they are free to trade any number of assets in a single transaction. Another important difference between our paper and those in the on-the-job search literature has to do with focus: given dramatic increases in trading speed, it is natural for us to study equilibrium volume, trading patterns, and prices as contact rates tend to infinity, while such analysis has no obvious counterpart in the labor market.

---

[8]For another example of endogenous network formation through search, see , e.g., Oberfield (2013).

[9]This literature has recently incorporated assets into the workhorse model of Lagos and Wright (2005) in order to study issues related to financial markets, liquidity, and asset pricing; see, e.g., Lagos (2010), Geromichalos, Licari, and Suárez-Lledó (2007), Lester, Postlewaite, and Wright (2012), and Li, Rocheteau, and Weill (2012).

[10]The reason firms can grow without bound in these models is that they are assumed to operate a constant returns to scale production technology.

2 The model

2.1 Preference, endowments, and matching technology

We consider a continuous-time, infinite-horizon model where 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. The utility flow an investor receives from holding a unit of the asset, which we denote by δ, differs across investors and, for each investor, changes over time. In particular, each investor receives i.i.d. preference shocks that arrive according to a Poisson process with intensity γ, whereupon the investor draws a new utility flow δ' from some cumulative distribution function $F(\delta')$. 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 decentralized or *over-the-counter* market in which each investor initiates contact with another randomly selected investor according to a Poisson process with intensity $\lambda/2$. 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 where 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. It turns out that the model becomes very tractable when we represent these distribution in terms of their cumulative functions. Thus, in what follows, we 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 δ. Assuming that initial types are randomly drawn from the cumulative distribution $F(\delta)$, the

following accounting identities must hold for all $t \geq 0$:[11]

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

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

Equation (1) highlights that the cross-sectional distribution of utility types in the population is constantly equal to $F(\delta)$, which is due to the facts 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 that accounts for the fact that the total measure of investors who own the asset must equal the total measure of assets in the economy. Given our previous assumptions, note that this implies $\Phi_{0,t}(1) = 1 - s$ 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. Since the cross-sectional distribution of types in the population is time-independent, this price will be constant in equilibrium, i.e., $p_t = p$ for all $t \geq 0$.

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

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

after integration by parts. This representation of an investor's objective makes it clear that, at each

---

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

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

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

with $x^+ \equiv \max\{0, x\}$. It now follows from (1) that the distribution of types among investors who do not own the asset is 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

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

and the equilibrium price of the asset has to equal the present value of the marginal investor's utility flow from the asset, so that $p^\star = \delta^\star / r$.[12]

3   Equilibrium with search frictions

In this section, we turn to the economy with search frictions. Our solution method goes beyond previous work by characterizing the equilibrium both in and out of steady state, for an arbitrary distribution of types, in closed form.

Our characterization proceeds in three steps. First, we derive the reservation value of an investor with asset holdings $q \in \{0, 1\}$ and utility type $\delta \in [0, 1]$ at any time $t \geq 0$, taking as given

---

[12]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.

the evolution of the joint distribution of asset holdings and utility types of potential trading partners that such an investor might meet. Importantly, we establish that, given any such distributions satisfying (1) and (2), reservation values are strictly increasing in utility type, so that an asset owner of type δ' and a non-owner of type δ have strictly positive gains from trade if and only if $\delta' \leq \delta$. Then, using the patterns of exchange implied by strictly increasing reservation values, we completely characterize the joint distribution of asset holdings and utility types that must prevail in equilibrium. Finally, given explicit solutions for reservation values and distributions, 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 by

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

the reservation value of an investor of utility type δ at time $t \geq 0$. Note that this reservation value is well defined for all utility types $\delta \in [0,1]$, and not only for utility types in the support of F. An application of Bellman's principle of optimality shows that

$$V_{1,t}(\delta) = \mathbb{E}_t\left[\int_t^\tau e^{-r(u-t)}\delta du + e^{-r(\tau-t)}\left(\mathbf{1}_{\{\tau=\tau_1\}}V_{1,\tau}(\delta) \right. \right. \tag{3}$$
$$+ \mathbf{1}_{\{\tau=\tau_\gamma\}} \int_0^1 V_{1,\tau}(\delta')dF(\delta')$$
$$\left. \left. + \mathbf{1}_{\{\tau=\tau_0\}} \int_0^1 \max\{V_{1,\tau}(\delta), V_{0,\tau}(\delta') + P_\tau(\delta,\delta')\}\frac{d\Phi_{0,\tau}(\delta')}{1-s} \right) \right],$$

where τ_γ is an exponentially distributed random variable with intensity γ that represents the arrival of a preference shock, τ_q is an exponentially distributed random variable with intensity λs if $q=1$ and $\lambda(1-s)$ if $q=0$, that represents the arrival of a meeting with an investor who owns $q \in \{0,1\}$

11

units of the asset, the expectation $\mathbb{E}_t[\cdot]$ 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 asset owner of utility type δ'. Substituting the price (4) into (3) and simplifying the resulting expression shows that

$$V_{1,t}(\delta) = \mathbb{E}_t\left[\int_t^\tau e^{-r(u-t)}\delta du + e^{-r(\tau-t)}\left(V_{1,\tau}(\delta) + \mathbf{1}_{\{\tau=\tau_\gamma\}}\int_0^1 (V_{1,\tau}(\delta') - V_{1,\tau}(\delta))dF(\delta')\right.\right.$$
$$\left.\left. + \mathbf{1}_{\{\tau=\tau_0\}}\int_0^1 \theta_1 \left(\Delta V_\tau(\delta') - \Delta V_\tau(\delta)\right)^+ \frac{d\Phi_{0,\tau}(\delta')}{1-s}\right)\right]. \tag{5}$$

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 fact that there are three possible events for the investor at the stopping time τ: he can receive a preference shock $(\tau = \tau_\gamma)$, in which case a new preference 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

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

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

dynamic programming equation

$$\Delta V_t(\delta) = \mathbb{E}_t \left[\int_t^\tau e^{-r(u-t)} \delta du + e^{-r(\tau-t)} \left(\Delta V_\tau(\delta) + \mathbf{1}_{\{\tau=\tau_\gamma\}} \int_0^1 (\Delta V_\tau(\delta') - \Delta V_\tau(\delta)) dF(\delta') \right. \right.$$

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

$$\left. \left. - \mathbf{1}_{\{\tau=\tau_1\}} \int_0^1 \theta_0 \left(\Delta V_\tau(\delta) - \Delta V_\tau(\delta') \right)^+ \frac{d\Phi_{1,\tau}(\delta')}{s} \right) \right]. \qquad (7)$$

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 second line, this 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 third line, this decreases her willingness to pay and hence her reservation value.

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

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

for all $\delta \in [0,1]$. The next proposition establishes the existence, uniqueness and some basic 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). Moreover, 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 space derivative. Given the reservation value function 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 distribution of future trading opportunities and the same distribution of future utility types. Thus, the only relevant difference between the two is their current utility types,

13

and this implies that the reservation value of an investor of type δ' is strictly larger than that of an investor of type $\delta < \delta'$. This monotonicity holds regardless of the distributions $\Phi_{q,t}(\delta)$ that investors take as given, and will greatly simplify the derivation of closed form solutions for both reservation values and the equilibrium distribution of asset holdings and types.

Integrating both sides of (7) against the distribution of the random time τ, and using the fact that reservation values are strictly increasing in δ to simplify the positive parts, we find 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 space derivative, 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* at type δ 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 the integral equation (9), changing the order of integration and differentiating both sides of the resulting equation with respect to t and δ shows that the local surplus satisfies the ordinary differential 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 $\mathbb{R}_+ \times [0, 1]$. A calculation provided in the appendix shows that, together with the requirements of boundedness and absolute continuity in time, this ordinary differential

equation uniquely pins down the local surplus as the present value of a perpetuity using an appropriate effective discount rate:

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

Finally, combining this explicit solution for the local surplus with equations (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 bounded solution to (7) is explicitly given by*

$$\Delta V_t(\delta) = \int_t^\infty e^{-r(u-t)} \bigg(\delta - \int_0^\delta \sigma_u(\delta')\left(\gamma F(\delta') + \lambda\theta_0\Phi_{1,u}(\delta')\right) d\delta' \tag{13}$$
$$+ \int_\delta^1 \sigma_u(\delta')\left(\gamma(1-F(\delta')) + \lambda\theta_1(1-s-\Phi_{0,u}(\delta'))\right) d\delta' \bigg) du,$$

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

We close this sub-section with several intuitive comparative static results regarding the impact of the economic environment on 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 (FOSD shift in $F(\delta')$), or if he expects to trade with higher-valuation counterparts (FOSD 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, the net effect is ambiguous and depends on all the parameters of the model.

Remark 1 Differentiating with respect to time on both sides of (9) shows that the reservation value function can be characterized as the unique bounded and absolutely continuous solution to the Hamilton-Jacobi-Bellman equation

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

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

This alternative characterization is not directly exploited in our derivation but it is nonetheless quite useful. In particular, in Section 5 we show that it implies a sequential representation of an investor's reservation values of the form

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

where $\hat{\delta}_t$ is a market-adjusted valuation process that accounts not only for the investor's changes of type but also for his trading opportunities. This representation allows us to analyze and interpret the price impact of search frictions.

3.2   The joint distribution of asset holdings and types

In this section, we provide a closed form characterization of the joint distribution of asset holdings and utility types that prevail in equilibrium. We then establish that this distribution converges strongly to the steady-state distribution from any initial conditions satisfying (1) and (2). Finally, we discuss several properties of the 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 of them is an asset owner with utility type δ' and the other is an non-owner with utility type $\delta' \geq \delta'$. Investors with the same type are indifferent between trading or not, but whether or not they trade is irrelevant for the distribution since the owner and the non-owner effectively exchange type. As a result, the rate of change in the measure of asset owners with utility type less than or

16

equal to δ satisfies

$$\dot{\Phi}_{1,t}(\delta) = \gamma\left(s - \Phi_{1,t}(\delta)\right)F(\delta) - \gamma\Phi_{1,t}(\delta)\left(1 - F(\delta)\right) - \lambda\Phi_{1,t}(\delta)\left(1 - s - \Phi_{0,t}(\delta)\right). \quad (15)$$

The first term in equation (15) 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 $F(\delta)$. 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 (15).[13]

Using condition (1), we can re-write equation (15) as a first order ordinary differential equation:

$$\dot{\Phi}_{1,t}(\delta) = -\lambda\Phi_{1,t}(\delta)^2 - \Phi_{1,t}(\delta)\left(\gamma + \lambda(1 - s - F(\delta))\right) + \gamma s F(\delta). \quad (16)$$

Importantly, this Ricatti equation holds for every $\delta \in [0,1]$ without imposing any regularity conditions on the distribution of utility types. Proposition 3 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

$$\Phi_1(\delta) = F(\delta) - \Phi_0(\delta) \equiv -\frac{1}{2}\left(1 - s + \gamma/\lambda - F(\delta)\right) + \frac{1}{2}\Lambda(\delta) \quad (17)$$

---

[13]Note that trading generates *no net inflow* into the set of owners with type less than δ. Indeed, such inflow requires that a non-owner with type $\delta' \leq \delta$ meets an owner with an even lower type $\delta'' < \delta'$. By trading, the previous owner of type δ'' leaves the set but the new owner of type δ' enters the same set, resulting in no net inflow.

