Asset Reallocation in Markets with Intermediaries Under Selling Pressure*

Swaminathan Balasubramaniam[†], Armando Gomes[‡], SangMok Lee[§]

February 16, 2022

Abstract

We study a search model of investors' asset trading, intermediated by financial institutions that are at risk of selling under pressure. The selling pressure leads to the development of a secondary market, where intermediaries can bail each other out. Interestingly, an increase in competing intermediaries can improve each intermediary's value, because the enhanced benefits of secondary trades can prevail the reduction in value from narrower buy-sell spreads due to more intense competition. The market exhibits search externality; suppressing investors' direct trading can improve welfare.

Keywords: Asset trading, OTC markets, Financial intermediation, Search and matching, Secondary market.

*We thank Ana Babus, Paco Buera, Mina Lee, Jason Roderick, and various audiences for helpful comments.

[†]Olin Business School, Washington University in St. Louis, Email: balasu.s@wustl.edu

[‡]Olin Business School, Washington University in St. Louis, gomes@wustl.edu

[§]Department of Economics, Washington University in St. Louis, sangmoklee@wustl.edu

1 Introduction

1.1 Overview

Financial intermediaries, which purchase assets and hold until resell, require external capital. The outside funding routinely put the intermediaries at the risk of selling assets under pressure. An example is a closed-ended private equity (PE) buyout or real estate fund that acquire portfolio assets and exit by selling them to provide liquidity to fund investors, all occurring within an average life span of 10-12 years.[1] When they approach the end of their life span, they sell assets under pressure, often to other PE funds in the secondary market (Arcot, Fluck, Gaspar, and Hege (2015)).

In general, intermediaries suffer from the risk of selling assets under pressure. Dealers in bond, derivatives and mortgage-backed securities markets trade under roll-over risks and margin constraints. Refinancing of external capital may come at a higher cost - either in terms of direct interest rate or opportunity cost. Moreover, dealers often purchase security with borrowed money by pledging the security as a collateral. Dealers are unable to borrow the entire price of the security and must fund the difference between the price and collateral value with their own capital. When the price of the security goes down, they must ensure additional capital to be available to satisfy the margin requirements or alternatively, sell the security. In derivatives trading, dealers face the possibility of a hedge position expiring at a loss, necessitating a cash payment when the expiring hedge is replaced with a new one.

We build a search-theoretic model of asset reallocation, where intermediaries (hereafter, *funds*) are at such risk of selling under pressure. Our model is similar in spirit to Duffie, Gârleanu, and Pedersen (2005) and Hugonnier, Lester, and Weill (2018). The search-and-bargaining features are suitable for capturing intermediaries' attempt to sell under pressure in a decentralized market. In particular, while trading portfolio assets, PE funds often take several months to close a transaction. We explore the impact of market characteristics, such

[1] A PE fund consists of General Partners (GPs) who have specialties specific industries and operating expertise such finance and marketing. The GPs raise capital from outside investors, called Limited Partners (LPs). After raising capital, a typical PE fund usually has around 10 to 12 years of life after its inception. The rationale for a finite life is in that funds' invest in private firms whose market value is unknown. Only after an asset is sold, GPs and LPs observe gains of the fund and can determine GPs' management compensation and LPs' share.

as search frictions and the number of funds, on welfare, fund valuations, trading volumes, and transaction prices. The model allows us to quantify how the characteristics of the funds' ecosystem determine their performance. Funds compete for intermediation opportunities but also provide each other with greater exit opportunities (i.e., secondary transactions). As such, funds may derive benefits from having additional funds in the ecosystem, in spite of lower spreads due to competition.

Our model comprises a continuum of investors and funds who hold either one or zero assets. The assets are reallocated over time through (investor-to-investor) direct trading, fund-investor trading, and (fund-to-fund) secondary trading, with heterogeneous search frictions. High-type investors can generate higher flow payoffs, so assets are reallocated from low-type investors to high-type investors, sometimes through transactions involved with funds. Each investor's type changes over time by an exogenous shock. Funds buy assets from low-type investors and sell to high-type investors through fund-investor transactions. While holding assets, funds may receive liquidity shocks. Due to liquidity constraints, funds incur a holding cost from owing assets, which puts them under selling pressure. The funds then try to sell the assets to high-type investors or other funds through secondary transactions.

We find a unique steady-state equilibrium in which all three kinds of transactions are active. We find that funds enhance market efficiency by alleviating search frictions and providing greater liquidity to the market. The first channel, whereby intermediaries facilitate asset transfers, is well-known. The second channel – the liquidity channel – is unique to our model and is related explicitly to funds' demand for liquidity due to their exit requirements.

Our first set of results focus on the (fund-to-fund) secondary market. Secondary transactions provide greater exit opportunities to funds, enhance fund value, and improve overall welfare (Proposition 3). Funds sourcing deals provide liquidity to funds at the exit phase, and conversely funds at the exit phase allow fund buyers to acquire assets more quickly. Secondary transactions offer a channel through which funds can complement each other (Proposition 4). As a result, perhaps surprisingly, funds' value can increase in the number of funds (Proposition 5). This mechanism explains that intermediaries can continue to perform well in spite of the consistently increasing number of them.

Secondary transactions are sometimes criticized as opportunistic behavior among fund managers passing sub-par assets to their counterparts while both parties collect management fees.[2] However, our model suggests that funds may generate high returns because of (and not in spite of) secondary transactions.

We study fast-search market approximation.[3] We derive various results on welfare, search externalities, and trade volumes (Proposition 7 and Proposition 8). Direct trades between low and high type investors clearly improves the corresponding owner's flow payoffs, but its welfare consequence is subtle. Investors' direct trading deprive funds of potential buyers, and funds find it harder to turn over their inventory quickly and intermediate the asset market efficiently. Slowing down investors' direct trading allow the exit-phase funds to off-load assets, reset the life cycle, and purchase new assets, all quickly, especially when funds have low search friction. This dynamics drive our results of asymptotic inefficiency and search externalities of a fast-search market equilibrium.

As a quantitative exercise to illustrate our model, we calibrate it to the US corporate acquisition market in the Supplemental Appendix. We view a corporate acquisition as a sale of assets among corporate investors and PE funds.[4] For example, General Electric became a conglomerate through a spate of acquisitions but has recently announced it will divest its healthcare and oil-service divisions, as it is no longer perceived to be generating the most value with these assets. The secondary transactions among funds is estimated to increase each fund value by 26%, and as Proposition 5 predicts fund value increases in the number of funds. The corporate acquisition market is an interesting new application of the OTC literature, but with some caveats. Unlike our model, private equity funds in practice can hold multiple assets, directly improve the fundamental value of the holding assets, and anticipate selling pressure to arrive at the end of their life cycle (Kaplan and Stromberg (2009)). We hope our model opens the door for future work that allows such unique features of PE funds in the corporate acquisition markets.

[2] See `https://www.economist.com/node/15580148`.

[3] A fast-search market represents a market with many participants, due to population normalization. If a type (i, j) pair of investors with population N_i and N_j meet at a Poisson rate L, the number of direct trading (say, per year) normalized by the total population $N = \sum_i N_i$ is $(LN_iN_j)/N = (LN)(N_i/N)(N_j/N) = (LN)n_in_j$. The last expression also arises in a fast-search market (i.e., large LN) with normalized population $\sum_i n_i = 1$.

[4] For alternative models of corporate acquisitions, see Jovanovic and Rousseau (2002), Rhodes-Kropf, Robinson, and Viswanathan (2005), Rhodes-Kropf and Robinson (2008), Eisfeldt and Rampini (2008), David (2017), and Almeida, Campello, and Hackbarth (2011).

1.2 Related Literature

We discuss only closely related papers on OTC markets with intermediation and refer others to Nosal and Rocheteau (2011) and references therein.[5]

Hugonnier, Lester, and Weill (2020) study an inter-dealer market and share some similarities with our direct and secondary trading markets. There are several important distinctions. First, our model is tailored to capture funds' experiencing selling pressure due to liquidity shocks; the shocks arrive contingent on holding assets and cease to exist when funds off-load assets. Our model is suitable for funds intermediating with external capital. Second, their inter-dealer market is singled out such that investors can trade only through dealers. In contrast, we are interested in concurrent operations of unrestricted interactions among investors and funds. The unrestricted interaction is relevant for applications such as the corporate acquisition market and the real estate market (Phillips and Zhdanov (2017)). Third, their results focus on trading patterns including intermediation chains, whereas our focus is on fund valuations and welfare.

We focus on steady-state equilibrium in which funds choose to intermediate between low and high type investors, due to their moderate flow payoffs. This self-selection of intermediaries has been studied in Neklyudov (2012), Uslu (2019), Nosal, Wong, and Wright (2016), Shen, Wei, and Yan (2021), Yang and Zeng (2018), and Farboodi, Jarosch, and Shimer (2017). It is often the case that mid-type investors, similar to our funds, choose to intermediate with comparative advantages in search skills. Atkeson, Eisfeldt, and Weill (2015) also study an OTC market with endogenous intermediation. For derivative swap contracts, investors with risky endowments may be unable to share the risk fully, because of a size limit on bilateral trades. Buy and sell prices do not reflect the aggregate risk, and the price dispersion incentivizes some banks to act as intermediaries.

