Meeting Technologies in Decentralized Asset Markets*

Maryam Farboodi† Princeton University

Gregor Jarosch‡ Princeton University

Robert Shimer§ University of Chicago

September 9, 2015

Abstract

We study decentralized trading networks where agents differ in both their time-varying taste for an asset and the constant frequency at which they meet others. We show that "fast" agents can endogenously arise as intermediators whose net valuation of an asset gets moderated through their exposure to others. We also show that allocating meetings in an ex-ante asymmetric fashion across agents generates higher welfare then a homogeneous distribution of meeting frequencies.

1 Introduction

This paper studies a search model of over-the-counter asset markets where agents are heterogeneous in terms of their meeting technology. As is standard in the literature, we focus on an environment where the underlying reason for the trade of an asset are differences in the current flow utility agents receive from holding the asset. Yet in our framework, agents differ along a second dimension, namely in terms of how often they meet others.

As in Duffie et al. (2005), Vayanos and Wang (2007), Weill (2008), Hugonnier et al. (2014), Shen and Yan (2014) and Neklyudov (2013) the (different) flow utilities for the asset captures differential liquidity needs or hedging motives of individuals. The meeting technology of an agent captures how effectively he can find trading partners, and we refer to it alternatively as "speed". By speed we mean a feature that reduces the time between the occurrence of a desire to trade and the execution of the trade.[1] This notion includes not only how interconnected an agents is, but also all technological innovations that make trading more convenient and more reliable, such as efficient data feeds, user-friendly software and reliable hardware.

*Preliminary and Incomplete – Not for Circulation.

†farboodi@princeton.edu

‡gjarosch@princeton.edu

§robert.shimer@gmail.com

[1]Which also depends on the specifics of the aggregate meeting technology.

There is ample evidence that there is a large heterogeneity in how interconnected financial institutions are, and frequencies at which they trade. The ones who are able to trade with other firms more rapidly are more willing to take the opposite side of a given trade as they are able to offload any given position faster or buy a given asset more quickly. As a result they effectively they act as “market makers” and provide liquidity. However, this liquidity provision comes at a cost of investing in a better meeting technology, for instance maintaining many interconnections, or investment in more efficient trading software and hardware, that can potentially lead to an arms race among financial institutions.

We ask two key normative questions in the paper. First, we ask for the efficient trading pattern when agents differ along these two dimensions. That is, we ask how much value the planners associates with an agent holding the asset which dictates the direction of trades. Second, we ask for the efficient distribution of meetings across agents when holding the aggregate meeting technology fixed.

Our findings are the following. First, under parameter restrictions, the efficient trading pattern is such that the fast agents act as intermediators, or market makers. That is, if two agents with a low current flow value of holding the asset meet, a planner dictates that the slow agent transfers the asset to the fast one. The reason is simple: The fast agent is more likely to meet yet another agent with a currently high taste for the asset. In turn, if two agents with a high flow value meet, a planner dictates that the fast agent transfers the asset to the slow one. The reason is that the fast agent is more likely to meet another agent with currently low taste for the asset so she can replace the asset at a faster pace. Importantly, such a trading pattern gives rise to an endogenous core-periphery structure, where the fast agents are responsible for more than proportional (to their speed) share of the volume of trades. To see this note that the above trading pattern implies that the social value of asset holdings is more ”moderate” for agents with higher speed. That is, when ordering agents according to the value a planner associates with them holding the asset slow agents are at the extremes. When they have a high flow value they are at the top of the ordering. When having low value, they are at the bottom. In turn, faster agents remain in the middle of the ordering even as their flow valuation is subject to the same iid taste shocks.[2] This implies that fast agents endogenously emerge as intermediators whereas slow agents sit at the extremes of the intermediation chain.

Next, we study the optimal allocation of “speed” across agents. Specifically, we endow a social planner with an aggregate meeting technology that governs the average flow rate of bilateral meetings in the economy. We then study the optimal time-invariant distribution of meeting rates across agents holding the aggregate meeting technology constant. We find that social planner prefers to have some asymmetry in meeting technology among agents, but he does not want them to be too asymmetric. To show this, we first contrast an economy with homogeneous speed with an extremely asymmetric economy where a subset of agents is effectively in autarky and the remainder have accordingly higher meetings rates and show

[2]This might be most easily understandable when considering an agents whose meeting rate is extremely fast. The value a planner associates with that agent holding the asset should be close independent of her current flow valuation simply because she finds another trading partner at such high pace.

that the symmetric economy has higher welfare. Yet we show next that the symmetric economy is dominated in terms of welfare by an asymmetric one where a subset of agents are slightly faster and the remainder accordingly slower. The intuition for this result is the following. In a symmetric world, the asset solely flows from low-taste to high-taste types. In an asymmetric world, this still happens yet there is additional trades that enhance welfare, namely trades across agents with identical taste. Finally, we show that the optimal distribution of speed across agents is continuous, that is we rule out mass points from the planners solution. The logic behind the result is similar to the result that rules out a symmetric solution: One can "split" any mass of agents with homogeneous speed into two groups with marginally higher and lower speed leaving their interactions with the rest of the economy unchanged yet yielding new gainful trading opportunities across the two subgroups.

Literature Review. Our trading model follows the recent literature on dynamic trading with search friction, initiated by seminal work of Duffie et al. (2005), followed by Lagos and Rocheteau (2009), Weill (2008), Hugonnier et al. (2014). As in these models, we focus on a market where trade happens in bilateral meetings where agents bargain over the terms of trade, such as in OTC derivatives markets. Moreover, investor valuations for the asset changes randomly over time. However, in all of the above models agents are either symmetric in how frequently they meet others, or there is an endogenously given market making sector who facilitate trade among agents, as in Duffie et al. (2005) and Neklyudov (2013). We add to this literature by showing how market makers arise endogenously who are willing to take the opposite side of trade from other investors and provide liquidity to the market.

Chang and Zhang (2015) provides an alternative model of intermediation, where agents differ on how volatile their taste for asset is, and agents with more moderate types act as intermediators. Compared to their model, we take a step further back and ask why some agents have more moderate taste for holding an asset compared to others. We find that being able to meet other market participants more frequently implies a higher option value of search for an agent, which in turn moderates his taste and make him an intermadiator. Finally, Pagnotta and Philippon (2015) also considers the effect of differential speed on efficiency of financial markets. Unlike our model, they focus on a centralized model of trade so bargaining and search frictions play no roe in their model.

The rest of the paper is organized as follows:

2 Model

Time is continuous, $t \in \mathbb{R}_+$. There is one type of asset in strictly positive net supply $M \in (0,1)$, and a measure one of heterogeneous agents with type $\theta = (\delta_s, \lambda) \in \Theta$. Agents are risk neutral and live forever. The first dimension, $\delta_s \in [\underline{\delta}, \bar{\delta}]$, captures an agent's current taste, or flow utility, for the asset, and is drawn identically and independently at random from a known distribution, to be shortly specified, at rate γ. In our analytical derivations we focus on a taste distribution with two tastes, $s \in \{l, h\}$, $\delta_l < \delta_h$.[3] Let f_l ($f_h = 1 - f_l$) denote the fraction of low (high) taste agents in the population.

Agent's asset holdings are restricted to $\{0, 1\}$, so at every instant of time, a fraction M of agents hold the asset, called the *owners* and a fraction $(1 - M)$ of agents are *non-owners*.