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

$$\Phi_{1,t}(\delta) = \Phi_1(\delta) + \frac{(\Phi_{1,0}(\delta) - \Phi_1(\delta)) \Lambda(\delta)}{\Lambda(\delta) + (\Phi_{1,0}(\delta) - \Phi_1(\delta) + \Lambda(\delta)) (e^{\lambda \Lambda(\delta)t} - 1)} \tag{18}$$

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

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 agents change type on average once a year; the initial distribution of utility types among asset owners is given by $\Phi_{1,0}(\delta) = sF(\delta)\mathsf{l}$; 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 towards investors with higher valuations: the distribution of utility types among owners improves in the FOSD sense. 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, equation (17) offers several natural comparative statics that we summarize in Corollary 2 below. 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 a first order stochastic dominant 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

18

FIGURE 1: Equilibrium distributions

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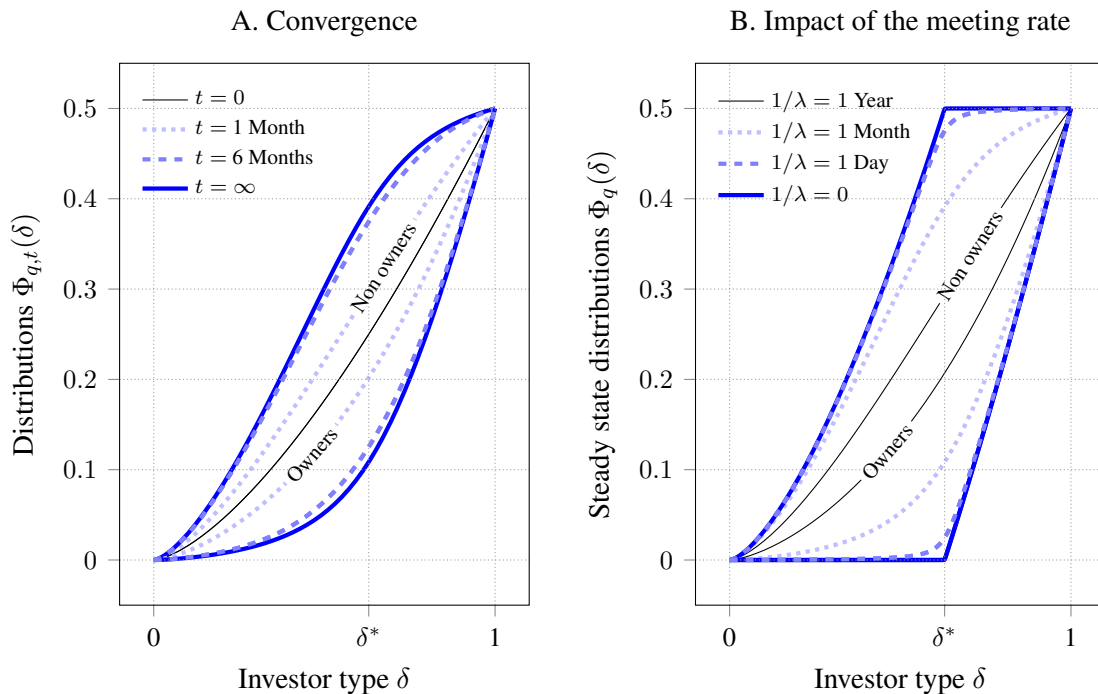
Notes. The left panel plots the cumulative distribution of types among non-owners (upper curves) and owners (lower curves) at different points in time. The right panel plots these distributions in the steady state, for different level of search frictions, indexed by the average inter-contact time, $1/\lambda$.

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.

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.*

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 (18), 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.

Theorem 1 *There exists a unique equilibrium. Moreover, given any initial conditions satisfying* (1) *and* (2)*, this equilibrium converges to the steady state given by*

$$r\Delta V(\delta) = \delta - \int_0^\delta \sigma(\delta')(\gamma F(\delta') + \lambda\theta_0\Phi_1(\delta'))d\delta' \qquad (19)$$
$$+ \int_\delta^1 \sigma(\delta')(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_0(\delta')))d\delta'$$

with the time-invariant local surplus

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

and the steady state distributions defined in (17)*.*

4 Trading patterns

In this section, we flesh out the basic predictions that emerge from our model. We structure the discussion by first focusing on the implications of our equilibrium characterization for the trading experience of an *individual* investor—in particular, the expected amount of time it will take him to trade, and the expected price at which this trade will take place. We then examine how these experiences at the individual level shape (and are shaped by) *aggregate* outcomes—in particular, the nature and extent of misallocation, the volume of trade, the structure of the network through which these trades are executed, and the distribution of bilateral prices.

After laying out the model's implications at both the individual and aggregate levels, we argue that a number of these implications find support in the empirical literature on OTC markets. Given the tractability of the model, the fact that it is also able to replicate several salient features of these markets suggests that it might be a useful structural framework for quantifying the size and impact

of trading frictions in decentralized markets. Finally, in addition to arguing that the model can account for many existing observations, we also highlight several novel predictions that have yet to be tested empirically.

4.1 Time to trade, prices, and an individual's utility type

In this section, we characterize the expected amount of time it takes to trade in the steady state and the expected price at which this trade will occur for an individual investor. Of particular importance, in Lemma 1 we establish that investors with the most to gain from trade tend to trade relatively quickly, and in Lemma 2 we establish that expected trading prices are increasing in the valuations of both the buyer and the seller.

While these results are fairly intuitive, their implications are far-reaching. For example, not only does Lemma 1 tell us that an asset owner with a low valuation will trade quickly, but it is also informative about his likely counterparty: since non-owners with high valuations also trade quickly, there will be few of them in equilibrium and, as a result, a non-owner with a low valuation is most likely to trade with an owner whose valuation is also relatively low. From Lemma 2, then, we anticipate that an owner with a low valuation will trade at a relatively low price, on average. More generally, the trading patterns described above imply that assets in a decentralized market are re-allocated to investors with higher valuations through chains of inframarginal trades. Absent preference shocks, each trade in this chain will be executed, on average, at slower and slower speeds but higher and higher prices. We establish these results formally below.

Time to trade. The steady state arrival rate of profitable trading opportunities for an owner with utility type δ is given by $\lambda_1(\delta) \equiv \lambda(1 - s - \Phi_0(\delta))$, which is the product of the arrival rate of a meeting and the probability that the investor meets a non-owner with utility type greater than or equal to his own type, δ. Since the latter probability is decreasing in δ, while the arrival rate of preference shocks is independent of δ, we naturally expect that the average time it takes an owner to trade is increasing in his utility type. Similarly, we have that the steady state arrival rate of profitable trading opportunities for a non-owner with utility type δ is given by $\lambda_0(\delta) \equiv \lambda\Phi_1(\delta)$;

again, since the probability $\Phi_1(\delta)$ is increasing in δ, we expect that the average time it takes a non-owner to trade is decreasing in his utility type.

To formalize this intuition, let $\eta_q(\delta)$ denote the expected amount of time that elapses before a trade occurs for an investor who currently has utility type $\delta \in [0, 1]$ and holds $q \in \{0, 1\}$ units of the asset. Our next result characterizes this expected waiting time in closed form for an arbitrary distribution of utility types, and offers several natural comparative statics.

Lemma 1 *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.

Figure 2 plots $\eta_0(\delta)$ and $\eta_1(\delta)$ for various values of the meeting intensity λ. Note that, as λ gets large, trading times fall (rise) sharply for non-owners (owners) with utility types very close to $\delta^\star$. Intuitively, as λ gets large, the allocation of the asset becomes closer to the frictionless allocation, *especially* at extreme values of δ (see Figure 1B). Hence, an owner with utility type $\delta \ll \delta^\star$ has essentially no willing counterparties for trade, and his expected waiting time remains high. An owner with utility type just above $\delta^\star$, on the other hand, has a much easier time finding a willing seller, and hence has a signficantly shorter waiting time before trade.[14]

Expected prices. Let $p_q(\delta)$ denote the price at which an investor with utility type $\delta \in [0, 1]$ holding $q \in \{0, 1\}$ units assets expects to trade, conditional on the arrival of a profitable trading opportunity. Because a non-owner with utility type $\delta \in [0, 1]$ only buys the asset from owners with utility types $\delta' \le \delta$, we have that the expected buying price is given by

$$p_0(\delta) = \theta_1 \Delta V(\delta) + \theta_0 \int_0^\delta \Delta V(\delta') \frac{d\Phi_1(\delta')}{\Phi_1(\delta)}. \tag{20}$$

---

[14]More formally, as Proposition 5 will confirm, the mass of potential counterparties for a non-owner with utility type $\delta < \delta^\star$ converges to zero for $\delta < \delta^\star$ when the meeting intensity is large. In contrast, the mass of potential counterparties for a non-owner with utility type $\delta > \delta^\star$, $\Phi_1(\delta) \simeq F(\delta) - (1 - s) > 0$, is uniformly bounded away from zero, which implies such a non-owner will trade quickly.

FIGURE 2: Expected time to trade

A. Non-owners
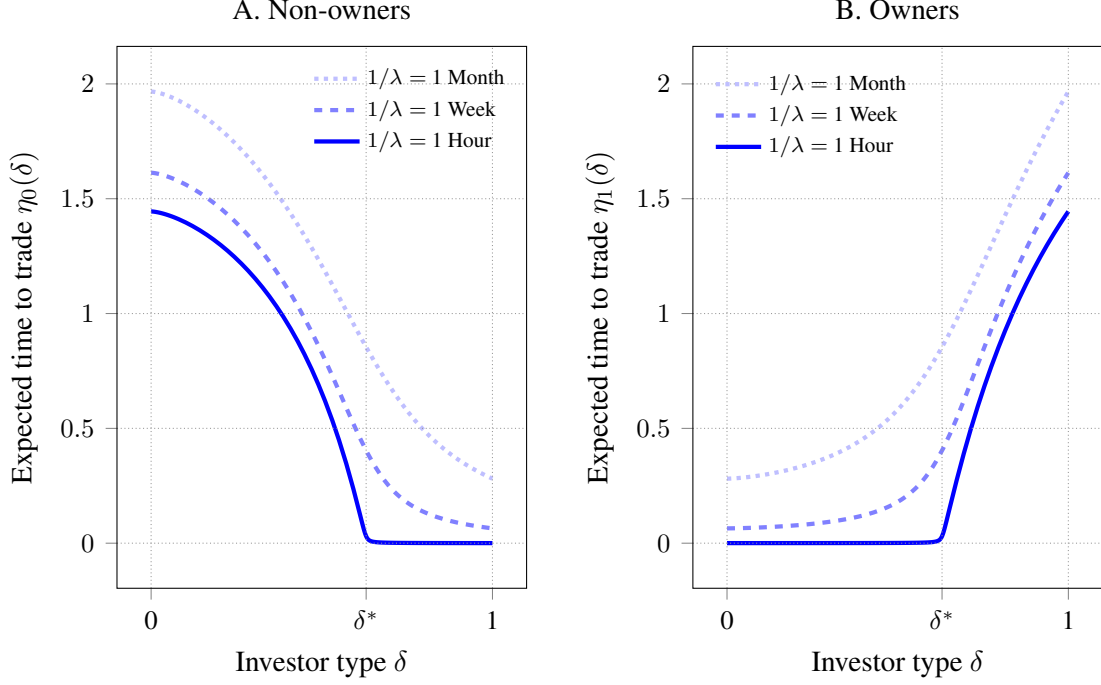
B. Owners

Notes. This figure plots the expected time to trade for non-owners (Panel A) and owners (Panel B) as functions of the investor's utility type when meetings happen, on average, every month (dotted), every week (dashed) and every hour (solid). The parameters we use in this figure are otherwise the same as in Figure 1.

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

$$p_1(\delta) = \theta_0 \Delta V(\delta) + \theta_1 \int_{\delta-}^{1} \Delta V(\delta') \frac{d\Phi_0(\delta')}{1 - s - \Phi_0(\delta-)}. \tag{21}$$

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 2 *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.*

The fact that expected trading prices are increasing in utility type is due to two reasons. First, investors with higher utility types have higher reservations values, so that buyers with high δ are

willing to pay more and sellers with high δ are unwilling to accept less. This effect is captured by the first terms in equations (20) and (21), and both of these terms are increasing in utility type by Proposition 1. Second, investors with higher utility types tend to trade with *other* investors who have high utility types, which further increases the expected transaction price. This effect is captured by the integrals in the expressions of the expected trading prices, and we establish in the Appendix that these integrals are also increasing in utility type.

The ambiguous effect of λ on the expected trading prices follows from the non-monotonic relationship between reservation values and the meeting rate that we discussed above and the fact that, as shown by Corollary 2, the mass of investors with whom a given buyer or seller can secure a trade is monotonically decreasing in the meeting rate.

4.2 Implications for aggregate outcomes

In this section, we examine how the patterns discussed above determine the nature and extent of misallocation in the market, the volume of trade, the structure of the network through which these trades occur, and the distribution of prices at which they are executed.

Misallocation. Since trade is decentralized and takes time, the asset cannot be allocated perfectly. Instead, there will be *misallocation*: some investors who would own the asset in a frictionless environment will not own it in the presence of search frictions, while some investors who would not own the asset in a frictionless environment will own it in the presence search frictions. However, even though search is random, the pattern of misallocation is not. In particular, we show that misallocation is most common among investors with utility types near the marginal type, $\delta^{\star}$.

