Heterogeneity in Decentralized Asset Markets*

Julien Hugonnier† Benjamin Lester‡ Pierre-Olivier Weill§

August 24, 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.

*We thank, for fruitful discussions and suggestions, Gadi Barlevy, Jaksa Cvitanic, Rudi Fahlenbrach, Semyon Malamud, Guillaume Rocheteau, and seminar participants at the Federal Reserve Bank of Philadelphia, the 2014 SaM Conference in Edinburgh, the 2014 conference on "Recent Advances in OTC Market Research" in Paris, Royal Holloway, UCL, CREST, the 2014 KW25 Anniversary conference, the 2014 Summer Workshop on Money, Banking, Payments and Finance at the Chicago Fed, Gerzensee, the Bachelier workshop, the 2013 AFFI Congress. The views expressed here are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This project was completed when Pierre-Olivier Weill was a visiting Professor at the Paris School of Economics, whose hospitality is gratefully acknowledged.

†Swiss Finance Institute and Ecole Polytechnique Fédérale de Lausanne.

‡Federal Reserve Bank of Philadelphia.

§UCLA and NBER.

1 Introduction

Many assets and durable goods trade in decentralized or "over-the-counter" (OTC) markets.[1] In order to study these markets, we consider 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 a new methodology that renders the analysis fully tractability: we provide a full characterization of the equilibrium, in closed form, both in and out of steady state. Hence, there is essentially no cost of allowing for more than two valuations. The benefits, however, are substantial.

First, a model with many types of investors can account for a number of stylized facts that have been established in recent empirical studies of OTC markets. For example, as in the data, the assets in our model are re-allocated through chains of inframarginal trades with each trade, on average, taking place at increasingly longer intervals and higher prices. These patterns of trade generate both a core-periphery network structure and a non-degenerate distribution of prices, which are both well-documented features of many OTC markets. Given the ease with which our model can account for many observed patterns of trade in these markets, along with its tractability, we argue that it constitutes a unified framework to study a number of important, emerging issues—both theoretical and empirical—regarding OTC markets.

Of equal importance, our model also generates a number of 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.

The structure of the paper is as follows. After providing a review of the related literature,

[1] Examples of assets that trade in OTC 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.

Section 2 provides a formal description of the environment. To facilitate comparison with the existing literature, our model starts with the basic building blocks of DGP. We assume that there is a fixed measure of agents in the economy and a fixed measure of indivisible shares of an asset. Agents have heterogeneous valuations for this asset, which can change over time, and they are allowed to hold either zero or one share. Each agent is periodically and randomly matched with another agent, 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 potentially makes the analysis considerably more complex, as the state of the economy now includes two (potentially) infinite-dimensional objects: the distributions of valuations among agents that hold zero and one assets, respectively, over time. However, in Section 3, we show that, using the proper methodology, the solution remains fully In particular, we characterize the unique equilibrium, in closed form, both in and out of steady state.

Then, in Section 4, we exploit our characterization of equilibrium to flesh out the basic, positive predictions that emerge from our model, and 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 agents with low valuations to those with high valuations through chains of inframarginal trades: an agent who owns an asset and has a low valuation tends to sell very quickly and then remain asset-less for a long duration; an agent who owns an asset and has a high valuation tends to sell very infrequently, but will typically buy a new one soon thereafter; and agents with moderate valuations tend to engage in both purchases and sales with a moderate frequency. 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, the expected duration of search, the expected valuation of his ultimate counterparty, and thus the expected price he pays or receives. These predictions are consistent with observations from certain OTC markets, such as the municipal bonds and the federal funds markets, but they could not be generated by more traditional models. More specifically, in models with many types of agents and frictionless centralized trades, the law of one price holds, and so the price that an investor pays is independent of the valuation of its counterparty. In a model of decentralized trade with only two types of agents, all sellers (and buyers) trade at the same average frequency and all trades occur at the same price.[2]

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 agents with "near-marginal" valuations tend to specialize in intermediation, a "core-periphery" trading network emerges endogenously, with transactions occuring at different prices at each node. We document that 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 (and is only getting faster). 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. Section 6 concludes.

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

[2]Several authors (Duffie et al., 2005; Weill, 2007; Lagos and Rocheteau, 2009) have studied models in which potentially many types of investors make periodic contacts with dealers who have access to a centralized market. In these models, trading times are constant across investors and prices are independent of the counterparty type.

in Duffie et al. (2005), who study how search frictions in OTC markets affect the bid-ask spread set by marketmakers who have access to a competitive interdealer market.[3] 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. Purely decentralized trade captures important features of reality; for example, Li and Schürhoff (2012a) show that, in the municipal bond market, even the inter-dealer market is bilateral, with intermediation chains that typically involve more than two transactions. In the literature, 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).[4] 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.[5] 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

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

[4]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.

[5]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 various types of 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.

asset markets and allow for agents to have more than two valuations. Perhaps the closest to 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.[6]

Our paper is also related to an important, growing literature that studies equilibrium asset pricing and exchange in exogenously specified trading networks. Examples include Gofman (2010), Babus and Kondor (2012), Malamud and Rostek (2012), Alvarez and Barlevy (2014), as well as Atkeson et al. (2012) whose framework blends elements of the search and of 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 in-

[6]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). Many of these insights were derived independently earlier 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).

vestors who value them most is to use an intermediation chain. In a 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.

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).[7] Naturally, understanding such liquidity premia is central to our analysis as well. A key issue in the latter literature is understanding how workers move from firm to firm through the process of on-the-job search; see, e.g., Burdett and Mortensen (1998) and Postel-Vinay and Robin (2002). The dynamics of these worker flows across firms, and the corresponding distributions and measures of misallocation, share much in common with the way that assets move across investors in our model.

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

[7]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).

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 $t \geq 0$:[8]

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

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

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

[8]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.

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 $p_t = p$, which is constant, in equilibrium, since the cross-sectional distribution of types in the population is time-independent.

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

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

after integration by part. This representation of an investor's objective makes it clear that, at each time t, optimal holdings satisfy:

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

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

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

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[9] $p^\star = \delta^\star / r$

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 two 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 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 gains from trade if and only if $\delta' \leq \delta$. The resulting patterns of exchange allow us to completely characterize the joint distribution of asset holdings and utility types that must prevail in equilibrium.

Given explicit solutions for reservation values and distributions, we then 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

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

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

the reservation value of an investor of utility type $\delta \in [0,1]$ at time $t \geq 0$. An application of Bellman's principle of optimality shows that

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

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

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

The first term on the right hand side of (3) accounts for the fact that an asset owner enjoys a constant flow of utility at rate δ until time τ. The three remaining terms capture the fact that there are three possible events for the investor at time τ: he can receive a preference shock ($\tau = \tau_\gamma$), in which case a new preference type is randomly 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 a non-owner satisfies

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

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

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

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 buyer 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 a seller 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.

We now use the contraction mapping theorem to establish existence, uniqueness and basic properties of bounded solutions to this equation.[10]

Lemma 1 *There exists a unique, uniformly bounded function* $\Delta V : \mathbb{R}_+ \times [0,1] \to \mathbb{R}$ *that satisfies* (7)*. This function is 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.*

The strict monotonicity in δ implies that, when an asset owner of type δ meets a non-owner of type

[10] The restriction to bounded solutions is natural in the context of our model and serves two purposes. First, it rules out bubbles by guaranteeing that the reservation values satisfy an appropriate transversality condition. Second, it allows us to achieve uniqueness. To see that such a restriction is necessary to guarantee uniqueness, it suffices to observe that if the function $f_t(\delta)$ is a solution to (7), then so is the function $f_t(\delta) + e^{rt}\beta$ for any constant β.

$\delta' > \delta$, 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, 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 the equilibrium distributions of types and asset holdings in the next section.

Integrating both sides of (7) with respect to the distribution of the random time τ, and using the fact that reservation values are strictly increasing in δ to simplify, we find that the reservation value function satisfies the integral equation

$$\begin{aligned}\Delta V_t(\delta) = \int_t^\infty e^{-(r+\gamma+\lambda)(u-t)}\bigg(\delta + \lambda \Delta V_u(\delta) + \gamma \int_0^1 \Delta V_u(\delta') dF(\delta') \\ + \lambda \int_\delta^1 \theta_1 \left(\Delta V_u(\delta') - \Delta V_u(\delta)\right) d\Phi_{0,u}(\delta') \\ - \lambda \int_0^\delta \theta_0 \left(\Delta V_u(\delta) - \Delta V_u(\delta')\right) d\Phi_{1,u}(\delta')\bigg) du.\end{aligned} \tag{8}$$

In addition, since Lemma 1 establishes that the reservation value function is absolutely continuous in $(t, \delta) \in \mathbb{R}_+ \times [0, 1]$ with a bounded space derivative, it can represented as

$$\Delta V_t(\delta) = \Delta V_t(0) + \int_0^\delta \sigma_t(\delta') d\delta' \tag{9}$$

for some nonnegative and uniformly bounded function $\sigma_t(\delta)$ that is itself absolutely continuous in time for almost every $\delta \in [0, 1]$. Since the gains from trade between a seller of type δ and a buyer of type $\delta + d\delta$ are approximately given by $\sigma_t(\delta) d\delta$, we naturally interpret this function as a measure of the *local surplus* at type δ in the decentralized market.

Substituting the representation (9) into the integral equation (8) and differentiating both sides of the resulting expression with respect to t and δ shows that the local surplus satisfies the ordinary differential equation

$$\big(r+\gamma+\lambda\theta_1(1-s-\Phi_{0,t}(\delta))+\lambda\theta_0\Phi_{1,t}(\delta)\big)\sigma_t(\delta)=1+\dot{\sigma}_t(\delta) \tag{10}$$

at almost every point of $\mathbb{R}_+\times[0,1]$. A calculation given in the appendix (see Lemma A.2) shows that, together with the requirements of uniform boundedness and absolute continuity in time, this equation pins down the local surplus as the present value of a perpetuity using an appropriate effective discount rate:

$$\sigma_t(\delta)=\int_t^\infty e^{-\int_t^u\big(r+\gamma+\lambda\theta_1(1-s-\Phi_{0,\xi}(\delta))+\lambda\theta_0\Phi_{1,\xi}(\delta)\big)d\xi}du \tag{11}$$

Combining the results above, we derive an explicit solution for the reservation value function and establish some basic comparative statics in Proposition 1.

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

$$\begin{aligned}\Delta V_t(\delta)=\int_t^\infty e^{-r(u-t)}\bigg(\delta-\int_0^\delta\sigma_u(\delta')\left(\gamma F(\delta')+\lambda\theta_0\Phi_{1,u}(\delta')\right))d\delta'\\ +\int_\delta^1\sigma_u(\delta')\left(\gamma(1-F(\delta'))+\lambda\theta_1(1-s-\Phi_{0,u}(\delta'))\right)d\delta'\bigg)du,\end{aligned} \tag{12}$$

where the local surplus $\sigma_t(\delta)$ is fully described by equation (11). Moreover, 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 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 increase 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 in general ambiguous and depends on all the parameters that characterize the decentralized market.

