Relationship dynamics in over-the-counter markets

Robin Lenoir

June 28, 2024

Abstract

I develop a model of relationship dynamics in OTC markets, focusing on how these relationships affect liquidity and trading behavior. Investors facing liquidity shocks trade assets through partners chosen from a personal network of dealers (randomly) met in the past. Dealers, heterogeneous in their execution speed, compete on fees within the network to secure trades. In equilibrium, investors trade with the fastest dealer in their network, but fees are determined by the speed of the second-fastest dealer. As agents meet new dealers, they can switch to better partners, concentrating trade among the fastest dealers. When trading repeatedly with the same dealer, equilibrium fees decrease over time. I demonstrate how heterogeneity in investor access to the market depends on their network, with those having faster dealers adjusting their asset holdings more after liquidity shocks. Finally, I show that increased dealer competition, captured by the meeting rate of new dealers, increases asset holding dispersion and improves liquidity.

1 Introduction

It is now well-established that, in many OTC markets, trade occurs through relationships. For example, Hendershott et al. 2020 shows that one-third of insurers on the bond market trade with a single dealer over a year. However, the impact of these relationships on market outcomes, particularly liquidity, is not well understood. Do relationships improve liquidity by reducing search frictions, or do they reduce competition and increase spreads? If investors access the market through relationships, do they trade differently based on *who* they trade with or for how long? How do agents choose their trading partners? Although relationships are often long-lasting, they are not eternal. How and when do agents switch partners? How do the formation and evolution of relationships affect trade and market outcomes? To address these questions, I develop a model of relationship dynamics in an OTC market intermediated by broker-dealers.

The model captures key features of OTC markets: (1) trade happens through existing relationships with chosen dealers, (2) not all dealers are equal—some are efficient (execute trades quickly) while others are not, and (3) relationships are long-lasting but not eternal—agents occasionally switch dealers. The model reproduces these features and studies their implications on the dynamics of relationships. It also examines how these dynamics create heterogeneous access to the market and trading behavior for investors. Finally, I study how they translate into observable market outcomes, such as dispersion in asset holdings or bid-ask spreads.

Building on Lagos and Rocheteau 2009, I consider a continuous-time model where agents trade an asset and face liquidity shocks in the form of preference changes for asset holdings. To trade the asset, agents must go through a dealer. Agents meet dealers over time at a Poisson rate and form an evolving list of dealers they can trade with– their "network." To ensure stationarity, I assume that investors lose their dealers' list at a Poisson rate, starting again from zero. Each time an investor wants to trade, they submit their order to a *single* dealer in their network, in exchange for a fee. The fee is determined via Bertrand competition between dealers in the network. Importantly, dealers are ex-ante heterogeneous in their execution speed. When choosing which dealer to trade with, the investor balances the speed of execution and the fee charged by the dealer. Trade orders submitted to dealers are executed at the dealers' rate on a perfectly competitive inter-dealer market, which provides the asset's price.

The model is solved using standard techniques from the continuous-time heterogeneous agent literature (see, for example, Achdou et al. 2022). I show that the model can be reduced to a two-dimensional state space, where agents only keep track of the two fastest dealers in their network. An investor will always choose to trade with the fastest dealer, and the fee paid will depend on the speed of the second-fastest dealer, which acts as an outside option. I characterize the rate at which agents change dealers (and how this depends on their current network). Furthermore, I show that the fee paid by agents will decrease over time if they keep the same dealer, as the outside option improves. On the other hand, fees may change in any direction after a switch. Finally, I characterize the stationary distribution of investors' current networks (i.e., who are their two fastest dealers) and illustrate numerically that it concentrates on the fastest dealers in the population, consistent with empirical evidence suggesting that most trades are concentrated among a few very well-connected dealers.

I then study how agents choose their asset holdings. I show that agents will adjust their asset holdings more or less depending on how accessible the market is for them (i.e., on the execution speed of their dealer). Agents will adjust their holdings more if they can trade with faster dealers. Agents stuck with slow dealers as their trading partners are afraid of being left with an asset they cannot sell quickly in case of a new liquidity shock. This creates an important source of heterogeneity in asset holdings as agents vary their holdings based on both their current preferences and their current network.

Finally, I study how dealer competition, captured by the meeting rate of new dealers, impacts market outcomes in equilibrium, focusing on liquidity in particular. I show numerically that dispersion in asset holdings increases with the meeting rate. As the meeting rate goes to infinity, investors essentially have access to all dealers, including the fastest ones. Therefore, they are more willing to match their current preference state, knowing they can readjust later. For an infinite meeting rate, competition between dealers is perfect, reducing fees (hence bid-ask spread) to zero—the market is perfectly liquid. When meeting rates are low, dealers have market power and can charge high fees, leading to high bid-ask spreads—an illiquid market. I illustrate numerically that the evolution is gradual, with fees reducing and converging to zero as the meeting rate increases.

Literature Review. My paper relates to two strands of the literature on OTC markets: search-and-matching literature focusing on price formation and liquidity, and literature focusing on trading relationships (but not on price formation).

The first strand inherits from the seminal papers by Duffie, Gârleanu, and Pedersen 2005 and Lagos and Rocheteau 2009. These papers use tools from search-and-matching to model frictions giving rise to bid-ask spreads and other forms of illiquidity (see Weill 2020 for an extensive review). Models in this literature do not focus on repeated relationships between investors and traders, unlike my paper. Agents can usually trade with any dealer (although they may face search frictions). Closest to my paper, are models using tools from directed search allowing for heterogeneity and competition between dealers. A prominent example is Lester, Rocheteau, and Weill 2015, in which dealers can post terms of trade (fee and execution speed) and investors choose which dealer to trade with. In contrast to my paper, investors have access to all dealers and do not necessarily form relationships. Other relevant works include Hugonnier, Lester, and Weill 2022 which study heterogeneity in dealers without forming relationships.

The second strand is mostly empirical and studies the structure of relationships in OTC markets. This literature primarily documents the stability of relationships in OTC markets and describes the structure of the network. Afonso, Kovner, and Schoar 2013 describes relationships in the overnight interbank lending OTC market. Di Maggio, Kermani, and Song 2017 and Li and Schürhoff 2019 study the structure of inter-dealer relationships in the corporate bonds market. Hendershott et al. 2020 studies client-dealer relationships. These last two studies are closest to my paper as they also encompass theoretical models to study the effect of relationships on market outcomes, including liquidity. Unlike my paper, these studies do not focus on heterogeneity between accessible dealers and the *evolution* of relationships over time. Sambalaibat 2022 proposes a dynamic model of network formation similar in spirit to my model, but focuses on the role of heterogeneity on the client side and on the role of dealer specialization. Furthermore, it does not study the evolution of relationships.