To see this, note that misallocation has two symptoms: some assets are owned by the "wrong" investors (with type $\delta < \delta^{\star}$), and some of the "right" investors (with type $\delta > \delta^{\star}$) do not own an asset. In the steady state equilibrium, these two symptoms are captured by the following measure of cumulative misallocation:

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

To measure misallocation at a specific type, one can simply calculate the Radon-Nikodym density $\frac{dM}{dF}(\delta)$ of the measure $M(d\delta)$ with respect to the measure $F(d\delta)$ induced by the distribution of utility types (see equation (47) in the appendix for an explicit expression). Intuitively, the value of this density at a given point $\delta \in [0,1]$ represents the *fraction* of type$-\delta$ investors whose asset holdings in the environment with search frictions differs from their holdings in the frictionless benchmark.

Lemma 3 *The misallocation density $\frac{dM}{dF}(\delta)$ achieves a global maximum at either $\delta_-^\star$ or $\delta^\star$.*

While the slopes of the expected waiting times $\eta_0(\delta)$ and $\eta_1(\delta)$ are sufficient to understand why misallocation peaks at $\delta^\star$, the *shape* of these functions tell us even more about the patterns of misallocation.[15] In particular, we noted in the discussion of Figure 2 that the expected time to trade is very small for owners with valuations below $\delta^\star$, and increases rapidly for owners in a neighborhood of $\delta^\star$ (the opposite is true for non-owners). As Figure 3 illustrates, this implies that misallocation is concentrated in a cluster around the marginal type. Importantly, note that this pattern of misallocation arises because trade is decentralized, and would not arise in a model in which investors trade through centralized markets, or in a model where all trades are executed by dealers who have access to centralized markets. Indeed, in either of these alternative environments, all investors would trade with the same intensity, so that the measure of misallocation described above would be constant across types.

Trading volume. The discussion above highlights the fact that, in a purely decentralized market, assets are reallocated over time through chains of bilateral and infra-marginal trades, a phenomenon that has been pointed out before in both the search and the network literatures. Our closed form characterization allows us to derive an explicit expression for the trading volume, which reveals the contribution of these inframarginal trades to aggregate volume.

---

[15]To understand why our measure of misallocation can be discontinuous at $\delta^\star$, note that $\frac{dM}{dF}(\delta)$ is equal to $d\Phi_1(\delta)$ for $\delta \in [0, \delta^\star]$, and equal to $d\Phi_0(\delta)$ for $\delta \in [\delta^\star, 1]$; the former is an increasing function that achieves a maximum at $\delta^\star$, while the latter is a decreasing function that achieves a (generally different) maximum at $\delta^\star$. Hence, $\frac{dM}{dF}(\delta)$ will typically be discontinuous at $\delta^\star$, with a maximum either at $\delta_-^\star$ or $\delta^\star$.

FIGURE 3: Misallocation and search frictions
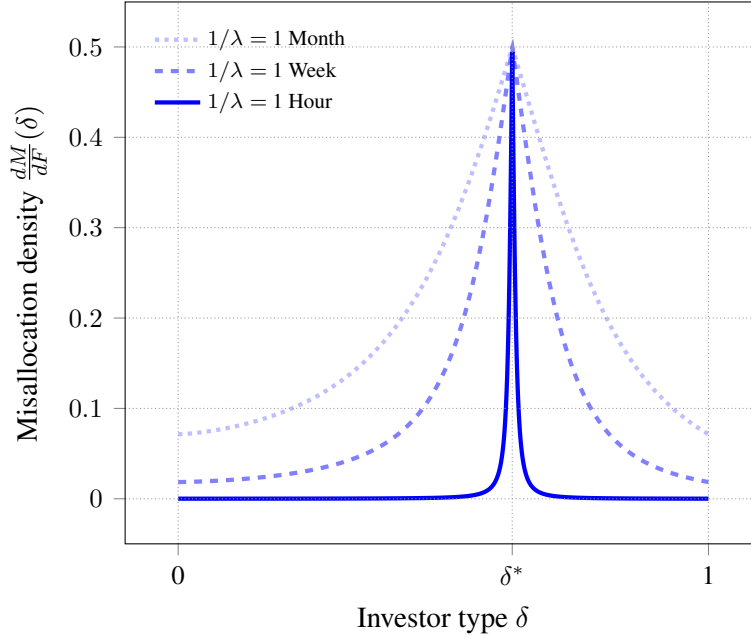
Notes. This figure plots the misallocation density as a function of the investor's utility type when meetings happen, on average, once every month (dotted), once every week (dashed), and once every hour (solid). The parameters we use in this figure are otherwise the same as in Figure 1.

To compute the trading volume generated in equilibrium, we need to specify a tie-breaking rule that determines the outcome of meetings between buyers and sellers with the same utility type. Assume for simplicity that, in the steady state equilibrium, such a meeting results in a trade with some constant probability $\pi \in [0, 1]$. Under this convention we can express 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 utility type exactly equal to δ. Using the integration by parts formula

for functions with jumps allows us to decompose the trading volume as

$$\vartheta(\pi) = \lambda\Phi_1(\delta^\star)(1 - s - \Phi_0(\delta^\star)) + \pi\lambda\Delta\Phi_0(\delta^\star)\Delta\Phi_1(\delta^\star) \tag{22}$$
$$+ \lambda \int_0^{\delta^\star} (\Phi_1(\delta_-) + \mathbf{1}_{\{\delta<\delta^\star\}}\pi\Delta\Phi_1(\delta))d\Phi_0(\delta)$$
$$+ \lambda \int_{\delta^\star}^1 (1 - s - \Phi_0(\delta) + \mathbf{1}_{\{\delta>\delta^\star\}}\pi\Delta\Phi_0(\delta))d\Phi_1(\delta).$$

The terms on the first line capture the volume that is generated by trades between owners with types in $[0, \delta^\star]$ and non-owners with types in $[\delta^\star, 1]$. These would be the only trades taking place in the equilibrium of a model with a frictionless exchange: the price would be set by the marginal type, so investors below $\delta^\star$ would only find it optimal to sell, while investors above $\delta^\star$ would only find it optimal to buy. With search frictions, however, there are additional trades because any investor may end up buying or selling, depending on who she meets. These inframarginal trades are captured by the integrals on the second and third lines of (22): the second line measures the trading volume generated by meetings in which a non-owner with type $\delta_0 \in [0, \delta^\star]$ buys the asset from an owner with types $\delta_1 \leq \delta_0$, while the third line measures the trading volume generated by meetings in which an owner with type $\delta_1 \in [\delta^\star, 1]$ sells to a non-owner with type $\delta_0 \geq \delta_1$.

Our next result provides an explicit expression for the equilibrium trading volume when the distribution of utility types is continuous.

Lemma 4 *If the distribution of utility types is continuous then the steady state trading volume is explicitly given by*

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

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 $\vartheta(0) < \vartheta_c < \vartheta(1)$.

The results in Lemma 4 have several interesting implications. First, note that ϑ_c is independent of $F(\delta)$: in a decentralized market with random matching, total trading volume depends on neither

the support nor the dispersion of private valuations in the market when $F(\delta)$ is continuous. This is because, in our environment, whether or not a trade occurs (and the quantity exchanged) depends only on the ordinal ranking of the two investors. The lemma also shows that, under the same continuity assumption, trading volume increases when investors find counterparties more quickly and when they change utility type more frequently, and is a dome shaped function of the asset supply with a maximum at the point $s = 1/2$. This last result is intuitive and well-known in this literature: trading volume peaks when the asset supply equates the number of potential buyers and sellers. Finally, note that when the distribution of utility types is continuous the equilibrium trading volume satisfies $\lim_{\lambda \to \infty} \vartheta_c = \infty$. By contrast, the equilibrium trading volume is finite in the frictionless benchmark (see Lemma A.1 in the appendix). Therefore, in our fully decentralized market, intermediation activity can generate arbitrarily large excess volume relative to the frictionless benchmark, as long as search frictions are sufficiently small.

Structure of the trading network.   Motivated by the emergence of several important OTC markets, a recent literature studies how the structure of the trading network in a decentralized market affects the efficiency with which assets and information are transferred, as well as the vulnerability of the market to negative shocks and contagion.[16]

In many of these studies, the trading network itself is taken as exogenous. In our model, even though meetings between investors are random, the topology of the trading network that emerges in equilibrium is not. In particular, as we noted above, misallocation is less prevalent at extreme types and tends to cluster around the marginal type, $\delta^\star$, as the meeting rate increases. Therefore, since investors outside of this cluster rarely have an occasion to trade, the share of trading volume accounted for by investors within this cluster grows. Hence, a *core-periphery* network structure emerges endogenously: 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."

---

[16]For models of information transmission in decentralized markets, see Duffie et al. (2009), Golosov et al. (2014), and Babus and Kondor (2012). For models of contagion and fragility see, among many others, Eisenberg and Noe (2001), Zawadowski (2013), Elliott, Golub, and Jackson (2014), Farboodi (2013), and Atkeson, Eisfeldt, and Weill (2014).

To illustrate this phenomenon, Figure 4 plots the contribution of each owner-nonowner pair to the equilibrium trading volume as measured by the Radon-Nykodym density

$$\kappa(\delta_0, \delta_1) \equiv \lambda \mathbf{1}_{\{\delta_1 \leq \delta_0\}} \frac{d\Phi_0}{dF}(\delta_0) \frac{d\Phi_1}{dF}(\delta_1).$$

As can be seen from the figure, agents with extreme utility types account for a very small fraction of total trades. For example, owners with low utility types may trade fast, but there are very few such owners in equilibrium, and hence we do not observe these owners trading very frequently. On the other hand, there are many owners with high utility types, but these investors trade very slowly, which is why 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. In fact, we can show that, when frictions are small, most of the volume is accounted for by the infra-marginal trades of investors in the cluster around the marginal type.

Lemma 5 *If the distribution of utility types is continuous then:*

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

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

Finally, we note that the trades executed by investors in the cluster around the marginal type resemble intermediation activity. Indeed, when the distribution of utility types is continuous, investors of type $\delta^\star$ have identical buying and selling intensities:

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

Hence, investors in the cluster around the marginal type are approximately equally likely to buy or sell the asset in equilibrium.

The distribution of realized prices. In contrast to models of frictionless exchange, or models of decentralized exchange with only two types of agents, our model with more than two types

FIGURE 4: Contribution to trading volume
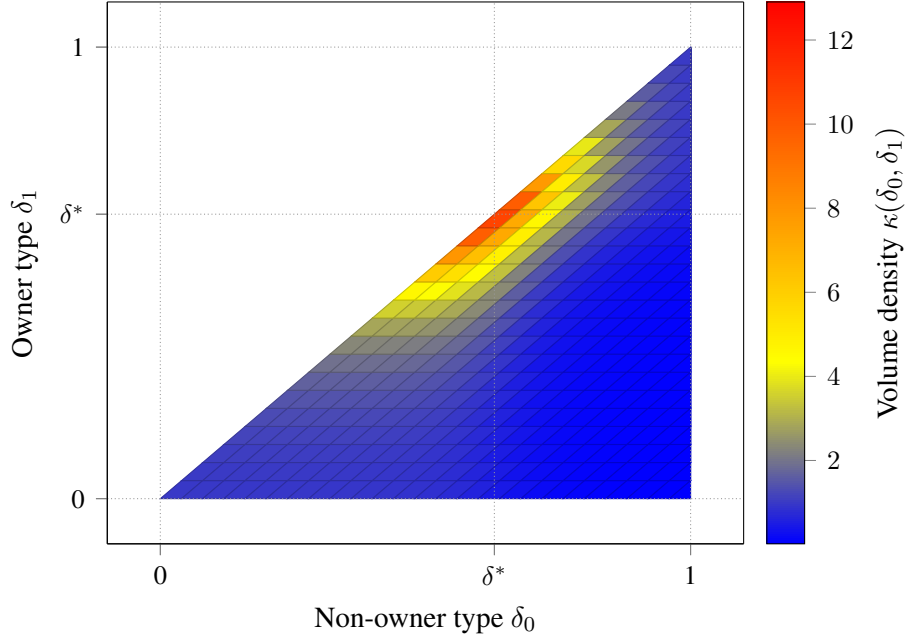
Notes. This figure plots the volume density $\kappa(\delta_0, \delta_1)$ as a function of the owner's and non-owner's type when meetings occur, on average, once a week. The parameters we use in this figure are otherwise the same as in Figure 1.

generates price dispersion. In particular, each transaction in the network described above occurs at a price

$$P(\delta, \delta') = \theta_0 \Delta V(\delta) + \theta_1 \Delta V(\delta')$$

in which the utility types $\delta \leq \delta'$ of the two counterparties are random variables on $[0, 1]$ drawn from the joint distribution described by

$$\mathbb{P}\left[\{\delta' \leq \delta_0\} \cap \{\delta \leq \delta_1\}\right] = \frac{G(\delta_0, \delta_1)}{\vartheta(1)},$$

where the denominator is the steady state equilibrium trading volume defined in equation (22) and the numerator is given by

$$G(\delta_0, \delta_1) = \int_{[0,1]^2} \mathbf{1}_{\{\delta' \leq \delta_0\}} \mathbf{1}_{\{\delta \leq \delta' \wedge \delta_1\}} \lambda d\Phi_0(\delta') d\Phi_1(\delta).$$

In words, the function $G(\delta_0, \delta_1)$ gives the probability that a non-owner with valuation less than δ_0 meets an owner with valuation less than δ_1 with whom a mutually beneficial trade can be agreed upon. When the distribution of types is continuous, the integrals above can be computed explicitly. However, even with this simplification, it remains quite difficult to characterize the properties of the distribution of prices analytically. Hence, we postpone studying the relationship between realized prices and trading speed to Section 5—where we study a region of the parameter space that admits analytical results—and focus here on the support of the distribution of realized prices.