The welfare effect of secondary transactions has been studied in other contexts. Gofman (2014) shows that better-connected intermediaries in financial markets can shorten intermediation chains and improve welfare. Pagano and Volpin (2012) study the bank-loan market where banks lend to consumers (primary issuance) while other banks provide liquidity by investing in the securitized loans (secondary market liquidity). High securitization activities

[5] Our model differs from interbank network models surveyed by Allen and Babus (2009). There, the focus is on lending and borrowing.

in the secondary market yield high loan issuance in the primary market, driven by greater transparency in the secondary market. Finally, Hochberg, Ljungqvist, and Lu (2007) study the benefits of PE funds' networks in sharing information about assets.

Our paper focuses on intermediaries' selling under pressure. However, in some applications, fund managers initiate secondary transactions, either under buying pressure when the fund is near the end of their investment phase with excess capital, or under selling pressure when the fund is close to the end of the fund's life (Arcot, Fluck, Gaspar, and Hege (2015), Degeorge, Martin, and Phalippou (2016), and Wang (2012)). In either case, the secondary market offers a channel through which funds can provide liquidity to each other.

The remainder of the paper is organized as follows. Section 2 introduces the formal model; Section 3 provides equilibrium properties; Section 4 discusses our main analysis of secondary transactions; Section 5 provides an analysis of a fast-search market; and Section 6 concludes.

2 Model

Time runs continuously in $t \in [0, \infty)$. Over time, two kinds of agents, **investors** and **funds** trade **assets**. Initially, a fraction of investors and funds are endowed with assets. The measures of investors n_v, funds n_f, and tradable assets n_a remain constant. All agents are risk neutral and infinitely lived, with time preferences determined by a constant discount rate r. Each agent holds one or zero assets. Hence, $n_a < n_v + n_f$. We normalize the total measure of investors as $n_v = 1$.

An investor that holds an asset generates either a high payoff flow u_h or a low payoff flow u_l ($< u_h$). An investor does not receive any payoff flow when not holding an asset. An investor's ability to create payoff flow switches from low to high with Poisson intensity ρ_u, or from high to low with intensity ρ_d. The arrival rate of this Poisson shock for each type of investor is independent of other investors. The set of investor types is $\mathcal{T}_v \equiv \{ho, lo, hn, ln\}$, where the letters h and l represent each investor's ability to generate payoffs, and the letters o and n denote whether an investor owns an asset or not.

A fund's life cycle consists of an investment phase, a harvesting phase, and an exit phase. A fund in the investment phase does not own assets and searches for an investor or a fund selling assets. After purchasing an asset, the fund enters the harvesting phase and creates payoff flow u_f. A fund in the harvesting phase sells its assets and starts a new life cycle (i.e., goes back to the investment phase),[6] or it receives a liquidity shock with intensity ρ_e and enters the exit phase. A fund in the exit phase incurs a holding cost and generates a lower payoff flow u_e $(< u_f)$. After selling the asset, the fund automatically starts a new life at the investment phase. We denote a fund in the investing phase by type fn (a fund non-owner), in the harvesting phase by type fo (a fund owner), and in the exiting phase by type fe (a fund that is exiting). The set of fund types is $\mathcal{T}_f \equiv \{fn, fo, fe\}$. We assume that funds generate moderate payoff flows, $u_l < u_e < u_f < u_h$,such that funds play the role of intermediaries by purchasing assets from low-type investors and selling them to high-type investors.[7]

Let $\mathcal{T} \equiv \mathcal{T}_v \cup \mathcal{T}_f$ denote the set of types with typical elements i, j, etc. The measure of type $i \in \mathcal{T}$ at time $t \in [0, \infty)$ is denoted by $\mu_i(t)$. Then,

$$\begin{aligned} \mu_{ho}(t) + \mu_{hn}(t) + \mu_{lo}(t) + \mu_{ln}(t) &= n_v (= 1), \\ \mu_{fn}(t) + \mu_{fo}(t) + \mu_{fe}(t) &= n_f, \\ \mu_{ho}(t) + \mu_{lo}(t) + \mu_{fo}(t) + \mu_{fe}(t) &= n_a. \end{aligned} \tag{1}$$

Agents meet each other over time and negotiate a trade. Two investors meet each other with intensity λ_d for (investor-to-investor) **direct trading**. An investor and a fund meet each other with intensity λ_f for a **fund-investor trading**. A fund in the exit phase (fe) and a fund in the investment phase (fn) meet each other with intensity λ_s for (fund-to-fund) **secondary trading**. The meeting rate between any pair of groups is linear in each group's population. That is, for any pair of investor types $i, j \in \mathcal{T}_v$ with measures μ_i and μ_j, the total meeting rate is $\lambda_d \mu_i \mu_j$. Similarly, the total meeting rate between an investor type $i \in \mathcal{T}_v$ and a fund type $j \in \mathcal{T}_f$ is $\lambda_f \mu_i \mu_j$, and the total meeting rate in the secondary market is $\lambda_s \mu_{fe} \mu_{fn}$. When two agents meet each other, they trade an asset instantaneously if and only

[6]General partners of PE funds often start a new fund around the liquidation of an existing fund.

[7]Funds add operational value through better corporate governance and reducing agency costs, but they are unable to improve operational value as much as corporate buyers, who can generate greater payoff through synergies. In financial markets, intermediating funds typically have lower cost of capital as compared to sellers but do not have hedging or portfolio diversification benefits.

if the gain from trade (which we explain later) is positive. The assumption of immediate trading upon meeting follows the literature on bargaining without asymmetric information.

We will find an equilibrium in which all tradings denoted by $\mathcal{M} \equiv \{lo\text{-}hn, lo\text{-}fn, fo\text{-}hn, fe\text{-}hn, fe\text{-}fn\}$ are active. That is, a lo-type investor sells an asset to a hn-type investor (lo-hn trade). Similarly, either a lo-type investor sells an asset to a fn-type fund (lo-fn trade), or a fund of type either fo or fe sells an asset to a hn-type investor (either fo-hn or fe-hn trade). In the secondary market, a fe-type fund sells to a fn-type fund (fe-fn trade). After all trades, the types change from 'o' to 'n' and vice versa. Overall, assets are transferred from low-type investors toward high-type investors, with a possible chain of trades among funds through secondary trades.

Figure 1 summarizes the model. Agent types are listed on the left column for owners and the right column for non-owners. An owner changes her type to one on the right column upon selling her asset; a non-owner changes her type to one on the left column upon purchasing an asset (a fund's type become fo after an asset purchase). The solid arrows represent asset reallocations from sellers to buyers. The vertical dashed arrows represent the exogenous type changes: high vs. low for investors, or a liquidity shock to fo-type funds.

An asset market with fund intermediation is a collection of exogenous variables $\theta \equiv (n, r, u, \rho, \lambda)$, where $n \equiv (n_v, n_f, n_a)$, $u \equiv (u_l, u_h, u_f, u_e)$, and $\lambda \equiv (\lambda_d, \lambda_f, \lambda_s)$. All exogenously given parameters are strictly positive.When necessary, we also include an initial population distribution $\mu(0)$ in the model.

3 Equilibrium

We are interested in a steady-state equilibrium with active trades. That is, in addition to investors trading assets amongst themselves, funds actively intermediate by buying and selling assets. In this section, we provide a condition on the market under which such a steady-state equilibrium uniquely exists.

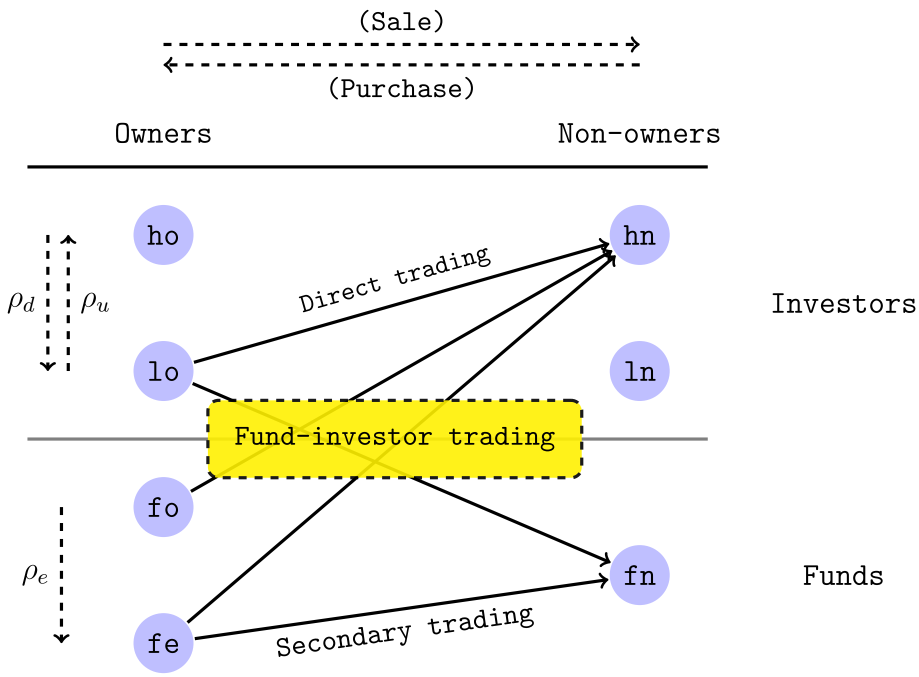

Figure 1: An asset market with fund intermediation.