Remark 1 (HJB equation) Differentiating with respect to time on both sides of (8) shows that the reservation value function can be equivalently characterized as the unique bounded and absolutely continuous solution to the Hamilton-Jacobi-Bellman equation

$$r\Delta V_t(\delta) = \delta + \gamma \int_0^1 (\Delta V_t(\delta') - \Delta V_t(\delta))\, dF(\delta') + \lambda \int_\delta^1 \theta_1 (\Delta V_t(\delta') - \Delta V_t(\delta))\, d\Phi_{0,t}(\delta')$$
$$- \lambda \int_0^\delta \theta_0 (\Delta V_t(\delta) - \Delta V_t(\delta'))\, d\Phi_{1,t}(\delta') + \Delta \dot{V}_t(\delta). \tag{13}$$

This alternative characterization is not directly exploited in our derivation of an explicit solution but it is nonetheless quite useful. In particular, we show in Section 5.1 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 one of them is an 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 change in the measure of asset owners with utility type less than δ can be written as:

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

The first term in equation (14) 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 (14).[11]

Using (1), we can re-write (14) as an ordinary differential equation (ODE):

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

Proposition 2 below provides an explicit expression for the unique solution to this Ricatti equation and shows that it converges to the steady state. To state the result, let

$$\Phi_1(\delta) = F(\delta) - \Phi_0(\delta) \equiv -\frac{1}{2}\left(1 - s + \gamma/\lambda - F(\delta)\right) + \frac{1}{2}\Lambda(\delta), \tag{16}$$

with

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

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

Proposition 2 *At any time $t \geq 0$ the measure of asset owners with utility type less than $\delta \in [0, 1]$*

[11] 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.

is explicitly given by

$$\Phi_{1,t}(\delta) = \Phi_1(\delta) + \frac{\left(\Phi_{1,0}(\delta) - \Phi_1(\delta)\right)\Lambda(\delta)}{\Lambda(\delta) + \left(\Phi_{1,0}(\delta) - \Phi_1(\delta) + \Lambda(\delta)\right)\left(e^{\lambda\Lambda(\delta)t} - 1\right)} \tag{17}$$

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

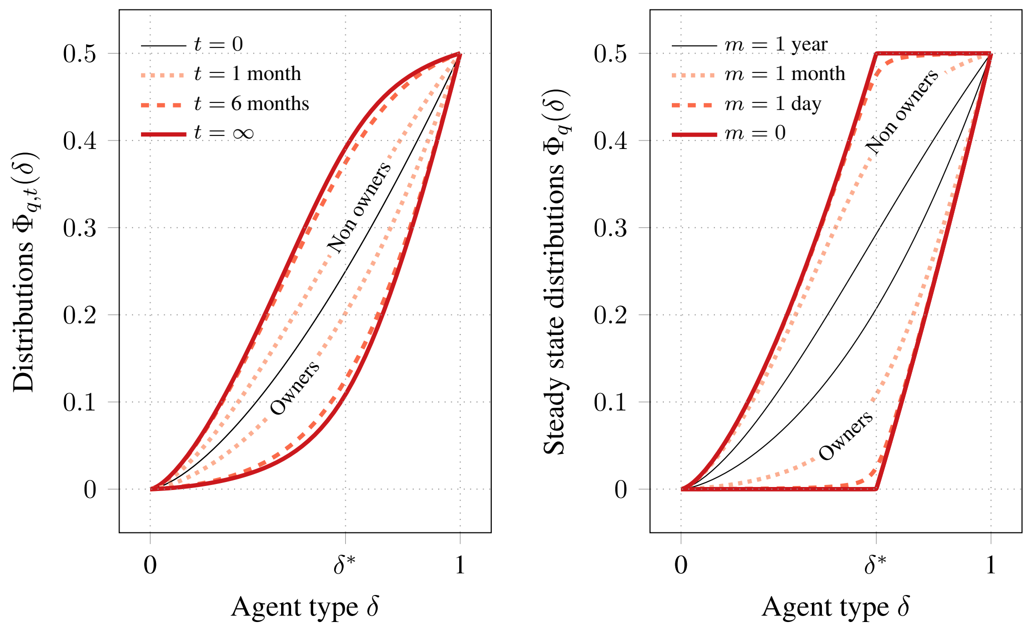
Note that we can solve the ODE for each δ, without imposing any regularity condition on the distribution as a function of δ. To illustrate the convergence of the equilibrium distributions to the steady state, we introduce a simple numerical example, which we will continue to use throughout the text. In this example, the discount rate is $r = 0.05$, the asset supply is $s = 0.5$, the arrival rate of meetings is $\lambda = 12$ so that agents meet others once a month on average, the arrival rate of preference shocks $\gamma = 1$ so that agents change type once a year on average, and the underlying distribution of types is given by $F(\delta) = \delta^\alpha$ with $\alpha = 1.5$ so that the marginal type is $\delta^\star = 0.6299$.

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 low valuation investors are less and less likely to hold the asset over time.

Finally, focusing on the steady state, equation (16) offers several natural comparative statics, summarized in Corollary 1 below. Intuitively, as preference shocks become less frequent (γ decreases) or trading opportunities become more frequent (λ increases), there is a first order stochastic dominant shift in $\Phi_1(\delta)$, which implies that the asset is being allocated to agents with higher valuations more efficiently.

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

FIGURE 1: Convergence to the steady state distributions



Notes. The left panel plots the cumulative distribution of types among non-owners (upper curves) and sellers (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, $m = 1/\lambda$.

3.3 Equilibrium

Definition 1 *An equilibrium is a reservation value function* $\Delta V_t(\delta)$ *and a pair of distributions* $\Phi_{0,t}(\delta)$ *and* $\Phi_{1,t}(\delta))$ *such that: the distributions satisfy* (1), (2) *and* (17), *and the reservation value function is uniformly bounded and satisfies* (7) *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 $\Phi_{0,t}(\delta)$ 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*

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

and the steady state distributions of equation (16).

4 Trading Patterns

In this section, we flesh out the basic, positive 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 with utility type δ—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 help to 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 for assets or durable goods. Finally, in addition to arguing that the model can account for many existing observations, we also highlight several novel predictions that emerge from the model 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 and the expected price at which this trade will occur for an investor of type δ with asset holdings $q \in \{0, 1\}$, along with a number of related comparative statics. Of particular importance, in Lemma 2 we establish that investors with the most to gain from trade tend to trade relatively quickly, and in Lemma 3 we establish that bilateral prices are increasing in the valuations of both the buyer and the seller.

While these results seem fairly intuitive, their implications are far-reaching. For example, not only does Lemma 2 tell us that an 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 relatively few of them in equilibrium, and hence a non-owner with a low valuation is most likely to trade with an owner whose valuation is also relatively low. From Lemma 3, 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 any 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 arrival rate of profitable trading opportunities for an asset owner with utility type δ is equal to the product of the arrival rate of meetings, λ, and the probability that the investor he meets is a non-owner with utility type greater than δ, or $[1 - s - \Phi_0(\delta)]$. Since the latter probability is decreasing in δ, while the arrival rate of preference shocks is independent of δ, it follows that the expected time it takes an asset owner to trade is decreasing in δ.

To formalize this intuition, let $\eta_q(\delta)$ denote the expected amount of time before a trade occurs for an investor with $q \in \{0, 1\}$ units of the asset and utility type δ. Lemma 2, below, characterizes $\eta_q(\delta)$ in closed form when $F(\delta)$ is continuous, and offers several natural comparative statics.

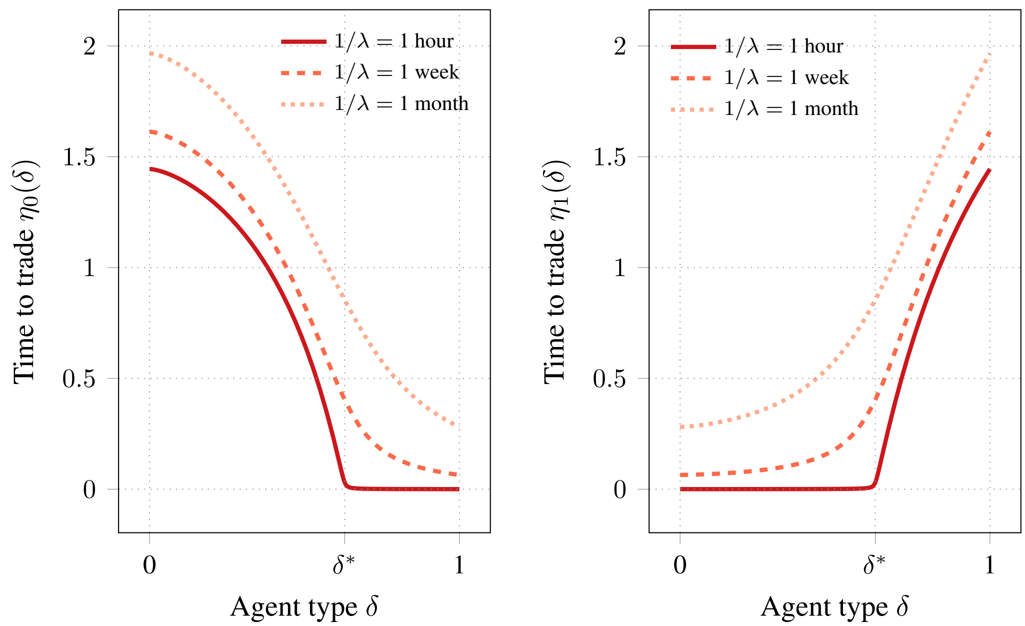
Lemma 2 *Suppose $F(\delta)$ is continuous, and let $\lambda_q(\delta) = \lambda q(1 - s - \Phi_0(\delta)) + \lambda(1 - q)\Phi_1(\delta)$. Then*

$$\eta_q(\delta) = \left[\frac{\gamma}{\lambda}\log\left(\frac{\gamma}{\gamma+\lambda}\right) + \left(1 - \frac{1+\frac{\gamma}{\lambda}}{\Phi_q(1)}\right)\log\left(1 - \frac{\Phi_q(1)}{1+\frac{\gamma}{\lambda}}\right)\right]^{-1}\frac{1}{\gamma + \lambda_q(\delta)}$$

for $q \in \{0, 1\}$. 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 λ, and illustrates the sharp differences between time to trade below and above the marginal type. Consider for instance the time to trade

FIGURE 2: Expected time to trade



Notes. This plots the expected waiting time for buyers and sellers as functions of the agents' type when meetings happen, on average, every month (dotted), every week (dashed) and every hour (solid). The model we use in this figure is otherwise the same as in Figure 1.

of a non-owner with utility type $\delta < \delta^\star$. This investors does not have many willing counterparties in the market: as Proposition 5 will confirm formally, since λ is large, the asset is well allocated so the measure of owners with type below δ, $\Phi_1(\delta)$ is close to zero, its frictionless counterpart. Therefore, the expected waiting time of this investor is large. By contrast, a non-owners with type $\delta > \delta^\star$ can trade very quickly because, when λ is large $\Phi_1(\delta) \simeq F(\delta) - (1-s) > 0$ is bounded away from zero.

Expected prices. Let $p_q(\delta)$ denote the expected price paid or received for an investor with $q = 0$ or $q = 1$ units of the asset, respectively, and utility type δ. Lemma 3, below, formally establishes that the bilateral price in any meeting is increasing in the valuation of either the buyer or the seller, while the effect of increasing the meeting rate on bilateral prices is ambiguous.

Lemma 3 *For any $\delta \in [0,1]$, the expected prices paid by a buyer or received by a seller are*

$$p_0(\delta) = \theta_1 \Delta V(\delta) + \theta_0 \int_0^\delta \Delta V(\delta') \frac{d\Phi_1(\delta')}{\Phi_1(\delta)} \text{ and } p_1(\delta) = \theta_0 \Delta V(\delta) + \theta_1 \int_\delta^1 \Delta V(\delta') \frac{d\Phi_0(\delta')}{1-s-\Phi_0(\delta)},$$

respectively. The expected price $p_q(\delta)$ is increasing in δ for $q \in \{0,1\}$, but can be non-monotonic in λ.

Note that the expected transaction price is increasing in δ for 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 corresponds to the first term in the expressions for $p_q(\delta)$ in Lemma 3, which is clearly increasing in δ. 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 corresponds to the second term in the expressions in Lemma 3, which we establish is also increasing in δ in the Appendix.