Lemma 6 *The spread between the highest and lowest realized price is*

$$\Delta V(1) - \Delta V(0) = \int_{\underline{\delta}}^{\overline{\delta}} \frac{d\delta}{r + \gamma + \lambda\theta_1\left(1 - s - \Phi_0(\delta)\right) + \lambda\theta_0\Phi_1(\delta)}.$$

The spread is decreasing in both the meeting rate, λ, and the arrival rate of preference shocks, γ.

Intuitively, as the arrival rate of preference shocks γ becomes large, the difference between the reservation values of any two utility types shrinks, as both investors are not likely to remain in their current state for very long. An increase in the meeting rate λ has a similar effect: the spread between the reservation values of any two investors narrows when they have access to other counterparties with greater frequencies.

4.3   Assessing the implications of the model

Reliable data from OTC markets has traditionally been difficult to find, since quantities and prices tend to be negotiated privately in these markets. However, several recent studies have documented a few basic facts about the patterns of trade and prices in certain OTC markets. In this section, we argue that our model can easily and naturally account for a number of these facts, and we draw attention to a few predictions that have not been tested in these data. This suggests that our model can be a useful structural framework to study these types of markets, both theoretically and quantitatively.

Predictions and stylized facts. In our model, agents with a stronger need to trade find willing counterparties more quickly, but trade at less favorable terms. These two predictions are consistent with observations made by Ashcraft and Duffie (2007) in the federal funds market, where banks borrow and lend reserves in bilateral meetings throughout the day in attempt to achieve a certain end-of-day balance. They proxy a bank's need to trade by its excess balances—in the language of our model, large excess balances correspond to a "low δ." In accordance with the results in Lemma 1, they find that a bank's probability of making a loan at a particular moment in time is increasing in its balances, while the probability that a bank borrows funds at any point in time is decreasing in its balances. Moreover, consistent with the results in Lemma 2, they find that the interest rate on a loan (the inverse of the price) is decreasing in the balances of both the borrower and the lender.

Another important prediction of our model is that assets are re-allocated via chains of infra-marginal trades. This, too, is a feature of many OTC markets. For example, Li and Schürhoff (2012) document such intermediation chains in the inter-dealer market for municipal bonds: when an asset is purchased by a dealer from a customer, it is bought and sold by a sequence of dealers until it reaches its final holder. Viswanathan and Wang (2004) find similar patterns of sequential trades among dealers in the foreign exchange markets. In the federal funds market, Afonso and Lagos (2012a) find that 40 percent of trades are "intermediated" during the day, i.e., lent from one bank to another, only to be lent again to a third bank.

Lemma 5 establishes the process of asset reallocation gives rise to a trading network in which a small measure of agents who engage in intermediation activity account for a large proportion of the overall trading volume. This prediction resembles the observed tendency of OTC markets to exhibit a "core-periphery" trading structure. See Bech and Atalay (2010) for evidence for the federal funds market; Li and Schürhoff (2012) for the interdealer municipal bonds market; Soramäki et al. (2007) for interbank flows across Fedwire, the large value transfer system operated by the Federal Reserve; Craig and Von Peter (2014), Boss et al. (2004), Chang et al. (2008) for foreign interbank markets; Peltonen et al. (2014) for the Credit Default Swap (CDS) market.

Lastly, price dispersion is also a well-known feature of OTC markets. Afonso and Lagos (2012a) find significant dispersion in the fed funds market, especially during times of distress,

while Jankowitsch et al. (2011) use data from the Trade Reporting and Compliance Engine (TRACE) to document substantial price dispersion in the US corporate bond market. Measures of dispersion are typically even higher in less liquid markets; see, for example, Gavazza (2011b)'s analysis of transaction prices in the decentralized market for commercial aircraft.

Additional testable predictions. Given the tractability of our analysis, and hence the availability of a variety of comparative statics, our model also generates a number of predictions that have not yet been tested in the data. For example, the results in Lemmas 1 and 2 imply that investors who purchase an asset at a relatively high price are more likely to have a relatively high valuation, and thus they are more likely to own the asset for a longer period of time than an investor who paid a relatively low price. By the same reasoning, an investor who sold an asset at a high price is likely to buy a new asset more quickly than an investor who sold an asset at a low price. To the best of our knowledge, this prediction has not been studied empirically in the context of an OTC asset market, but is highly relevant in some contexts. For example, in the housing market, the realized payoff from a mortgage depends on the tenure of the homeowner. Hence, even though mortgage lenders are typically wary when a buyer pays a relatively high price for a house (because of resale considerations), our model provides a reason why such a mortgage could be a relatively profitable loan (because of a longer expected tenure in the house).

5 The price impact of search frictions

In this section we ask how search frictions affect realized prices, paying particular attention to understanding how the answer depends on the degree of heterogeneity among market participants. To obtain analytic comparative statics, we focus on the case of small frictions, $\lambda \simeq \infty$. We obtain two main findings. First, heterogeneity magnifies the price impact of search frictions. Second, search frictions have larger effects on price levels than on their dispersion. Therefore, quantifying the level of frictions based on observed price dispersion may be misleading, since frictions can have substantial effects on price levels even when dispersion appears very small.

5.1 A sequential representation of reservation values

To derive and interpret the asymptotic behavior of realized prices as frictions vanish, we rely on a sequential representation of the investors' reservation values that we obtain by exploiting the Hamilton-Jacobi-Belman equation (14). Related representations have also been derived by Hugonnier (2012) and Kiefer (2012) in alternative models of decentralized exchanges.

Proposition 4 *The reservation value function can be represented as*

$$\Delta V_t(\delta) = \mathbb{E}_{t,\delta} \left[\int_t^\infty e^{-r(s-t)} \hat{\delta}_s ds \right] \tag{23}$$

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

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

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

We refer to the process $\hat{\delta}_t$ as the market valuation process because it not only takes into account investors' physical changes of types, but also their future trading opportunities. To describe this process, suppose that an investor's valuation is $\hat{\delta}_t = \delta \in [0,1]$ at time $t \geq 0$. Then, during the time interval $[t, t+dt]$, the market valuation process can change for three reasons. First, there is a preference shock that arrives with intensity γ, in which case a new valuation is drawn according to $F(\delta')$. Second, there is a purchase opportunity which arrives with intensity $\lambda \theta_0 \Phi_{1,t}(\delta)$, in which case a new valuation is drawn from the truncated support $[0, \delta]$ according to the conditional distribution $\Phi_{1,t}(\delta')/\Phi_{1,t}(\delta)$. Note that the market valuation process creates transitions towards lower types to account for the option value of searching for sellers and buying at a lower price. Symmetrically, the market valuation process creates transitions towards higher types to account for the option value of searching for buyers and selling at a higher price: the third reason the market valuation process can change is that sale opportunities arise with intensity $\lambda \theta_1 (1 - s - \Phi_{0,t}(\delta))$, in which case a new valuation is drawn from the truncated support $(\delta, 1]$ according to the conditional distribution $(\Phi_{0,t}(\delta') - \Phi_{0,t}(\delta))/(1 - s - \Phi_{0,t}(\delta))$.

In the frictionless benchmark, the market valuation process is constant and given by $\hat{\delta}_t = \delta^\star$: because investors can trade instantly at the constant price $\delta^\star/r$, the market valuation process must equal the valuation of the marginal investor at all times and for all investors. In a decentralized market, the market valuation process 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, conditional on finding a trading partners, investors do not always trade with the marginal type. Instead, the terms of trade are random and depend on the equilibrium distributions of utility types among potential trading partners. Note that this second channel is only active if there are more than two types, because otherwise a single price gets realized in bilateral meetings.

To relate the explicit solution of Proposition 2 to the sequential representation of Proposition 4, we show in the Appendix that one can rewrite (23) as

$$\Delta V_t(\delta) = \int_0^1 (\delta'/r) d\Psi_t(\delta'|\delta) \tag{24}$$

with the integrator defined by

$$\Psi_t(\delta'|\delta) \equiv \mathbb{E}_{t,\delta}\left[\int_t^\infty re^{-r(u-t)}\mathbf{1}_{\{\hat{\delta}_u \leq \delta'\}}du\right].$$

Comparing this expression to equation (13), we conclude that

$$\Psi_t(\delta'|\delta) = \int_t^\infty re^{-r(u-t)}\sigma_u(\delta')\left(\gamma F(\delta') + \lambda\theta_0\Phi_{1,u}(\delta')\right)du$$

if $\delta' \leq \delta$, and

$$\Psi_t(\delta'|\delta) = \int_t^\infty re^{-r(u-t)}\left(1 - \sigma_u(\delta')\left(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_{0,u}(\delta'))\right)\right)du$$

otherwise. The interpretation of this nondecreasing function is clear: $\Psi_t(\delta'|\delta) \in [0,1]$ measures the discounted expected amount of time that the market valuation process spends in the interval $[0, \delta']$ given that it started from $\delta \in [0,1]$ at time $t \geq 0$. In accordance with this interpretation we will refer to this function as the discounted occupation measure of the market valuation process.

5.2 The frictionless limit

As a first step in our study of the asymptotic price impact of search frictions, we establish two intuitive, but important, results about the economy as $\lambda \to \infty$. First, we show that the allocation becomes approximately frictionless, as an asset owned by an investor with utility type $\delta < \delta^\star$ will be reallocated very quickly upstream to an investor with utility type $\delta > \delta^\star$. Second, we show that, as a result of this increasingly efficient allocation process, the reservation value of all investors converges to the frictionless equilibrium price, $\delta^\star/r$.

Proposition 5 *As $\lambda \to \infty$, the equilibrium allocation and the reservation value function converge to their frictionless counterparts in that $\lim_{\lambda \to \infty} \Phi_0(\delta) = \Phi_0^\star(\delta)$, $\lim_{\lambda \to \infty} \Phi_1(\delta) = \Phi_1^\star(\delta)$, 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 price, consider the market valuation process. By Proposition 5 we know that the equilibrium allocation becomes approximately efficient as $\lambda \to \infty$, which implies that it becomes very easy for an investor with utility type $\delta < \delta^\star$ to sell his asset, but a lot more difficult for this investor to buy. Accordingly, we show in Appendix A.4 that the trading intensities of such an investor satisfy $\lim_{\lambda \to \infty} \lambda_1(\delta) = \infty$ and

$$\lim_{\lambda \to \infty} \lambda_0(\delta) = \lim_{\lambda \to \infty} \lambda \Phi_1(\delta) = \frac{\gamma s F(\delta)}{1 - s - F(\delta)}.$$

Thus, it follows from Proposition 4 that, starting from below the marginal type, the market valuation process moves up very quickly as the meeting frequency increases. Similarly, for $\delta > \delta^\star$, the trading intensities satisfy $\lim_{\lambda \to \infty} \lambda_0(\delta) = \infty$ and

$$\lim_{\lambda \to \infty} \lambda_1(\delta) = \lim_{\lambda \to \infty} \lambda(1 - s - \Phi_0(\delta)) = \frac{\gamma(1 - s)(1 - F(\delta))}{F(\delta) - (1 - s)}.$$

This shows that, as $\lambda \to \infty$, it becomes very easy for an investor with utility type $\delta > \delta^\star$ to buy an asset but a lot more difficult to sell; hence, beginning above the marginal type, the market valuation process moves 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

36

now follows from the sequential representation (23) that all reservation values converge to the frictionless equilibrium price $p^\star = \delta^\star/r$.

5.3 Price levels near the frictionless limit

By the same logic, in order 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. Doing so yields the following result.