3.1 Existence and Uniqueness

We first derive steady-state population measures. To ease our exposition, we denote the steady-state populations of high- and low-type investors by n_h and n_l. Since we normalized the total measure of investors as $n_v = 1$, from the rates of exogenous type changes,

$$n_h = \frac{\rho_u}{\rho_u + \rho_d} \quad \text{and} \quad n_l = \frac{\rho_d}{\rho_u + \rho_d}.[8]$$

A hn-type investor switches its type to ho upon purchasing an asset from either a lo-type investor or a fo- or fe-type fund. As such, hn-type investors become ho-type at the rate of $(\lambda_v \mu_{lo} + \lambda_f \mu_{fo} + \lambda_f \mu_{fe})\mu_{hn}$. On the other hand, as a result of exogenous type changes, hn-type investors switch their types to ln at the rate of $\rho_d \mu_{hn}$; similarly, ln-type investors

[8]Later, a reader can confirm the values of n_h and n_l by adding population equations (μ-ho) and (μ-hn), or (μ-lo) and (μ-ln), and applying $n_h + n_l = n_v = 1$.

switch their types to hn at the rate of $\rho_u \mu_{ln}$. Thus,

$$\dot{\mu}_{hn}(t) = -(\lambda_d \mu_{lo}(t) + \lambda_f \mu_{fo}(t) + \lambda_f \mu_{fe}(t))\mu_{hn}(t) - \rho_d \mu_{hn}(t) + \rho_u \mu_{ln}(t). \tag{μ-hn}$$

The population measures for other types change over time by similar processes:

$$\dot{\mu}_{ho}(t) = (\lambda_d \mu_{lo}(t) + \lambda_f \mu_{fo}(t) + \lambda_f \mu_{fe}(t))\mu_{hn}(t) - \rho_d \mu_{ho}(t) + \rho_u \mu_{lo}(t), \tag{μ-ho}$$

$$\dot{\mu}_{ln}(t) = (\lambda_d \mu_{hn}(t) + \lambda_f \mu_{fn}(t))\mu_{lo}(t) - \rho_u \mu_{ln}(t) + \rho_d \mu_{hn}(t), \tag{μ-ln}$$

$$\dot{\mu}_{lo}(t) = -(\lambda_d \mu_{hn}(t) + \lambda_f \mu_{fn}(t))\mu_{lo}(t) - \rho_u \mu_{lo}(t) + \rho_d \mu_{ho}(t), \tag{μ-lo}$$

$$\dot{\mu}_{fn}(t) = \lambda_f \left(\mu_{hn}(t)\mu_{fo}(t) + \mu_{hn}(t)\mu_{fe}(t) - \mu_{lo}(t)\mu_{fn}(t)\right), \tag{μ-fn}$$

$$\dot{\mu}_{fo}(t) = (\lambda_f \mu_{lo}(t) + \lambda_s \mu_{fe}(t))\mu_{fn}(t) - \lambda_f \mu_{hn}(t)\mu_{fo}(t) - \rho_e \mu_{fo}(t), \tag{μ-fo}$$

$$\dot{\mu}_{fe}(t) = -(\lambda_f \mu_{hn}(t) + \lambda_s \mu_{fn}(t))\mu_{fe}(t) + \rho_e \mu_{fo}(t). \tag{μ-fe}$$

Let $P(\theta)$ denote the above system of population equations (μ-hn)-(μ-fe). A real-vector $\mu \equiv (\mu_i)_{i\in\mathcal{T}}$ with each $\mu_i \geq 0$ is a steady-state solution of $P(\theta)$ if the right-hand sides of the equations, with $\mu_i(t)$ replaced by μ_i for each $i \in \mathcal{T}$, are equal to zero.

A steady-state population measure of every type is non-vanishing. As such, all kinds of meetings, hence all tradings, occur with strictly positive rates subject to positive gains.

Lemma 1. *If μ is a steady-state solution of $P(\theta)$, then $\mu_i > 0$ for all $i \in \mathcal{T}$.*

The intuition of the lemma is simple. Strictly positive rates $\lambda = (\lambda_d, \lambda_f, \lambda_s)$ and $\rho = (\rho_u, \rho_d, \rho_e)$ allow the mass of investors and funds to flow across all types. Given $0 < n_a < n_v + n_f$, some fraction of agents are owners and others are non-owners. Investors who own assets may have their types changing between high and low exogenously and non-owners have similar probabilistic type changes. As transaction rates $\lambda = (\lambda_d, \lambda_f, \lambda_s)$ are all strictly positive, some investors can buy or offload assets after their types change. However, not all of the investors can do so within any fixed time period. A similar idea holds for funds.

A steady-state solution of $P(\theta)$ uniquely exists. The solution is asymptotically stable in in that if the population measure $\mu(t)$ is perturbed locally from the steady-state solution μ, then it converges back to μ.

Proposition 1. *(Steady-state Population Measures) Take any market θ.*

1. *(Existence and Uniqueness) There exists a unique steady-state solution μ of $P(\theta)$.*

2. *(Asymptotic Stability) Let $\mu(t)$ be a dynamic solution of $P(\theta)$ with initial condition $\mu(0)$. For any $\epsilon > 0$, there exists $\delta > 0$ such that, if $\|\mu(0) - \mu\| < \delta$, then $\|\mu(t) - \mu\| \leq \epsilon$ for all t, and $\mu(t) \to \mu$ as $t \to \infty$.*

The proof uses the Poincare-Hopf index theorem (Simsek, Ozdaglar, and Acemoglu, 2007), which is new to the OTC market literature. The index theorem generalizes the Intermediate Value Theorem.

To get an intuition, we set $(\mu_i)_{i \in \mathcal{T}_v} \approx 0$, while satisfying the population equations $P(\theta)$ but violating $n_v = 1$. A small increase of μ_{lo} (or μ_{hn}) increases the supply (resp., demand) of assets for investors' direct trading $\lambda_d \mu_{lo} \mu_{hn}$, and in turn the number of investors that are rightly holding or not-holding assets (μ_{ho} and μ_{ln}). The increased populations of μ_{ho} and μ_{ln} lead to more inflow $\rho_d \mu_{ho}$ of agents back to the aggregate supply and the inflow $\rho_u \mu_{ln}$ to the aggregate demand for direct trading. That is, all four investor-type populations increase. Taking into account how investor-type populations are related to fund-type populations (elaborated in Section A.2 in Appendix), we find a unique supply μ_{lo} and demand μ_{hn} that yield $\mu_{lo} + \mu_{hn} + \mu_{ln} + \mu_{ho} = n_v = 1$, by the index theorem. The second part of the proposition on stability is due to a classical result in dynamical systems. If all eigenvalues of the linearized system at the steady-state solution have negative real parts, then the solution is asymptotically stable (Hirsch and Smale, 1973).

We define a steady-state equilibrium via a recursive equation of certain values. The sources of value to all agents in our model are two-fold: flow payoffs while holding assets and gains from trade. Let v_{hn} denote the expected **value** of time-discounted future payoffs for a type-hn investor. The value is defined implicitly by

$$r v_{hn} = \lambda_d \mu_{lo} g_{lo-hn} + \lambda_f \mu_{fo} g_{fo-hn} + \lambda_f \mu_{fe} g_{fe-hn} - \rho_d \left(v_{hn} - v_{ln} \right), \tag{v-hn}$$

where each g_{lo-hn}, g_{fo-hn}, and g_{fe-hn} denotes the investor's **gain from trade** (in fact, an equal share of the gain, which we define later). The meeting rate for a direct trading, taking into account the population of sellers, is $\lambda_d \mu_{lo}$, and the gain from a trade is g_{lo-hn}. Two other terms are defined similarly for the cases of trading with either a fo- or fe-type fund. The investor changes its type from high to low with rate ρ_d, in which case it loses value equivalent to $v_{hn} - v_{ln}$.

The values for other types are defined implicitly as follows:

$$rv_{ho} = u_h - \rho_d \left(v_{ho} - v_{lo}\right), \tag{v-ho}$$

$$rv_{ln} = \rho_u \left(v_{hn} - v_{ln}\right), \tag{v-ln}$$

$$rv_{lo} = u_l + \lambda_d \mu_{hn} g_{lo-hn} + \lambda_f \mu_{fn} g_{lo-fn} + \rho_u \left(v_{ho} - v_{lo}\right), \tag{v-lo}$$

$$rv_{fn} = \lambda_f \mu_{lo} g_{lo-fn} + \lambda_s \mu_{fe} g_{fe-fn}, \tag{v-fn}$$

$$rv_{fo} = u_f + \lambda_f \mu_{hn} g_{fo-hn} - \rho_e \left(v_{fo} - v_{fe}\right), \tag{v-fo}$$

$$rv_{fe} = u_e + \lambda_f \mu_{hn} g_{fe-hn} + \lambda_s \mu_{fn} g_{fe-fn}. \tag{v-fe}$$

A term representing flow payoffs is included for each owner type (the payoff flow is zero for non-owners).