My paper is also indebted to the competitive search literature from labor and trade. The mechanism under which newly met dealers compete in Bertrand competition with pre-existing relationships is inspired by models of labor search with sequential bargaining. The main model was developed by Cahuc, Postel-Vinay, and Robin 2006. Kaas and Kircher 2015, and Lise, Meghir, and Robin 2016 provide extensions. Fontaine, Martin, and Mejean 2023 is a recent paper that uses this framework in the context of international trade.

2 Model

2.1 Environment

The core of the model is borrowed from Lagos and Rocheteau 2009. Time is continuous and extends indefinitely. The economy is populated by two types of infinitely-lived agents: investors and dealers, each in unit mass. There is one asset in the economy, available in fixed supply B. Agents trade the asset using a numéraire consumption good.

Investors. Investors derive utility flow from holding the asset. The benefits from asset holding are heterogeneous and vary randomly, as is standard in this literature. Let k denote an investor's preference state, and let $u_k(a)$ represent the utility flow they receive from the asset in that state. I assume u_k is strictly increasing, strictly concave, that $u_k \to \infty$ as $a \to \infty$ and $u_k \to 0$ as $a \to 0$. I assume that k evolves according to a continuous-time Markov chain with generator Q[1]. This can be interpreted, for example, as agents receiving (positive or negative) liquidity shocks. The use of the Markov chain allows the probability of receiving specific shocks (and their timing) to depend on the current state. Investors also derive utility from consuming the numéraire good, which I assume enters linearly into the utility function. Denoting c as the consumption of the numéraire good, the total utility flow of an investor is $u_k(a) + c$. Finally, investors discount future utility at a rate of r.

Trade. The presence of liquidity shocks implies that agents need to readjust their asset positions over time, motivating them to buy or sell the asset on the market. However, investors do not have direct access to the market and must trade through dealers. Dealers are risk-neutral and have no preference for the asset. The inter-dealer market is competitive, and the asset price is taken as given. When an investor wants to trade, they can immediately turn to a dealer (not *any* dealer, as will be explained below). However, dealers face friction in executing trades, represented by an idiosyncratic "speed" s_i, which is the Poisson rate of execution. Dealers are heterogeneous in their speeds, and the distribution of speeds is characterized by a full support distribution on $[\underline{s}, \bar{s}]$, with a density $f(s)$. A dealer's speed is fixed over time and exogenous. In exchange for facilitating the trade, the dealer receives a flat fee determined endogenously (see below).

[1] The generator Q is assumed to be time-homogeneous, irreducible, and aperiodic. The state space is finite, and the transition rates ensure that the chain is ergodic.

Relationships. Investors can only trade with dealers they already know. Formally, investors maintain a list of dealers they have met in the past—their personal "network". They start without knowing any dealers and meet new ones at a rate α. I assume that investors who haven't met any dealers can trade (for free) at $\underline{s}$. To ensure stationarity, I assume that agents lose their entire network at rate β. This can be interpreted as a change in ownership of the investor's asset portfolio, resulting in the loss of accumulated contacts[2]. I denote the set of dealers in an investor's network by $\mathcal{N}$. When investors want to trade, they can turn to any dealer on their list to request a trade. To choose which dealer to trade with, investors have the different available dealers compete *à la Bertrand* for the fee they will charge. For simplicity, I assume that the fee is a flow fee, paid continuously to the preferred dealer[3]. In the next section, I show that the dealer with the highest speed wins the competition, and the fee depends on the speed of the *second-highest speed* dealer. Importantly, the investor can only choose *one* dealer to request the trade from. However, they can switch dealers at any time, even if the trade has not been executed yet (though in that case, the original dealer will abandon the trade execution).

2.2 Equilibrium Structure

The model described above is a dynamic heterogeneous agent model in continuous time. As such, it shares a standard structure with many heterogeneous agent models, such as those described in Achdou et al. 2022. In particular, we can solve for the equilibrium of the model by solving a system of differential equations: (1) a Hamilton-Jacobi-Bellman (HJB) equation for the optimal choices of investors given their states, (2) a Forward Kolmogorov equation for the distribution of investors over possible states, and (3) a market clearing condition. In our model, the state space comprises the preference state k and the investors' dealer network $\mathcal{N}$. In the next section, I show that it is not necessary to keep track of the entire dealer network; instead, it is sufficient to keep track of the two highest speeds in the network. Before that, however, I detail the equilibrium equations in general. We will return to the details of these equations in the subsequent sections.

Hamilton-Jacobi-Bellman Equation. Let $V_k(a, \mathcal{N}, t)$ be the maximum expected discounted utility attainable by an investor with holdings a, preference state k, and dealer network $\mathcal{N}$. Denote p_t as the price of the asset at time t. It is straightforward to show that the HJB takes the following form:

$$\begin{aligned} rV_k(a, \mathcal{N}, t) = u_k(a) &+ \eta^*(V_k(a^*, \mathcal{N}, t) - V_k(a, \mathcal{N}, t) - p_t(a^* - a)) - \phi^* \\ &+ (QV)_k(a, \mathcal{N}, t) \\ &+ (\mathcal{A}V_k)(a, \mathcal{N}, t) \\ &+ V_t, \end{aligned}$$

[2]Alternatively, we can assume that investors die at rate β and are replaced by new investors who have not met any dealers.

[3]As we will see in the next section, the fee is a percentage of the gain-from-trade, and an investor who doesn't want to adjust their trade would pay a 0 fee. Alternative assumptions could include paying the fee upon the realization of a trade. My model can easily accommodate such situations.

where a^* is the optimal asset holding, η^* is the speed of the in-network dealer chosen, and ϕ^* is the fee paid to the dealer. $(QV)_k(a, \mathcal{N}, t)$ is the term that accounts for changes in the preference state, formally it is the k-th entry of the vector obtained by multiplying the generator Q by the vector $V(a, \mathcal{N}, t)$, stacking the different preference states. The term $(\mathcal{A}V_k)(a, \mathcal{N}, t)$ accounts for changes in the network. In particular, $\mathcal{A}$ is an operator (semi-group) that describes the evolution of the network. The next section is dedicated to understanding this operator.