The second dimension of heterogeneity, $\lambda \in [\underline{\lambda}, \bar{\lambda}]$ determines the rate at which agents meet each other. λ is time invariant. Denote the distribution of "speed" λ in the population by $G(\lambda)$, with corresponding pdf $dG(\lambda) = g(\lambda)d\lambda$.

Moreover, let $\mu_{s,m}(\lambda)(t)$ be the fraction of individuals with speed λ, who are in preference state s and have asset holdings $m \in \{0, 1\}$ at time t; $\sum_{m=0,1} \int_s \mu_{s,m}(\lambda)(t)ds = 1, \forall\ \lambda,\ t$. We focus on steady state, so we will drop the time index t for the most part of the paper.

2.1 Matching Technology

The matching technology determines how frequently an agent of type θ meets another agent θ' as a function of their own type and the type distribution in the economy. Agents match via a generalization of telegraph matching function, allowing for heterogeneity. The rate at which an individual $\theta = (\delta_s, \lambda)$ meets some other individual is simply given by λ. For consistency, it must then be the case that she is more likely to meet an agent with a high λ. For instance, conditional on meeting someone, the probability of meeting some asset owner is given by

$$\int \frac{\lambda'}{\Lambda} \left(\sum_{s'=l,h} \mu_{s',1}(\lambda') \right) dG(\lambda')$$

where Λ captures the average speed,

$$\Lambda = \sum_{m=0,1} \int \lambda' \left(\sum_{s'=l,h} \mu_{s',m}(\lambda') \right) dG(\lambda')$$

Likewise, conditional on meeting someone, the probability of meeting an asset non-owner with low taste is given by

$$\int \frac{\lambda'}{\Lambda} \mu_{l,0}(\lambda') dG(\lambda')$$

Thus, we work with a matching technology where each agents meeting rate is linear in her own speed λ but she draws partners according to their speed relative to the population. It is easy to verify that this collapses to the standard telegraph matching function if λ is identical across agents.

[3]This is sufficiently general for our purposes as two taste shocks are sufficient to capture all the trade patterns that arise with a continuous taste shock distribution.

3 Example

We start with a simple but illustrative example. Suppose there are two preference states, h and l. Individuals in state h get one util each instant that they hold the asset. Individuals in state l and individuals not holding the asset get flow utility zero. Individuals switch preference states according to a Poisson process with arrival rate γ, which we normalize to $\gamma = 1$.[4] The aggregate endowment of the asset is $M = 1/2$; this means that there is just enough of the asset for every individual in state H to hold it. However, because of search frictions, this cannot happen. Instead, assume individuals meet in pairs according to the symmetric matching technology.

First suppose all individuals meet another individual in the economy at a rate λ that is common across all individuals in the economy. Let $\mu_{s,m}$ denote the fraction of individuals in preference state s with m units of the asset in steady state. Assuming individuals trade if and only if an individual in the high state without the asset meets an individual in the low state with the asset, this solves the following system of equations:

$$
\begin{aligned}
(\gamma + \lambda\mu_{h,0})\mu_{l,1} &= \gamma\mu_{h,1} \\
\gamma\mu_{l,0} &= (\gamma + \lambda\mu_{l,1})\mu_{h,0} \\
\mu_{h,0} + \mu_{l,0} &= 1 - M \\
\mu_{h,1} + \mu_{l,1} &= M
\end{aligned}
$$

The first equation equates the inflows and outflows from the state $(s, m) = (l, 1)$. Individuals exit this state when they are in state $(l, 1)$ and either experience a preference shock (rate γ) or meet an individual in state $(h, 0)$ and trade. They enter this state when they are in state $(h, 1)$ and have a preference shock. The second equation equates inflows and outflows from state $(l, 0)$. The third notes that a fraction $1 - M$ of the population doesn't hold the asset and the fourth that the complementary fraction holds the asset.

Solve this using $M = 1/2$ and $\gamma = 1$ to get

$$
\mu_{h,0} = \mu_{l,1} = \frac{\sqrt{2+\lambda} - \sqrt{2}}{\sqrt{2}\lambda},
$$

with $\mu_{h,1} = \mu_{l,0} = 1/2 - \mu_{h,0}$. A notion of inefficiency is the size of $\mu_{h,0}$, the fraction of individuals who would value the asset but do not hold it. This is always positive, because of the search frictions, and declines as the arrival rate of meetings grows. For large values of λ, this is well-approximated by $1/\sqrt{2\lambda}$, and in particular the mismatch rate declines with the square root of the matching efficiency.

We now consider how intermediation may improve efficiency in this economy. Our key assumption is that the average meeting rate λ is fixed, but we do not distribute the individual meetings uniformly across the population. Instead, we give half the meetings to a small fraction ε of the population (the intermediaries) and the other half of the meetings to the rest

[4] This is a slight deviation from the rest of the paper where we assume that individuals take an i.i.d. taste draw at rate γ.

of the population. Moreover, search cannot be directed towards or away from intermediaries. Conditional on meeting someone, there is a fifty percent probability that the individual is an intermediary and a fifty percent probability that they are a regular trader. Depending on trading patterns, the individual may or may not be holding the asset. For small values of ε, intermediaries make many more contacts per unit of time than do regular traders, but we stress that the total number of meetings per person in the economy is fixed at λ per unit of time.

We also assume, as their name suggests, that intermediaries intermediate trade when they meet a regular trader. That means that if the intermediary holding the asset meets a regular trader who wants the asset, in state $(h, 0)$, he gives him the asset, regardless of the intermediaries preferences. If an intermediary not holding the asset meets a regular trader who does not want the asset, in state $(l, 1)$, he takes the asset, regardless of the intermediaries preferences. Meetings between two intermediaries or two regular traders are as above, with trades occurring according to preferences.

Now let λ_1 and λ_2 denote the trading speed of regular traders and intermediaries, with $(1-\varepsilon)\lambda_1 = \varepsilon\lambda_2 = \lambda/2$, so in fact half the meetings are with each type of trader. Also let $\mu_{s,m}(\lambda_1)$ denote the fraction of the population that is a regular trader in state (s, m) and $\mu_{s,m}(\lambda_2)$ denote the corresponding fraction that is an intermediary in that state. We can write the steady state equations for each of the eight states and types of traders. We focus here on the regular traders. First, consider a regular trader in state $(l, 1)$:

$$\left(\gamma + \lambda_1\left(\frac{\lambda_1\mu_{h,0}(\lambda_1) + \lambda_2(\mu_{h,0}(\lambda_2) + \mu_{l,0}(\lambda_2))}{\lambda}\right)\right)\mu_{l,1}(\lambda_1) = \gamma\mu_{h,1}(\lambda_1).$$

The left hand side is the rate of exiting that state. This can happen either because of a preference shock, at rate γ, or a meeting, at rate λ_1. A fraction $\lambda_1\mu_{h,0}(\lambda_1)/\lambda$ of the meetings are with another regular trader who is in state $(h, 0)$. A fraction $\lambda_2(\mu_{h,0}(\lambda_2) + \mu_{l,0}(\lambda_2))/\lambda$ of the meetings are with intermediaries who don't hold the asset. All these meetings, and no other meetings, result in trade. The right hand side is the rate of entering this state. This happens only with a preference shock. The steady state equation for a regular trader in state $(h, 0)$ is symmetric:

$$\left(\gamma + \lambda_1\left(\frac{\lambda_1\mu_{l,1}(\lambda_1) + \lambda_2(\mu_{l,1}(\lambda_2) + \mu_{h,1}(\lambda_2))}{\lambda}\right)\right)\mu_{h,0}(\lambda_1) = \gamma\mu_{l,0}(\lambda_1).$$

For regular traders with the correct asset holdings, the expressions are a bit different:

$$\gamma\mu_{l,0}(\lambda_1) = \gamma\mu_{h,0}(\lambda_1) + \lambda_1\left(\frac{\lambda_1\mu_{h,0}(\lambda_1) + \lambda_2(\mu_{h,0}(\lambda_2) + \mu_{l,0}(\lambda_2))}{\lambda}\right)\mu_{l,1}(\lambda_1)$$
$$\gamma\mu_{h,1}(\lambda_1) = \gamma\mu_{l,1}(\lambda_1) + \lambda_1\left(\frac{\lambda_1\mu_{l,1}(\lambda_1) + \lambda_2(\mu_{l,1}(\lambda_2) + \mu_{h,1}(\lambda_2))}{\lambda}\right)\mu_{h,0}(\lambda_1).$$

The first equates the outflow and inflow from state $(l, 0)$ and the second from state $(h, 1)$ for a regular trader.