Proposition 6 *Assume that the distribution of utility types is twice continuously differentiable with a strictly positive derivative. Then, for all $\delta \in [0,1]$,*

$$\Delta V(\delta) = p^\star + \frac{\pi/r}{F'(\delta^\star)}\left(\frac{1}{2} - \theta_0\right)\left(\frac{\gamma s(1-s)}{\theta_0\theta_1}\right)^{\frac{1}{2}}\frac{1}{\sqrt{\lambda}} + o\left(\frac{1}{\sqrt{\lambda}}\right), \tag{25}$$

where the remainder satisfies $\lim_{\lambda \to \infty} \sqrt{\lambda} o(1/\sqrt{\lambda}) = 0$.

The idea of the proof consists in centering the market valuation process around its frictionless limit $\delta^\star$ and scaling the result by its convergence rate, which turns out to be $\sqrt{\lambda}$. This delivers an auxiliary process $\hat{x}_t \equiv \sqrt{\lambda}(\hat{\delta}_t - \delta^\star)$ whose nondegenerate limit distribution determines the correction term in the expansion of the reservation values, because

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

for all $\delta \in [0,1]$ as a result of Proposition 4. To understand why the convergence rate of the market valuation process is equal to $\sqrt{\lambda}$, recall that the asset is almost perfectly allocated as $\lambda \to \infty$, so that buyers become very scarce below the marginal type while sellers become very scarce above the marginal type. Consistent with this observation, we show in Appendix A.4 that, when the distribution of utility types is continuous, the trading intensities

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

of investors in a neighborhood of $\delta^\star$ converge to infinity at rate $\sqrt{\lambda}$. This rate is much slower than the rate λ that prevails away from the marginal type, and pins down the convergence rate of the market valuation process in a neighborhood of the marginal type.

The first term in the expansion follows directly from Proposition 5, since all reservation values converge to the frictionless equilibrium price $p^\star = \delta^\star/r$. The second term is a correction term that determines how reservation values deviate from the frictionless equilibrium price, and we argue that this term depends on three key features of our decentralized market model. The first key feature is the average time that it takes investors with near-marginal valuations to find counterparties, which we have just argued is in order $1/\sqrt{\lambda}$. The second key feature of the market 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 equilibrium price, irrespective of the other features of the market.

The third feature of the market that matters for reservation values is the amount of heterogeneity that exists in a neighborhood of the marginal type. If $F'(\delta^\star)$ 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 $F'(\delta^\star)$ 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

$$F'(\delta^\star) \propto \frac{p}{F(rp) - 1} \frac{d(1 - F(rp))}{dp} \bigg|_{p=p^\star}$$

is proportional to the elasticity of the Walrasian demand at the frictionless equilibrium price. It is very intuitive that an elastic demand speeds up the convergence by mitigating the non competitive forces at play in the search model.

To further emphasize the role of heterogeneity, consider heuristically what happens when the continuous distribution of utility types is taken to approximate a discrete distribution. In such a case the cumulative distribution function $F(\delta)$ will approach a step function that is vertical at the

38

FIGURE 5: Continuous vs. discrete distribution

A. Price deviation
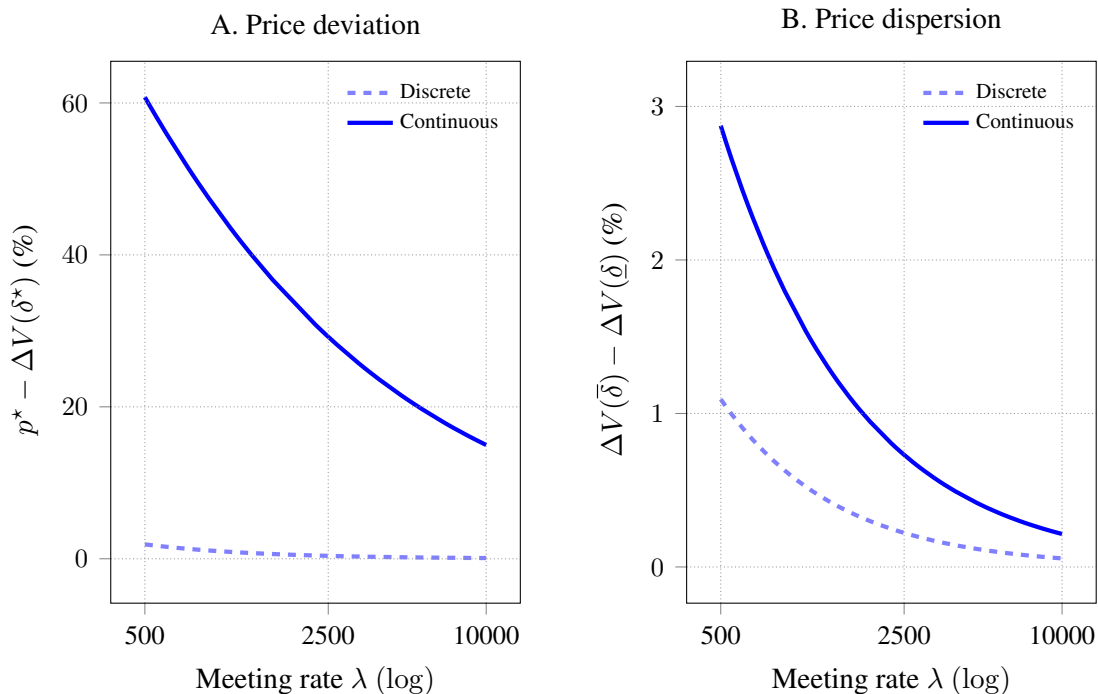
B. Price dispersion

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 agent as the continuous distribution of the base case model.

marginal type. As a result, we will have that the derivative $F'(\delta^\star)$ will approach infinity and it follows from (25) 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 7 *When the distribution of utility types is discrete the convergence rate of reservation values is generically equal to $1/\lambda$.*

Taken together, Propositions 6 and 7 show that existing results based on search models with discrete distributions of types can substantially underestimate the price impact of frictions (see, e.g., Vayanos and Weill (2008) Weill (2008), and Praz (2013) for asymptotic expansions in such models). We illustrate this finding in the left panel of Figure 5 where we compare the price

deviation $p^\star - \Delta V(\delta^\star)$ implied by the continuous distribution of our baseline model to the price deviation implied by a discrete distribution that delivers identical marginal and average investors. The figure shows that when investors find counterparties twice a day on average (i.e., $\lambda = 500$), the price discount is 60% with a continuous distribution, and only about 2% with the corresponding discrete distribution. When investors find counterparties twenty times per day on average (i.e., $\lambda = 10000$), the discount is approximately 15% with a continuous distribution, but it is now indistinguishable from zero with a discrete distribution.

5.4 Price dispersion near the frictionless limit.

An important implication of Proposition 6 is that, to a first-order approximation, there is no price dispersion. This can be seen by noting that the correction term in equation (25) does not depend on the investor's utility type, and implies that 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 8 *Assume that the distribution of utility types is twice continuously differentiable with a strictly positive derivative. Then*

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

where the remainder satisfies $\lim_{\lambda\to\infty} |\lambda O(1/\lambda)| < \infty$. By contrast, with a discrete distribution of utility types, the convergence rate is generically equal to $1/\lambda$.

Comparing the results of Propositions 6 and 8 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 the price dispersion and, yet, have large impact on the price level.

40

This finding is illustrated in Figure 5. 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 (i.e, $\lambda = 500$), the price discount implied by our baseline model is about 60%, but the corresponding price dispersion is twenty times smaller (about 3%). One can also see from the figure that, in accordance with the result of Proposition 8, price dispersion is larger with a continuous distribution of utility types than with a discrete distribution.[17]

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. Using new techniques, we show that there is no loss in tractability from this generalization, while the benefits are substantial. For one, the model is able to account for many of the key empirical facts recently reported in studies of OTC markets, which suggests it could provide a unified framework to study a number of important issues, whether they be theoretical or quantitative. Moreover, the model generates a number of new results which underscore the importance of heterogeneity in decentralized markets.

---

[17]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.

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, 2012a.

Gara Afonso and Ricardo Lagos. Trade dynamics in the market for federal funds. *FRB of New York Staff Report*, (549), 2012b.

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 G. Atkeson, Andrea L. Eisfeldt, and Pierre-Olivieri Weill. The market for otc credit derivatives. Working paper, UCLA, 2012.

Andrew G. Atkeson, Andrea L. Eisfeldt, and Pierre-Olivier Weill. Entry and exit in otc derivatives markets. Working paper UCLA, 2014.

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 4. 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*, pages 257–273, 1998.

Eui Jung Chang, Eduardo José Araujo Lima, Solange M Guerra, and Benjamin M 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 inter-mediation chains. Working paper, HEC and PSE, 2014.

Ben Craig and Goetz Von Peter. Interbank tiering and money center banks. *Journal of Financial Intermediation*, 2014.

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.

Darrell Duffie, Semyon Malamud, and Gustavo Manso. Information percolation with equilibrium search dynamics. *Econometrica*, 77(1513-1574), 2009.

Larry Eisenberg and Thomas H. Noe. Systemic risk in financial systems. *Management Science*, February:236–249, 2001.

Matthew Elliott, Benjamin Golub, and Matthew O. Jackson. Financial network and contagion. *American Economic Review*, Forthcoming, 2014.

Maryam Farboodi. Intermediation and voluntary exposure to counterparty risk. Working paper, University of Chicago, 2013.

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, 2011a.

Alessandro Gavazza. The role of trading frictions in real asset markets. *The American Economic Review*, 101(4):1106–1143, 2011b.

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.

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

Mikhail Golosov, Guido Lorenzoni, and Aleh Tsyvinski. Decentralized trading with private information. *Econometrica*, 82:1055–1091, 2014.

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

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

Matthew O Jackson. *Social and economic networks*. Princeton University Press, 2010.

Rainer Jankowitsch, Amrut Nashikkar, and Marti G Subrahmanyam. Price dispersion in otc markets: A new measure of liquidity. *Journal of Banking & Finance*, 35(2):343–357, 2011.

Patrick Kiefer. A search-based pricing model of credit default swaps. Working paper, UCLA, 2012.

Nobuhiro Kiyotaki and Randall Wright. A search-theoretic approach to monetary economics. *American Economic Review*, 83:63–63, 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. *Federal Reserve Bank of Philadelphia, working paper 14-9*, 2014.

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.

Jianjun Miao. A search model of centralized and decentralized trade. *Review of Economic dynamics*, 9(1):68–92, 2006.

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.

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 T. Reid. *Riccati Differential Equations*. Mathematics in science and engineering. Academic Press, 1972.

Ji Shen and Hongjun Yan. A search model of the aggregate demand for safe and liquid assets. *Available at SSRN*, 2014.

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

Daniel F. Spulber. Market making by price-setting firms. *The Review of Economic Studies*, 63(4): pp. 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.

Dimitri Vayanos and Tan Wang. Search and endogenous concentration of liquidity in asset markets. *Journal of Economic Theory*, 136(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.

S Viswanathan and James JD Wang. Inter-dealer trading in financial markets*. *The Journal of Business*, 77(4):987–1040, 2004.

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

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

Adam Zawadowski. Entangled financial systems. Working paper, Boston University, 2013.

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 *minimum* 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* $\vartheta^\star = \gamma \max\{sF(\delta^\star_-), (1-s)(1-F(\delta^\star))\}$.