Also note that the effect of λ on expected prices is ambiguous, which follows from the non-monotonic relationship between reservation values and meeting rates discussed above.

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 the asset in the presence of search frictions. However, even though search is random, the pattern of misallocation is not. In particular, misallocation is most common among investors who have valuations 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. These two symptoms are captured by the following measure of cumulative misallocation:

$$M(\delta) = \mathbf{1}_{\{\delta \leq \delta^*\}}\Phi_1(\delta) + \mathbf{1}_{\{\delta > \delta^*\}}\left[1 - s - \Phi_0(\delta)\right].$$

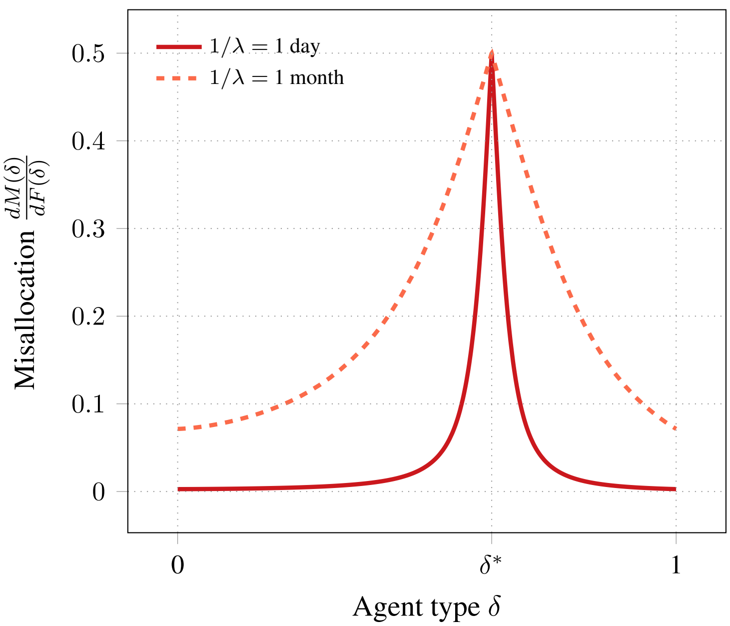
To measure misallocation at a specific utility type δ, one can simply calculate $\frac{dM(\delta)}{dF(\delta)}$, which is the fraction of type-δ investors whose asset holdings in the environment with search frictions differs from their holdings in the frictionless benchmark. Since the speed with which an owner trades is decreasing in δ, the likelihood that an investor owns an asset—and hence $d\Phi_1(\delta)$—is increasing in δ. By the same reasoning, an investor with a higher δ is less likely to be a non-owner. The next result follows immediately.

Lemma 4 *The function $\frac{dM(\delta)}{dF(\delta)}$ achieves a maximum at $\delta^\star$.*

While the slopes of $\eta_0(\delta)$ 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. In particular, we have noted in Figure 2 that the time to trade is very small small for owners with valuations below $\delta^\star$, and start increasing 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. Finally, note that this pattern of misallocation arises because trade is fully decentralized, and would not arise in a model in which investors trade through dealers. Indeed, in such an environment, all investors would contact dealers and trade with the same intensity, so that the measure of misallocation described above would be constant across types.

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.

FIGURE 3: Misallocation and search frictions



Notes. This figure plots misallocation as a function of the agents' type when meetings happen, on average, once every month (dashed) and once every day (solid). The model we use in this figure is otherwise the same as in Figure 1.

To see this, note that we can express trading volume as:

$$\vartheta = \int_0^1 \lambda \Phi_1(\delta) d\Phi_0(\delta) = \lambda \int_0^{\delta^\star} d\Phi_0(\delta)\Phi_1(\delta) + \lambda \int_{\delta^\star}^1 d\Phi_0(\delta)\Phi_1(\delta)$$
$$= \lambda \Phi_1(\delta^\star) \left[1 - s - \Phi_0(\delta^\star)\right] + \lambda \int_0^{\delta^\star} d\Phi_0(\delta)\Phi_1(\delta) + \lambda \int_{\delta^\star}^1 d\Phi_1(\delta) \left[1 - s - \Phi_0(\delta)\right], \quad (18)$$

where the second line follows from integration by parts. The first term is the volume that is due to trades between owners of type $[0, \delta^\star)$ and non-owners of type $[\delta^\star, 1]$. In a model with a frictionless exchange, or in a model in which agents trade through dealers who have access to a frictionless exchange, these would be the only trades taking place in equilibrium. Indeed, the price in the frictionless market would be set by the marginal type, so an investor below $\delta^\star$ would only find it optimal to sell, and an investor above $\delta^\star$ would only find it optimal to buy. In our environment with search frictions, there are additional trades because any investor may end up buying or selling depending on who she meets. These additional trades are captured by the second and third terms in

(18): the second term is the measure of trades between investors of type $\delta < \delta^\star$ who purchase the asset when they meet owners with lower types; and the third term is the measure of trades between investors of type $\delta > \delta^\star$ who sell the asset when they meet non-owners with higher types.

Lemma 5 *If $F(\delta)$ is continuous, then the steady state trading volume is*

$$\vartheta = \gamma s(1-s)\left[\left(1+\frac{\gamma}{\lambda}\right)\log\left(1+\frac{\lambda}{\gamma}\right)-1\right], \tag{19}$$

and is strictly increasing in λ and γ. If $F(\delta)$ has an atom, then ϑ is strictly in between the maximum and the minimum trading volume consistent with bilateral optimality.

To build some intuition about the result we note that, at any point in time, one can always rank investors in terms of their valuation using the ordinal index $q \in [0,1]$, and let the valuation of an investor of rank q using the quantile function, $\Delta(q) = \inf\{\delta' \in [0,1] : F(\delta') \geq q\}$. One can show that the trading volume ϑ obtains when agents trade according to their rank q, i.e. when non-owners buy from lower-ranked owners. Clearly, since $\Delta(q)$ is increasing, this trading pattern is consistent with bilateral optimality. When the distribution is continuous, $\Delta(q)$ is strictly increasing, and this trading pattern is the *only one* consistent with bilateral optimality, and so the trading volume must equal ϑ. By contrast, when the distribution has an atom, then $\Delta(q)$ has flat spots and there are other trading patterns consistent with bilateral optimality. For instance, if investors in the same flat spot always trade regardless of their rank, then the trading volume is strictly larger than ϑ, and if they never trade it is strictly smaller than ϑ.

Equation (19) offers some natural comparative statics: volume increases when investors find counterparties more quickly and when they change type more frequently.

Finally, note that $\lim_{\lambda\to\infty}\vartheta = \infty$. By contrast, in the frictionless benchmark volume is finite (see Lemma A.1 in the Appendix). Therefore, in our OTC market, intermediation activity can generate arbitrarily large excess volume, relative to the frictionless benchmark, as long as frictions are small enough.

Structure of the Trading Network. Given the emergence of several important OTC markets, a recent literature has emerged to study how the structure of the trading network in an OTC market affects the efficiency with which assets and information are transferred in these markets, as well as the vulnerability of these markets to negative shocks and contagion.[12] 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 values of δ and tends to cluster around the marginal type, $\delta^\star$. 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."

To illustrate this phenomenon, Figure 4 plots the intensity of sales by each owner-nonowner pair. It shows that agents with extreme utility types account for a very small fraction of total trades. For example, owners with low valuations 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 valuations, but these investors trade very slow, which is why they do not account for many trades. Only in the cluster of investors with near-marginal valuations 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, then most of the volume is accounted for by the infra-marginal trades of investors in the cluster near $\delta^\star$.

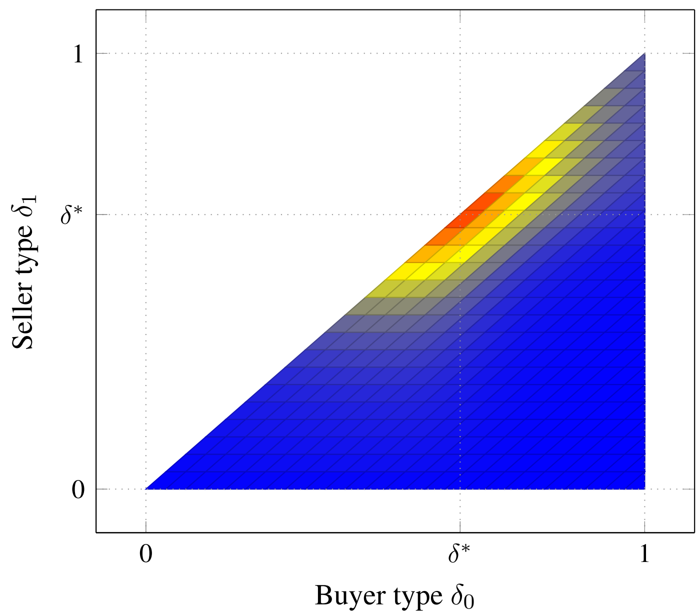
Proposition 3 *If $F(\delta)$ is differentiable then, as λ goes to infinity, for any $\varepsilon > 0$:*

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

Finally, we note that the volume created by investors in the cluster near $\delta^\star$ resemble intermediation activity. Indeed, when $\delta = \delta^\star$, when $F(\delta)$ is continuous, the selling intensity is exactly equal to

[12] 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).

FIGURE 4: Contribution to trading volume



Notes. This figure plots the contribution to volume $1_{\{\delta_1 \le \delta_0\}} \lambda d\Phi_0(\delta_0) d\Phi_1(\delta_1)$ as a function of the buyers' and sellers' types when meetings occur, on average, once a week. The model we use in this figure is the same as in Figure 1.

the buying intensity:

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

Hence, investors in the cluster near $\delta^\star$ are approximately equally likely to buy and sell assets.

The Distribution of Prices. In contrast to models of frictionless exchange, or models of decentralized exchange with only two types of agents, our model with many types generates price dispersion. In particular, each transaction in the network described above occurs at a price $P(\delta, \delta')$ in which the utility types $\delta \le \delta'$ are random variables drawn from a joint distribution described by

$$\text{Prob}\left[\{\delta' \le \delta_0\} \cap \{\delta \le \delta_1\} \right] = \frac{G(\delta_0, \delta_1)}{\vartheta},$$

where the denominator is the steady state equilibrium trading volume defined in equation (18) and

$$G(\delta_0,\delta_1) = \int\int \mathbf{1}_{\{\delta'\leq\delta_0\}}\mathbf{1}_{\{\delta\leq\delta'\wedge\delta_1\}}\lambda d\Phi_0(\delta')d\Phi_1(\delta)$$
$$= \lambda\Phi_1(\delta_1)(\Phi_0(\delta_0)-\Phi_0(\delta_1))^+ + \int \mathbf{1}_{\{\delta'\leq\delta_0\wedge\delta_1\}}\lambda\Phi_1(\delta')d\Phi_0(\delta').$$

In words, $G(\delta_0,\delta_1)$ 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 price levels 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 transaction price in the support of the distribution is given by*

$$\Delta V(1) - \Delta V(0) = \int_0^1 \frac{d\delta}{r+\gamma+\lambda\theta_1\left(1-s-\Phi_0(\delta)\right)+\lambda\theta_0\Phi_1(\delta)}.$$

The spread is decreasing in both γ and λ.

Intuitively, as 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 λ has a similar effect: the reservation values of any two investors narrows when they have access to other counterparties with greater frequencies.

4.3 Assessing the Model's Implications

Though reliable data from OTC markets has traditionally been difficult to find, since quantities and prices tend to be negotiated privately in these markets, 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, which suggests that it might be a useful structural framework to study these types of markets, both theoretically and quantitatively. Finally, in addition to highlighting several implications of our model that are consistent with existing empirical observations, we draw attention to a few predictions that emerge from our model that have *not* been tested in the data, but are clearly interesting and relevant.