We assume that buyers and sellers in each trading have equal bargaining power so that the transaction prices (that we will characterize in the next subsection) will be set to ensure an equal division of gain from trade between a buyer and a seller.[9] Then, each trading partner's gain is:

$$\begin{aligned}
g_{lo-hn} &\equiv (1/2)(v_{ho} + v_{ln} - v_{lo} - v_{hn}),\\
g_{lo-fn} &\equiv (1/2)(v_{fo} + v_{ln} - v_{lo} - v_{fn}),\\
g_{fo-hn} &\equiv (1/2)(v_{ho} + v_{fn} - v_{fo} - v_{hn}),\\
g_{fe-hn} &\equiv (1/2)(v_{ho} + v_{fn} - v_{fe} - v_{hn}),\\
g_{fe-fn} &\equiv (1/2)(v_{fo} + v_{fn} - v_{fe} - v_{fn}) = (1/2)(v_{fo} - v_{fe}).
\end{aligned}$$

Let $V(\theta)$ denote the above system of value equations (v-hn)-(v-fe), with μ being replaced by the unique steady-state solution of $P(\theta)$. The following proposition ensures that the

[9] Our qualitative results do not depend on the assumption of equal bargaining power. Ahern (2012) observes that the dollar gains of trades are often equally split between buyers and sellers.

values $v \equiv (v_i)_{i \in \mathcal{T}}$ are well-defined:

Proposition 2. *There exists a unique solution of $V(\theta)$.*

We have characterized the unique steady-state population measures μ and the values v, assuming that all tradings are active. If the unique steady-state solution (μ, v) results in positive trade gains, we call it a **steady-state equilibrium**.[10]

We provide the conditions for positive trade gains as (20) and (21) in Appendix. We argue below that such conditions hold if the differences $u_f - u_l$ and $u_h - u_f$ are sufficiently large.

Gains from certain trades are trivially positive. For example, $g_{fe-fn} \geq 0$ is immediate from $u_f > u_e$: secondary transactions bail out funds under liquidity constraints. The gains from a direct trading by investors and a fund-investor trading are related as $g_{lo-hn} = g_{lo\text{-}fn} + g_{fo-hn}$: a direct trade and indirect trade through fund intermediation result in the same total gains. In a similar vein, gains in a fund-investor trading and a secondary trading are related as $g_{fe-hn} = g_{fo-hn} + g_{fe-fn}$. Thus, it remains to make sure that two gains g_{lo-fn} and g_{fo-hn} are positive, for example, by large $u_f - u_l$ and $u_h - u_f$.

The conditions for $g_{lo-fn} \geq 0$ and $g_{fo-hn} \geq 0$ (i.e., (20) and (21) in Appendix) can be intuitively expressed if (i) $\lambda_d < \lambda_f < \lambda_s$, (ii) $\rho = (\rho_u, \rho_d, \rho_e)$ is close to zero relative to $\lambda = (\lambda_d, \lambda_f, \lambda_s)$, and (iii) assets are somewhat scarce ($n_h < n_a < n_h + n_f$).[11] First, $\mu_{hn} \approx 0$ as there are hardly any type changes from ln to hn, and $g_{lo-fn} \geq 0$ by (21) becomes equivalent to:

$$u_f + 0 \geq u_l + \lambda_s \mu_{fe} g_{fe-fn}.$$

Each side of the inequality considers a steady-state in which selling investors (lo) and fund buyers (fn) either trade or do not trade. On the left-hand side (i.e., trading in steady state), the total payoff flows after trading is likely to be u_f for a long time; a liquidity shock rarely

[10] The uniqueness of a steady-state equilibrium is not universal in the literature on OTC, or more broadly search-and-bargain, markets with intermediation. A unique steady-state equilibrium appears in Duffie, Gârleanu, and Pedersen (2005), but multiple equilibria appear more commonly with financial market applications: e.g., Vayanos and Weill (2008) and Trejos and Wright (2016).

[11] Conditions (i) and (ii) are natural. Condition (iii) is also likely to hold. For otherwise, there would be almost no direct or secondary trading: if $n_a < n_h$ almost all assets are held by high-type investors, and if $n_a > n_h + n_f$ almost all high-type investors and funds hold assets.

arrives, and buying investors (hn) for funds' assets are scarce. On the right-hand side (i.e., no trading in steady state), an lo-type investor receives payoff flow u_l for an extended period of time, as she does not trade with a fund buyer (fn), and buying investors (hn) are scarce. Fund buyers (fn) purchase assets mostly from exiting funds (fe) with the gain g_{fe-fn}. Next, $\mu_{fe} \approx 0$ because a non-negligible number of fn-type funds try to purchase assets, partly from fe-type funds, whereas liquidity shocks hardly arrive to fo-type funds. Hence, the above condition for $g_{lo-fn} \geq 0$ always holds. Last, $g_{lo-fn} \geq 0$ implies $g_{fo-hn} \geq 0$.[12]

3.2 Equilibrium Properties

The transaction prices are determined in a manner so that buyers and sellers equally share the gains from trades (the equal bargaining power assumption). For a direct trading,

$$p_{lo-hn} \equiv (1/2)(v_{ho} + v_{lo} - v_{hn} - v_{ln}),$$

so that a buyer's gain $(v_{ho} - p_{lo-hn}) - v_{hn}$ and a seller's gain $(v_{ln} + p_{lo-hn}) - v_{lo}$ are the same. The prices for other trades are similarly determined.[13] We establish below various relationships among transaction prices and spreads:

Lemma 2. *(Equilibrium Prices)*

1. $p_{fo-hn} \geq p_{fe-hn} \geq p_{fe-fn}$*: funds sell at a lower price during the exit phase than in the harvesting phase, and at an even lower price in secondary trading.*

2. $p_{fo-hn} \geq p_{lo-hn} \geq p_{lo-fn}$*: funds buy assets at a lower price and sell at a higher price than investors.*

A fund that manages to sell assets before receiving liquidity shocks can generate positive profits, from payoff flows u_f and the positive spread $p_{fo-hn} - p_{lo-fn}$. If a fund suffers a liquidity shock before finding a buyer, the fund may incur losses ex-post as the spread at the exit phase $p_{fe-hn} - p_{lo-fn}$ can be negative.

[12]For a proof, take $\rho \to 0$ in (19) where $g_1 \equiv g_{fo-hn}$ and $g_2 \equiv g_{lo-fn}$.

[13]To be more precise, $p_{lo-fn} = (1/2)(v_{fo} + v_{lo} - v_{ln} - v_{fn})$, $p_{fo-hn} = (1/2)(v_{ho} + v_{fo} - v_{fn} - v_{hn})$, $p_{fe-hn} = (1/2)(v_{ho} + v_{fe} - v_{fn} - v_{hn})$, and $p_{fe-fn} = (1/2)(v_{fo} + v_{fe} - 2v_{fn})$.

We define welfare as $W \equiv \sum_{i \in \mathcal{T}} \mu_i v_i$. The investors' welfare and the funds' welfare are defined similarly as $W_v \equiv \sum_{i \in \mathcal{T}_v} \mu_i v_i$ and $W_f \equiv \sum_{i \in \mathcal{T}_f} \mu_i v_i$. The welfare are naturally related to the investors' and funds' payoff flows and gains from trades as follows:

Lemma 3. *(Equilibrium Welfare) For any market θ with a steady-state equilibrium (μ, v),*

$$
\begin{aligned}
rW &= \mu_{ho}u_h + \mu_{fo}u_f + \mu_{fe}u_e + \mu_{lo}u_l, && (2)\\
rW_v &= \mu_{ho}u_h + \mu_{lo}u_l + \underbrace{\lambda_f \mu_{lo}\mu_{fn} p_{lo-fn}}_{\textit{sales to funds}} - \underbrace{\lambda_f \mu_{hn}(\mu_{fo}p_{fo-hn} + \mu_{fe}p_{fe-hn})}_{\textit{purchases from funds}}, \quad \textit{and} && (3)\\
rW_f &= \mu_{fo}u_f + \mu_{fe}u_e + \underbrace{\lambda_f \mu_{fo}\mu_{hn} p_{fo-hn} + \lambda_f \mu_{fe}\mu_{hn} p_{fe-hn}}_{\textit{sales to investors}} - \underbrace{\lambda_f \mu_{fn}\mu_{lo} p_{lo-fn}}_{\textit{purchases from investors}}.
\end{aligned}
$$

The first two terms for the investors' welfare W_v represent payoff flows to ho- and lo-type investors. The next term represents the inflow from selling assets to funds. Only lo-type investors sell assets to funds with the total rate $\lambda_f \mu_{lo}\mu_{fn}$ and at price p_{lo-fn}. The last term represents the hn-type investors' payments to funds: p_{fo-hn} to fo-type funds with the aggregate rate of $\lambda_f \mu_{hn}\mu_{fo}$, or p_{fe-hn} to fe-type funds with the aggregate rate of $\lambda_f \mu_{hn}\mu_{fe}$. A similar interpretation explains the funds' welfare W_f.

Finally, we obtain an expression for the average time to sell for investors and funds.