The steady state equations for an intermediary are a bit different, since intermediaries sometimes trade even when they are holding the asset that they want:

$$\left(\gamma+\lambda_2\left(\frac{\lambda_1\mu_{h,0}(\lambda_1)+\lambda_2\mu_{h,0}(\lambda_2)}{\lambda}\right)\right)\mu_{l,1}(\lambda_2)=\gamma\mu_{h,1}(\lambda_2)+\lambda_2\frac{\lambda_1\mu_{l,1}(\lambda_1)}{\lambda}\mu_{l,0}(\lambda_2)$$
$$\left(\gamma+\lambda_2\left(\frac{\lambda_1\mu_{l,1}(\lambda_1)+\lambda_2\mu_{l,1}(\lambda_2)}{\lambda}\right)\right)\mu_{h,0}(\lambda_2)=\gamma\mu_{l,0}(\lambda_2)+\lambda_2\frac{\lambda_1\mu_{h,0}(\lambda_1)}{\lambda}\mu_{h,1}(\lambda_2)$$
$$\left(\gamma+\lambda_2\frac{\lambda_1\mu_{l,1}(\lambda_1)}{\lambda}\right)\mu_{l,0}(\lambda_2)=\gamma\mu_{h,0}(\lambda_1)+\lambda_2\left(\frac{\lambda_1\mu_{h,0}(\lambda_1)+\lambda_2\mu_{h,0}(\lambda_2)}{\lambda}\right)\mu_{l,1}(\lambda_2)$$
$$\left(\gamma+\lambda_2\frac{\lambda_1\mu_{h,0}(\lambda_1)}{\lambda}\right)\mu_{h,1}(\lambda_2)=\gamma\mu_{l,1}(\lambda_1)+\lambda_2\left(\frac{\lambda_1\mu_{l,1}(\lambda_1)+\lambda_2\mu_{l,1}(\lambda_2)}{\lambda}\right)\mu_{h,0}(\lambda_2).$$

For example, the last equation states that an intermediary with the asset in the high state can exit because of a preference shock or because of meeting a regular trader in the same preference state without the asset. An intermediary enters the high state with the asset either through a preference shock or by meeting someone in the low state without the asset.

We close the model by noting that the total number of each type of trader is given exogenously,

$$\sum_{s,m}\mu_{s,m}(\lambda_1)=1-\varepsilon \text{ and } \sum_{s,m}\mu_{s,m}(\lambda_2)=\varepsilon,$$

and the total supply of the asset is given by M:

$$\sum_{s,t}\mu_{s,1}^t=M.$$

We can then solve this system of equations for the number of individuals in each state.

Using the normalization $\gamma=1$ and setting the total asset supply to $M=\frac{1}{2}$, we can solve the steady state equations explicitly. The resulting solution is messy, and so we focus here just on the limiting behavior as the number of intermediaries converges to zero. This gives

$$\mu_{h,0}(\lambda_1)=\mu_{l,1}(\lambda_1)=\sqrt{\frac{16}{\lambda^2}+\frac{4}{\lambda}+\frac{1}{16}}-\frac{1}{4}-\frac{4}{\lambda},$$

with $\mu_{h,1}(\lambda_1)=\mu_{l,0}(\lambda_1)=1/2-\mu_{h,0}(\lambda_1)$. The $\mu_{s,m}(\lambda_2)$ all converge to zero because there are no intermediaries. For large values of λ, the measure of individuals in the high preference state without the asset is well approximated by $\mu_{h,0}(\lambda_1)\approx 4/\lambda$. The speed of converge in the economy without intermediation was only proportional to $\sqrt{\lambda}$, so intermediation raises the speed that asset holds converges towards the frictionless limit from $\sqrt{\lambda}$ to λ convergence.

Why does intermediation improve the quality of matching? A key observation is that intermediaries are less likely to hold their desired asset position than are regular traders. The fact that they have many more meetings is outweighed by the fact that they frequently trade away from their desired asset position. In fact, although the number of intermediaries vanishes in the limit as ε converges to zero, the share of intermediaries holding the undesired asset position is well-behaved in the limit:

$$\frac{\mu_{h,0}(\lambda_2)}{\mu_{h,0}(\lambda_2)+\mu_{h,1}(\lambda_2)} = \frac{\mu_{l,1}(\lambda_2)}{\mu_{l,1}(\lambda_2)+\mu_{l,0}(\lambda_2)} = \frac{1}{2}+\frac{8}{\lambda}-\sqrt{\frac{64}{\lambda^2}+\frac{16}{\lambda}+\frac{1}{4}}+\frac{4}{\sqrt{\lambda}}\sqrt{1+\frac{8}{\lambda}-\sqrt{\frac{64}{\lambda^2}+\frac{16}{\lambda}+\frac{1}{4}}}$$

For large λ, this is well approximated by $\sqrt{8/\lambda}$. This highlights that intermediation is useful because it makes it easier for the regular traders to obtain the desired asset holdings, not because intermediaries themselves have preferences well-aligned with their asset holdings.

This example illustrates the potential usefulness of intermediation, but opens as many questions as it resolves. First, in this example, intermediation is only valuable when λ is large. For $\lambda < 6$, the symmetric model with pairwise exchange is more efficient than the outcome with intermediation. Is intermediation only valuable in economies with moderate search frictions? We find that there is always some value to a small amount of intermediation, but when λ is small, the intermediaries should optimally get less than half of the total search intensity. Only in the limit as λ converges to infinity is it optimal to give exactly half the aggregate search intensity to the intermediaries.

Second, we have assumed a particular pattern of trades, with the intermediaries always respecting the preferences of the regular traders. But one could imagine other trading patterns. We find that this trading pattern is optimal when the supply of the asset is equal to the number of traders in the high preference state, but other trading patterns may be optimal more generally.

Third, we have assumed that there are only two speeds of traders. Is there a value to having more speeds? We find numerically that the answer is yes. Indeed, our preliminary results suggest that for a given average number of meetings λ, the efficiency of exchange is maximized if no two traders have the same meeting efficiency. This ensures that there are gains from trade in meetings even when two individuals have the same preferences, since assets can be transferred to take advantage of trading efficiency.

Fourth, can these outcomes be decentralized, particularly in an environment where individuals must choose their speed ex ante. Superficially, there is some tension coming from the fact that intermediaries are less likely to hold their desired asset than are regular traders. How then can one convince some traders to invest in becoming intermediaries if this hurts their position in the asset market? The answer must come through prices. Intermediaries are compensated not by their ability to hold the desired asset position, but by their ability to buy low and quickly turn around and sell high. Bid-ask spreads must be key to any decentralization of this outcome.