Facts about OTC markets. One of the most important OTC markets is the interbank market for overnight loans—i.e., the *federal funds market*—where banks borrow and lend excess reserves in bilateral meetings throughout the day in attempt to achieve a certain end-of-day balance. In a recent study, Ashcraft and Duffie (2007) document several key patterns in this market that are consistent with the comparative static results with respect to δ that we reported in Section 4.1.

In the federal funds market, those banks with excess reserves have incentive to lend—in the language of our model, they have a "low δ"—while those banks with few reserves have incentive to borrow—in our model, they have a "high δ." In accordance with the results in Lemma 2, Ashcraft and Duffie (2007) find that a bank's probability of making a loan at a particular moment in time is increasing in the bank's balances, while the probability that a bank borrows funds at any point in time is decreasing in their balances. Moreover, consistent with the results in Lemma 3, 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.

As in our model, these trading patterns imply that assets in the fed funds market are re-allocated through chains of inframarginal trades: Ashcraft and Duffie (2007) cite a significant number of loans made by lenders who have balances in the lower deciles. Looking at the same market, Afonso and Lagos (2012a) lend further support to this pattern of re-allocation, as they 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. Afonso and Lagos (2012a) also document that this process of reallocation ultimately produces a trading network in which a small number of banks account for a large proportion of the overall trading volume.

The finding that OTC markets tend to exhibit a "core-periphery" network structure extends well beyond the fed funds market in the US. Li and Schürhoff (2012b) find a similar network structure in the interdealer market for municipal bonds, while Soramäki et al. (2007) show that there is a core-periphery structure for interbank flows across Fedwire, the large value transfer system operated by the Federal Reserve. Moreover, this network structure has also been documented within foreign interbank markets, such as Germany (Craig and Von Peter, 2014), Austria (Boss et al., 2004), and Brazil (Chang et al., 2008).

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; for example, Gavazza (2011b) studies the decentralized market for commercial aircraft and reports that transaction prices in his sample exhibit a coefficient of variation of approximately 0.8.

Taken together, the trading patterns cited above point strongly towards a model with both bilateral trade and many types of market participants. In particular, a model with either frictionless exchange or only two types of agents will predict a constant frequency of trades across investors, no role for the type of "endogenous intermediation" that Afonso and Lagos (2012a) highlight, and no price dispersion.

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 2 and 3 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 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 of small frictions, $\lambda \simeq \infty$. This is a relevant region of the parameter space to study since frictions are thought to be small in financial markets. We obtain two main findings. First, heterogeneity magnifies the price impact of search frictions. In particular, we show that the price impact is much larger when the distribution of valuation is continuous than when it is discrete. Second, search frictions have larger effects on price levels than on their dispersion. Therefore, quantifying the price impact of search frictions based on observed dispersion can 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 behavior of bilateral prices when $\lambda \simeq \infty$, we rely on a sequential representation of reservation values (see Hugonnier, 2012; Kiefer, 2012, for earlier work deriving related representations). From (13), it follows that:

Proposition 4 *Reservation values can be represented as*

$$\Delta V_t(\delta) = \mathbb{E}_{t,\delta}\left[\int_t^\infty e^{-rs}\hat{\delta}_s ds\right] \tag{20}$$

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

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

a function of the joint equilibrium distribution of types and asset holdings.

We call the process $\hat{\delta}_t$ "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 precisely, suppose that an investor's valuation at time t is δ. Then, during $[t, t+dt]$, the market valuation process can change for three reasons. With intensity γ, there is a change of type, in which case the new type is drawn according $F(\delta')$. With intensity $\lambda\theta_0\Phi_{1,t}(\delta)$, there is a purchase opportunity, at which point the the new type is drawn from the support $[0, \delta]$ according to the CDF $\Phi_{1,t}(\delta')/\Phi_{1,t}(\delta)$. Notice that the market valuation process creates transitions towards lower types in order to account for the option value of searching for sellers and bargain low buying prices. The intensity of these transitions is scaled down by the investor's bargaining power θ_0. Symmetrically, the market valuation process creates transition towards higher types in order to account for the option value of searching for buyers and bargain high prices: with intensity $\lambda\theta_1(1 - s - \Phi_0(\delta))$, there is a sale opportunities, at which point new type is drawn from the support $[\delta, 1]$ according to the CDF $(\Phi_{0,t}(\delta') - \Phi_{0,t}(\delta))/(1 - s - \Phi_{0,t}(\delta))$.

Note as well that, in a Walrasian market, the corresponding market valuation process is simply $\hat{\delta}_t = \delta^\star$: since investors can trade instantly at the constant price $\delta^\star/r$, the market valuation process must equal the valuation of the marginal investor $\delta^\star$ at all times and for all investors. In a decentralized market, the market- valuation process differs for two reasons. First, since meetings are not instantaneous, an owner must enjoy his private utility flow until he finds a counterparty. Second, conditional on finding a counterparty, investors are not trading against the same "marginal type"; the terms of trade are random and depend on the distributions of potential counterparties, $\Phi_{0,t}(\delta)$ and $\Phi_{1,t}(\delta)$.

Finally we note that one can rewrite (20) as

$$\Delta V_t(\delta) = \int_0^1 (\delta'/r)\, d\Psi_t(\delta'|\delta)$$

where

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

In words, the CDF $\Psi_t(\delta' \mid \delta)$ is a discounted occupation measure: it is the discounted amount of time that the market valuation process, $\hat{\delta}_t$, spends visiting a valuation less than δ'.

5.2 The frictionless limit

We start by establishing two intuitive but important results about the economy as $\lambda \to \infty$. First, the allocation becomes approximately frictionless, as an asset owned by an investor with $\delta < \delta^\star$ will be reallocated very quickly upstream to an investor with $\delta > \delta^\star$. Second, as a result, all reservation values converge to the frictionless price, $\delta^\star/r$.

Proposition 5 *As $\lambda \to \infty$, the asset 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 $\lim_{\lambda\to\infty} \Delta V(\delta) = \delta^\star/r$ for every $\delta \in [0,1]$.*

To understand why $\Delta V_t(\delta)$ converges to $\delta^\star/r$, consider the sequential representation (20). If $\hat{\delta} < \delta^\star$ we show in Appendix A.4 that:

$$\lim_{\lambda\to\infty} \lambda\left[1 - s - \Phi_0(\hat{\delta})\right] = +\infty \text{ and } \lim_{\lambda\to\infty} \lambda\Phi_1(\hat{\delta}) = \frac{\gamma s F(\hat{\delta})}{F(\delta^\star) - F\hat{\delta})}.$$

Hence, as $\lambda \to \infty$, the selling intensity goes to infinity but the buying intensity has a finite limit. Because the asset is almost perfectly allocated, it is much faster to find a buyer than a seller. As a result, the market valuation process, $\hat{\delta}_t$, drifts up very quickly, reflecting the fact that sellers can negotiate high prices when they meet buyers. Similarly, if $\hat{\delta}_t > \delta^\star$ then, as $\lambda \to \infty$, the selling intensity $\lambda\left[1 - s - \Phi_0(\hat{\delta}_t)\right]$ has a finite limit and the buying intensity $\lambda\Phi_1(\hat{\delta}_t)$ goes to infinity, so that $\hat{\delta}_t$ drifts down very quickly. Taken together, these observations imply that, as $\lambda \to \infty$, the market valuation process converges to $\delta^\star$. This implies the last statement of the Proposition: starting from any δ, the expected discounted present value of the market valuation process is approximately equal to $\delta^\star/r$.

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 $\delta^\star$. Our main result is:

Proposition 6 *Suppose that $F(\delta)$ is twice continuously differentiable with a strictly positive derivative. Then, as $\lambda \to \infty$, the discounted occupation measure of $\hat{x}_t \equiv \sqrt{\lambda}\left(\hat{\delta}_t - \delta^\star\right)$, converges pointwise to*

$$\Psi^\star(x) = \frac{\theta_0 g(x)}{\theta_0 g(x) + \theta_1 g(-x)},$$

where $x \in (-\infty, \infty)$ and $g(x)$ is the positive solution to $g(x)^2 - g(x)F'(\delta^\star)x - \gamma s(1-s) = 0$. As a result, reservation values admit the approximation:

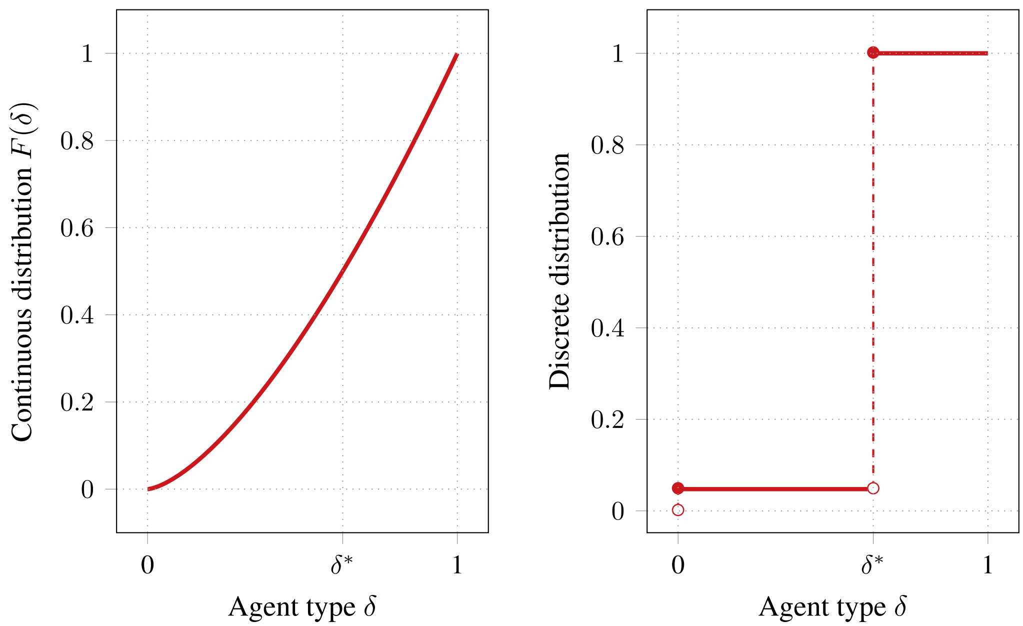
$$\begin{aligned}\Delta V(\delta) &= \frac{\delta^\star}{r} + \int_{\mathbb{R}} (x/r)d\Psi^\star(x)\frac{1}{\sqrt{\lambda}} + o\left(\frac{1}{\sqrt{\lambda}}\right) \\ &= \frac{\delta^\star}{r} + \frac{\pi}{F'(\delta^\star)}\left(\frac{1}{2} - \theta_0\right)\left(\frac{\gamma s(1-s)}{\theta_0\theta_1}\right)^{\frac{1}{2}}\frac{1}{\sqrt{\lambda}} + o\left(\frac{1}{\sqrt{\lambda}}\right)\end{aligned}$$

The idea of the proof is to center the market valuation process around its limit, $\delta^\star$, and to scale it up by its speed of convergence, which turns out to be $\sqrt{\lambda}$. This delivers a new process, $\hat{x}_t$, whose non-degenerate limit distribution determines the second term of the asymptotic expansion.

To understand why the speed of convergence of the market valuation process is in order $\sqrt{\lambda}$, recall that as $\lambda \to \infty$, the asset is almost perfectly allocated. As a result, it becomes harder and harder to find willing trading counterparties as $\hat{\delta}_t$ approaches $\delta^\star$: buyers become very scarce as $\hat{\delta}_t \uparrow \delta^\star$, and sellers become very scarce as $\hat{\delta}_t \downarrow \delta^\star$. Consistent with this observation, we show in Appendix A.4 that buying and selling intensities in the cluster, $\lambda\Phi_1(\delta^\star) = \lambda\left[1 - s - \Phi_0(\delta^\star)\right]$, both converge to infinity slowly at a speed in order $\sqrt{\lambda}$ instead of λ. These trading intensity pin down the speed of convergence of the market valuation process in a neighborhood of $\delta^\star$.