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 throughout this appendix let

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

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

$$
T_t[f](\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)}(\delta + (\gamma+\lambda)f_u(\delta) + \mathcal{O}_u[f](\delta))du. \tag{26}
$$

Assume that $\Delta V_t(\delta) = T_t[\Delta V](\delta)$ is a fixed point that satisfies (8). Since the right hand side of (26) 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

$$\Delta V_t(\delta) = e^{-r(H-t)}\Delta V_H(\delta) + \int_t^H e^{-r(u-t)}(\delta + \mathcal{O}_u[\Delta V](\delta))du \tag{27}$$

$$= \lim_{H \to \infty} \int_t^H e^{-r(u-t)}(\delta + \mathcal{O}_u[\Delta V](\delta))du \tag{28}$$

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 (26) is absolutely continuous in time this assumption implies that

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

By definition we have that

$$\Delta V_u(\delta) \geq \Delta V_u(\delta'), \qquad t \leq u \leq H^* \tag{29}$$

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

$$\mathcal{O}_u[\Delta V](\delta) \leq \mathcal{O}_u[\Delta V](\delta'), \qquad t \leq u \leq H^*. \tag{30}$$

To proceed further we distinguish two cases depending on whether the constant H^* is finite or not. Assume first that it is finite. In this case it follows from (27) that we have

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

and combining this identity with (30) then gives

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

$$= \int_t^{H^*} e^{-r(u-t)}(\delta + \mathcal{O}_u[\Delta V](\delta'))du + e^{-r(H^*-t)}\Delta V_{H^*}(\delta') < \Delta V_t(\delta') \qquad (31)$$

where the equality follows by continuity, and the second inequality follows from the fact that $\delta < \delta'$. Now assume that $H^* = \infty$ so that (29) and (30) hold for all $u \geq t$. In this case (28) 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 (31) delivers the required contradiction and establishes that $\Delta V_t(\delta)$ is non decreasing. To see that it is strictly increasing rewrite (26) as

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

with the operator

$$\mathcal{M}_u[f](\delta) = \lambda \eta f_u(\delta) + \gamma \int_0^1 f_u(\delta')dF(\delta') + \lambda\theta_0 \int_0^1 \min\left\{f_u(\delta'), f_u(\delta)\right\} d\Phi_{1,t}(\delta') \qquad (32)$$

$$+ \lambda\theta_1 \int_0^1 \max\left\{f_u(\delta'), f_u(\delta)\right\} d\Phi_{0,t}(\delta'),$$

and the constant $\eta = 1 - s\theta_0 - (1-s)\theta_1 > 0$. Because $\mathcal{M}_u[f](\delta)$ is increasing in $f_u(\delta)$ and the given solution is nondecreasing in space we have that

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

for any $0 \leq \delta \leq \delta' \leq 1$, and the required strict monotonicity follows. To conclude the proof it now only 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 (28) 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 increase 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. 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 uniform bound with (32) 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 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 $\mathcal{X}_k \subseteq \mathcal{X}$ denote the set of uniformly bounded functions $f : \mathcal{S} \to \mathbb{R}$ that are nonnegative and non decreasing in space with

$$0 \leq f_t(\delta') - f_t(\delta) \leq \frac{1}{r + \gamma}(\delta' - \delta) \equiv k(\delta' - \delta) \tag{33}$$

for all $0 \leq \delta \leq \delta' \leq 1$ and $t \geq 0$. Let further $\mathcal{X}_k^*$ 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. Because $\mathcal{X}_k$ is closed in $\mathcal{X}$, it suffices to prove that T maps $\mathcal{X}_k$ into $\mathcal{X}_k^*$. Fix and arbitrary function $f \in \mathcal{X}_k$. Since this function is nonnegative it follows from (32) that the function $T_t[f](\delta)$ is nonnegative. On the other hand, using the inequalities in (33) 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 (32) 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 $\mathcal{X}_k$ 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 (26) and

changing the order of integration shows that

$$T_t[f](\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)}\bigg(\delta + (\lambda+\gamma)f_u(\delta) - \int_0^\delta \phi_u(\delta')(\gamma F(\delta') + \lambda\theta_0\Phi_{1,u}(\delta'))d\delta' \quad (34)$$

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

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

$$V_{q,t}(\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)}(\lambda V_{q,u}(\delta) + \mathcal{C}_{q,u}(\delta) + \gamma \int_0^1 V_{q,u}(\delta')dF(\delta'))du. \quad (35)$$

with the uniformly bounded functions defined by

$$\mathcal{C}_{q,t}(\delta) = q\delta + \int_0^1 \lambda\theta_q\left((2q-1)(\Delta V_t(\delta') - \Delta V_t(\delta))\right)^+ d\Phi_{1-q,t}(\delta'). \quad (36)$$

Because the right hand side of (35) 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

$$\dot{V}_{q,t}(\delta) = rV_{q,t}(\delta) - \mathcal{C}_{q,t}(\delta) - \gamma \int_0^1 (V_{q,t}(\delta') - V_{q,t}(\delta))dF(\delta') \quad (37)$$

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.4 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 uniformly bounded. Combining these properties with (37) 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

$$V_{q,t}(\delta) = \mathbb{E}_{t,\delta}\left[\int_t^\infty e^{-r(u-t)} \mathcal{C}_{q,u}(\delta_u) du\right].\tag{38}$$

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 (38) 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

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

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

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

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

Lemma A.4 *The functions $\mathcal{C}_{q,t}(\delta)$ are increasing in $\delta \in [0,1]$.*

Proof. For $q = 0$ the result follows immediately from (36) and the fact that the reservation value function is increasing in $\delta \in [0,1]$. Assume that now that $q = 1$. Using the fact that the reservation value function is increasing and integrating by parts on the right hand side of (36) gives

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

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

Lemma A.5 *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*

$$\sigma_t(\delta) = \int_t^\infty e^{-\int_t^u R_\xi(\delta)d\xi} du.\tag{39}$$

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 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. ∎

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

$$\dot{f}_t(\delta) = rf_t(\delta) - \delta + \int_0^\delta \sigma_t(\delta')(\gamma F(\delta') + \lambda\theta_0\Phi_{1,t}(\delta'))d\delta' \tag{40}$$
$$- \int_\delta^1 \sigma_t(\delta')(\gamma(1 - F(\delta')) + \lambda\theta_1(1 - s - \Phi_{0,t}(\delta')))d\delta'$$

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

$$f_t(\delta') - f_t(\delta) = \int_\delta^{\delta'} \sigma_t(\delta'')d\delta'', \qquad (\delta, \delta') \in [0, 1]^2 \tag{41}$$

and it follows that $f_t(\delta)$ is strictly increasing in space. Using this monotonicity in conjunction with (41) and integrating by parts on the right hand side of (40) shows that

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

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

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

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 (26) 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 (32) and, by inspection, this mapping is increasing in $f_t(\delta)$ and decreasing in r. Furthermore, it follows from equation (34) 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. 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 Ricatti 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 denumerator of (18) 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*

$$\lim_{\phi \to 0} \Phi_1(\delta) = sF(\delta)$$

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

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

$$\frac{\partial \Phi_1(\delta)}{\partial s} = \frac{\Phi_1(\delta) + \phi F(\delta)}{\Lambda(\delta)} \tag{42}$$

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

$$\frac{\partial \Phi_1(\delta)}{\partial \phi} = \frac{sF(\delta) - \Phi_1(\delta)}{\Lambda(\delta)} = \frac{s(1-s)F(\delta)(1-F(\delta))}{(\phi + \Phi_1(\delta) + (1-s)(1-F(\delta)))\Lambda(\delta)} \tag{43}$$

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

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

is also increasing in ϕ and it now follows from the first equality in (43) that

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

The expressions for the limiting values follow by sending ϕ to zero and ∞ 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 3. 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*

$$\ell(x) \equiv -\frac{1}{2}(1 - s + \phi - x) + \frac{1}{2}\sqrt{(1 - s + \phi - x)^2 + 4s\phi x}, \tag{44}$$

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 a direct calculation:

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

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

Proof of Lemma 1. Consider an agent of ownership type q and denote his utility type process by δ_t. The next time that 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

$$
\begin{aligned}
\eta_q(\delta) &= \mathbb{E}\left[\int_0^\sigma e^{-\int_0^t \lambda_q(\delta_s)ds}dt + e^{-\int_0^\sigma \lambda_q(\delta_s)ds}\eta_q(\delta_\sigma)\right] \qquad (45)\\
&= \mathbb{E}\left[\int_0^\sigma e^{-\lambda_q(\delta)t}dt + e^{-\lambda_q(\delta)\sigma}\eta_q(\delta_\sigma)\right] = \frac{1}{\gamma + \lambda_q(\delta)}\left(1 + \gamma\int_0^1 \eta_q(\delta')dF(\delta')\right)
\end{aligned}
$$

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 (45) 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 (45) gives

$$\eta_q(\delta) = \frac{1}{\gamma + \lambda_q(\delta)} \left(1 - \gamma \int_0^1 \frac{dF(\delta')}{\gamma + \lambda_q(\delta')} \right)^{-1}. \tag{46}$$

Now assume that the cumulative distribution function $F(\delta)$ is continuous. Combining Proposition 3 with Lemma A.8 and the change of variable formula for Stieljes 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 (44) 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 (46) 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 (46) 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 (43). This shows that $\lambda_q(\delta)$ is an increasing function of λ and the desired result now follows from (46) 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 (42) 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 (46) by noting that the function $F(\delta)$ does not depend on s. ∎

Proof of Lemma 2. Because the reservation value function

$$\Delta V(\delta) = \Delta V(0) + \int_0^\delta \sigma(\delta')d\delta'$$

is absolutely continuous in $\delta \in [0,1]$ it follows from an integration by parts that the expected buyer price can be written as

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

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

and the required monotonicity follows by observing that the local surplus $\sigma(\delta')$ 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. ∎

Proof of Lemma 3. Let **A** denote the set of atoms of the distribution $F(\delta)$. With this definition we have

that the Radon-Nykodim density is given by

$$m(\delta) = \frac{dM}{dF}(\delta) \equiv \begin{cases} \mathbf{1}_{\{\delta \notin \mathbf{A}\}}\ell'(F(\delta)) + \mathbf{1}_{\{\delta \in \mathbf{A}\}}\frac{\Delta\ell(F(\delta))}{\Delta F(\delta)}, & \text{if } \delta < \delta^\star, \\ \mathbf{1}_{\{\delta \notin \mathbf{A}\}}(1 - \ell'(F(\delta))) + \mathbf{1}_{\{\delta \in \mathbf{A}\}}\left(1 - \frac{\Delta\ell(F(\delta))}{\Delta F(\delta)}\right), & \text{otherwise.} \end{cases} \quad (47)$$

To establish the result we need to show that $m(\delta)$ is increasing on $[0, \delta^\star)$ and decreasing on $[\delta^*, 1]$. 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 (47). ∎

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, that 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 towards $X(q)$. 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 Lemma 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 $\ell(x)$ shown in Lemma A.8 one obtains that the functions $G_1(x)$ and $G_0(x)$ satisfy the identity:

$$G_1(x) = \frac{\phi G_0(x)}{1 + \phi - G_0(x)} \quad (48)$$

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 (48) can be written equivalently as:

$$G_1(G_0^{-1}(y)) = \frac{\phi y}{1 + \phi - y}. \quad (49)$$

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 from Lemma (A.9) that the trading volume can 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 (49) 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. ∎

Proof of Lemma 5. When $F(\delta)$ is continuous, there are no atom in the distribution of types and hence we obtain from (22) that:

$$
\vartheta_c = \lambda \Phi_1(\delta^\star)(1 - s - \Phi_0(\delta^\star)) + \lambda \int_0^{\delta^\star - \varepsilon} \Phi_1(\delta) d\Phi_0(\delta) + \lambda \int_{\delta^\star + \varepsilon}^1 (1 - s - \Phi_0(\delta)) d\Phi_1(\delta) \tag{50}
$$
$$
+ \lambda \int_{\delta^\star - \varepsilon}^{\delta^\star} \Phi_1(\delta) d\Phi_0(\delta) + \lambda \int_{\delta^\star}^{\delta^\star + \varepsilon} (1 - s - \Phi_0(\delta)) \, d\Phi_1(\delta).
$$

We show that all the terms on the first line remain bounded as $\lambda \to \infty$. We note that the first term is equal to $\lambda \Phi_1(\delta^\star) (1 - s - F(\delta^\star) + \Phi_1(\delta^\star)) = \lambda \Phi_1(\delta^\star)^2$ since $F(\delta^\star) = 1 - s$ when the distribution of type is continuous. From Lemma A.8, $\Phi_1(\delta^\star)$ solves:

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

Hence, the first term on the first line of (50) remains bounded as $\lambda \to \infty$. Turning to the second term, we note that:

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

where the inequality follows from the fact that, for $\delta \in [0, \delta^\star - \varepsilon]$, $\Phi_1(\delta) \leq \Phi_1(\delta^\star - \varepsilon)$, and that $\Phi_0(\delta^\star - \varepsilon) \leq F(\delta^\star - \varepsilon)$. From Lemma A.8, $\Phi_1(\delta^\star - \varepsilon)$ solves:

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

Taken together, this shows that the second term on the first line of (50) remain bounded as $\lambda \to \infty$. Proceeding similarly, one can show that the third term on the first line of (50) remains also bounded as $\lambda \to \infty$. The result then follows from the observation that $\lim_{\lambda \to \infty} \vartheta_c = \infty$. ∎