Lemma 4. *(Time to Sell) Let τ_{sv} and τ_{sf} denote the time to sell for investors and funds. Then,*

$$
\begin{aligned}
E[\tau_{sv}] &= \frac{1}{\lambda_d \mu_{hn} + \lambda_f \mu_{fn}}, && (4)\\
E[\tau_{sf}] &= \frac{1}{\lambda_f \mu_{hn} + \rho_e} + \frac{\rho_e}{\lambda_f \mu_{hn} + \rho_e}\left(\frac{1}{\lambda_f \mu_{hn} + \lambda_s \mu_{fn}}\right). && (5)
\end{aligned}
$$

For example, each seller-buyer meeting arrives according to a Poisson process, so the time until the first meeting by a selling investor follows an exponential distribution with parameter $\lambda_d \mu_{hn} + \lambda_f \mu_{fn}$.

4 Secondary Market: Main Results

Our main results focus on the secondary transactions which can be a new source of deals and an exit channel for funds.

We first observe that secondary trading provides liquidity and improves welfare. The result is as expected. Each secondary trade bails a fund out of liquidity constraints and offers fund buyers more transaction opportunities, at no cost to any other types of agents. Thus, a more liquid secondary market attenuates the effects of fund liquidity shocks and improves the overall welfare.

Proposition 3. *1. While liquidity shocks reduce funds' values* ($v_{fe} < v_{fo}$)*, their influence is mitigated by more liquid secondary market* ($\frac{\partial(v_{fo}-v_{fe})}{\partial\lambda_s} < 0$) *and vanishes when the secondary market becomes completely liquid* ($\lim_{\lambda_s\to\infty}(v_{fo} - v_{fe}) = 0$).

2. The welfare increases in the secondary market liquidity ($\frac{\partial W}{\partial\lambda_s} = \frac{\partial\mu_{fo}}{\partial\lambda_s}(\frac{u_f-u_e}{r}) > 0$).

Intuitively, a faster secondary market allows fe-type funds to exit and re-enter as fn-type quickly, resulting in more funds in the investment phase (fn) moving to the harvesting phase (fo). The welfare increases by $\frac{u_f-u_e}{r}$ for each unit measure of population shifted from μ_{fe} to μ_{fo}.

Our main finding is that each fund's value may increase in the number of funds. Funds complement each other through secondary trades. Although funds compete for intermediation opportunities, self-interested funds help each other out through secondary trades. When there is a greater number of funds in the market, exiting funds under pressure can sell more quickly. At the same time, funds looking to buy assets benefit from additional trade opportunities from exiting funds. The mutual benefits of secondary trades among funds can be so large as to dominate the reduction in value from narrower buy-sell spreads due to the increased competition.

Proposition 4. *There exists* $\overline{n}_f$ *such that if* $n_f < \overline{n}_f$*, then there is a complementarity between the secondary market liquidity and the number of funds* ($\frac{\partial^2 v_{fn}}{\partial\lambda_s\partial n_f} > 0$).

The closed-form expression of the complementarity ($\frac{\partial^2 v_{fn}}{\partial\lambda_s\partial n_f}$) at $n_f \approx 0$ in the proof of Proposition 4 shows that the complementarity is strong if there is an oversupply of assets relative to the number of buying investors. For otherwise, exiting funds could easily sell assets through the secondary market, and the benefits of secondary trades would be insignificant. Moreover, the complementarity is stronger if the liquidity shocks are more severe either in intensity ($u_f - u_e$) or frequency (ρ_e).

Proposition 5. *There exists $\overline{n}_f$ and a function $\overline{\lambda}_s(n_f)$ such that, if there are not too many funds ($n_f < \overline{n}_f$) and the secondary market is liquid enough ($\lambda_s > \overline{\lambda}_s(n_f)$), then funds' value increases in their number ($\frac{\partial v_{fn}}{\partial n_f} > 0$).*[14]

When the number of funds increases, competition amongst them in chasing intermediation opportunities becomes stiffer, and the value reduction due to competition may ultimately cancel out the benefit from complementarities. Figure 2 shows how the value of fn-type fund (i.e., a direct competitor of a new entrant) responds to n_f and λ_s, while other parameters are fixed at the values calibrated for the corporate acquisition market (Supplemental Appendix). The vertical line at $n_f = n_f^*$ and the curve for $\lambda_s = \lambda_s^*$ correspond to the observed number of funds and the calibrated search rate for secondary trades.[15] The figure confirms Proposition 5. At the calibrated values of parameters, the benefit from complementarity among funds due to secondary trades supersedes competitive pressures. However, for large n_f, lower spreads due to increased competition for intermediation opportunities turn out to be the dominant force impacting fund valuation.

When the number of funds increases, an increasing fraction of funds exit through secondary trading rather than sales of assets to investors (Figure 3). While it is obvious that the number of secondary trades increase with the growth in funds, our model explains a tandem increase in the *share* of exits through secondary trades.

[14]The matching technology plays only a limited role in Proposition 5. Note that in our model, the rate of meetings between funds of type fe and fn is $\lambda_s \mu_{fn} \mu_{fe}$, and so scaling up the total number of funds n_f by k while holding the type distribution fixed scales up the meeting rate by k^2 and the individual fund's meeting intensities by k. The qualitative result ($\frac{\partial v_{fn}}{\partial n_f} > 0$) holds in general. Observe that the derivative $\frac{\partial v_{fn}}{\partial n_f}$ at $n_f \approx 0$ in Figure 2 gets indefinitely large as λ_s increases. For any matching technology with increasing returns to scale, if λ_s is sufficiently large and n_f is small, then $\frac{\partial v_{fn}}{\partial n_f} > 0$.

[15]The calibrated search rates $\lambda = (\lambda_d, \lambda_f, \lambda_s)$ tend to be very large due to our normalization of $n_v = 1$ and motivates us to study a fast-search market. See Section 5 for detailed discussions.

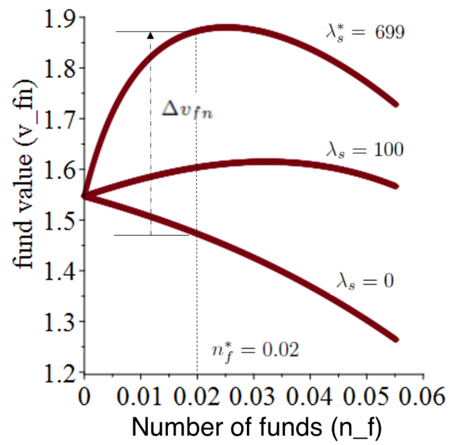

Figure 2: The y-axis shows v_{fn} for various values of n_f. All other parameters are fixed at calibrated values (see Supplemental Appendix). The topmost curve is for the calibrated value of λ_s while the bottom two curves are based on the counter-factual parameter values. The observed number of funds is $n_f^* = 0.02$ (normalized by the number of investors). The increase in fund valuation Δv_{fn} represents the contribution by secondary trades, i.e., an increase of λ_s from 0 to the calibrated value $\lambda_s^* = 699$.

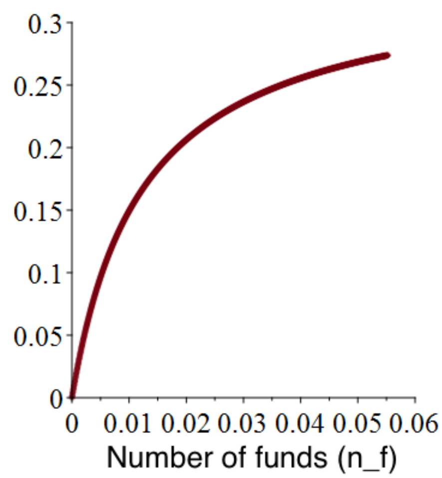

Figure 3: The share of fund exits by secondary trading $\left(\frac{\lambda_s \mu_{fe} \mu_{fn}}{\lambda_f \mu_{fe} \mu_{hn} + \lambda_s \mu_{fe} \mu_{fn}}\right)$ for various values of n_f.

5 Asymptotic Results on Welfare and Trading

We are interested in quantitative equilibrium properties, such as welfare and trading patterns, given parameter values that represent real markets. It is common that markets for corporate acquisitions, bonds, derivatives, and mortgage-backed securities have many buyers, sellers, and intermediaries. As such, we consider a set of parameter values that represents a large number of agents.

A market with many participants corresponds to *fast-search* – large search rates $\lambda = (\lambda_d, \lambda_f, \lambda_s)$ – in our model. The reason is that we normalized the total number of investors as $n_v = 1$ and proportionally re-scaled the number of funds and trade volumes. Let N_v be the total number of investors before normalization, with N_i for $i \in \mathcal{T}_v$ being the number of type-i investors. If each pair of investors meet at a Poisson rate l_d, the total number of direct trading (say, per year), with normalization, would be $(l_d N_{lo} N_{hn})/N_v = (l_d N_v)(N_{lo}/N_v)(N_{hn}/N_v) = (l_d N_v)\mu_{lo}\mu_{hn}$.