The first term of the expansion, $\delta^\star/r$, follows from Proposition 5: all reservation values converge to the Walrasian price. The second term of the expansion determines how reservation values and prices deviate from their Walrasian limit. We argue that this term depends on several key features of the market.

FIGURE 5: Distributions



Notes. The left panel plots the continuous cumulative distribution of types $F(\delta)$ that we use throughout our examples whereas the right hand side plots a two point distribution that is constructed to have the same mean and to induce the same marginal agent.

The first key feature of the market that affects reservation values 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, θ_0 and θ_1, which determine whether the asset is sold at a discount or at a premium: if $\theta_0 > \frac{1}{2}$, then the asset is sold at a discount in all bilateral meetings, relative to the Walrasian price, and vice versa if $\theta_0 < \frac{1}{2}$.

Finally, the third feature of the market that matters is the size of gains from trade in a typical match, which depends on the amount of heterogeneity in valuations between investors in the cluster near $\delta^\star$. If $F'(\delta^\star)$ is very small, then investors' valuations are not highly concentrated near $\delta^\star$, the gains from trade are large, and prices deviate significantly from their Walrasian limit. Alternatively, if $F'(\delta^\star)$ is very large, then investors' valuations are highly concentrated around $\delta^\star$, the gains from trade are small,and prices remain closer to their Walrasian limit.

To further emphasize the role of heterogeneity, consider heuristically what happens when the continuous distribution $F(\delta)$ becomes closer and closer to a discrete distribution. Then, at $\delta^\star$, the CDF $F(\delta)$ will become closer and closer to a step function, which is vertical at the marginal type: $F'(\delta^\star) \simeq \infty$. As we argued in the previous paragraph, the asymptotic expansion shows that the deviation from the Walrasian price will be very small. This can be confirmed directly by working out the asymptotic expansion with discrete instead of continuous types.

Proposition 7 *When $F(\delta)$ is discrete, generically, bilateral prices converge to their frictionless counterpart in order $1/\lambda$.*

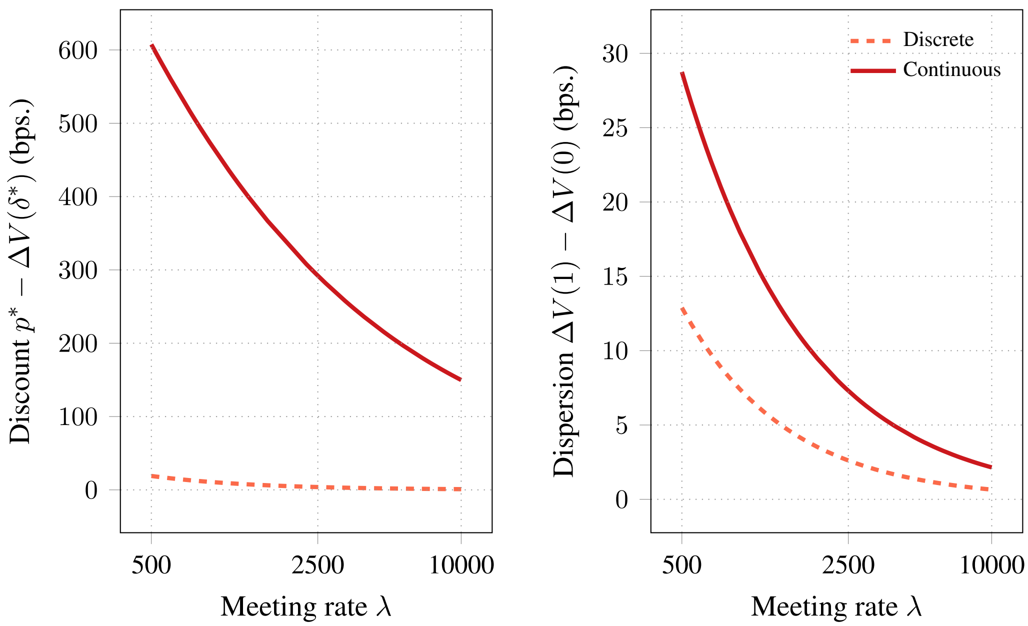
Taken together, Proposition 6 and 7 show that existing results based on discrete-type models can substantially underestimate the price impact of frictions (see Vayanos and Weill, 2008; Weill, 2008; Praz, 2013, for asymptotic approximations in discrete type models). We illustrate this finding on the left panel of Figure 6, in which we compare the reservation value $\Delta V(\delta^\star)$ for a continuous vs. a two types discrete distribution, with identical marginal and average investor. One sees that, when investors find counterparties twice a day on average ($\lambda = 500$), the price discount is above 6% with a continuous distribution, and far below 1% with the corresponding discrete distribution. When investors find counterparties ten times per day on average ($\lambda = 2500$), the discount is between 2% and 3% with continuous distribution, but it is now indistinguishable from zero with a discrete distribution.

Price dispersion near the frictionless limit. A second implication of Proposition 6 is that, to a first-order approximation, there is no price dispersion: the coefficient multiplying the converge at rate, $1/\sqrt{\lambda}$, is independent of δ. To obtain further results about price dispersion, one thus needs to work out higher order terms.

Proposition 8 *Suppose that $F(\delta)$ is twice continuously differentiable with a strictly positive derivative. Then, price dispersion admits the asymptotic expansion:*

$$\Delta V(1) - \Delta V(0) = \int_0^1 \sigma(\delta)\, d\delta = \frac{1}{2\theta_0\theta_1 F'(\delta^\star)} \frac{\log(\lambda)}{\lambda} + O\left(\frac{1}{\lambda}\right).$$

FIGURE 6: Continuous vs. discrete distribution



Notes. This figures plots the price discount relative to the Walrasian equilibrium (left panel) and price dispersion (right panel) as functions of the meeting rate for the base case model of Figure 1 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 he base case model.

With a discrete distribution, the rate of convergence is generically in order $1/\lambda$*.*

The finding that price dispersion vanishes at a rate $\log(\lambda)/\lambda$, which is faster than the price discount or premium has an important empirical implication: inferring the impact of the search frictions based on the observable level of price dispersion can be misleading. In particular, frictions can have a very small impact on dispersion and, yet, have large impact on the price discount or premium.

This finding is illustrated in Figure 6. Comparing the left and the right panel, one sees clearly that price dispersion converges to zero much faster than the price discount. For instance, when investors meet counterparties twice a day on average ($\lambda = 500$), then the price discount is about 6%, but price dispersion is about 30 basis points, twenty times smaller. One also see in the picture that price dispersion is larger with a continuous than with a discrete distribution. We confirm this in Appendix A.5 where we show that, with a discrete distribution, price dispersion converges to zero generically at a speed $1/\lambda$.

5.3 The welfare cost of frictions near the frictionless limit

Finally, one may wonder what is the asymptotic impact of frictions on welfare. To answer this question we note that the welfare cost of misallocation can be written:

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

The two terms account for the two type of misallocations 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. The equality follows from integration by part. We obtain:

Proposition 9 *Suppose that the distribution of utility type, $F(\delta)$, is twice continuously differentiable with a strictly positive derivative. Then, the welfare cost of misallocation admits the approximation:*

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

With a discrete distribution, the rate of convergence is generically in order $1/\lambda$.

The proposition shows that misallocation also has a larger welfare impact when the distribution is continuous, than when it is discrete. It also shows that the welfare cost of frictions may be accurately measured by the observed amount of price dispersion.

6 Conclusion

To be added.

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.

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.

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.

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.

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, 2012a.

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

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.

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 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. URL http://ejde.math.txstate.edu.

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.

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 study the total volume of trade that occurs at each instant in the frictionless benchmark equilibrium of Section 2.2, which we denote by $\vartheta^\star$. 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 benchmark equilibrium of Section 2.2, the minimum volume is:*

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

Proof of Lemma A.1. 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

Proof of Lemma 1. To establish the first part we need to prove the existence and uniqueness of a fixed point of the map

$$T_t[f](\delta) = \mathbb{E}_t\Bigg[\int_t^\tau e^{-r(u-t)}\delta du + e^{-r(\tau-t)}\Bigg(f_\tau(\delta) + \mathbf{1}_{\{\tau=\tau_\gamma\}}\int_0^1 (f_\tau(\delta') - f_\tau(\delta))dF(\delta') + \mathbf{1}_{\{\tau=\tau_0\}}\theta_1 \int_0^1 \big(f_\tau(\delta') - f_\tau(\delta)\big)^+ \frac{d\Phi_{0,\tau}(\delta')}{1-s} - \mathbf{1}_{\{\tau=\tau_1\}}\theta_0 \int_0^1 \big(f_\tau(\delta) - f_\tau(\delta')\big)^+ \frac{d\Phi_{1,\tau}(\delta')}{s}\Bigg)\Bigg] \tag{21}$$

in the Banach space $\mathcal{X}$ of uniformly bounded functions from $\mathcal{S} \equiv \mathbb{R}^+ \times [0,1]$ to $\mathbb{R}$ equipped with the usual sup norm

$$\|f\|_\infty = \sup_{(t,\delta)\in\mathcal{S}} |f_t(\delta)|.$$

Since the cumulative distribution functions $F(\delta)$ and $\Phi_{q,t}(\delta)$ are non decreasing and bounded it is easily seen that that T maps $\mathcal{X}$ into itself and integrating on both sides of equation (21) with respect to the conditional distribution

$$\mathbb{P}[\tau = \tau_\gamma > u | \tau > t] = \gamma e^{-(\gamma+\lambda)(u-t)}$$
$$\mathbb{P}[\tau = \tau_q > u | \tau > t] = \lambda[qs + (1-q)(1-s)]e^{-(\gamma+\lambda)(u-t)}$$

of the stopping time τ shows that

$$\begin{aligned} T_t[f](\delta) = \int_t^\infty e^{-\rho(u-t)} \Bigg(\delta + \lambda f_u(\delta) + \gamma \int_0^1 f_u(\delta') dF(\delta') \\ + \lambda\theta_1 \int_0^1 \left(f_u(\delta') - f_u(\delta)\right)^+ d\Phi_{0,u}(\delta') \\ - \lambda\theta_0 \int_0^1 \left(f_u(\delta) - f_u(\delta')\right)^+ d\Phi_{1,u}(\delta') \Bigg) du, \end{aligned} \tag{22}$$

where $\rho \equiv r + \lambda + \gamma$. Equivalently,

$$\begin{aligned} T_t[f](\delta) = \int_t^\infty e^{-\rho(u-t)} \Bigg(\delta + \eta f_u(\delta) + \gamma \int_0^1 f_u(\delta') dF(\delta') \\ + \lambda\theta_1 \int_0^1 \max\left(f_u(\delta'), f_u(\delta)\right) d\Phi_{0,u}(\delta') \\ + \lambda\theta_0 \int_0^1 \min\left(f_u(\delta'), f_u(\delta)\right) d\Phi_{1,u}(\delta') \Bigg) du \end{aligned} \tag{23}$$

where $\eta \equiv \lambda(1 - \theta_0 s - \theta_1(1-s)) > 0$. Now, a direct calculation shows that we have

$$\sup_{(u,\delta,\delta') \in \mathcal{S} \times [0,1]} \left| \min\left(f_u(\delta'), f_u(\delta)\right) - \min\left(g_u(\delta'), g_u(\delta)\right) \right| \leq \|f - g\|_\infty$$
$$\sup_{(u,\delta,\delta') \in \mathcal{S} \times [0,1]} \left| \max\left(f_u(\delta'), f_u(\delta)\right) - \max\left(g_u(\delta'), g_u(\delta)\right) \right| \leq \|f - g\|_\infty$$

and combining these bounds with (23) and the nonnegativity of η we deduce that

$$\|T[f] - T[g]\|_\infty \leq \left(1 - \frac{r}{\rho}\right) \|f - g\|_\infty .$$

Since by definition $\rho > r$ this shows that T is a contraction on the Banach space $\mathcal{X}$ and the existence of a unique fixed point follows from the Banach fixed point theorem.