At this stage, we remark that neither a^*, η^*, nor ϕ^* depend on the current asset holdings a. This is a standard result in the literature, attributed to the linearity of the utility function in the consumption of the numéraire good[4]. This implies that we can ignore a as a state variable, particularly in the Forward Kolmogorov equation.

Forward Kolmogorov Equation. Let $\mu(k, \mathcal{N}, t)$ be the distribution of investors over the preference and network state space. The evolution of this distribution is given by the Kolmogorov forward equation:

$$\partial_t \mu = Q^* \mu(k, \mathcal{N}, t) + (\mathcal{A}^* \mu)(k, \mathcal{N}, t),$$

where Q^* and $\mathcal{A}^*$ are the adjoints of Q and $\mathcal{A}$, respectively. As before, the term $Q^*\mu(k, \mathcal{N}, t)$ represents the k-th entries of the vector obtained by multiplying the generator Q by the vector $\mu(\cdot, \mathcal{N}, t)$, stacking over the different preference states, and $(\mathcal{A}^*\mu)(k, \mathcal{N}, t)$ is the value of $(\mathcal{A}^*\mu(k, \cdot, t))$ (which is a function, because $\mathcal{A}^*$ is an operator on functions) evaluated at $\mathcal{N}$. We will describe $\mathcal{A}^*$ in detail in the next section.

Market Clearing Condition. Finally, the model is closed by a market clearing condition. At each period, a total number of $\int \eta^*(k, \mathcal{N}, t)\, d\mu$ dealers access the market with a trade request volume of $\int a^*(k, \mathcal{N}, t)\eta^*(k, \mathcal{N}, t)\, d\mu$, representing demand. The asset supply on the inter-dealer market is $\int B\eta^*(k, \mathcal{N}, t)\, d\mu$. We can write the market clearing condition as:

$$\int a^*(k, \mathcal{N}, t)\, d\mu = B.$$

I prefer the following notation. Let $\mu_k(\mathcal{N}, t)$ be the distribution over possible networks of dealers for an investor with preference k. Let $n(t) = (n_k(t))_t$ be the distribution over preferences states. The market clearing condition can be written as:

$$\sum_k n_k(t) \int a^*(k, \mathcal{N}, t)\, d\mu_k = B.$$

3 Relationship Dynamics

In this section, I describe how agents select dealers to trade with and how this choice evolves. I will show that it is sufficient to keep track of the two highest speeds in the investor network. Trade occurs with the dealer who has the highest speed, and the fee paid depends on the

[4] I describe this result in more detail in Section 4

speed of the second-highest-speed dealer. Then, I will demonstrate how investors move within this two-dimensional space: when and how they switch the dealer they trade with and how the fee they pay evolves.

3.1 Dealer Competition

At each period t, agents submit a trade order to a dealer of their choice (the order can be of size 0), in exchange for a flow fee ϕ set by the dealer. The choice of dealers and the determination of fees represent a game akin to an auction: dealers submit bids in the form of pairs (s_i, ϕ_i), where speed s_i is exogenous and fee ϕ_i is endogenous. The investor selects the bid that maximizes their expected utility. Note that this problem is repeated in each period and is effectively static. The game depends on $\mathcal{N}$, the set of dealers in the network, i.e., the players (excluding the investor) in the game. It is useful to relabel dealers such that $\mathcal{N} = \{1, 2, \ldots, |\mathcal{N}|\}$, with $s_1 \geq s_2 \geq \ldots \geq s_{|\mathcal{N}|}$ representing the speeds of the dealers in the network in *decreasing* order (i.e., s_1 is the highest speed).

The Investor's Problem. Let $\bar{V} = V_k(a^*, \mathcal{N}, t) - V_k(a, \mathcal{N}, t) - p_t(a^* - a)$ be the gain from trade by choosing the optimal asset holdings at time t. It is clear from the HJB that the investor will choose the bid (s_i, ϕ_i), $i \in \mathcal{N}$, that maximizes $s\bar{V} - \phi$, which is a static choice.

The Dealer's Problem. Dealers are risk-neutral and maximize their immediate profit ϕ. They will choose the highest possible fee while ensuring they are selected by the investor. In effect, dealer i's maximization problem is:

$$\max_{\phi} \phi \quad \text{s.t.} \quad s_i\bar{V} - \phi_i \geq s_j\bar{V} - \phi_j, \quad \forall j \in \mathcal{N}.$$

This resembles a standard Bertrand competition problem. One can check that a Nash equilibrium[5] to this game is:

- The investor chooses the dealer with the highest speed s_1.
- $\phi_j = 0$ for all $j \neq 2$.
- $\phi_1 = s_1\bar{V} - s_2\bar{V}$.

This implies that the investor will always choose the dealer with the highest speed, and the fee paid will be $(s_1 - s_2)\bar{V}$. Finally, note that:

$$s_1\bar{V} - \phi_1 = s_2\bar{V},$$

which means that, from the investor's perspective, the solution is as if they were trading with the dealer with the second-highest speed, but gaining the full benefits from the trade (without paying a fee).

[5] This is the unique equilibrium provided we assume that the investor chooses the highest-cost dealer when indifferent between two bids.

3.2 Moving in the Dealer Space

In this subsection, I describe how to characterize the way investors' networks evolve. As shown above, the relevant value to keep track of is s_2, the second-highest speed in the network. However, future values of s_2 also depend on the current value of s_1. Suppose an investor whose two fastest dealers are s_1 and s_2 meets a new dealer. We can be in one of three situations:

1. The new dealer has a speed $s_1 > s' > s_2$, i.e., the new second-highest speed is s'. In this case, the investor will continue trading with the current dealer at speed s_1, but the fee will change (due to the improved outside option). From the investor's point of view, it will be as if they trade at s' with no fee.

2. The new dealer has a speed $s' > s_1$. The investor will start trading with this dealer, and their new second-highest speed will be $s_2 = s_1$ (i.e., s_2 is replaced by s_1).

3. The new dealer has a speed $s' < s_2$. The investor will ignore the new dealer and continue trading with s_1 with fee set by s_2.