We set up a formal fast-search market as follows. Given any exogenous parameters $\theta \equiv (n, r, u, \rho, \lambda)$, we increase meeting rates $(\lambda_d, \lambda_f, \lambda_s)$, while preserving the relative ratios. That is, we consider a sequence of markets $\theta^\kappa \equiv (n, r, u, \rho, \kappa\lambda)$, where $\kappa\lambda = (\kappa\lambda_d, \kappa\lambda_f, \kappa\lambda_s)$. We analyze the steady-state solution (μ^κ, v^κ) in the limit as κ increases to infinity. To ease expositions, we assume a *regular* environment: $n_a \notin \{n_h, n_h + n_f\}$. The fast-search market is well-defined as the steady-state population measures μ^κ converge (the convergence of v^κ, as a linear function of μ^κ, follows immediately (see (15)). The speed of convergence is $O(1/\kappa)$ which gives a precise sense of how closely a fast-search equilibrium would approximate an equilibrium of the calibrated market (in our calibration, $\kappa \approx 103,000$):

Proposition 6. *(Convergence and Convergence Speed) Take any regular environment θ. For any $i \in \mathcal{T}$, the population limit $\mu_i^* \equiv \lim_{\kappa\to\infty} \mu_i^\kappa$ and the convergence speed $\mu_i^{**} \equiv \lim_{\kappa\to\infty} \kappa(\mu_i^\kappa - \mu_i^*)$ exist.*

The closed-form expressions for the population limits and the convergence speeds are in Appendix.

5.1 Inefficiency and Search Externality

We compare the fast-search equilibrium welfare W^* against two extreme situations: an autarkic economy with no functioning market or fund intermediation, and a centralized economy with a planner moving assets across agents without search friction.[16] In the autarkic economy, an n_a fraction among $n_v(=1)$ corporations hold assets with no trades, resulting in the welfare $\underline{W}$ such that $r\underline{W} = n_a(n_h u_h + n_l u_l)$. In the centralized economy, a planner solves

$$r\overline{W} \equiv \max_{\mu \in \mathbb{R}_+^{\mathcal{T}}} \mu_{ho}u_h + \mu_{fo}u_f + \mu_{fe}u_e + \mu_{lo}u_l,$$
$$\text{subject to} \quad \mu_{ho} + \mu_{hn} = n_h, \quad \mu_{lo} + \mu_{ln} = n_l, \quad \text{and} \quad (1).$$

The maximum welfare $\overline{W}$ takes into account exogenous type changes ρ_u and ρ_d but ignores search frictions. The liquidity shock ρ_e imposes no restriction to the planner who can transfer assets between funds instantaneously. The planner can set aside an ϵ mass of funds as type fn and transfer assets to them when some other funds receive a liquidity shock. The mass of fo and fe type funds remain the same. Consequently, the mass arbitrarily close to n_f of assets can be held by fo type funds.

The population measures $\overline{\mu}$ that achieves the maximum welfare is such that $\overline{\mu}_{ho} = \min\{n_a, n_h\}$, $\overline{\mu}_{fo} = \min\{(n_a - n_h)^+, n_f\}$, $\overline{\mu}_{lo} = (n_a - n_f - n_h)^+$, and $\overline{\mu}_i = 0$ for $i \neq ho, fo, lo$. In essence, assets are allocated to high-type investors up to their steady-state population n_h; any remaining assets are given to funds up to n_f; and the still-remaining assets are given to low-type investors. The maximum welfare satisfies $r\overline{W} = \overline{\mu}_{ho}u_h + \overline{\mu}_{fo}u_f + \overline{\mu}_{lo}u_l$.[17]

We compare the efficient allocation $\overline{\mu}$ (the upper part of Table 1) with the fast-search equilibrium population μ^* (the lower part of Table 1).

Proposition 7. *(Fast-search Market: Welfare) As $\kappa \to \infty$,*

[16] The welfare of W^κ converges to W^* at the same speed $O(1/\kappa)$ as μ^κ, because the welfare W^κ is a linear aggregation of the population measures $(\mu_i^\kappa)_{i \in \mathcal{T}}$.

[17] The maximum welfare $\overline{W}$ is also achieved by a planner who is under the search friction, like agents, but can choose not to execute some transactions. The planner's problem is $rW_p(\lambda) \equiv \sup_{0 \le \lambda_p \le \lambda} \mu_{ho}u_h + \mu_{fo}u_f + \mu_{fe}u_e + \mu_{lo}u_l$, subject to μ being a solution of $P(n, r, u, \rho, \lambda_p)$. Fast search allows the planner to achieve the maximum welfare approximately: i.e., $\overline{W} = \lim_{\lambda \to \infty} W_p(\lambda)$. An intuition will become clear after Proposition 7. The planner slows down the investors' direct trading, eliminates search externalities, and increases the welfare to the maximum.

	A. $n_a < n_h$	B. $n_h < n_a < n_h + n_f$	C. $n_h + n_f < n_a$
$\overline{\mu}_{ho} =$	n_a	n_h	n_h
$\overline{\mu}_{fo} =$	0	$n_a - n_h$	n_f
$\overline{\mu}_{lo} =$	0	0	$n_a - n_f - n_h$
$\overline{\mu}_{fe} =$	0	0	0
$\mu^*_{ho} =$	n_a	n_h	n_h
$\mu^*_{fo} =$	0	$n_a - n_h$	$< n_f$
$\mu^*_{lo} =$	0	0	$n_a - n_f - n_h$
$\mu^*_{fe} =$	0	0	> 0

Table 1: The population under fast search (μ^*) and the efficient allocation ($\bar{\mu}$).

A. *If $n_a < n_h$, then $W^* = \overline{W}$, which is independent of u_f, u_e, λ_s, and λ_d.*

B. *If $n_h < n_a < n_h + n_f$, then $W^* = \overline{W}$, which is strictly increasing in u_f and independent of u_e, λ_s, and λ_d.*

C. *If $n_h + n_f < n_a$, then W^* is strictly less than $\overline{W}$, strictly increasing in u_f, u_e, and λ_s, and strictly decreasing in λ_d.*

The characterization of the fast-search equilibrium welfare depends on the number of assets (n_a) relative to the number of potential buyers (n_h, n_f). A sufficiently large number of assets ($n_a > n_h + n_f$) gives rise to an inefficient fast-search equilibrium.

Suppose that the fast-search market has sufficiently many potential buyers ($n_a < n_h + n_f$) as in Cases A and B (the first two columns in Table 1). Fast search allows investors and funds to quickly transfer assets from low-type investors (lo) and exiting funds (fe) to high-type investors (hn) and, in Case B, also to funds at the investment phase (fn). Accordingly, the steady-state population μ^* equals the efficient allocation $\overline{\mu}$ and achieves the maximum welfare ($W^* = \overline{W}$).

The comparative statics of the welfare becomes trivial: the maximum welfare is dependent on payoff flows (e.g., u_f) only if the corresponding type's population (resp., μ_{fo}) is non-zero. In either case, the impact of a liquidity shock is zero. Funds transfer assets without holding any inventory (Case A) just like, e.g., in Rubinstein and Wolinsky (1987); or, they hold assets, but funds under liquidity constraints transfer assets to others through speedy secondary trades (Case B), as is the case in Duffie, Gârleanu, and Pedersen (2005). Under a surplus of tradable assets relative to potential buyers ($n_h + n_f < n_a$), as in Case C (the third column in Table 1), the equilibrium is more interesting because, counter-intuitively, slowing down investors' direct trading *improves* the welfare ($\frac{\partial W^*}{\partial \lambda_d} < 0$). Since investors or funds on demands can quickly find sellers and purchase assets, there is negligible left-over high-type investors or fund non-owners. Hence, a significant fraction of exiting funds (fe) will find it difficult to offload their assets, and the welfare loss is $r(\overline{W} - W^*) = \mu^*_{fe}(u_f - u_e) > 0$.

The inefficiency is a result of investors' search externalities on funds. A direct trading by investors take away selling opportunities from exit-phase funds and leads them to suffer from liquidity constraints for a longer period of time. If the investors' direct trading were absent, an exit-phase fund could off-load an asset, reset its type, and purchase another asset, all quickly under fast search. This alternative scenario results in a more efficient asset allocation.

5.2 Trade Volumes

In this section, we characterize the transaction volumes and how they respond to exogenous parameters, such as search frictions and transition rates. For each submarket $m \in \mathcal{M}$, with a seller's type s and the buyer's type b, the **trade volume** is $\eta^\kappa_m \equiv (\kappa\lambda_m)\mu^\kappa_s \mu^\kappa_b$.