Denote by $\overline{\mathcal{X}} \subseteq \mathcal{X}$ the subset of functions $f \in \mathcal{X}$ that are nonnegative, strictly increasing in δ and absolutely continuous on S (see Sremr (2010) for a precise definition) with a uniformly bounded space derivative; and let $f \in \overline{\mathcal{X}}$ be fixed. Since the function $f_t(\delta)$ is nonnegative it follows from (23) that the function $T_t[f](\delta)$ is nonnegative. On the other hand, a direct calculation shows that the map

$$\delta \longmapsto \{\min, \max\} \left(f_u(\delta'), f_u(\delta)\right)$$

is increasing for any $(u, \delta') \in \mathcal{S}$. Using this property we deduce from (23) that

$$T_t[f](\delta') - T_t[f](\delta) \geq \frac{\delta' - \delta}{\rho} + \eta \int_t^\infty e^{-\rho(u-t)} \left(f_u(\delta') - f_u(\delta)\right) du \geq \frac{\delta' - \delta}{\rho}$$

for all $\delta \leq \delta'$ and it follows that the function $T_t[f](\delta)$ is strictly increasing in δ. Since f is absolutely continuous and strictly increasing in δ with a bounded derivative it follows by the fundamental theorem of calculus we have

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

for all $t \geq 0$, almost every $\delta, \delta' \in [0,1]^2$ and some nonnegative function $f' \in \mathcal{X}$. Substituting this identity into (22) and changing the order of integration shows that

$$T_t[f](\delta) = \int_t^\infty e^{-\rho(u-t)} \Bigg(\delta + (\gamma + \lambda) f_u(\delta) - \int_0^\delta f_u'(x)(\gamma F(x) + \lambda\theta_0 \Phi_{1,u}(x))dx + \int_\delta^1 f_u'(x)(\gamma(1 - F(x)) + \lambda\theta_1(1 - s - \Phi_{0,u}(x)))dx\Bigg) du \tag{24}$$

and it now follows from Sremr (2010, Theorem 3.1) that $T_t[f](\delta)$ is absolutely continuous on S. Finally, using the above expression and Lebesgue's differentiation theorem shows that

$$T_t'[f](\delta) = \int_t^\infty e^{-\rho(u-t)}(1 + (\rho - R_u(\delta))f_u'(\delta))du.$$

and it follows that $\|T'[f]\|_\infty \leq \text{const} \cdot \|f'\|_\infty$. Since $f \in \overline{\mathcal{X}}$ was arbitrary the above results show that T maps $\overline{\mathcal{X}}$ into itself and it follows that its unique fixed point necessarily lies in $\overline{\mathcal{X}}$. ■

Lemma A.2 *For any fixed $\delta \in [0,1]$ the unique solution to (10) that is both absolutely continuous in time and uniformly bounded is explicitly given by*

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

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

Proof. Fix an arbitrary $\delta \in [0,1]$ and assume that $\sigma_t(\delta)$ is a uniformly bounded solution solution to (10) 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 1. Let the local surplus $\sigma_t(\delta)$ be as above and consider the absolutely continuous function defined by

$$\begin{aligned} f_t(\delta) = \int_t^\infty e^{-r(u-t)}\Bigg(\delta - \int_0^\delta \sigma_u(\delta')[\gamma F(\delta') + \lambda\theta_0\Phi_{1,u}(\delta')]d\delta' \\ + \int_\delta^1 \sigma_u(\delta')[\gamma(1-F(\delta')) + \lambda\theta_1(1-s-\Phi_{0,u}(\delta'))]d\delta'\Bigg)du. \end{aligned} \tag{25}$$

Using the uniform boundedness of $\sigma_t(\delta)$, $F(\delta)$ and $\Phi_{q,t}(\delta)$ we deduce that $|f_t(\delta)| \le \text{const}\cdot\delta$ and it follows that $f \in \mathcal{X}$. On the other hand, Lebesgue's differentiation theorem shows that the function $f_t(\delta)$ is *almost* everywhere differentiable in both the time and the space variable with

$$\begin{aligned} \dot f_t(\delta) = rf_t(\delta) - \delta + \int_0^\delta \sigma_t(\delta')(\gamma F(\delta') + \lambda\theta_0\Phi_{1,t}(\delta'))d\delta' \\ - \int_\delta^1 \sigma_t(\delta')(\gamma(1-F(\delta')) + \lambda\theta_1(1-s-\Phi_{0,t}(\delta')))d\delta' \end{aligned}$$

and

$$\begin{aligned} f_t'(\delta) &= \int_t^\infty e^{-r(u-t)}\left(1 - \sigma_u(\delta)[\gamma + \lambda\theta_1(1-s-\Phi_{0,u}(\delta)) + \lambda\theta_0\Phi_{1,u}(\delta)]\right)du \\ &= \int_t^\infty e^{-r(u-t)}\left(r\sigma_u(\delta) - \dot\sigma_u(\delta)\right)du = \sigma_t(\delta) \end{aligned}$$

where the second equality follows from (10) and the third follows from integration by parts and the uniform boundedness of the local surplus. In particular, the fundamental theorem of calculus implies that

$$f_t(\delta') - f_t(\delta) = \int_\delta^{\delta'} \sigma_t(\delta'')d\delta'', \qquad (\delta, \delta') \in [0,1]^2 \tag{26}$$

and, since the local surplus is strictly positive, we have that $f_t(\delta)$ is strictly increasing in δ. Using this property in conjunction with (26) and integrating by parts on the right hand side of (25) gives

$$\begin{aligned}\dot{f}_t(\delta) &= rf_t(\delta) - \delta - \gamma \int_0^1 (f_t(\delta') - f_t(\delta))dF(\delta') \\ &\quad - \lambda\theta_1 \int_0^1 (f_t(\delta') - f_t(\delta))^+ d\Phi_{0,t}(\delta') + \lambda\theta_0 \int_0^1 (f_t(\delta) - f_t(\delta'))^+ d\Phi_{1,t}(\delta')\end{aligned}$$

and combining this expression with the definition of the constant ρ and an application of the integration by parts formula shows that

$$\begin{aligned}f_t(\delta) - e^{-\rho(H-t)} f_H(\delta) &= \int_t^H e^{-\rho(u-t)}(\rho f_u(\delta) - \dot{f}_u(\delta))du \\ &= \int_t^H e^{-\rho(u-t)} \Bigg(\delta + \lambda f_u(\delta) + \gamma \int_0^1 f_t(\delta')dF(\delta') \\ &\qquad + \lambda\theta_1 \int_0^1 (f_t(\delta') - f_t(\delta))^+ d\Phi_{0,t}(\delta') \\ &\qquad - \lambda\theta_0 \int_0^1 (f_t(\delta) - f_t(\delta'))^+ d\Phi_{1,t}(\delta')\Bigg) du.\end{aligned}$$

for any $0 \le H < \infty$. Taking limits as on both sides and using the dominated convergence theorem in conjunction with the strict positivity of ρ and the uniform boundedness of f then gives

$$\begin{aligned}f_t(\delta) = \int_t^\infty e^{-\rho(u-t)} \Bigg(&\delta + \lambda f_u(\delta) + \gamma \int_0^1 f_u(\delta')dF(\delta') \\ &+ \lambda\theta_1 \int_0^1 \left(f_u(\delta') - f_u(\delta)\right)^+ d\Phi_{0,u}(\delta') \\ &- \lambda\theta_0 \int_0^1 \left(f_u(\delta) - f_u(\delta')\right)^+ d\Phi_{1,u}(\delta')\Bigg) du.\end{aligned}$$

Comparing this expression to (22) we conclude that $f_t(\delta) = T_t[f](\delta) \in \mathcal{X}$ and the desired result now follows from the uniqueness part of Lemma 1. The comparative static results follow directly from Lemma A.3 below, we omit the details. ■

Lemma A.3 *The function $\Delta V_t(\delta)$ is increasing in θ_1 and decreasing in r, θ_0, $F(\delta)$ and $\Phi_{q,t}(\delta)$*

Proof. By Lemma 1 we have that $\Delta V_t(\delta)$ is the unique fixed point of the contraction mapping $T : \overline{\mathcal{X}} \to \overline{\mathcal{X}}$ defined by (23) and, by inspection, this mapping is increasing in f and decreasing in r. On the other hand, it follows from equation (24) 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.4 below. ■

Lemma A.4 *Let $\mathcal{C} \subseteq \mathcal{X}$ and assume that the mapping $A[\cdot;\alpha] : \mathcal{C} \to \mathcal{C}$ 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

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

Iterating this relation gives

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

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 2. 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 standard Ricatti equation with constant coefficients whose unique solution can be found in any differential equation textbook, see for example Reid (1972).

Let us now turn to the convergence part. Using equations (1) and (2) together with the definition of the functions $\Lambda(\delta)$ and $\Phi_q(\delta)$ shows that the term

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

that appears in the denumerator of (17) is nonnegative for all $\delta \in [0,1]$. Since $\lambda\Lambda(\delta) > 0$ this implies that the function $|\Phi_{1,t}(\delta) - \Phi_1(\delta)|$ is decreasing in time with

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

and the proof is complete. ■

Lemma A.5 *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.5. A direct calculation shows that

$$\frac{\partial\Phi_1(\delta)}{\partial s} = \frac{\Phi_1(\delta) + \phi F(\delta)}{\Lambda(\delta)} \tag{27}$$

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

$$\frac{\partial\Phi_1(\delta)}{\partial\phi} = \frac{sF(\delta) - \Phi_1(\delta)}{\Lambda(\delta)} = \frac{s(1-s)F(\delta)(1-F(\delta))}{(\phi + \Phi_1(\delta) + (1-s)(1-F(\delta)))\Lambda(\delta)} \tag{28}$$

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 (28) 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 1. The result follows directly from Lemma A.5. ■

Proof of Theorem 1. The result follows directly from the definition, Lemma 1 and Proposition 2. We omit the details. ■

A.3 Proofs Omitted in Section 4

The following lemma follows immediately from the equation defining the steady state distribution of utility types among asset owners:

Lemma A.6 *The steady state distributions of types satisfy* $\Phi_1(\delta) = F(\delta) - \Phi_0(\delta) = \ell(F(\delta))$ *where the bounded function*

$$\ell(x) \equiv -\frac{1}{2}(1 - s + \gamma/\lambda - x) + \frac{1}{2}\sqrt{(1 - s + \gamma/\lambda - x)^2 + 4s(\gamma/\lambda)x} \tag{29}$$

is the unique positive solution to $\ell^2 + (1 - s + \gamma/\lambda - x)\ell - s(\gamma/\lambda)x = 0$

Proof of Lemma 2. 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

$$\begin{aligned}\eta_q(\delta) &= \mathbb{E}\left[\int_0^\sigma e^{-\int_0^t \lambda_q(\delta_s)ds} dt + e^{-\int_0^\sigma \lambda_q(\delta_s)ds}\eta_q(\delta_\sigma)\right] \\ &= \mathbb{E}\left[\int_0^\sigma e^{-\lambda_q(\delta)t} dt + e^{-\lambda_q(\delta)\sigma}\eta_q(\delta_\sigma)\right] = \frac{1}{\gamma + \lambda_q(\delta)}\left(1 + \gamma\int_0^1 \eta_q(\delta')dF(\delta')\right)\end{aligned} \tag{30}$$