4 Normative Analysis

This section studies two normative questions. First, we ask for the optimal trading pattern, that is we ask how a planner ranks agents across the two dimensions of heterogeneity. This ranking, which determines the optimal trading pattern, reflects the social value the planner associates with an agent (s,λ) holding the asset. We set up the problem for a continuous distribution of speed and discuss the optimal trading pattern in a numerical example. In order to make analytical progress we then restrict attention to a simpler two-by-two case. The second subchapter studies the optimal distribution of speed.

4.1 The Optimal Trading Pattern - Planner's Problem

In the entire section we focus on two tastes, $s \in \{l,h\}$. A social planner seeks to maximize the discounted sum of utilities,

$$\int e^{-\rho t}\left(\sum_s \delta_s \left(\int \mu_{s,1}(\lambda)(t)dG(\lambda)\right)\right)dt, \tag{1}$$

Note that

$$\sum_s \Big(\int \mu_{s,1}(\lambda)(t)dG(\lambda)\Big) = M$$

As a result

$$\int \mu_{l,1}(\lambda)(t)dG(\lambda) = M - \int \mu_{h,1}(\lambda)(t)dG(\lambda) \tag{2}$$

Which simplifies equation 1 to

$$\int e^{-\rho t}\left[(\delta_h-\delta_l)\Big(\int \mu_{h,1}(\lambda)(t)dG(\lambda)\Big)+\delta_l M\right]dt,$$

which is equivalent to maximizing

$$\int e^{-\rho t}\left(\int \mu_{h,1}(\lambda)(t)dG(\lambda)\right)dt \tag{3}$$

or, in steady state,

$$\int \mu_{h,1}(\lambda)(t)dG(\lambda) \tag{4}$$

It follows that maximizing welfare in the environment with two different taste levels amounts to maximizing the share of asset owners that currently have high taste. This confirms the notion of inefficiency taken in chapter (3).

When two individuals meet and only one of them holds the asset, the planner must decide who holds the asset after the meeting. Let $N(s,\lambda|s',\lambda')$ denote the probability that (s,λ) holds the asset after a meeting with (s',λ'), so $N(s,\lambda|s',\lambda') = 1 - N(s',\lambda'|s,\lambda)$. Note that this does not depend on who holds the asset going into the meeting.

Let Λ denote the average speed in the economy. With this notation, we can write the law of motion for μ, suppressing the time dependence:

$$\begin{aligned}\dot{\mu}_{s,0}(\lambda) = \gamma_{\tilde{s}s}\mu_{\tilde{s},0}(\lambda) + \frac{\lambda}{\Lambda}\mu_{s,1}(\lambda)\int \lambda'\left(\sum_{s'}\mu_{s',0}(\lambda')N(s',\lambda'|s,\lambda)\right)dG(\lambda')\\ - \gamma_{s\tilde{s}}\mu_{s,0}(\lambda) - \frac{\lambda}{\Lambda}\mu_{s,0}(\lambda)\int \lambda'\left(\sum_{s'}\mu_{s',1}(\lambda')N(s,\lambda|s',\lambda')\right)dG(\lambda')\end{aligned}$$

and

$$\begin{aligned}\dot{\mu}_{s,1}(\lambda) = \gamma_{\tilde{s}s}\mu_{\tilde{s},1}(\lambda) + \frac{\lambda}{\Lambda}\mu_{s,0}(\lambda)\int \lambda'\left(\sum_{s'}\mu_{s',1}(\lambda')N(s,\lambda|s',\lambda')\right)dG(\lambda')\\ - \gamma_{s\tilde{s}}\mu_{s,1}(\lambda) - \frac{\lambda}{\Lambda}\mu_{s,1}(\lambda)\int \lambda'\left(\sum_{s'}\mu_{s',0}(\lambda')N(s',\lambda'|s,\lambda)\right)dG(\lambda')\end{aligned} \tag{5}$$

where $\tilde{s} = l\ (h)$ when $s = h\ (l)$. Adding these up imply that

$$\sum_m \dot{\mu}_{s,m}(\lambda) = \gamma_{\tilde{s}s}\sum_m \mu_{\tilde{s},m}(\lambda) - \gamma_{\tilde{s}s}\sum_m \mu_{s,m}(\lambda).$$

We assume throughout that the left hand side of this equation is zero, so preferences are in steady state all the time. Since $\sum_{s,n}\mu_{s,m}(\lambda) = 1$, this implies

$$\sum_m \mu_{s,m}(\lambda) = \frac{\gamma_{\tilde{s}s}}{\gamma_{\tilde{s}s}+\gamma_{s\tilde{s}}}.$$

We can write $\gamma_{ss'} = \gamma f_{s'}$, $\forall\ s, s'$, which would imply f_s fraction of agents have taste s.

Let $V(\mu)$ denote the value of the planner when the state is μ. Using subscripts to denote partial derivatives, we can write the value function as

$$\begin{aligned}\rho V(\mu) = \max_{N(s,\lambda|s',\lambda')}\sum_s \Bigg[\delta_s\left(\int \mu_{s,1}(\lambda)dG(\lambda)\right)\\ + \gamma\int\left(\sum_m V_{\mu_{s,m}(\lambda)}\big(\sum_{s'}\mu_{s',m}(\lambda)f_s - \mu_{s,m}(\lambda)f_{s'}\big)\right)dG(\lambda)\\ + \int\left(\big(V_{\mu_{s,0}(\lambda)} - V_{\mu_{s,1}(\lambda)}\big)\frac{\lambda}{\Lambda}\mu_{s,1}(\lambda)\int\lambda'\left(\sum_{s'}\mu_{s',0}(\lambda')N(s',\lambda'|s,\lambda)\right)dG(\lambda')\right)dG(\lambda)\\ + \int\left(\big(V_{\mu_{s,1}(\lambda)} - V_{\mu_{s,0}(\lambda)}\big)\frac{\lambda}{\Lambda}\mu_{s,0}(\lambda)\int\lambda'\left(\sum_{s'}\mu_{s',1}(\lambda')N(s,\lambda|s',\lambda')\right)dG(\lambda')\right)dG(\lambda)\Bigg]\end{aligned}$$

The first derivative with respect to a typical value of $N(s,\lambda|s',\lambda') = 1 - N(s',\lambda'|s,\lambda)$ is

$$\big(V_{\mu_{s,1}(\lambda)} + V_{\mu_{s',0}(\lambda')} - V_{\mu_{s,0}(\lambda)} - V_{\mu_{s',1}(\lambda')}\big)\big(\mu_{s,0}(\lambda)\mu_{s',1}(\lambda') + \mu_{s,1}(\lambda)\mu_{s',0}(\lambda')\big)\frac{\lambda\lambda'}{\Lambda}dG(\lambda)dG(\lambda').$$

The first term is the gain from allocating the asset optimally across the two agents. The remaining terms reflect the frequency of meetings between these two agents and are always nonnegative.

With some abuse of notation, let $V_{s,v}$ be the marginal *net* value the planner assigns to a taste s speed v agent holding the asset, $V_{s,v} = V_{\mu_{s,1}(\lambda_v)} - V_{\mu_{s,0}(\lambda_v)}$. It follows that the optimal control is bang-bang and so satisfies

$$N(s,\lambda|s',\lambda') = \begin{cases} 1 & \\ 0 & \end{cases} \quad \text{if } V_{s,\lambda} \gtrless V_{s',\lambda'}.$$

That is, when (s,λ) meets (s',λ') and exactly one of the two has the asset, then (s,λ) leaves the meeting with the asset if the marginal value to the planner of an additional (s,λ) with the asset and an additional (s',λ') without the asset exceeds the alternative configuration, an additional (s',λ') with the asset and an additional (s,λ) without the asset.