Formally, we want to keep track of a stochastic process $X_t = (s_1(t), s_2(t))$. This process is Markovian and can be described by its generator $\mathcal{A}$. It is straightforward to show that, denoting $f(\cdot)$ the pdf of dealers' speeds (in the full population), and $F(\cdot)$ the associated CDF, the generator[6] is characterized by

$$\begin{aligned}\mathcal{A}h(s_1, s_2) = \alpha\Big(\int_{s_1}^{\bar{s}} f(s)h(s, s_1)\, ds + \int_{s_2}^{s_1} f(s)h(s_1, s)\, ds - (1 - F(s_2))h(s_1, s_2)\Big) \\ - \beta(h(\underline{s}, \underline{s}) - h(s_1, s_2)),\end{aligned}$$

for any continuous function $h : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$. $\mathcal{A}h(s_1, s_2)$ represents the infinitesimal change in $h(s_1, s_2)$ at (s_1, s_2). We can interpret it as follows:

- The terms $-\beta h(s_1, s_2)$ and $-\alpha(1 - F(s_2))h(s_1, s_2)$ represent flows out of the current state, either due to a shock β (losing the entire network) or because a new dealer with $s > s_2$ is met.

- The term $\alpha \int_{s_1}^{\bar{s}} f(s)h(s, s_1)\, ds$ represents the flow of agents meeting a dealer with speed $s > s_1$. This will lead to a change in the highest speed, becoming s, and the second-highest speed becoming s_1.

- The term $\alpha \int_{s_2}^{s_1} f(s)h(s_1, s)\, ds$ represents the flow of agents meeting a dealer with speed $s_2 < s < s_1$. This will lead to a change in the second-highest speed, becoming s, while the highest speed remains s_1.

The generator will enter the HJB, which we can finally rewrite, keeping only s_1 and s_2 as state variables:

$$\begin{aligned}rV_k(s_1, s_2, t) = u_k(a) + s_2\left(V_k(s_1, s_1, t) - V_k(s_1, s_2, t) - p_t(s_1 - s_2)\right) \\ + (QV)_k(s_1, s_2, t) + (\mathcal{A}V_k)(s_1, s_2, t) + V_t.\end{aligned}$$

[6]For a Markov processes a generator is defined as $\lim_{\Delta \to 0} \frac{\mathbb{E}[f(X_\Delta)|X_0=x]-f(x)}{\Delta}$. for appropriate (see e.g. Pavliotis 2014 for details) functions f of the value of the process. Generators are very useful in describing Markov processes as they capture infinitesimal expected changes in the process.

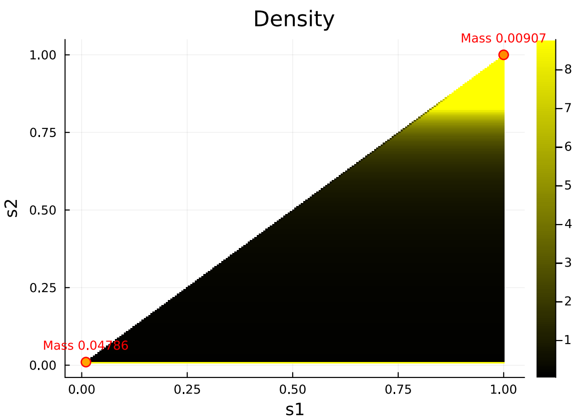

Figure 1: Stationary distribution of dealers' speeds

3.3 The Stationary Distribution

The formula for the generator $\mathcal{A}$ is also useful for describing the evolution of the distribution of dealers' speeds. Let $\mu_k(s_1, s_2, t)$ denote the distribution of investors' effective networks (the speeds of their two best dealers) when they have preference k. This is the distribution of the random variable X_t. Since the process for X_t is Markovian and time-homogeneous, we can use the Kolmogorov Forward Equation to describe the evolution of μ_k. Denoting $\mathcal{A}^*$ as the adjoint of $\mathcal{A}$, we have:

$$\partial_t \mu_k = \mathcal{A}^* \mu_k(s_1, s_2, t),$$

which can be solved from an initial condition. I am particularly interested in the *stationary* distribution μ_k, which solves

$$\mathcal{A}^* \mu_k = 0.$$

Notice that $\mathcal{A}$ is irreducible[7] (on the space $s_1 > s_2$), so the stationary distribution exists and is unique.

We can solve for this distribution numerically. To do so, I discretize the space (s_1, s_2), transforming $\mathcal{A}$ into a matrix. I can then use standard numerical linear algebra techniques to solve for the stationary distribution. Note that since the choice of dealers and network evolution do not depend on asset holdings, we do not need to solve the full model to compute the stationary distribution. The code I wrote to solve the model is available upon request. I solve the model for the $\alpha = 0.2$, $\beta = 0.01$ and speed is uniformly between 0.01 and 1. Figure 1, plots the distribution. Observe first the point mass at $\underline{s}$, corresponding to the investor "starting anew" with no network. Besides this, the distribution concentrates on the lowest costs, both in s_1 and s_2 (note the graph is in terms of density, but the point masses indicate actual probabilities). This occurs because investors with slower dealers tend to leave them, as they are likely to find a faster dealer. This is true for both s_1 and s_2. On the other hand, if a dealer is already very fast, it is unlikely that the investor will find a faster one.

This result is consistent with empirical evidence showing that most trades tend to be concentrated among a few dealers. In my model, these would be the most efficient dealers. Furthermore, while I assumed dealers' speed is exogenous, if one were to endogenize it, the

[7]From any state, we can attain any other state with positive probability in any time interval.

model would likely remain consistent. Dealers who concentrate trades would, in effect, be faster, as they have more investors to propose trades to[8].

3.4 Observable Dynamics

In this subsection, I describe how the above translates into potentially observable dynamics.

Observed Speed. Figure 2 uses the stationary distribution to compute the distribution of trade execution rates that may be observed (or inferred) in some datasets. Formally, this is simply the marginal distribution over s_1. We can see that the distribution concentrates at high average execution rates, precisely because of the forces described above: agents tend to accumulate towards the faster dealers. We also observe a mass at the 0.01 rate, corresponding to traders just starting anew. It is not clear if this would be consistent with observations. In any case, we have some flexibility to fit the data better by changing the exogenous distribution of dealer speeds.

Switches. We can use the stationary distribution to get information on the frequency of switches. According to the model, agents switch dealers when they meet a faster dealer, which happens at rate $\alpha \times (1 - F(s_1))$. Agents switch less frequently if their current speed is higher. Figure 3 shows the distribution of switch rates implied by the stationary distribution. We can see that the distribution decreases exponentially, consistent with a large mass of agents already trading at low cost and therefore rarely switching dealers.