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 (30) 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 (30) gives

$$\eta_q(\delta) = \frac{1}{\gamma + \lambda_q(\delta)} \left(1 - \gamma \int_0^1 \frac{dF(\delta')}{\gamma + \lambda_q(\delta')}\right)^{-1}. \tag{31}$$

Now assume that the cumulative distribution function $F(\delta)$ is continuous. Combining Proposition 2 with Lemma A.6 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 (29) 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 (31) 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 (31) 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.5 below) this shows that $\lambda_q(\delta)$ is concave in the meeting rate and it follows that

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

where the second equality results from Lemma A.5 and the fact that

$$\begin{aligned}\lim_{\lambda\to\infty} \lambda \frac{\partial \Phi_1(\delta)}{\partial \lambda} &= -\lim_{\phi\to 0} \phi \frac{\partial \Phi_1(\delta)}{\partial \phi} \\ &= -\lim_{\phi\to 0} \frac{\phi s(1-s)F(\delta)(1-F(\delta))}{(\phi + \Phi_1(\delta) + (1-s)(1-F(\delta)))\Lambda(\delta)} = 0\end{aligned}$$

due to (28). This shows that $\lambda_q(\delta)$ is an increasing function of λ and the desired result now follows from (31) 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.5 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 (27) 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 (31) by noting that the function $F(\delta)$ does not depend on s. ■

Proof of Lemma 3. Since the reservation value function

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

is always absolutely continuous in δ it follows from an integration by parts that the expected buyer price can be written as

$$\begin{aligned} p_0(\delta) &= \theta_1 \Delta V(\delta) + \theta_0 \left(\Delta V(0) + \int_0^\delta \sigma(\delta') \left[1 - \frac{\Phi_1(\delta')}{\Phi_1(\delta)} \right] d\delta' \right) \\ &= \Delta V(\delta) - \theta_0 \int_0^\delta \sigma(\delta') \frac{\Phi_1(\delta')}{\Phi_1(\delta)} d\delta' = \Delta V(0) + \int_0^\delta \sigma(\delta') \left[1 - \theta_0 \frac{\Phi_1(\delta')}{\Phi_1(\delta)} \right] d\delta'. \end{aligned}$$

and the required monotonicity in δ follows by observing that the local surplus $\sigma(\delta')$ is nonnegative and that the function

$$1 - \theta_0 \frac{\Phi_1(\delta')}{\Phi_1(\delta)}$$

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 in δ follows by observing that the function

$$1 - \theta_1 \frac{1 - s - \Phi_0(\delta')}{1 - s - \Phi_0(\delta)}$$

is decreasing in δ. ■

Proof of Lemma 4. Using the same notation as in the proof of Lemma 2 we have that

$$\frac{\partial M_q(\delta)}{\partial F(\delta)} = \frac{\partial((1-q)F(\delta) + (2q-1)\ell(F(\delta))}{\partial F(\delta)} = 1 - q + (2q-1)\ell'(F(\delta))$$

and the desired result follows by noting that the function $\ell(x)$, which is defined by equation (29), is convex. Indeed, a direct calculation shows that we have

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

and the required conclusion follows from the fact that $s \in [0,1]$ and $\phi > 0$. ■

Lemma A.7 *Let $H : [0,1] \to [0,1]$ be a cumulative probability distribution function. If U is uniformly distributed then the random variable* $\inf\{x \in [0,1] : H(x) \geq U\}$ *is distributed according to $H(\cdot)$.*

Proof of Lemma A.7. 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 5. 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.6 one obtains that the functions $G_1(x)$ and $G_0(x)$ satisfy the identity:

$$G_1(x) = \frac{\phi G_0(x)}{1 + \phi - G_0(x)} \tag{32}$$

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 (32) can be written equivalently as:

$$G_1(G_0^{-1}(y)) = \frac{\phi y}{1 + \phi - y}. \tag{33}$$

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.7) 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

$$\begin{aligned}\frac{\vartheta(\pi)}{\lambda s(1-s)} &= \mathbb{P}\left[G_0^{-1}(u_0) > G_1^{-1}(u_1)\right] = \mathbb{P}\left[u_1 < G_1(G_0^{-1}(u_0))\right] \\ &= \mathbb{E}\left[G_1(G_0^{-1}(u_0))\right] = \int_0^1 G_1(G_0^{-1}(x))dx = \int_0^1 \frac{\phi x}{1+\phi-x}dx = \frac{\vartheta^\star}{\lambda s(1-s)}\end{aligned}$$

where we used formula (33) 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

$$\begin{aligned}\frac{\vartheta(0)}{\lambda s(1-s)} &= \mathbb{P}\left[\Delta(G_0^{-1}(u_0)) > \Delta(G_1^{-1}(u_1))\right] \\ &< \mathbb{P}\left[G_0^{-1}(u_0) > G_1^{-1}(u_1)\right] = \frac{\vartheta^\star}{\lambda s(1-s)} \\ &= \mathbb{P}\left[G_0^{-1}(u_0) \geq G_1^{-1}(u_1)\right] < \mathbb{P}\left[\Delta(G_0^{-1}(u_0)) \geq \Delta(G_1^{-1}(u_1))\right] = \frac{\vartheta(1)}{\lambda s(1-s)}\end{aligned}$$

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. ■

A.4 Allocation near the frictionless limit

From equation (16), it follows that:

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

Since $\Phi_0(\delta) + \Phi_1(\delta) = \Phi_0^\star(\delta) + \Phi_1^\star(\delta)$, it thus follows that $\lim_{\lambda\to\infty} \Phi_0(\delta) = \Phi_0^\star(\delta)$ as well.

Next, we derive the rate at which $\Phi_1(\delta)$ and $1-s-\Phi_0(\delta)$ converge to their frictionless limits. Recall the inflow-outflow steady-state equation:

$$\gamma F(\delta)\left[s - \Phi_1(\delta)\right] = \gamma \Phi_1(\delta)\left[1 - F(\delta)\right] + \lambda \Phi_1(\delta)\left[1 - s - \Phi_0(\delta)\right]. \tag{34}$$

We have just shown that when $\delta < \delta^\star$, $\Phi_1(\delta) \to 0$ and $1 - s - \Phi_0(\delta) \to F(\delta^\star) - F(\delta)$. Therefore, for $\delta < \delta^\star$, the asymptotic distribution of owners is

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

By symmetry, for $\delta > \delta^\star$, we obtain

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

Finally, we derive what happens at the point $\delta^\star$. For this we observe that $1 - s - \Phi_0(\delta^\star) = 1 - s - F(\delta^\star) + \Phi_1(\delta^\star) = \Phi_1(\delta^\star)$, since $1 - s = F(\delta^\star)$. Plugging this back into (34) shows that $\lambda \Phi_1(\delta^\star)^2$ has a non-zero limit, which can be written as:

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

A.5 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

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

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

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

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 Proposition 5 and 6. After evaluating (12), at $\delta = \delta^\star$ and at the steady state, we obtain that

$$r\Delta V(\delta^\star) = \delta^\star - \frac{D(\lambda)}{\sqrt{\lambda}} + \frac{P(\lambda)}{\sqrt{\lambda}},$$

where

$$D(\lambda) = \sqrt{\lambda}\int_0^{\delta^\star} \sigma(\delta)\left\{\gamma F(\delta) + \lambda\theta_0\Phi_1(\delta)\right\} d\delta = \sqrt{\lambda}\int_0^{\delta^\star} \frac{\gamma F(\delta) + \lambda\theta_0\Phi_1(\delta)}{r + \gamma + \lambda\theta_0\Phi_1(\delta) + \lambda\theta_1\left[1 - s - \Phi_0(\delta)\right]}$$

$$\begin{aligned} P(\lambda) &= \sqrt{\lambda}\int_{\delta^\star}^1 \sigma(\delta)\left\{\gamma\left[1 - F(\delta)\right] + \lambda\theta_1\left[1 - s - \Phi_0(\delta)\right]\right\} d\delta \\ &= \sqrt{\lambda}\int_{\delta^\star}^1 \frac{\gamma\left[1 - F(\delta)\right] + \lambda\theta_1\left[1 - s - \Phi_0(\delta)\right]}{r + \gamma + \lambda\theta_0\Phi_1(\delta) + \lambda\theta_1\left[1 - s - \Phi_0(\delta)\right]} d\delta. \end{aligned}$$

To study the asymptotic behavior of $D(\lambda)$ and $P(\lambda)$, we make the change of variable $x = \sqrt{\lambda}\,(\delta - \delta^\star)$. We obtain:

$$\begin{aligned} D(\lambda) &= \int_{-\delta^\star\sqrt{\lambda}}^0 \frac{\frac{\gamma}{\sqrt{\lambda}}F\left(\delta^\star + x/\sqrt{\lambda}\right) + \theta_0 g_1(x)}{\frac{r+\gamma}{\sqrt{\lambda}} + \theta_0 g_1(x) + \theta_1 g_0(x)} dx \\ P(\lambda) &= \int_0^{\delta^\star\sqrt{\lambda}} \frac{\frac{\gamma}{\sqrt{\lambda}}\left[1 - F\left(\delta^\star + x/\sqrt{\lambda}\right)\right] + \theta_1 g_0(x)}{\frac{r+\gamma}{\sqrt{\lambda}} + \theta_0 g_1(x) + \theta_1 g_0(x)} dx \end{aligned}$$

where

$$g_1(x) \equiv \sqrt{\lambda}\Phi_1\left(\delta^\star + x/\sqrt{\lambda}\right) \text{ and } g_0(x) \equiv \sqrt{\lambda}\left[1 - s - \Phi_0\left(\delta^\star + x/\sqrt{\lambda}\right)\right].$$

To apply the Dominated Convergence Theorem, we first study the limit of the integrand. For this we use the following Lemma:

Lemma A.8 *Assume that $F(\delta)$ is differentiable at $\delta^\star$ and denote by $g(x)$ the positive solution to the quadratic equation*

$$g^2 - gF'(\delta^\star)x - \gamma s(1-s) = 0. \tag{35}$$

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

Proof. Equation (15), evaluated at the steady state, implies that $g_1(x)$ is the unique positive solution to the quadratic equation:

$$g^2 + \left[\frac{\gamma}{\sqrt{\lambda}} + \sqrt{\lambda}\left(F(\delta^\star) - F\left(\delta^\star + x/\sqrt{\lambda}\right)\right)\right] g - \gamma s F\left(\delta^\star + x/\sqrt{\lambda}\right) = 0. \tag{36}$$

Because this quadratic equation is positive at $g = 1$ we have that $g_1(x) \leq 1$ and it follows that $g_1(x)$ has a well-defined limit as $\lambda \to \infty$. Going to the limit (36) shows that this limit is given by the unique positive solution to (35). Next, we note that

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

Plugging this into (36) then shows that $g_0(x)$ is the unique positive solution to the quadratic equation

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

and the result follows from the same arguments as before. ■

Lemma A.9 *Suppose that $F'(\delta)$ is continuous and strictly positive. Then, there exists a constant some $K \geq 0$ such that:*

$$g_1(x) \leq -\frac{K}{x} \quad \text{if } x \leq 0, \quad \text{and} \quad g_0(x) \leq \frac{K}{x} \quad \text{if } x \geq 0.$$

Moreover, for any $\bar{x} > 0$, there is some $k > 0$ such that, for all λ large enough:

$$g_1(x) \geq kx \quad \text{if } x \geq 0, \quad \text{and} \quad g_0(x) \geq -kx \quad \text{if } x \leq 0.$$

Proof. Recall that $g_1(x)$ is the positive root of (36). Thus $g_1(x) \leq -K/x$ if and only if the second-order polynomial on the left-hand side of (36) positive when evaluated at $-K/x$, that is if and only if:

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

For this inequality to hold for all $x \leq 0$ and all $\lambda > 0$, it is sufficient that:

$$-\frac{K}{x}\sqrt{\lambda}\left[F(\delta^\star) - F\left(\delta^\star + x/\sqrt{\lambda}\right)\right] - \gamma \geq 0.$$

Using Taylor Theorem, we can write $F(\delta^\star) - F(\delta^\star + x/\sqrt{\lambda}) = -x/\sqrt{\lambda}F'\left[\hat{\delta}(x)\right]$ for some $x \in [\delta^\star + x/\sqrt{\lambda}, \delta^\star]$. Plugging this back into the above inequality we obtain that $g_1(x) \leq -K/x$ if $KF'\left[\hat{\delta}(x)\right] - \gamma \geq 0$, and so we can choose

$$K \geq \frac{\gamma}{\min_{\delta \in [0,1]} F'(\delta)}.$$

Note that the right-hand side is less than infinity because $F'(\delta)$ is assumed to be strictly positive for all $\delta \in [0,1]$. One obtain the same constant K when applying the same calculations to $g_0(x)$ and $x \geq 0$.