Thus, what determines the optimal pattern of trade is the marginal net values which we derive next.

4.1.1 Marginal Social Values

We can write the marginal social net values as[5]

$$V_{s,v} = \delta_s + \gamma\left(V_{s',v} - V_{s,v}\right) + \lambda_v S_{s,v} \tag{6}$$

where $S_{s,v}$ captures the *net* option value of meeting other agents,

$$\begin{aligned} S_{s,\lambda} = & \int \frac{\lambda'}{\Lambda}\left(\sum_{s'} \mu_{s',0}\left(\lambda'\right)\left(\max\left\{V_{s',\lambda'} - V_{s,\lambda}, 0\right\}\right)\right) dG\left(\lambda'\right) \\ & - \int \frac{\lambda'}{\Lambda}\left(\sum_{s'} \mu_{s',0}\left(\lambda'\right)\left(\max\left\{V_{s,\lambda} - V_{s',\lambda'}, 0\right\}\right)\right) dG\left(\lambda'\right) \end{aligned}$$

This expression captures option value of transferring the asset to an agent with higher $V_{s,v}$ net of the foregone option value of receiving the asset from an agent with lower $V_{s,v}$.

Note that equation (6) determines the socially optimal pattern of trade. That is, it maps the two dimensions of heterogeneity (s,v) into a hierarchy which captures which meetings should result in a transaction. Of course, $\mu_{s',0}$ is an endogenous object that depends on $V_{s,v}$ which makes solving for $V_{s,v}$ in closed form intractable. We next study several numerical example with a continuous distribution $G(\lambda)$ and then turn to a simpler two-by-two case to characterize the optimal trading pattern analytically.

[5]Here, we set $f_s = f_h = 1/2$. The generalization is straightforward.

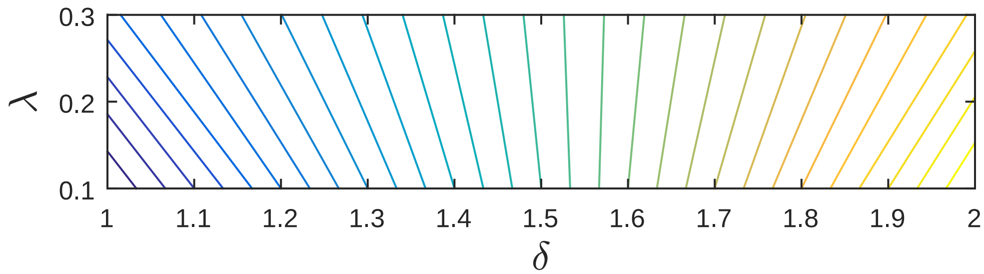

Figure 1: Social Iso-Net-Value Curves

4.1.2 Characterization - Numerical Example

This section briefly discusses several trading patterns that emerge in numerical examples. To that end, we solve (6)jointly with the associated steady state distribution of asset holdings which is determined by standard inflow-outflow equalities. The pattern which most commonly arises in numerical examples is plotted in figure 1:[6] We plot Iso-Value Curves across the taste-speed space, that is the planner associates the same net value with agents at different points on one of the lines holding the asset. Curves further to the right represent a higher valuation by the planner.

It follows that the following pattern of trade emerges: If two agents with a low δ meet the planner wants the asset to flow towards the fast one. The reason is that the latter is faster at intermediating the asset towards agents with a high flow valuation. In turn, if the two agents have a high δ, the planner requires that the asset flows to the slower type. The reason is that the faster agent has a larger option value of meeting low-valuation asset owners from whom she can take on the asset. It follows that fast agents endogeneously arise as intermediators in the efficient pattern of trade. That is, their "interconnectedness" dampens the volatility of the net valuation the planner associates with a fast agent holding the asset relative to slow agents, although their flow valuation follows exactly the same stochastic process

Note that the pattern in figure 1 corresponds to the one assumed in the numerical example in section 3. Fast individuals intermediate the asset from slow individuals with low current valuation to slow individuals with high current valuation.

It is important however, to point out, that this pattern is not always the efficient one. We find that two other patterns may arise. First, the planner may demand that the asset always flows to the faster agents as depicted in figure 2. This happens whenever the asset is scarce. That is, even among the high valutation agents the asset get traded towards the fast ones. If she experiences a negative taste shock, the agent with higher speed finds non-owners with high valuation faster. If the asset is very scarce this comes to dominate the fact that she is also faster at finding low type owners, simply because it is hard to find any asset owners.

Figure 3 depicts exactly the opposite situation. If the asset is very abundant it is hard to find non-owners of the asset with a currently high valuation and easy to find owners that have low flow value. Thus, for a given level of δ the planner requires the asset to be traded

[6]When generating the figures we use many different taste levels δ simply for expositional purposes.

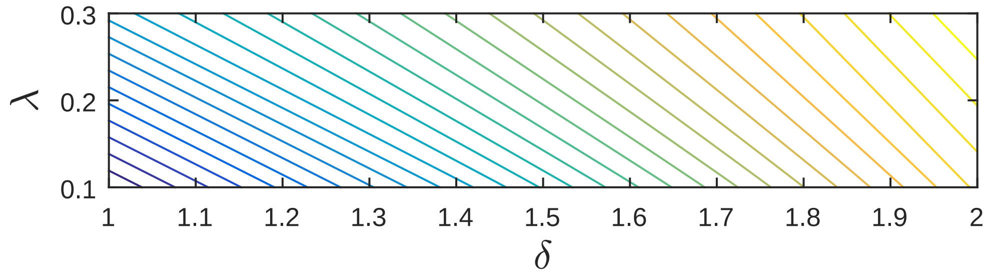

Figure 2: Social Iso-Net-Value Curves if Asset is Scarce

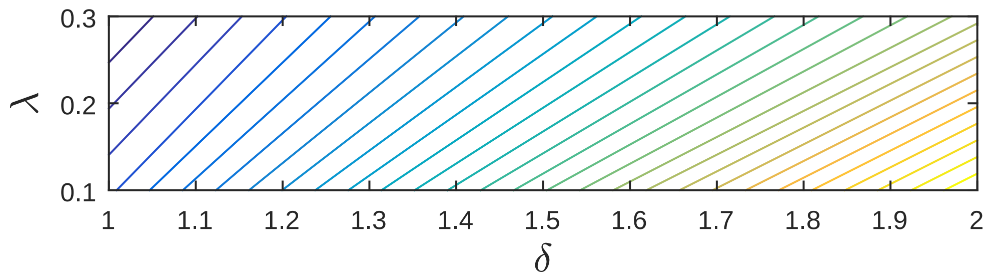

Figure 3: Social Iso-Net-Value Curves if Asset is Abundant

toward the slow agent. In the event of a taste shock, the agent with higher speed is faster at finding an owner with low flow value. If the asset is very abundant this comes to dominate the fact that she is also faster at finding high type non-owners, simply because it is hard to find any non-owners.

4.1.3 Characterization - Two-by-Two Case

Assume that the social planner can choose at most two levels of speed, λ_v, $v \in \{S, F\}$, $\lambda_S \leq \lambda_F$. Let g_S and $g_F = 1 - g_S$ denote the fraction of slow and fast agents, respectively.