A.4 Proofs omitted in Section 5

Proof of Proposition 4. Using Remark 1 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 Itó'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 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 Equation (24). Using equation (23) in conjunction with Fubini's theorem and the well-known expectation identity

$$\mathbb{E}_{t,\delta}[\hat{\delta}_u] = \int_0^1 \mathbb{P}_{t,\delta}[\hat{\delta}_u > x]dx = 1 - \int_0^1 \mathbb{P}_{t,\delta}[\hat{\delta}_u \le x]dx$$

we obtain that

$$r\Delta V_t(\delta) = 1 - \int_t^\infty re^{-r(u-t)}\left(\int_0^1 \mathbb{P}_{t,\delta}[\hat{\delta}_u \le x]dx\right) du$$

$$= 1 - \int_0^1 \left(\int_t^\infty re^{-r(u-t)}\mathbb{P}_{t,\delta}[\hat{\delta}_u \le x]du\right) dx = 1 - \int_0^1 \Psi_t(x|\delta)dx$$

and equation (24) now follows from an integration by parts. ∎

Proof of Proposition 5. From equation (17), it follows that we have

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

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

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

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

$$\lim_{\lambda\to\infty} k_q(\delta') = \frac{\theta_q\Phi^\star_{1-q}(\delta')}{\theta_0\Phi^\star_1(\delta') + \theta_1\Phi^\star_0(\delta')} = \mathbf{1}_{\{q=0\}}\mathbf{1}_{\{\delta\geq\delta^\star\}} + \mathbf{1}_{\{q=1\}}\mathbf{1}_{\{\delta<\delta^\star\}}$$

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:

$$\gamma F(\delta)\left(s - \Phi_1(\delta)\right) = \gamma\Phi_1(\delta)\left(1 - F(\delta)\right) + \lambda\Phi_1(\delta)\left(1 - s - \Phi_0(\delta)\right). \tag{51}$$

By Proposition 5 we have that $\Phi_1(\delta) \to 0$ and $\Phi_0(\delta) \to F(\delta) < 1 - s$ for all utility types $\delta < \delta^*$ as the meeting frequency becomes infinite, and it thus follows from (51) that for $\delta < \delta^\star$ the distribution of utility types among asset owners admits the approximation

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

Similarly, by Proposition 5 we have that $\Phi_1(\delta) \to F(\delta) - 1 + s > 0$ and $\Phi_0(\delta) \to 1 - s$ for all utility types $\delta > \delta^*$ as the meeting frequency becomes infinite, and it thus follows from (51) that for $\delta > \delta^\star$ the distribution of utility types among non owners admits the approximation

$$1 - s - \Phi_0(\delta) = \frac{\gamma(1 - s)(1 - F(\delta))}{F(\delta) - (1 - s)}\left(\frac{1}{\lambda}\right) + o\left(\frac{1}{\lambda}\right). \tag{52}$$

To derive the convergence rate at the point $\delta = \delta^\star$ assume first that the distribution of utility types crosses 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 (51) evaluated at the marginal type and letting $\lambda \to \infty$ on both sides shows

that the equilibrium distributions admit the approximation given by

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

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 (52) also holds at the marginal type. ∎

Proof of Proposition 6. Assume without loss of generality that the support of the distribution of utility types is the interval $[0, 1]$. Evaluating (19) at $\delta^\star$ and making the change of variable $x = \sqrt{\lambda}(\delta' - \delta^\star)$ in the two integrals shows that

$$r\sqrt{\lambda}(\Delta V(\delta^\star) - p^\star) = P(\lambda) - D(\lambda) \tag{53}$$

where the functions on the right hand side are defined by

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

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

with

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

Letting the meeting rate $\lambda \to \infty$ on both sides of equation (53) and using the convergence result established by Lemma A.12 below we obtain that

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

$$= \int_{0}^{\infty} \frac{(1 - 2\theta_0)g(x)g(-x)dx}{(\theta_0 g(x) + \theta_1 g(-x))(\theta_0 g(-x) + \theta_1 g(x))} dx$$

$$= \int_{0}^{\infty} \frac{\gamma s(1-s)(1-2\theta_0)dx}{\gamma s(1-s) + \theta_0\theta_1(xF'(\delta^\star))^2} = \frac{\pi}{F'(\delta^\star)}\left(\frac{1}{2} - \theta_0\right)\left(\frac{\gamma s(1-s)}{\theta_0\theta_1}\right)^{\frac{1}{2}}$$

where the function

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

is the unique positive solution to (54), 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

$$\int_0^\infty \frac{dx}{a + x^2} = \frac{\pi}{2\sqrt{a}}, \qquad a > 0.$$

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 8. ∎

Lemma A.10 *Assume that the conditions of Proposition 6 hold and denote by $g(x)$ the positive solution to the quadratic equation*

$$g^2 - gF'(\delta^\star)x - \gamma s(1 - s) = 0. \tag{54}$$

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 (16) at the steady state shows that the function $g_1(x)$ is the unique positive solution to the quadratic equation given by

$$g^2 + \left[\frac{\gamma}{\sqrt{\lambda}} + \sqrt{\lambda}\left(F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right) \right] g - \gamma s F(\delta^\star + x/\sqrt{\lambda}) = 0. \tag{55}$$

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 (55) that this limit is given by the positive solution to (54). 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 (55) 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.11 *Assume that the conditions of Proposition 6 hold. Then*

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

$$\begin{aligned} g_1(x) &\leq K/|x|, & x \in I_-^\lambda &\equiv [-\delta^\star\sqrt{\lambda}, 0], \\ g_0(x) &\leq K/|x|, & x \in I_+^\lambda &\equiv [0, (1 - \delta^\star)\sqrt{\lambda}]. \end{aligned} \tag{56}$$

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

$$g_1(x) \geq k|x|, \qquad x \in I_+^\lambda \cap [\bar{x}, \infty), \tag{57}$$

$$g_0(x) \geq k|x|, \qquad x \in I_-^\lambda \cap (-\infty, -\bar{x}].$$

for all sufficiently large λ.

Proof. Recall that the function $g_1(x)$ is the positive root of (55). Thus we have that equation (56) 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

$$\min_{x \in I_-^\lambda} \left\{ \frac{K}{|x|} \sqrt{\lambda} \left[F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) \right] - \gamma s(1 - s) \right\} \geq 0. \tag{58}$$

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

$$F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) = -\frac{xF'(\hat{\delta}(x))}{\sqrt{\lambda}} \tag{59}$$

and substituting this expression into (58) shows that a sufficient condition for the validity of equation (56) is that we have

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

Because the derivative of the distribution of utility types is assumed to be strictly positive on the whole interval $[0, 1]$ we have that K^* is finite and equation (56) follows. One obtains the same constant when applying the same calculations to the function $g_0(x)$ over the interval $x \in I_+^\lambda$.

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 (55) we have that (57) 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 (59) then shows that a sufficient condition for the validity of equation (57) is given by

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

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^* to be strictly positive. One obtains the same constant when applying the same calculations to the function $g_0(x)$ over the interval $x \in I_-^\lambda \cap (-\infty, -\bar{x}]$. ■

Lemma A.12 *Assume that the conditions of Proposition 6 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.10.

Proof. By Lemma A.10 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

$$\lim_{\lambda \to \infty} H(x; \lambda) = \frac{\theta_0 g(x)}{\theta_0 g(x) + \theta_1 g(-x)}. \tag{60}$$

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.11 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 to apply the dominated convergence theorem, and the result for $D(\lambda)$ now follows from (60). The result for the other integral follows from identical calculations. We omit the details. ■

Proof of Proposition 7. In this section we show that, with a discrete distribution of types, then generically equilibrium objects converge to their frictionless counterpart in order $1/\lambda$.

To see this assume that there are I types $\delta_1 < \delta_2 < \ldots < \delta_I$. Let the marginal type be the $m \in \{1, \ldots, I\}$ such that:

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

We let $\delta_0 \equiv 0$ and $\delta_{I+1} \equiv 1$. We assume further that $1 - F(\delta_m) < s$, which occurs generically when the distribution of types is restricted to be discrete. In this case, the same algebraic manipulations as in the text show that:

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

and, for all $\delta \in [\delta_i, \delta_{i+1})$, $\Phi_1(\delta) = \Phi_1(\delta_i)$. Likewise, the local surplus is equal to:

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

and, for all $\delta \in [\delta_i, \delta_{i+1})$, $\sigma(\delta) = \sigma(\delta_i)$. Therefore, for all $i < m$:

$$\Delta V(\delta_m) - \Delta V(\delta_i) = \int_0^{\delta_m} \sigma(\delta)\, d\delta = \sum_{j=i}^{m-1} (\delta_{j+1} - \delta_j)\,\sigma(\delta_j) = \sum_{j=i}^{m-1} \frac{\delta_{j+1} - \delta_j}{\theta_1\left[1 - s - F(\delta_j)\right]} + o\left(\frac{1}{\lambda}\right).$$

Likewise, for all $i > m$:

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

Finally, we calculate the reservation value $\Delta V(\delta_m)$ using formula (13)

$$r\Delta V(\delta_m) = \delta_m - \int_0^{\delta_m} \sigma(\delta)\left\{\gamma F(\delta) + \lambda\theta_0\Phi_1(\delta)\right\}\, d\delta + \int_{\delta_m}^1 \sigma(\delta)\left\{\gamma\left[1 - F(\delta)\right] + \lambda\theta_1\left[1 - s - \Phi_0(\delta)\right]\right\}\, d\delta$$

$$= \delta_m - \sum_{i=0}^{m-1} \frac{(\delta_{i+1} - \delta_i)\left\{\gamma F(\delta_i) + \lambda\theta_0\Phi_1(\delta_i)\right\}}{r + \gamma + \lambda\theta_0\Phi_1(\delta_i) + \lambda\theta_1\left[1 - s - \Phi_0(\delta_i)\right]}$$

$$+ \sum_{i=m}^{I} \frac{(\delta_{i+1} - \delta_i)\left\{\gamma\left[1 - F(\delta_i)\right] + \lambda\theta_1\left[1 - s - \Phi_0(\delta_i)\right]\right\}}{r + \gamma + \lambda\theta_0\Phi_1(\delta_i) + \lambda\theta_1\left[1 - s - \Phi_0(\delta_i)\right]}$$

$$= \delta_m - \frac{1}{\lambda}\sum_{i=0}^{m-1} \frac{(\delta_{i+1} - \delta_i)\,\gamma F(\delta_i)\left[1 - F(\delta_i) - s(1 - \theta_0)\right]}{\left[1 - s - F(\delta_i)\right]^2}$$

$$+ \frac{1}{\lambda}\sum_{i=m}^{I} \frac{(\delta_{i+1} - \delta_i)\,\gamma\left[1 - F(\delta_i)\right]\left[F(\delta_i) - (1 - s)(1 - \theta_1)\right]}{\left[F(\delta_i) - (1 - s)\right]^2} + o\left(\frac{1}{\lambda}\right),$$

where the last equality follows after using the asymptotic expansions for $\Phi_1(\delta)$ and $1 - s - \Phi_0(\delta) = 1 - s - F(\delta) + \Phi_1(\delta)$. All in all, these calculations establish the claim that bilateral price levels all converge to their Walrasian counterpart, δ_m/r at a speed in order $1/\lambda$.