Fees. Finally, we may be interested in the distribution of fees paid by investors. In the model, the fee paid by an investor is $(s_1 - s_2)\bar{V}$. Unfortunately, the gains from trade $\bar{V}$ are not observable, making it difficult to use this for calibration. To gain insights we can focus on $(s_1 - s_2)$ instead, which represents the percentage of gains from trade paid as fees. It is interesting to understand how the fee evolves as agents repeat trades with a dealer. In the model, agents keep the same dealer but renegotiate the price (as a percentage of gains from trade) if they meet agents with speeds $s \in [s_2, s_1]$. The price is always renegotiated *down*. The rate at which the price is renegotiated (conditional on maintaining the relationship) is given by $\alpha \frac{F(s_2)}{1-F(s_1)}$. Notice that during the relationship, s_1 stays constant but s_2 decreases, so renegotiations become less frequent and eventually stabilize around 0. If we can find some way to control gains from trade, we may use this fact for calibration. Finally, note that the fee (as a percentage) paid after a switch could move in any direction, which is most likely consistent with the data.

4 Impact on Trading

Having characterized how agents choose their dealers and how these relationships evolve in the model, I now turn to the question of how agents choose their asset holdings and how this choice translates into aggregates. The main insight is that the core forces identified in classic

[8] Hopefully, I can study such extensions in the future.

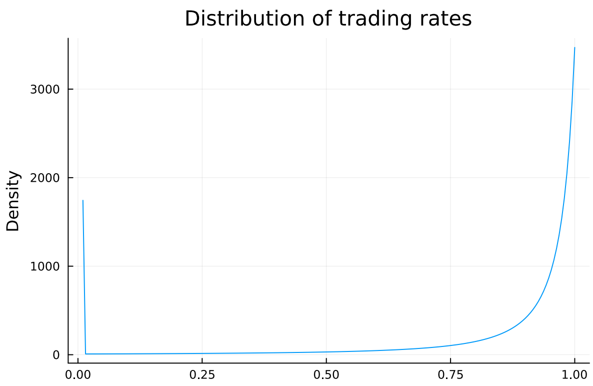

Figure 2: Observed distribution of dealers' speed

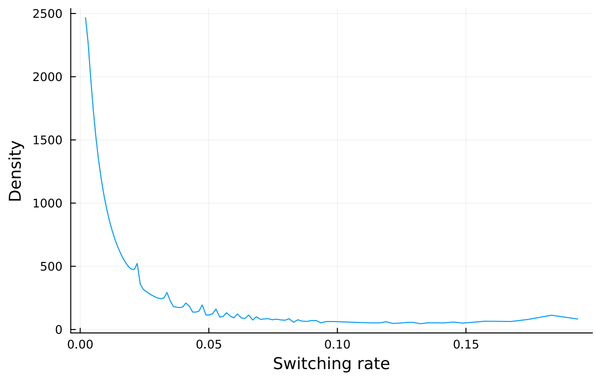

Figure 3: Distribution of switch rates

papers from the literature (Lagos and Rocheteau 2009) are still at play: agents smooth less between preference states if they can trade at a higher speed. Since trading speeds are heterogeneous and their distribution is endogenous (as determined in the previous section), this implies interesting aggregate dynamics in the model.

4.1 Optimal Asset Holdings

I start by characterizing the optimal asset holdings of an investor with preference k and network (s_1, s_2) (recall from the previous section that it is sufficient to keep track of the two highest speeds). For simplicity in this early version of the paper, we focus on characterizing optimal asset holdings in a stationary environment. As in Lagos and Rocheteau 2009, we can see from the HJB that the asset holdings an investor readjusts to do not depend on the current holding. It is as if each time the agent trades, they sell all of their assets and buy the optimal quantity they currently want. The selling of the asset creates a lump-sum transfer that does not impact their optimal decision. I obtain the following Lemma using techniques similar to those in Lagos and Rocheteau 2009:

Lemma 1. *Fixing the asset price p to be stationary, an investor with preference k and network (s_1, s_2) chooses each time they trade:*

$$a_k^*(s_1, s_2) = \arg\max_{a'>0} \left\{ \bar{u}_k(a', s_1, s_2) - p(v(s_1, s_2) - 1)a \right\},$$

where $\bar{u} = (\bar{u}_k(a', s_1, s_2))_k$ (the stacked vector of $\bar{u}_k(a')$) solves[9]*:*

$$Q\bar{u} + (\mathcal{A} - r - s_2)\bar{u} = u,$$

with $u = (u_k(a'))_k$. Furthermore, $v(s_1, s_2)$ is a function that solves the functional equation:

$$(\mathcal{A}v)(s_1, s_2) - (s_2 + r)v(s_1, s_2) = -s_2 p.$$

[9] I have a slight abuse of notation. Q acts on the k space so $Q\bar{u}$ is a standard matrix multiplication. $(\mathcal{A} - r - s_2)$ acts on the (s_1, s_2) space so $(\mathcal{A} - r - s_2)\bar{u}$ means the vector stacking $\{((\mathcal{A} - r - s_2)\bar{u}_k)(a', s_1, s_2)\}_k$

Proof. Following the same steps as in the proof of Lemma 1 in Lagos and Rocheteau 2009, we can show that the agent's problem can be written as:

$$\max_{a'} \left\{ \bar{u}_k(a', s_1, s_2) - \mathbb{E}_k[e^{rT}p|s_1, s_2]a + pa \right\},$$

where the expectation is over the next trading time T (it accounts for the randomness of trade execution and speed changes). As in Lagos and Rocheteau 2009, $\bar{u}_k(a', s_1, s_2) = \mathbb{E}_k[e^{rT}u_k(a)]$.