In this section, take the two distinct speed levels, $\lambda_S < \lambda_F$, as well as distribution of agent speeds, $\{g_S, g_F\}$, as given. The goal is to characterizes the optimal trading rule which determines whether agent (δ', λ'), when owning the asset, hands the asset over to agent (δ, λ) who does not own the asset; $N(\delta, \lambda|\delta', \lambda') = 1$, or not.

We represent the different trading rules that might emerge as optimal as chains of inequalities that determine $N(\delta, \lambda|\delta', \lambda')$. Note that agents of identical speed always trade towards the agent with higher flow utility.

1. Intermediation

$$V_{l,S} < V_{l,F} < V_{h,F} < V_{h,S} \tag{7}$$

This trading pattern corresponds to the one where the fast agents act as intermediaries, as used in section 3 and plotted in figure 1.

2. Trade towards the fast agents

   This corresponds to figure 2 but spans 2 cases:

   (a)
   $$V_{l,S} < V_{l,F} < V_{h,S} < V_{h,F} \tag{8}$$
   At same flow value asset flows toward fast agent, but low fast sells to high slow.

   (b)
   $$V_{l,S} < V_{h,S} < V_{l,F} < V_{h,F} \tag{9}$$
   Asset flows toward fast agent even when a high slow owner meets a low fast non-owner: in this case fast agents are *"absorbing"* the asset and the asset is either held solely by the fast agents or all the fast agents hold an asset which they never trade and the remaining assets are traded between the slow types.

3. Trade towards the slow agents

   This corresponds to figure 3 but again spans 2 cases:

   (a)
   $$V_{l,F} < V_{l,S} < V_{h,F} < V_{h,S}$$
   At same flow value asset flows toward slow agent, but low slow sells to high fast.

   (b)
   $$V_{l,F} < V_{h,F} < V_{l,S} < V_{h,S}$$
   Asset flows toward slow agent even when a high fast owner meets a low slow non-owner: in this case slow agents are *"absorbing"* the asset.

The following proposition characterizes the optimal trading rule when the total supply of asset is equal to fraction of agents with high taste for the asset.

Proposition 1. *If $M = f_h$, the optimal trading pattern is given by the chain of inequalities (7).*

Proof. See Appendix □

Proposition 1 thus establishes that the planner requires the fast agents to take on a role as intermediaries. If agents with identical taste meet, the fast agents receives the asset if both agents have low taste and vice versa if both agents have high taste. Intuitively, this happens because of option value considerations: If agents have low idiosyncratic taste the gains from finding a non-owner with high taste are large and thus the planner wants the fast agent to acquire the asset. In turn, if agents have high idiosyncratic taste, the gains from finding an owner with low taste are large and the agents thus wants the slow agent to acquire the asset.

4.2 Optimal Speed Distribution

In this section we consider planner's optimal choice of speed distribution.

4.2.1 2 Speed Levels

The planner is endowed with some average speed level Λ. He chooses a speed distribution for a continuum of measure one of agents (with at most two speed levels) to maximize welfare as defined in equation (1), such that the average speed of population is Λ: $\int_\lambda \lambda dG(\lambda) = \Lambda$. The next proposition shows that the planner chooses some intermediate level of speed asymmetry among agents. Note that that the planner can choose any asymmetric speed distribution as long as the average speed remains constant Λ.

Proposition 2. *The welfare maximizing speed distribution requires an intermediate degree of heterogeneity, that is $\lambda_f^\star > \lambda_s^\star > 0$ and $g_s^\star > 0, g_f^\star > 0$.*

This proof proceeds in two steps. First we show that the planner prefers a symmetric distribution to an extremely asymmetric one where one group of agents is in autarky. Then we show that a perturbation of symmetric speed distribution improves welfare, which along with the first part completes the proof.

Step 1 Consider an average speed level λ and construct the following family of speed distributions. Choose a fraction x of the population and give it speed $\frac{\lambda}{x}$. Give the remaining $(1-x)$ fraction of population speed zero, so that average speed in population remains λ. This family of speed distributions is features "extreme asymmetry" in the sense that given (λ, x), the two groups of agents have speeds as different as possible.

Next, find the division of total amount of asset between the fast and slow (zero speed) sub-populations which maximizes total welfare, $x\mu_{h,1}(\frac{\lambda}{x}, m_1) + (1-x)\mu_{h,1}(0, m_2)$. In this division, m_1 denotes the per capita amount of asset own by non-zero speed agents, so $xm_1 + (1-x)m_2 = M$. Note that $x = 1$ corresponds to the symmetric speed distribution where every agent has speed λ.

Next, we show that among all members of this family (indexed by x), the symmetric distribution maximizes welfare (at the corresponding optimal division of asset). In other words

$$\hat{x} = \text{argmax}_x \max_{m1} \{x\mu_{h,1}(\frac{\lambda}{x}, m_1) + (1-x)\mu_{h,1}(0, m_2)\}. \tag{10}$$

$$\text{s.t. } xm_1 + (1-x)m_2 = M \tag{11}$$

We want to show that $\hat{x} = 1$. This is sufficient condition for the social planner not to choose an "extreme-asymmetric" speed distribution.

To do so, we solve the *unconstrained* version of problem 10 and show that $x = 1$ maximizes it. The relevant constraints are $m_1 > 0$, $m_1 \leq 1$ and $m_1 \leq \frac{M}{x}$ (i.e. $m_2 > 0$). Note that at $x = 1$ (symmetric equilibrium) none of these constraints bind, while for other x's they can potentially bind. So if $x = 1$ is the argmax of the unconstrained problem, it is also the argmax of the constrained problem. This completes the first step of the proof.

Step Two The next step is to show that a small pertubation to the the symmetric speed distribution, which makes some agents faster and some slower, holding the aggregate speed constant, is welfare-improving.

Strategy 1 We use the following approach: Conjecture a perturbation in which half of agents are fast and half are slow, $g = \{g_S, g_F\} = \{\frac{1}{2}, \frac{1}{2}\}$. Next consider the flow equations for μ's which characterize an "intermediation" trading rule as defined above. Impose $\lambda_F = \lambda_S = \lambda$ and compute the equilibrium stocks. In other words, endow all agents with same speed λ, but label half of them S and the other half F, and then force the S and F agents to trade in a fan trading pattern.

With some abuse of notation, let $\mu_{s,m}(F)$ ($\mu_{s,m}(F)$) denote the fraction of agents labeled F (S) who are in state s with ownership status m, while both groups have the same speed λ. The first thing to note is that

$$\frac{1}{2}\mu_{s,m}(S) + \frac{1}{2}\mu_{s,m}(F) = \mu_{s,m}(\lambda), \quad s \in \{l, h\},\ m \in \{0, 1\} \tag{12}$$

where $\mu_{s,m}(\lambda)$ is the corresponding fraction in the symmetric equilibrium.