Proposition 8. *(Fast-search Market: Trade Volumes) For each submarket $m \in \mathcal{M}$, trade volume in the limit $\eta^*_m \equiv \lim_{\kappa\to\infty} \eta^\kappa_m$ is given by:*

	A. $n_a < n_h$	*B.* $n_h < n_a < n_h + n_f$	*C.* $n_h + n_f < n_a$
$\eta^*_{lo-hn} =$	$\lambda_d \mu^{**}_{lo} \mu^*_{hn}$	0	$\lambda_d \mu^*_{lo} \mu^{**}_{hn}$
$\eta^*_{lo-fn} =$	$\lambda_f \mu^{**}_{lo} \mu^*_{fn}$	$\lambda_f \mu^{**}_{lo} \mu^*_{fn}$	$\lambda_f \mu^*_{lo} \mu^{**}_{fn}$
$\eta^*_{fo-hn} =$	$\lambda_f \mu^{**}_{fo} \mu^*_{hn}$	$\lambda_f \mu^*_{fo} \mu^{**}_{hn}$	$\lambda_f \mu^*_{fo} \mu^{**}_{hn}$
$\eta^*_{fe-hn} =$	0	0	$\lambda_f \mu^*_{fe} \mu^{**}_{hn}$
$\eta^*_{fe-fn} =$	0	$\lambda_s \mu^{**}_{fe} \mu^*_{fn}$	$\lambda_s \mu^*_{fe} \mu^{**}_{fn}$

where (i) μ^* denotes the population limit, and (ii) for type i with $\mu^*_i = 0$, $\mu^{**}_i \equiv \lim_{\kappa\to\infty} \kappa\mu^\kappa_i$ denotes the convergence speed.

With a large number of high-type investors (Case A), secondary trades are unnecessary for funds; if the deficit of high-type investors is supplemented by funds (Case B), selling investors resort to fund buyers and there are no investors' direct trading; with an excess supply of tradable assets (Case C), there are transactions in all submarkets.

The trade volumes under fast search follow from the convergence and the convergence speed of population measures (Proposition 6). For each submarket $m \in \mathcal{M}$, because of fast search, the steady-state measure of either buyers or sellers vanishes: i.e., $\mu_b^* = 0$ or $\mu_s^* = 0$. If $\mu_s^* = 0$, then $\eta_m^\kappa = (\kappa\lambda_m)\mu_b^\kappa \mu_s^\kappa = \lambda_m \mu_b^\kappa(\kappa\mu_s^\kappa) \to \lambda_m \mu_b^* \mu_s^{**}$ as $\kappa \to \infty$. Proposition 8 suggests that all submarkets are active under fast search only if the market has an excess supply of tradable assets ($n_a > n_h + n_f$).

The trade volumes also identify the main drivers of the convergences of certain population measures. For example, $\mu_{fo}^* = 0$ in Case A could be due to the fact that (i) funds can rarely purchase assets because of a vanishingly small number of selling investors (lo), or (ii) funds do acquire assets, but quickly re-sell to buying investors (hn). Proposition 8 implies the latter case; funds buy/sell a significant number of assets from/to investors in the fast-search market and there are no secondary transactions (like middlemen in Rubinstein and Wolinsky (1987)). Similarly, the vanishing number of selling investors (lo) and buying investors (hn) in Case B is the result of an efficient fund-investor trading rather than an efficient market for investors' direct trading – the number of investors' direct transactions (η_{lo-hn}^*) is indeed vanishingly small.

Table 2 summarizes a comparative static analysis for trade volumes relative to search frictions and transition rates.[18]

Most results are intuitive. If investors' types are more volatile (i.e., larger ρ_u and ρ_d with a fixed ratio), assets will be transferred across agents frequently, ultimately from low-type investors to high-type investors, with possible fund intermediations. Fast search among investors (i.e., higher λ_d) allows them to transact directly (i.e., higher η_{lo-hn}^*), resulting in fewer intermediation opportunities for funds. The parameters for the secondary market (λ_s

[18] The results follow directly from the expressions of μ^* and μ^{**} in Table 1 and Lemmas 7 and 8 in Appendix, so we omit. As an example, take η_{lo-hn}^* and λ_d. Note that $\eta_{lo-hn}^* = \lambda_d \mu_{lo}^* \mu_{hn}^{**}$. In Cases A and B, $\mu_{lo}^* = 0$, so η_{lo-hn}^* is independent of λ_s. In Case C, $\mu_{hn}^{**} = \frac{\rho_u \mu_{ln}^*}{\lambda_d \mu_{lo}^* + \lambda_f(\mu_{fo}^* + \mu_{fe}^*)} = \frac{\rho_u \mu_{ln}^*}{\lambda_d \mu_{lo}^* + \lambda_f n_f}$. Note that μ_{lo}^* and μ_{ln}^* are independent of λ_d. Thus, the volume η_{lo-hn}^* is increasing in λ_d because $\frac{\partial}{\partial \lambda_d}\left(\frac{\lambda_d}{\lambda_d \mu_{lo}^* + \lambda_f n_f}\right) = \lambda_f n_f > 0$.

	λ_d	λ_f	λ_s	(ρ_u, ρ_d)(with a fixed ratio)	ρ_e
η^*_{lo-hn}	$+$	$-$	0	$+$	0
η^*_{lo-fn}	$-$	$+$	0	$+$	0
η^*_{fo-hn}	$-$	$+$	$+$	$+$	$-$
η^*_{fe-hn}	$-$	$+$	$-$	$+$	$+$
η^*_{fe-fn}	$-$	$-$	$+$	$+$	$+$

Table 2: Comparative statics of trade volumes. Each $+$ (or $-$) indicates the corresponding volume to be non-decreasing (resp., non-increasing) in the parameter, and 0 indicates that the volume is independent of the parameter.

and ρ_e) only shift the population measures between fo and fe. Therefore, these parameters do not affect the volume of investors' direct trading (η^*_{lo-hn}) and only shift volumes between two kinds of fund-investor transactions: η^*_{fo-hn} and η^*_{fe-hn}.

The positive response of η^*_{fe-hn} to λ_f is perhaps surprising. On one hand, fast search between funds and investors (i.e., higher $\kappa\lambda_f$) orchestrates more transactions between exiting funds (fe) and buying investors (hn). On the other hand, as funds are able to sell assets before receiving liquidity shocks, fewer funds enter the exit phase, which could potentially reduce the trade volume between exiting funds and buying investors. It turns out that the former effect of λ_f dominates the latter.

6 Conclusion

We provide a search-based model of asset trading with fund intermediation. Funds in our model intermediate between buyers and sellers at risk of selling assets under pressure, possibly to other funds. The model shows that secondary transactions can make substantial contribution to fund values. Our paper offers a novel explanation on persistent intermediators' returns when the number of intermediaries increases, despite increased competition. A well-lubricated private market for corporate acquisitions can partly explain the recent shift in firm ownership from public to private, predicated by Jensen (1991). Lower liquidity costs in a market for buying and selling private firms help make PE ownership a form of governance that may indeed eclipse public companies.

A Appendix

A.1 Proof of Lemma 1:

Take any market $\theta = (n, r, u, \rho, \lambda)$ with $n_f > 0$. Let $\mu \in \mathbb{R}_+^{\mathcal{T}}$ be a steady-state solution of $P(\theta)$.

First, we show that $\mu_{ho} > 0$ and $\mu_{lo} > 0$. It is clear that $\mu_{ho} = 0$ if and only if $\mu_{lo} = 0$. If $\mu_{ho} = 0$, then $\mu_{lo} = 0$, as only ho-type investors flow in type lo; conversely, if $\mu_{lo} = 0$, then $\mu_{ho} = 0$ as the inflow to the type lo must be zero. Suppose, toward contradiction, that $\mu_{ho} = \mu_{lo} = 0$. That is, all investors are non-owners. The in-flow from hn-type investors to type ho must be zero, so it must be that $\mu_{fo} = \mu_{fe} = 0$. Then, $\mu_{ho} + \mu_{lo} + \mu_{fo} + \mu_{fe} = 0$, a contradiction to $n_a > 0$.

Second, we show that $\mu_{ln} > 0$ and $\mu_{hn} > 0$. As before, it is clear that $\mu_{ln} = 0$ if and only if $\mu_{hn} = 0$. If $\mu_{ln} = 0$, then $\mu_{hn} = 0$ as only ln-type investors can flow in type hn; conversely, if $\mu_{hn} = 0$, then $\mu_{ln} = 0$ as the inflow to the type hn must be zero. Suppose, toward contradiction, that $\mu_{ln} = \mu_{hn} = 0$. That is, all investors are owners, which implies that some funds are non-owners: $\mu_{fn} = n_v + n_f - n_a > 0$. Since $\lambda_f \mu_{lo} \mu_{fn} > 0$, some lo-type investors change their types and flow into type ln by trading with PE funds, a contradiction to $\mu_{ln} = 0$.

Lastly, we consider funds. Suppose that $\mu_{fn} = 0$. As the inflow to type fo becomes zero, it must be that $\mu_{fo} = 0$, which in turn leads to no inflow by liquidity shocks to type fe: i.e., $\mu_{fe} = 0$. Such case contradicts to $n_f > 0$. When $\mu_{fn} > 0$, given strictly positive population μ_{lo}, the inflow of type-fn funds to type fo is strictly positive: $\lambda_f \mu_{lo} \mu_{fn} > 0$. As such, $\mu_{fo} > 0$, which in turn creates a strictly positive inflow by liquidity shocks to type fe: $\mu_{fe} > 0$.