I start by characterizing v. Remember that we have characterized the evolution of the network by a (Markov) stochastic process $X_t = (s_1(t), s_2(t))$. Define $v(s_1, s_2) = \mathbb{E}_k[e^{rT}|X_0 = (s_1, s_2)]$. Observe that:

$$\begin{aligned} v(s_1, s_2) = \mathbb{E}_k[e^{rT}p|s_1, s_2] &= \mathbb{E}_k[\mathbb{E}_k[e^{rT}p|X_{\Delta t}]|X_0] \\ &= (1 - s_2\Delta t)\mathbb{E}_k[\mathbb{E}_k[e^{rT-\Delta t}|X_0 = X_{\Delta t}]|X_0 = s_1, s_2] + s_2\Delta t e^{-r\Delta t} + o(\Delta t) \\ &= (1 - s_2\Delta t)\mathbb{E}_k[v(X_{\Delta t})|X_0 = s_1, s_2] + s_2\Delta t e^{-r\Delta t} + o(\Delta t). \end{aligned}$$

The first line is the law of iterated expectations. The second line writes out explicitly the outer expectation about the execution or not of the trade at time Δt (but *not* changes in the speed of dealers, which is still in the expectation). The third line uses the Markovian nature of X_t to shift the inner expectation and write it as $v(X_{\Delta t})$. Rearranging terms, we get:

$$\frac{(1 - e^{r\Delta t})}{\Delta t} = e^{r\Delta t}\left(\frac{\mathbb{E}_k[v(X_{\Delta t})|s_1, s_2] - v(s_1, s_2)}{\Delta t} - s_2\mathbb{E}_k[v(X_{\Delta t})|s_1, s_2] + s_2\right)$$

Taking Δt to 0, and noting that $\lim_{\Delta t \to 0} \frac{\mathbb{E}_k[v(X_{\Delta t})|s_1,s_2] - v(s_1,s_2)}{\Delta t} = \mathcal{A}v$, by definition of the generator, we get the functional equation in the lemma.

We can use a similar approach for $\bar{u}_k$ and check it is characterized by the following recursive equation:

$$(r + s_2)\bar{u}_k(a', s_1, s_2) = u_k(a') + \sum_{k'} Q_{kk'}u_{k'}(a) + \mathcal{A}\bar{u}_k(a, s_1, s_2).$$

Stack the vector $\bar{u}_k(a')$ to get $\bar{u}(a')$.

$\square$

Notice that $v \in [0, 1]$. As long as $p > 0$, the assumptions we made on u_k guarantee the existence of a unique interior solution. We can easily solve v and $\bar{u}$ numerically to obtain a^*. In Figure 4, I fix a price $p = 1$ and illustrate the optimal choice for CRRA asset preferences $u_k(a) = k \times \frac{a^{1-\sigma}}{1-\sigma}$, with $k \in \{1, 2\}$[10]. The choice is presented in the (s_1, s_2) space. We can observe that agents in $k = 2$ buy more than in $k = 1$, which makes sense. The more important observation is that agents in $k = 2$ increase their holdings with speed, while agents in $k = 1$ decrease theirs. This is because agents who can trade faster (because they have faster dealers) can afford to smooth less between states. After all, they are not afraid of getting stuck in a mismatch. Figure 5 captures this intuition again by showing the size of the optimal adjustment of a "mismatched" agent (i.e., the difference between optimal holdings

[10] I assume $\alpha = 0.1$ and $\beta = 0.01$, and dealers' idiosyncratic speed is uniformly distributed between 0.01 and 1.

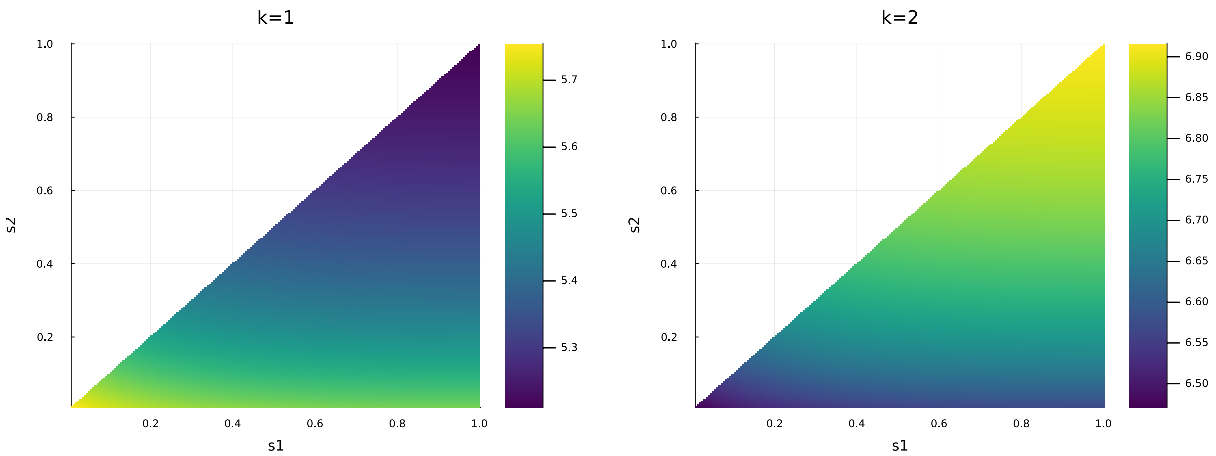

Figure 4: Optimal asset holdings for different preferences.

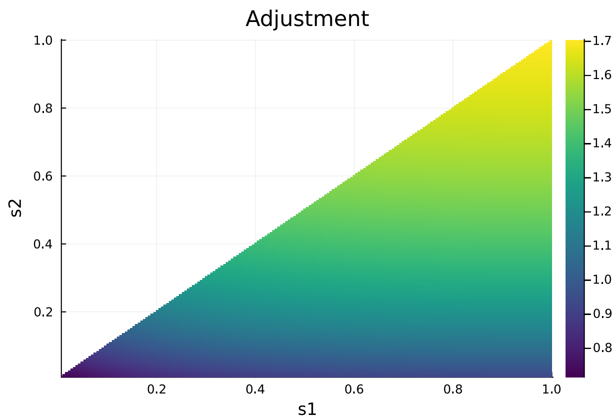

Figure 5: Optimal adjustment holdings for different preferences.

for the two preference states holding the network constant). Notice that changes in holdings mostly depend on s_2. Remember investors trade at speed s_1 but pay a fee $s_1 - s_2$ on their gains from trade, making it as if they were accessing the market at rate s_2. s_1 also matters but to a lower extent, as it captures expected changes in s_2.

4.2 General Equilibrium

Having characterized the optimal asset holdings of agents, we now have all the necessary ingredients to describe the general equilibrium of the model. Again, I focus on stationary equilibrium. Thanks to the simplifications made in the previous sections, we can write the general equilibrium as the following system of equations.