Now to show that a intermediation trading pattern with asymmetric speed improves welfare, we need to show that

$$\left(\frac{\partial \mu_{h,1}(\lambda - \epsilon)}{\partial \epsilon} + \frac{\partial \mu_{h,1}(\lambda + \epsilon)}{\partial \epsilon}\right)|_{\epsilon=0} > 0. \tag{13}$$

To compute the above partial derivatives totally differential the system of equations that characterize the steady state stocks μ and evaluate it at $\epsilon = 0$. Equation 13 boils down to

$$\mu_{l,1}(F)\mu_{h,0}(F) > \mu_{l,1}(S)\mu_{h,0}(S) \tag{14}$$

as defined above. To show that this equation always holds, consider the following: assume the above exercise, except that after labeling agents as S and F we force all low agents to sell to all high agents and to no-one else (irrespective of their label). So the S and F populations will be identical, $\mu_{s,m}(F) = \mu_{s,m}(S) \quad \forall\ s,\ m$; which means the above inequality will hold with equality. Now enforce that low S owners sell to low F non-owners and high F owners sell to high S non-owners. As equation 12 holds as long as the two labeled populations have the same trading speed, we will have $\mu_{l,1}(F) > \mu_{l,1}(S)$. A symmetric argument for the high types implies $\mu_{h1}(F) < \mu_{h1}(S)$. Note that because of our choice of perturbation ($\frac{1}{2}$ of agents being each F and S), $\mu_{h,1}(F) + \mu_{h,0}(F) = \mu_{h,1}(S) + \mu_{h,0}(S) = \frac{1}{2}f_h$. This implies $\mu_{h,0}(F) > \mu_{h,0}(S)$ which along $\mu_{l,1}(F) > \mu_{l,1}(S)$ implies equation 14 and completes the proof.

The above two together show that social planner chooses some intermediate asymmetric distribution. So the planner does not want all the market participants to have the same speed, but he does not want some very fast and some extremely slow agents either.

Strategy 2 Consider a market where agents have two different tastes and a single homogeneous group. The following argues that redistributing meetings across two groups of agents such that the aggregate number of meetings is constant is welfare improving. The proof has two parts,

1. We show that the following is welfare neutral relative to the homogeneous case: In the heterogeneous case, maintain the assumption that the only transactions are from low types to fast types. That is, there is no trade between agents of identical taste yet different speed. To do so, note that we can write the inflow outflow equations under this trading pattern (for instance for the stock of slow-high non-owners) in the following form,[7]

$$\begin{aligned}&\gamma\big(\,(1-f_l)\,\mu_{l,0}(\lambda_S)\\&-f_l\mu_{h,0}(\lambda_S)\big)-\mu_{h,0}(\lambda_S)\frac{(\lambda-\varepsilon)}{\Lambda}\big(\mu_{l,1}(\lambda_S)\,(\lambda-\varepsilon)+\mu_{l,1}(\lambda_F)\,(\lambda+\varepsilon)\,\big)=0\end{aligned} \tag{15}$$

   Take the system of three inflow outflow equations, for $\mu_{h,0}(\lambda_S)$ as in (15) and the equivalents for $\mu_{l,0}(\lambda_S)$ and $\mu_{l,1}(\lambda_F)$, which jointly with 5 adding up constraints implicitly determine the equilibrium stocks. Totally differentiate that system of equations with respect to ε and evaluate at the symmetric equilibrium$\varepsilon=0$. We find that

$$-\frac{\partial\mu_{h,1}(\lambda_S)}{\partial\varepsilon}=\frac{\partial\mu_{h,1}(\lambda_F)}{\partial\varepsilon}>0$$

   That is, the share of high-type asset holders is unchanged. It follows that, locally, a pertubation around a symmetric equilibrium where agents of the same taste are not allowed to trade with each other is welfare neutral.

2. Show that this trading pattern is not efficient. Note that it implies

$$V_{l,S}=V_{l,F}<V_{h.s}=V_{h,F}$$

   From (16), it is clear that this cannot hold. Consider the equality $V_{l,S}=V_{l,F}$. Both low types have the same flow utility and the same expected gains from trade conditional on meeting high-type trading partners. Yet, fast types meet high-type trading partners more often. Thus, this trading pattern cannot be efficient and for that reason the efficient trading pattern must generate strictly higher welfare than the homogeneous equilibrium.

4.3 Ruling Out Mass Points

In general, an very similar proof should rule out any mass point: Suppose a mass point exists

[7]Note that here $\gamma_{ss'}=\gamma(1-f_{s'})\;\forall s,s'$.

1. Perturb the distribution of speed such that a share of the mass receives marginally higher speed and a share of the mass receives marginally lower speed, holding average speed constant. Do not let the two sets of agents that previously constituted the mass point trade with each other. Show that welfare is unchanged.

2. Show that not allowing the two sets of agents to engange in trade is strictly welfare enhancing for the same reasons as above.

References

Chang, B. and S. Zhang (2015). Endogenous market-making and formation of trading links. Technical report.

Du, S. and H. Zhu (2014). Welfare and optimal trading frequency in dynamic double auctions. Technical report, National Bureau of Economic Research.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005). Over-the-counter markets. *Econometrica 73*(6), 1815–1847.

Hugonnier, J., B. Lester, and P.-O. Weill (2014). Heterogeneity in decentralized asset markets. Technical report, National Bureau of Economic Research.

Lagos, R. and G. Rocheteau (2009). Liquidity in asset markets with search frictions. *Econometrica 77*, 403–426.

Neklyudov, A. V. (2013). Bid-ask spreads and the decentralized interdealer markets: Core and peripheral dealers. Technical report, Working Paper, University of Lausanne.

Pagnotta, E. and T. Philippon (2015). Competing on speed. Technical report, National Bureau of Economic Research.

Shen, J. and H. Yan (2014). A search model of the aggregate demand for safe and liquid assets. *Available at SSRN 2384895*.

Shimer, R. and L. Smith (2001). Matching, search, and heterogeneity. *Advances in Macroeconomics 1*(1).

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

Weill, P.-O. (2008). Liquidity premia in dynamic bargaining markets. *Journal of Economic Theory 140*(1), 66–96.

A Proof of Proposition 1

Here is a sketch of the proof. The following argument rules out case 2a when the share of high types is equal to the endowment of the asset. To do so, plug

$$(\rho+\gamma)\,V_{s',v}=\delta_{s'}+\gamma S_{s,v}+\lambda_v S_{s',v}$$

into equation (6) to get

$$\frac{\rho\,(\rho+2\gamma)}{\rho+\gamma}V_{s,v}=\delta_s+\frac{\gamma}{\rho+\gamma}\delta_{s'}+\lambda_v\left(S_{s,v}+\frac{\gamma}{\rho+\gamma}S_{s',v}\right)\tag{16}$$

For pattern (8) to hold we need that

$$V_{h,F}>V_{h,S}$$

and thus, from (16)

$$\lambda_F\left(S_{h,F}+\frac{\gamma}{\rho+\gamma}S_{l,F}\right)>\lambda_S\left(S_{h,S}+\frac{\gamma}{\rho+\gamma}S_{l,S}\right)$$

I next show that

$$S_{l,F}<-S_{h,F}$$

and

$$-S_{l,S}<S_{h,S}$$

which implies that inequality (8) is violated. To so, write the net option value of buying and selling that is implied by the above trading pattern,

$$S_{l,F}=\mu_{h,0}(\lambda_F)g_F\left(V_{h,F}-V_{l,F}\right)+\mu_{h,0}(\lambda_S)g_S\left(V_{h,S}-V_{l,F}\right)-\mu_{l,1}(\lambda_S)g_F\left(V_{l,F}-V_{l,S}\right)$$

and

$$S_{h,F}=-\mu_{h,1}\left(\lambda_S\right)g_S\left(V_{h,F}-V_{h,S}\right)-\mu_{l,1}\left(\lambda_F\right)g_F\left(V_{h,F}-V_{l,F}\right)-\mu_{l,1}\left(\lambda_S\right)g_S\left(V_{h,F}-V_{l,S}\right)$$