Finally, we study the welfare cost of misallocation:

$$
\begin{aligned}
C &= \int_0^{\delta_m} \Phi_1(\delta_m)\, d\delta + \int_{\delta_m}^1 \left[1 - s - \Phi_0(\delta) \right] d\delta \\
&= \frac{1}{\lambda} \sum_{i=0}^{m-1} \frac{(\delta_{i+1} - \delta_i)\gamma s F(\delta_i)}{1 - s - F(\delta_i)} + \frac{1}{\lambda} \sum_{i=m}^{I} \frac{(\delta_{i+1} - \delta_i)\gamma(1 - s)\left[1 - F(\delta_i)\right]}{F(\delta_i) - (1 - s)} + o\left(\frac{1}{\lambda}\right),
\end{aligned}
$$

using the asymptotic expansions for $\Phi_1(\delta)$ and $1 - s - \Phi_0(\delta)$. This expression also goes to zero in order $1/\lambda$. $\blacksquare$

Proof of Proposition 8. The first intermediate result is:

Lemma 7 *As λ goes to infinity:*

$$
\lambda \int_0^{\delta^\star} \sigma(\delta)\, d\delta = \int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} + O(1) \tag{61}
$$

$$
\lambda \int_{\delta^\star}^1 \sigma(\delta)\, d\delta = \int_{\delta^\star}^1 \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta^\star - \delta) + 1 - s - \Phi_0(\delta)} + O(1). \tag{62}
$$

Proof. We start with (61), noting that:

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

where we used that $\Phi_0(\delta) = F(\delta) - \Phi_1(\delta)$, and $F(\delta^\star) = 1 - s$. Using this expression to calculate the difference between the left- and the right-hand side of (61), we obtain:

$$
\left| \int_0^{\delta^\star} \left(\lambda\sigma(\delta) - \frac{1}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} \right) d\delta \right| \leq \int_0^{\delta^\star} \frac{\theta_1 \left| F'(\delta^\star)(\delta^\star - \delta) - \left[F(\delta^\star) - F(\delta)\right] \right|}{\theta_1^2 F'(\delta^\star)\left[\delta^\star - \delta\right]\left[F(\delta^\star) - F(\delta)\right]}\, d\delta.
$$

In the right-hand side integral, under our assumption that $F(\delta)$ is twice continuously differentiable, we can use the Taylor Theorem to extend the integrand by continuity at $\delta^\star$, with value $\frac{F''(\delta^\star)}{2\theta_1 F'(\delta^\star)}$. Thus, the integrand is bounded, establishing (61). Turning to equation (62), we first note that:

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

69

where we used that $\Phi_1(\delta) = F(\delta) - F(\delta^\star) + F(\delta^\star) - \Phi_0(\delta) = F(\delta) - F(\delta^\star) + 1 - s - \Phi_0(\delta)$ since $F(\delta^\star) = 1 - s$. The rest of the proof is identical as the one for (61). ∎

Next, we obtain a lower bound for the integral on the right-hand side of (61) by bounding $\Phi_1(\delta)$ above by $\Phi_1(\delta^\star)$:

Lemma A.13 *As $\lambda \to \infty$:*

$$\int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} \geq \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1) \tag{63}$$

$$\int_{\delta^\star}^1 \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta - \delta^\star) + 1 - s - \Phi_0(\delta)} \geq \frac{\log(\lambda)}{2\theta_0 F'(\delta^\star)} + O(1). \tag{64}$$

Proof. For (63), this follows by noting that $\Phi_1(\delta) \leq \Phi_1(\delta^\star)$, and integrating directly:

$$\int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta^\star)} = \left[-\frac{1}{\theta_1 F'(\delta^\star)} \log\left(\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)[\delta^\star - \delta] + \Phi_1(\delta^\star) \right) \right]_0^{\delta^\star}$$

$$= O(1) - \frac{1}{\theta_1 F'(\delta^\star)} \log\left[\sqrt{\frac{\gamma s(1-s)}{\lambda}} + o\left(\frac{1}{\sqrt{\lambda}}\right) \right] = \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1),$$

where the second equality follows from plugging in the asymptotic expansion of $\Phi_1(\delta^\star)$ derived in Section 7. For (64), this follows from the same manipulation: first by noting that $1 - s - \Phi_0(\delta) \leq 1 - s - \Phi_0(\delta^\star)$, and integrating directly. ∎

Next we establish the reverse inequality:

Lemma A.14 *As $\lambda \to \infty$:*

$$\int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} \leq \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1) \tag{65}$$

$$\int_{\delta^\star}^1 \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta - \delta^\star) + 1 - s - \Phi_0(\delta)} \leq \frac{\log(\lambda)}{2\theta_0 F'(\delta^\star)} + O(1). \tag{66}$$

Proof. For (65), let us break down the integral into an integral over $\left[0, \delta^\star - 1/\sqrt{\lambda}\right]$, and an integral over $\left[\delta^\star - 1/\sqrt{\lambda}, \delta^\star\right]$. The first integral can be bounded above by:

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

The second term can be bounded above by:

$$\int_{\delta^\star - 1/\sqrt{\lambda}}^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)}$$

$$= \frac{1}{\theta_1 F'(\delta^\star)} \log\left(\frac{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)\alpha(\lambda) + \Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)}{\frac{r+\gamma}{\lambda} + \Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)}\right)$$

$$\rightarrow \frac{1}{\theta_1 F'(\delta^\star)} \log\left(\frac{\theta_1 F'(\delta^\star) + g(-1)}{g(-1)}\right) = O(1),$$

where $g(-1)$ is the limit of $\sqrt{\lambda}\Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)$ as shown in Lemma A.10. For (66), the result follows from identical algebraic manipulations. ∎

∎

B The welfare cost of frictions near the frictionless limit

In this section, we study the asymptotic impact of frictions on welfare. To answer this question we start by noting that in the context of our decentralized market model the welfare cost of misallocation can be defined as:

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

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 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 9 *Assume that the distribution of utility types is twice continuously differentiable with a strictly positive derivative. Then*

$$C(\lambda) = \frac{\gamma s(1-s)}{2F'(\delta^\star)} \frac{\log(\lambda)}{\lambda} + O\left(\frac{1}{\lambda}\right).$$

where the remainder satisfies $\lim_{\lambda \to \infty} |\lambda O(1/\lambda)| < \infty$. By contrast, with a discrete distribution of utility types, the convergence rate is generically equal to $1/\lambda$.

Proof of Proposition 9. Recall that

$$C = \int_0^{\delta^\star} \Phi_1(\delta)\, d\delta + \int_{\delta^\star}^1 [1 - s - \Phi_0(\delta)]\, d\delta.$$

71

Let us start with the first integral. We note that, from the quadratic equation for $\Phi_1(\delta)$:

$$\lambda\Phi_1(\delta) = \frac{\gamma s F(\delta)}{\frac{\gamma}{\lambda} + \Phi_1(\delta) + F(\delta^\star) - F(\delta)}.$$

Given the above formula, and using the same arguments as in the surplus calculation, in Section A.4, we note that:

$$\left| \int_0^{\delta^\star} \left(\lambda\Phi_1(\delta) - \frac{\gamma s F(\delta^\star)}{\frac{\gamma}{\lambda} + \Phi_1(\delta) + F'(\delta^\star)[\delta^\star - \delta]} \right) d\delta \right| = O(1).$$

Then, we also note that:

$$\int_{\delta^\star - \frac{1}{\sqrt{\lambda}}}^{\delta^\star} \frac{\gamma s F(\delta^\star)}{\frac{\gamma}{\lambda} + \Phi_1(\delta) + F'(\delta^\star)[\delta^\star - \delta]} d\delta \leq \frac{\gamma s F(\delta^\star)}{F'(\delta^\star)} \log\left(\frac{\frac{\gamma}{\lambda} + \Phi_1\left(\delta^\star - \frac{1}{\sqrt{\lambda}}\right) + F'(\delta^\star)\frac{1}{\sqrt{\lambda}}}{\frac{\gamma}{\lambda} + \Phi_1\left(\delta^\star - \frac{1}{\sqrt{\lambda}}\right)} \right) = O(1),$$

because $\Phi_1\left(\delta^\star - \frac{1}{\sqrt{\lambda}}\right) = \frac{g(-1)}{\sqrt{\lambda}} + o(1)$. So, we find that:

$$\int_0^1 \lambda\Phi_1(\delta)\, d\delta = \int_0^{\delta^\star - \frac{1}{\sqrt{\lambda}}} \frac{\gamma s F(\delta^\star)}{\frac{\gamma}{\lambda} + \Phi_1(\delta) + F'(\delta^\star)[\delta^\star - \delta]} d\delta + O(1).$$

To obtain a lower bound for the integral, we can bound $\Phi_1(\delta)$ above by $\Phi_1\left(\delta^\star - \frac{1}{\sqrt{\lambda}}\right)$, and to obtain an upper bound, we can bound $\Phi_1(\delta)$ below by zero. In both case, we can compute the integral explicitly and we find that the upper and the lower bound can both be written as:

$$\frac{\gamma s F(\delta^\star)}{2F'(\delta^\star)} \log(\lambda) + O(1) = \frac{\gamma s (1 - s)}{2F'(\theta^\star)} \log(\lambda) + O(1).$$

We then go through the same algebraic manipulations to characterize the asymptotic behavior of the second integral, $\int_{\delta^\star}^1 [1 - s - \Phi_0(\delta)]\, d\delta$, and the result follows. ∎

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 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 Ricatti 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

$$\lambda\Phi_{1,t}(\delta) = \lambda(F_t(m) - \Phi_{0,t}(\delta)) = \frac{\dot{Y}_{+,t}(\delta) - A(\delta)\dot{Y}_{-,t}(\delta)}{Y_{+,t}(\delta) - A(\delta)Y_{-,t}(\delta)} \tag{67}$$

with

$$\begin{aligned}
Y_{\pm,t}(\delta) &= e^{-\lambda Z_\pm(\delta)t}W_{\pm,t}(\delta) \\
Z_\pm(\delta) &= \frac{1}{2}\left(1 - s + \gamma/\lambda - F(\delta)\right) \pm \frac{1}{2}\Lambda(\delta) \\
W_{\pm,t}(\delta) &= M_1\left(\frac{\lambda}{\gamma}Z_\pm(\delta), 1 \pm \frac{\lambda}{\gamma}\Lambda(m); e^{-\gamma t}\frac{\lambda}{\gamma}(F(\delta) - F_0(\delta))\right)
\end{aligned} \tag{68}$$

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* (67) *satisfies* $\lim_{t\to\infty} \Phi_{q,t}(\delta) = \Phi_q(\delta)$ *for any initial distributions* $F_0(\delta)$ *and* $F_{1,0}(\delta)$.

Proof. Straightforward algebra shows that (67) 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 (68) 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 Market-makers

Assume that in addition to a continuum of agents the market also includes a unit mass of competitive market makers who have access to a frictionless interdealer market and keep no inventory. An agent contacts market makers 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 $\delta \leq w_t$ are willing to sell, while only those non owners of type $\delta \geq w_t$ are willing to buy. Since market makers 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 market makers must net out to zero because they keep no inventory. The total mass of owners who contact market makers to sell is $\alpha\Phi_{1,t}(w_t)$. On the other hand, the total mass of non owners who contact market makers 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

$$\Phi_{1,t}(w_t) - (1 - \pi_{1,t})\Delta\Phi_{1,t}(w_t) = 1 - s - \Phi_{0,t}(w_t) + \pi_{0,t}\Delta\Phi_{0,t}(w_t) \tag{69}$$

where $\pi_{q,t} \in [0, 1]$ denotes the probability with which market makers 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 all $t \geq 0$, 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. This choice has by construction 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

$$\Delta V_t(\delta) = \mathbb{E}_t \left[\int_t^\sigma e^{-r(u-t)} \delta du + e^{-r(\sigma-t)} \left(\Delta V_\sigma(\delta) + \mathbf{1}_{\{\sigma=\tau_\alpha\}} z \left(\Delta V_\sigma(\delta^\star) - \Delta V_\sigma(\delta) \right) \right. \right. \tag{70}$$

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

$$+ \mathbf{1}_{\{\sigma=\tau_0\}} \theta_1 \int_0^1 \left(\Delta V_\sigma(\delta') - \Delta V_\sigma(\delta) \right)^+ \frac{d\Phi_{0,\sigma}(\delta')}{1-s}$$

$$\left. \left. - \mathbf{1}_{\{\sigma=\tau_1\}} \theta_0 \int_0^1 \left(\Delta V_\sigma(\delta) - \Delta V_\sigma(\delta') \right)^+ \frac{d\Phi_{1,\sigma}(\delta')}{s} \right) \right].$$

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 market makers is isomorphic to that which would prevail in an environment where there are no market makers, 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 market makers.

Lemma D.1 *There exists a unique function that satisfies* (70) *subject to* (8)*. This function is uniformly*

76

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 market makers, 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 (69) 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 (16) 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 market makers execute orders from marginal agents. To solve this Ricatti differential equation let

$$\Phi_1(\delta) = -\frac{1}{2}\left(1 - s + \gamma/\lambda + \alpha/\lambda - F(\delta)\right) + \frac{1}{2}\Psi(\delta) \tag{71}$$

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 3 and Corollary 2 for the model with market makers.

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*

$$\Phi_{1,t}(\delta) = \Phi_1(\delta) + \frac{(\Phi_{1,0}(\delta) - \Phi_1(\delta))\,\Psi(\delta)}{\Psi(\delta) + (\Phi_{1,0}(\delta) - \Phi_1(\delta) + \Psi(\delta))\left(e^{\lambda\Psi(\delta)t} - 1\right)}. \tag{72}$$

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 3. ∎

Corollary D.3 *The steady state measure $\Phi_1(\delta)$ is increasing in γ, decreasing in λ and 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 5. ∎

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 (72), and the reservation value function satisfies (7) subject to (8) given the distributions.*

As in the benchmark model without market maker 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 market makers. Moreover, given any initial conditions satisfying (1) and (2) this equilibrium converges to the steady state distributions of equation (71) and the reservation value function of equation (19) 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 market makers, and verify that the corresponding predictions are qualitatively similar to those of the benchmark model.