Definition 1. *A stationary equilibrium consists of a price p, value functions $V_k(s_1, s_2, a)$, a generator for network dynamics $\mathcal{A}$, distributions over networks μ_k, fees $\phi(s_1, s_2)$, and asset holdings $a^*(s_1, s_2)$ such that:*

- $\mathcal{A}$ *is an operator acting on functions* $h : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$, *and is characterized by:*

$$\mathcal{A}h(s_1, s_2) = \alpha\Big(\int_{s_1}^{\infty} f(s)h(s, s_1)\, ds + \int_{s_2}^{s_1} f(s)h(s_1, s)\, ds - (1 - F(s_2))h(s_1, s_2)\Big) \\ - \beta(h(\underline{s}, \underline{s}) - h(s_1, s_2)),$$

- *The HJB is satisfied:*

$$rV_k(s_1, s_2, a) = u_k(a) + s_2\left(V_k(s_1, s_1, a) - V_k(s_1, s_2, a) - p(a^*(s_1, s_2) - a)\right) \\ + (QV)_k(s_1, s_2, a) + (\mathcal{A}V_k)(s_1, s_2, a),$$

- *The Forward Kolmogorov equation is satisfied:*

$$\mathcal{A}^*\mu_k(s_1, s_2) = 0,$$

  where $\mathcal{A}^*$ *is the adjoint of* $\mathcal{A}$.

- *The market clearing condition is satisfied:*

$$\sum_k n_k \int a_k^*(s_1, s_2)\, d\mu_k = B,$$

  where n_k *is the number of agents with preference* k *under the stationary distribution associated with* Q.

- *Optimal asset holdings are characterized by:*

$$a_k^*(s_1, s_2) = \arg\max_{a'>0} \left\{\bar{u}_k(a', s_1, s_2) - p(v(s_1, s_2) - 1)a\right\},$$

  with $\bar{u}_k$ *and* v *characterized in Lemma 1.*

- *The fee is determined by the dealer competition:*

$$\phi(s_1, s_2, a) = (s_1 - s_2)\left(V_k(s_1, s_1, a_k^*(s_1, s_2)) - V_k(s_1, s_2, a) - p(a_k^*(s_1, s_2) - a)\right).$$

The stationary distribution μ_k is independent of the price, and we can solve it in isolation (as done in Section 3). Asset holdings can be solved without solving the HJB, as the value function doesn't enter (this is as in Lagos and Rocheteau 2009). The two points above imply that we can solve for the price without solving the HJB, which is a significant simplification. In numerical implementation, we can solve for the stationary distribution first and then iterate (with a bisection, e.g.) on optimal asset holdings to find the price that clears the market. The only point that would require solving the HJB is to characterize the exact fee, which depends on the gains from trade.

Unfortunately as of yet, I was not able to complete the proof for the existence and uniqueness of equilibrium, but I outlined the main ideas below.

Proposition 1. *There exists a unique stationary equilibrium for the model.*

Proof.

- The irreducibility of $\mathcal{A}$ and Q implies that the stationary distribution μ_k and $(n_k)_k$ exist and are unique. The unique optimal solution for a^* has been discussed in Section 4.
- Now I would need to prove that $p \to 0$ implies (from the formula in Lemma 1) that $a^* \to \infty$ for all k. This shouldn't be too difficult under the assumption that $u_k(a) \to \infty$ as $a \to \infty$.
- Similarly, I would need to prove that $p \to \infty$ implies $a^* = 0$ for all k. Again, the right assumption on u_k should guarantee it.
- It follows that aggregate asset demand increases in p and goes from 0 to ∞. The intermediate value theorem implies that a solution p to the market clearing condition exists is unique.
- It would remain to prove a unique solution to the HJB. This is a standard result in the literature, and I would need to adapt the proof to my model.

□

4.3 Market Liquidity

In this section, I numerically study some implications of the model in stationary equilibrium. I am particularly interested in studying the effect of dealer competition, through the meeting rate of new dealers α, on measures of market liquidity.

4.3.1 Dispersion

I start by studying the dispersion in asset holdings. In a perfectly liquid market, where there is no implicit cost for readjustment, the dispersion should be maximal as agents choose assets only based on their current preferences. However, with frictions to trade, agents will smooth more between states, and the dispersion will be lower. In Figure 6, I compute the variance of asset holdings in the population and plot it against α. We can observe that as α increases, the dispersion in asset holdings increases, as anticipated. The increase is exponential, capturing the exponential effect of increases in the meeting rate of new dealers. Figure 7 proposes another interesting measure of liquidity, capturing the same idea: the average absolute difference in asset holdings between two agents with the same network but different preferences. Again, adjustments increase with α.

4.3.2 Bid-Ask Spread

A traditional measure of market liquidity is the bid-ask spread. In my model, the bid-ask spread is determined by the fee paid by the investor. The fee is given by:

$$\phi(s_1, s_2, a) = (s_1 - s_2)\left(V_k(s_1, s_1, a_k^*(s_1, s_2)) - V_k(s_1, s_2, a) - p(a_k^*(s_1, s_2) - a)\right).$$

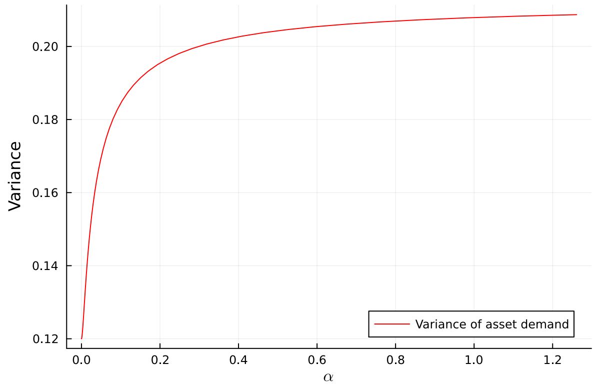

Figure 6: Variance in asset holdings.

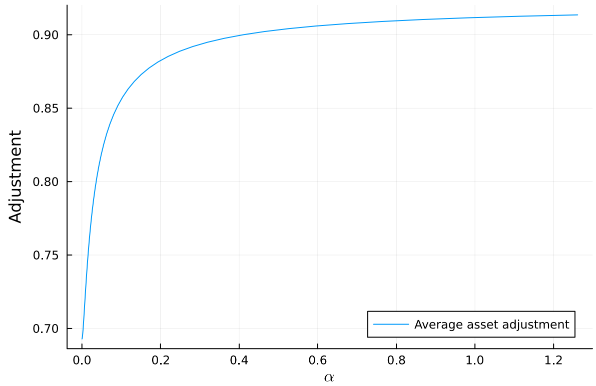

Figure 7: Average adjustment in asset holdings.