A sufficient condition for $S_{l,F}<-S_{h,F}$ is

$$\begin{gathered}\mu_{l,1}\left(\lambda_F\right)g_F\left(V_{h,F}-V_{l,F}\right)+\mu_{l,1}\left(\lambda_S\right)g_S\left(V_{h,F}-V_{l,S}\right)>\\ \mu_{h,0}\left(\lambda_F\right)g_F\left(V_{h,F}-V_{l,F}\right)+\mu_{h,0}\left(\lambda_S\right)g_S\left(V_{h,S}-V_{l,F}\right)\end{gathered}$$

The LHS is strictly larger than $\Big(\mu_{l,1}\left(\lambda_F\right)g_F+\mu_{l,1}\left(\lambda_S\right)g_S\Big)\left(V_{h,F}-V_{l,F}\right)$ and the RHS is strictly smaller than $\Big(\mu_{h,0}\left(\lambda_F\right)g_F+\mu_{h,0}\left(\lambda_S\right)g_S\Big)\left(V_{h,F}-V_{l,F}\right)$. Thus, if

$$\mu_{l,1}\left(\lambda_F\right)g_F+\mu_{l,1}\left(\lambda_S\right)g_S=\mu_{h,0}\left(\lambda_F\right)g_F+\mu_{h,0}\left(\lambda_S\right)g_S,$$

which is the case when the amount of asset is equal to the high types, we have established that $S_{l,F}<-S_{h,F}$.

Equivalently, we have that

$$S_{l,S} = \mu_{h,0}(\lambda_F) g_F (V_{h,F} - V_{l,S}) + \mu_{h,0}(\lambda_S) g_S (V_{h,S} - V_{l,S}) + \mu_{l,0}(\lambda_F) g_F (V_{l,F} - V_{l,S})$$

and

$$S_{h,S} = \mu_{h,0}(\lambda_F) g_F (V_{h,F} - V_{h,S}) - \mu_{l,1}(\lambda_S) g_S (V_{h,S} - V_{l,S}) - \mu_{l,1}(\lambda_F) g_F (V_{h,S} - V_{l,F})$$

In order to show that $-S_{l,S} < S_{h,S}$ it is sufficient that

$$\begin{aligned} &\mu_{h,0}(\lambda_F) g_F (V_{h,F} - V_{l,S}) + \mu_{h,0}(\lambda_S) g_S (V_{h,S} - V_{l,S}) > \\ &\mu_{l,1}(\lambda_S) g_S (V_{h,S} - V_{l,S}) + \mu_{l,1}(\lambda_F) g_F (V_{h,S} - V_{l,F}) \end{aligned}$$

The LHS is strictly larger than $\Big(\mu_{h,0}(\lambda_F) g_F + \mu_{h,0}(\lambda_S) g_S\Big)(V_{h,S} - V_{l,S})$ whereas the RHS is strictly smaller than $\Big(\mu_{l,1}(\lambda_S) g_S + \mu_{l,1}(\lambda_F) g_F\Big)(V_{h,S} - V_{l,S})$. Thus, under the same condition as above we have established that $-S_{l,S} < S_{h,S}$ and thus $V_{h,F} < V_{h,S}$.

The proof to rule out case 3a works equivalently.

Next, consider the pattern where everyone is trading towards the fast types, i.e. case 2b. Here, there are two cases. First, if there is more assets than fast types, $M > g_F$, in steady state all the fast types hold the asset with the remaining asset held by the slow types. In this case, a violation of the trading rule where an asset gets handed towards the slow agents should increase the mass of slow, high asset holders by at least the share of high types (this lower bound corresponds to the case where the slow types have speed 0). In turn, the reduction in the share of high asset holders should be strictly smaller than that as their speed is strictly positive.

When $M < g_F$ we proceed by computing the direction of change in social planner objective, by reallocating the asset from fast agent to slow agents when trading happens according to case 2b, and finding bounds on this change. The goal is to show that this reallocation improves welfare. This would imply that the planner would prefer at least some fast agents to sell to slow agents, which then along with the fact that social planner solution is bang-bang (as argued in section 4.1) completes the proof.

Recall that the fast agents are absorbing, i.e. in steady state they hold all the asset and the slow agents hold nothing. However, not all the fast agents have a unit of asset at every point in time. Consider the following thought experiment. Consider the two speed economy, where F (S) agents hold m_1 (m_2) units of asset per capita, and force agents of each speed only to trade with each other.

The social planner objective can be written as

$$\begin{aligned} \max_{m_1} \;\; & g_F \mu_{h,1}(\lambda_F, m_1) + (1 - g_F)\mu_{h,1}(\lambda_S, m_2) \\ \text{s.t.} \;\; & g_F m_1 + (1 - g_F) m_2 = M \end{aligned} \tag{17}$$

where $\mu_{h,1}(\lambda_F, m_1)$ $\Big(\mu_{h,1}(\lambda_S, m_2)\Big)$ is the fraction of high type owners associated with an economy with the corresponding symmetric speed and asset endowment of m_1 (m_2).[8]

[8] We are slightly abusing the notation to explicitly show the dependence of μ's on the per capita asset holding of each group.

Note that the steady state of case 2b is a special case of the above thought experiment where $m_1 = \bar{m}_1 = \frac{M}{g_F}$, so $m_2 = \bar{m}_2 = \frac{M - g_F \bar{m}_1}{1 - g_F} = 0$. We want to show that the derivative of the planner objective with respect to m_1 at $\bar{m}_1$ is negative, which in turn implies that if we force a fast owner agents to sell to a slow non-owner agent, that improves welfare, which would imply the trading rule where asset always flows to fast agents is sub-optimal. This requires:

$$g_F \frac{\partial \mu_{h,1}(\lambda_F, M/g_F)}{\partial m_1} - (1 - g_F) \frac{g_F}{1 - g_F} \frac{\partial \mu_{h,1}(\lambda_S, 0)}{\partial m_1} < 0$$

which boils down to

$$\frac{\partial \mu_{h,1}(\lambda_F, M/g_F)}{\partial m_1} < \frac{\partial \mu_{h,1}(\lambda_S, 0)}{\partial m_1}$$

The next thing to note is that $\partial \Big(\frac{\partial \mu_{h,1}(\lambda,0)}{\partial m_1} \Big) / \partial \lambda > 0$, so a sufficient condition for the above inequality is to show it at $\lambda_S = 0$, where we know $\frac{\partial \mu_{h,1}(\lambda,0)}{\partial m_1} = f_h$. So he sufficient condition is

$$\frac{\partial \mu_{h,1}(\lambda_F, M/g_F)}{\partial m_1} < \frac{\partial \mu_{h,1}(0, 0)}{\partial m_1} = f_h \tag{18}$$

For each pair (λ_F, g_F), construct a pair (λ, x)=($\lambda_F g_F$,g_F). Now in order to show the above, for a range of parameters, we will use a result that we prove in the next section. The idea is that for a given pair (λ, x), we divided the economy into two groups of agents: x fraction with speed $\frac{\lambda}{x}$ and $(1-x)$ fraction with speed zero, who (by construction of one group having speed zero) do not trade with each other. Then we find the optimal per capita division of assets between the two groups, m_1, which would maximize social planner objective in (17) substituting x for g_F.

Under the assumption that $M = f_h$ we find that the optimal division of (per capita) asset is interior, i.e. $0 \leq m_1 \leq \min\{1, \frac{M}{x}\}$. Thus, we get that equation 18 holds and we have the desired result.