Now let us turn to the second part of the proposition and fix some $\bar{x} > 0$. As before, we recall that $g_1(x)$ is the positive root of (36). Thus $g_1(x) \geq kx$ if and only if the above second-order polynomial on the left-hand side of (36) is negative when evaluated at kx, that is if and only if:

$$\begin{aligned}
& k^2x^2 + kx\left\{\frac{\gamma}{\sqrt{\lambda}} + \sqrt{\lambda}\left[F(\delta^\star) - F\left(\delta^\star + x/\sqrt{\lambda}\right)\right]\right\} - \gamma s F\left(\delta^\star + x/\sqrt{\lambda}\right) \leq 0 \\
\Leftrightarrow \quad & k^2x^2 + kx\left\{\frac{\gamma}{\sqrt{\lambda}} - xF'\left[\hat{\delta}(x)\right]\right\} - \gamma s F\left(\delta^\star + x/\sqrt{\lambda}\right) \leq 0
\end{aligned}$$

where we move from the first to the second line using Taylor Theorem as before, and where $\hat{\delta}(x) \in [0,1]$. Dividing through by x^2, we obtain that a sufficient condition for this inequality to hold for all $x \geq \bar{x}$ is:

$$k^2 - k \min_{\delta \in [0,1]} F'(\delta) + \frac{\gamma}{\sqrt{\lambda}\bar{x}} \leq 0.$$

If we pick any $k < \min_{\delta \in [0,1] F'(\delta)}$, then the above inequality holds for all λ large enough. One obtain the same constant k when applying the same calculations to $g_0(x)$ and $x \leq -\bar{x}$. ■

Using the above preliminary results, we obtain:

Lemma A.10 *As $\lambda \to \infty$:*

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

Proof. The integrand of $D(\lambda)$ can be written:

$$\mathbf{1}_{\{x \in [-\delta^\star\sqrt{\lambda}, 0]\}} \frac{\frac{\gamma}{\sqrt{\lambda}}F\left(\delta^\star + x/\sqrt{\lambda}\right) + \theta_0 g_1(x)}{\frac{r+\gamma}{\sqrt{\lambda}} + \theta_0 g_1(x) + \theta_1 g_0(x)}.$$

By Lemma A.8, this integrand converges pointwise to $\frac{\theta_0 g(x)}{\theta_0 g(x)+\theta_1 g(-x)}$. Now fix any $\bar{x} > 0$ and λ large enough. On the interval $[-\bar{x}, 0]$, we can bound the integrand above by 1. On the interval $\left[-\delta^\star\sqrt{\lambda}, -\bar{x}\right]$, we use Lemma A.9 to bound the integrand above by:

$$\frac{\frac{\gamma}{\sqrt{\lambda}} - \frac{\theta_0 K}{x}}{-\frac{\theta_0 K}{x} - \theta_1 kx} = \frac{-\frac{\gamma x}{\sqrt{\lambda}} + \theta_0 K}{\theta_0 K + \theta_1 kx^2} \leq \frac{\gamma\delta^\star + \theta_0 K}{\theta_0 K + \theta_1 kx^2},$$

where we used that $x \geq -\delta^\star\sqrt{\lambda}$. On the interval $(-\infty, -\delta^\star\sqrt{\lambda}]$, the integrand is zero and so the previous bound holds as well. Taken together, we have bounded the integrand above by a positive function, integrable over $(-\infty, 0]$. This allows to apply the Dominated Convergence Theorem (ADD REFERENCE), and the result follows. The result for $P(\lambda)$ follows from identical calculations. ■

After making the change of variable $y = -x$ in the limit of $D(\lambda)$, we obtain

$$\lim_{\lambda\to\infty} D(\lambda) = \int_0^\infty \frac{\theta_0 g(-x)}{\theta_0 g(-x) + \theta_1 g(x)}\, dx.$$

Therefore:

$$\begin{aligned}\lim_{\lambda\to\infty} P(\lambda) - D(\lambda) &= \int_0^\infty \frac{\theta_1 g(-x)}{\theta_0 g(x) + \theta_1 g(-x)}\, dx - \int_0^\infty \frac{\theta_0 g(-x)}{\theta_0 g(-x) + \theta_1 g(x)}\, dx \\ &= \int_0^\infty \frac{(1-2\theta_0) g(x) g(-x)\, dx}{\left[\theta_0 g(x) + \theta_1 g(-x)\right]\left[\theta_0 g(-x) + \theta_1 g(x)\right]}.\end{aligned}$$

Now solving the quadratic equation for $g(x)$ we obtain that:

$$g(x) = \frac{1}{2}\left[F'(\delta^\star)x + \sqrt{4\gamma s(1-s) + (xF'(\delta^\star))^2}\right].$$

Plugging this formula in the above and simplifying, we obtain that:

$$\begin{aligned}\lim_{\lambda\to\infty} P(\lambda) - D(\lambda) &= \int_0^\infty \frac{\gamma s(1-s)(1-2\theta_0)}{\gamma s(1-s) + \theta_0\theta_1 (xF'(\delta^\star))^2} dx \\ &= \frac{\pi\sqrt{\gamma}}{F'(\delta^\star)}\left(\frac{1}{2} - \theta_0\right)\left(\frac{s(1-s)}{\theta_0\theta_1}\right)^{\frac{1}{2}}\end{aligned}$$

since $\int_0^\infty \frac{dx}{1+x^2} = \frac{\pi}{2}$. Collecting the results we conclude that

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

This expansion applies to all reservation values and price because, by Proposition 8 show below:

$$\Delta V(\delta) = \Delta V(\delta^\star) + [\Delta V(\delta) - \Delta V(\delta^\star)] \text{ and } \lim_{\lambda\to\infty} \sqrt{\lambda}\,[\Delta V(\delta) - \Delta V(\delta^\star)] = 0.$$

■

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 $i < m$:

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

Likewise, for all $i > m$:

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

Finally, we calculate the reservation value $\Delta V(\delta_m)$ using formula (12)

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

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

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 $1/\lambda$. ■

Proof of Proposition 8. The first intermediate result is:

Lemma 7 *As λ goes to infinity:*

$$\lambda\int_0^{\delta^\star} \sigma(\delta)\,d\delta = \int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star-\delta) + \Phi_1(\delta)} + O(1) \tag{37}$$

$$\lambda\int_{\delta^\star}^1 \sigma(\delta)\,d\delta = \int_{\delta^\star}^1 \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta^\star-\delta) + 1-s-\Phi_0(\delta)} + O(1). \tag{38}$$

Proof. We start with (37), 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 (37), we obtain:

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

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 (37). Turning to equation (38), we first note that:

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

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 (37). ■

Next, we obtain a lower bound for the integral on the right-hand side of (37) by bounding $\Phi_1(\delta)$ above by $\Phi_1(\delta^\star)$:

Lemma A.11 *As $\lambda \to \infty$:*

$$\int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} \geq \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1) \tag{39}$$

$$\int_{\delta^\star}^1 \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta - \delta^\star) + 1 - s - \Phi_0(\delta)} \geq \frac{\log(\lambda)}{2\theta_0 F'(\delta^\star)} + O(1). \tag{40}$$

Proof. For (39), 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 (40), 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.12 *As $\lambda \to \infty$:*

$$\int_0^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1(\delta)} \leq \frac{\log(\lambda)}{2\theta_1 F'(\delta^\star)} + O(1) \tag{41}$$

$$\int_{\delta^\star}^{1} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_0 F'(\delta^\star)(\delta - \delta^\star) + 1 - s - \Phi_0(\delta)} \leq \frac{\log(\lambda)}{2\theta_0 F'(\delta^\star)} + O(1). \tag{42}$$

Proof. For (41), 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:

$$\begin{aligned}
&\int_{\delta^\star - 1/\sqrt{\lambda}}^{\delta^\star} \frac{d\delta}{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)(\delta^\star - \delta) + \Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)} \\
=&\frac{1}{\theta_1 F'(\delta^\star)} \log\left(\frac{\frac{r+\gamma}{\lambda} + \theta_1 F'(\delta^\star)\alpha(\lambda) + \Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)}{\frac{r+\gamma}{\lambda} + \Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)}\right) \\
\to&\frac{1}{\theta_1 F'(\delta^\star)} \log\left(\frac{\theta_1 F'(\delta^\star) + g(-1)}{g(-1)}\right) = O(1),
\end{aligned}$$

where $g(-1)$ is the limit of $\sqrt{\lambda}\Phi_1\left(\delta^\star - 1/\sqrt{\lambda}\right)$ as shown in Lemma A.8. For (42), the result follows from identical algebraic manipulations. ■

■

Proof of Proposition 9. Recall that

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

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.5, we note that:

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

Then, we also note that:

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

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

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

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, $\int_{\delta^\star}^1 \left[1 - s - \Phi_0(\delta)\right] d\delta$, and the result follows. ■

B Non stationary initial conditions

Assume that the initial distribution of 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 1 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 types in the whole population. Since each agent in the population draws a new type from $F(\delta)$ with intensity γ we have that this distribution satisfies

$$\dot{F}_t(\delta) = \gamma(F(\delta) - F_t(\delta)).$$

Solving this equation then shows that the cumulative distribution of 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 differential 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

$$\lambda\Phi_{1,t}(\delta) = \lambda(F_t(m) - \Phi_{0,t}(\delta)) = \frac{\dot{Y}_{+,t}(\delta) - A(\delta)\dot{Y}_{-,t}(\delta)}{Y_{+,t}(\delta) - A(\delta)Y_{-,t}(\delta)} \tag{43}$$

with

$$\begin{aligned} Y_{\pm,t}(\delta) &= e^{-\lambda Z_\pm(\delta)t}W_{\pm,t}(\delta) \\ Z_\pm(\delta) &= \frac{1}{2}\left(1 - s + \gamma/\lambda - F(\delta)\right) \pm \frac{1}{2}\Lambda(\delta) \\ W_{\pm,t}(\delta) &= M_1\left(\frac{\lambda}{\gamma}Z_\pm(\delta), 1 \pm \frac{\lambda}{\gamma}\Lambda(m); e^{-\gamma t}\frac{\lambda}{\gamma}(F(\delta) - F_0(\delta))\right) \end{aligned} \tag{44}$$

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 B.1 *The equilibrium distributions defined by* (43) *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 (43) 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 (44) and the definition of the steady state distribution $\Phi_1(\delta)$. ■

Given the joint distribution of types and asset holdings the unique equilibrium can be computed by substituting the equilibrium distributions into (**??**), (11) and (12) and the same arguments as in the stationary case show that this equilibrium converges to the same steady state equilibrium as in Theorem 1.