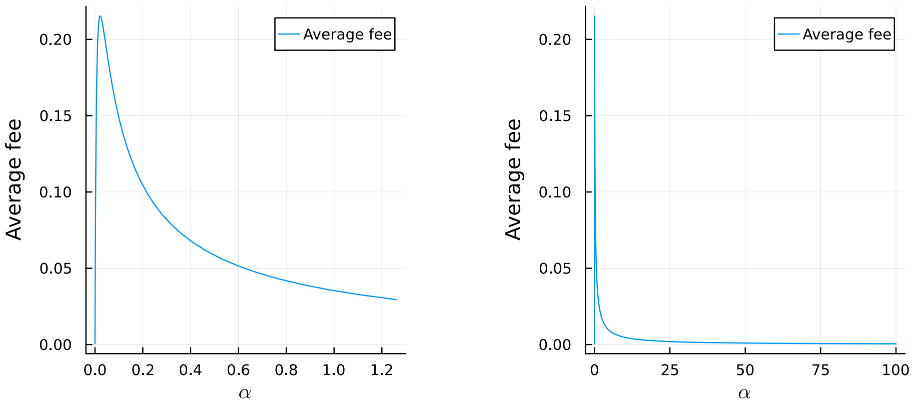

Figure 8: Average fee paid by investors.

The exact fee is complicated to compute, as it depends on the gains from trade. However, it is always a percentage of gains from trades. The size of this percentage increases with the *difference* in speed between the two fastest dealers (i.e., the current speed and the outside option). In Figure 8, I plot the average percentage of gains from trade paid by investors in the population as a function of α. The left-hand side shows fees for a reasonable rate α. We can observe that the fee generally decreases with α, capturing improved liquidity. However, there is a sharp increase at the beginning for very low values of α. At a very low value of α, the stationary distribution will concentrate on the lowest speed, where the fee is 0. The right-hand side increases sizes for α and illustrates the convergence of the fee towards 0 at very high rates.

It should be noted, however, that this is only a partial view, as the fee is also determined by the size of gains from trade, which is determined in equilibrium and depends on α. I haven't been able to formally solve this issue yet. I anticipate that gains from trades should be higher if dispersion is higher—i.e. when agents adjust their asset holdings more. This would counteract the effect mentioned above and make the fee increase with α. It is therefore likely that the fee is non-monotonic in α.

5 Conclusion

I present a dynamic model of relationship dynamics in OTC markets. My model describes the dynamics of relationships in detail, including switching rates and the evolution of fees within and across dealers for a given investor. These dynamics are interesting in themselves and can also be used as moments to fit data. A calibrated model would be useful for studying the more relevant question of heterogeneity in access to markets and its implications for market liquidity. As described in the last section, the heterogeneous access to the market created by relationship trading implies that otherwise similar agents will have different asset holdings depending on their dealers. This heterogeneity makes the market more or less liquid for different agents, an important dimension of market liquidity not captured by traditional measures. Finally, I show how the ease of finding new dealers increases overall market liquidity, suggesting that studying the durability of relationships is crucial to understanding market liquidity.

My model is a first step in this direction. In future work, I plan to extend the model in several ways. First, an interesting and not necessarily complicated extension would make dealer speeds endogenous, depending on how many trades they accumulate. My model suggests that trade concentrates on a small part of the dealer distribution, indicating that a few dealers accumulate most of the trade, suggesting a role for granularity. An interesting question would be to study what happens if those dealers face negative shocks (e.g., foreclosure), how investors reallocate to other dealers, and the impact on price and liquidity. To tackle such questions, I may need to consider a finite number of dealers. Finally, it would also be very interesting to study all of these questions if trade is bilateral rather than going through dealers.

References

Achdou, Yves et al. (Jan. 2022). "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach." In: *The Review of Economic Studies* 89.1, pp. 45–86.

Afonso, Gara, Anna Kovner, and Antoinette Schoar (May 2013). *Trading Partners in the Interbank Lending Market*. SSRN Scholarly Paper. Rochester, NY.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin (2006). "Wage Bargaining with On-the-Job Search: Theory and Evidence." In: *Econometrica* 74.2, pp. 323–364.

Di Maggio, Marco, Amir Kermani, and Zhaogang Song (May 2017). "The Value of Trading Relations in Turbulent Times." In: *Journal of Financial Economics* 124.2, pp. 266–284.

Duffie, Darrell, Nicolae Gârleanu, and Lasse Heje Pedersen (2005). "Over-the-Counter Markets." In: *Econometrica* 73.6, pp. 1815–1847.

Fontaine, François, Julien Martin, and Isabelle Mejean (2023). *Frictions and Adjustments in Firm-to-Firm Trade.*

Hendershott, Terrence et al. (2020). "Relationship Trading in Over-the-Counter Markets." In: *The Journal of Finance* 75.2, pp. 683–734.

Hugonnier, Julien, Benjamin Lester, and Pierre-Olivier Weill (2022). "Heterogeneity in Decentralized Asset Markets." In: *Theoretical Economics* 17.3, pp. 1313–1356.

Kaas, Leo and Philipp Kircher (Oct. 2015). "Efficient Firm Dynamics in a Frictional Labor Market." In: *American Economic Review* 105.10, pp. 3030–3060.

Lagos, Ricardo and Guillaume Rocheteau (2009). "Liquidity in Asset Markets With Search Frictions." In: *Econometrica* 77.2, pp. 403–426.

Lester, Benjamin, Guillaume Rocheteau, and Pierre-Olivier Weill (2015). "Competing for Order Flow in OTC Markets." In: *Journal of Money, Credit and Banking* 47.S2, pp. 77–126.

Li, Dan and Norman Schürhoff (2019). "Dealer Networks." In: *The Journal of Finance* 74.1, pp. 91–144.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (Jan. 2016). "Matching, Sorting and Wages." In: *Review of Economic Dynamics*. Special Issue in Honor of Dale Mortensen 19, pp. 63–87.

Pavliotis, Grigorios A. (2014). *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations*. Vol. 60. Texts in Applied Mathematics. New York, NY: Springer.

Sambalaibat, Batchimeg (May 2022). *Endogenous Specialization and Dealer Networks*. SSRN Scholarly Paper. Rochester, NY.

Weill, Pierre-Olivier (Aug. 2020). "The Search Theory of Over-the-Counter Markets." In: *Annual Review of Economics* 12.Volume 12, 2020, pp. 747–773.