A.2 Proof for Part 1 of Proposition 1

We reduce the number of variables and population equations in $P(\theta)$ by imposing some necessary conditions for a steady-state solution. Note that any steady-state solution μ must satisfy $\mu_{ho} + \mu_{hn} = n_h \equiv \frac{\rho_u}{\rho_u + \rho_d}$ and $\mu_{lo} + \mu_{ln} = n_l \equiv \frac{\rho_d}{\rho_u + \rho_d}$ (which we can obtain by adding (μ-ho) and (μ-hn), or (μ-lo) and (μ-ln), and apply $n_v = 1$). If we substitute $\mu_{ho} = n_h - \mu_{hn}$

and $\mu_{ln} = n_l - \mu_{lo}$ into (μ-ho)-(μ-fe), then we are left with the following three linearly independent equations:[19]

$$(\lambda_d \mu_{hn} + \lambda_f \mu_{fn})\mu_{lo} + \rho_u \mu_{lo} - \rho_d \mu_{ho} = 0, \qquad \text{(from } (\mu\text{-lo}))$$
$$(\lambda_d \mu_{lo} + \lambda_f \mu_{fo} + \lambda_f \mu_{fe})\mu_{hn} + \rho_d \mu_{hn} - \rho_u \mu_{ln} = 0, \qquad \text{(from } (\mu\text{-hn}))$$
$$-(\lambda_f \mu_{hn} + \lambda_s \mu_{fn})\mu_{fe} + \rho_e \mu_{fo} = 0. \qquad \text{(from } (\mu\text{-fe}))$$

We re-write the first two equations with respect to μ_{lo} and μ_{hn}.

$$\mu_{fo} + \mu_{fe} = n_a - \mu_{ho} - \mu_{lo} = n_a - (n_h - \mu_{hn}) - \mu_{lo} \quad \text{and} \tag{6}$$
$$\mu_{fn} = n_f - (\mu_{fo} + \mu_{fe}) = n_f - n_a + n_h - \mu_{hn} + \mu_{lo}, \tag{7}$$

Then,

$$(\lambda_d \mu_{hn} + \lambda_f (n_f - n_a + n_h - \mu_{hn} + \mu_{lo}))\mu_{lo} + \rho_u \mu_{lo} - \rho_d (n_h - \mu_{hn}) = 0, \tag{8}$$
$$(\lambda_d \mu_{lo} + \lambda_f (n_a - n_h + \mu_{hn} - \mu_{lo}))\mu_{hn} + \rho_d \mu_{hn} - \rho_u (n_l - \mu_{lo}) = 0. \tag{9}$$

We show below that there exists a unique solution (μ_{lo}, μ_{hn}) of (8)-(9) such that (i) $0 \leq \mu_{lo} \leq n_l$, (ii) $0 \leq \mu_{hn} \leq n_h$, and (iii) $n_a - n_f - n_h \leq \mu_{lo} - \mu_{hn} \leq n_a - n_h$ (for $0 \leq \mu_{fn} \leq n_f$). Other population measures will be determined by $\mu_{ho} = n_h - \mu_{hn}$, $\mu_{ln} = n_l - \mu_{lo}$, and

$$\mu_{fn} = n_f - n_a + n_h - \mu_{hn} + \mu_{lo}. \tag{10}$$

We find the last two populations (μ_{fe}, μ_{fo}) by solving

$$\mu_{fo} = -\mu_{fe} + (n_f - \mu_{fn}), \quad \text{(from } \mu_{fn} + \mu_{fo} + \mu_{fe} = n_f)$$
$$\mu_{fo} = \frac{\lambda_f \mu_{hn} + \lambda_s \mu_{fn}}{\rho_e} \mu_{fe}. \quad \text{(from } (\mu\text{-fe}))$$

[19] Any other equation in $P(\theta)$ is redundant, as it depends linearly on (μ-lo), (μ-hn), and (μ-fe). Each sum of the right-hand sides of (μ-ho) and (μ-hn), or (μ-lo) and (μ-ln) equals zero, which allow us to delete (μ-ho) and (μ-ln) without changing the solution set. The sum of the right-hand sides of (μ-fn), (μ-fo), and (μ-fe) equals zero, so we can delete (μ-fn). Last, the sum of the right-hand sides of (μ-ho), (μ-lo), (μ-fo), and (μ-fe) equals zero, so we can delete (μ-fo).

The unique solution is

$$\mu_{fe} = \frac{\rho_e(n_f - \mu_{fn})}{\lambda_f\mu_{hn} + \lambda_s\mu_{fn} + \rho_e}, \quad \text{and} \quad \mu_{fo} = \frac{(\lambda_f\mu_{hn} + \lambda_s\mu_{fn})(n_f - \mu_{fn})}{\lambda_f\mu_{hn} + \lambda_s\mu_{fn} + \rho_e}. \tag{11}$$

Therefore, it remains to prove the following claim:

Claim 1. *Let* $X(\theta) \equiv \{(x_1, x_2) \in \mathbb{R}^2 : 0 \le x_1 \le n_l, 0 \le x_2 \le n_h, 0 \le g_{fn}(x) \le n_f\}$, *where* $g_{fn}(x) \equiv n_a - n_h + x_2 - x_1$. *Also, define* $F \equiv (F_{lo}, F_{hn}) : \mathbb{R}^2 \to \mathbb{R}^2$ *by*

$$\begin{aligned} F_{lo}(x) &\equiv (\lambda_d x_2 + \lambda_f(n_f - g_{fn}(x)))x_1 + \rho_u x_1 - \rho_d(n_h - x_2), \\ F_{hn}(x) &\equiv (\lambda_d x_1 + \lambda_f g_{fn}(x))x_2 + \rho_d x_2 - \rho_u(n_l - x_1). \end{aligned}$$

Then, there exists a unique solution of $F(x) = 0$ *in* $X(\theta)$.

We apply the Poincare-Hopf index theorem, a version in Simsek, Ozdaglar, and Acemoglu (2007, p.194); see also Hirsch (2012). First, $X(\theta)$ is non-empty, compact, and convex.[20] The boundary of $X(\theta)$ is

$$\partial X(\theta) \equiv \{(x_1, x_2) \in X(\theta) : x_1 = 0, x_1 = n_l, x_2 = 0, x_2 = n_h, g_{fn}(x) = 0, \text{ or } g_{fn}(x) = n_f\}.$$

Second, the function $F(x)$ is continuously differentiable at every $x \in \mathbb{R}^2$. Third, the determinant of the Jacobian matrix of F is strictly positive for every interior point of $X(\theta)$: for each $x \in \mathbb{R}^2$,

$$\begin{aligned} \nabla F(x) &\equiv \begin{bmatrix} \frac{\partial F_{lo}}{\partial x_1} & \frac{\partial F_{lo}}{\partial x_2} \\ \frac{\partial F_{hn}}{\partial x_1} & \frac{\partial F_{hn}}{\partial x_2} \end{bmatrix} \\ &= \begin{bmatrix} (\lambda_d x_2 + \lambda_f(n_f - g_{fn}(x))) + \lambda_f x_1 + \rho_u & (\lambda_d - \lambda_f)x_1 + \rho_d \\ (\lambda_d - \lambda_f)x_2 + \rho_u & (\lambda_d x_1 + \lambda_f g_{fn}(x)) + \lambda_f x_2 + \rho_d \end{bmatrix}, \end{aligned}$$

[20] The Poincare-Hopf index theorem also requires $X(\theta)$ to be a 2-dimensional smooth manifold, which a reader can easily verify by applying the identify function to the definition of a smooth manifold in Simsek, Ozdaglar, and Acemoglu (2007, p.193).

and for any interior point $x \in X(\theta)\backslash\partial X(\theta)$,

$$\begin{aligned} det(\nabla F(x)) \geq &(\lambda_d x_2 + \lambda_f x_1)(\lambda_d x_1 + \lambda_f x_2) + \rho_u \lambda_d x_1 + \rho_d \lambda_d x_2 \\ &- (\lambda_d - \lambda_f)^2 x_1 x_2 - (\lambda_d - \lambda_f)(\rho_d x_2 + \rho_u x_1) \\ =&\lambda_d \lambda_f (x_1^2 + x_2^2) + 2\lambda_d \lambda_f x_1 x_2 + \lambda_f(\rho_d x_2 + \rho_u x_1) > 0. \end{aligned} \tag{12}$$

Last, we show that, for every boundary point $x \in \partial X(\theta)$, the vector $F(x) \in \mathbb{R}^2$ points strictly outward of $X(\theta)$. We partition the boundary $\partial X(\theta)$ into six *faces* (i.e., flat surfaces) of $X(\theta)$. For each face, we find an outward normal vector $\mathbf{n} \in \mathbb{R}^2$ and show that the angle between $\mathbf{n}$ and $F(x)$ is acute (i.e., ≤ 90) at any point x in the face:

1. ($x_1 = 0$ and $0 \leq x_2 < n_h$) $\mathbf{n} = (-1, 0)$ is an outward normal vector, and $\mathbf{n} \cdot F(x) = \rho_d(n_h - x_2) > 0$.

2. ($x_2 = 0$ and $0 \leq x_1 < n_l$) $\mathbf{n} = (0, -1)$ is an outward normal vector, and $\mathbf{n} \cdot F(x) = \rho_u(n_l - x_1) > 0$.

3. ($x_1 = n_l$ and $0 \leq x_2 \leq n_h$) $\mathbf{n} = (1, 0)$ is an outward normal vector, and

$$\begin{aligned} \mathbf{n} \cdot F(x) &= (\lambda_d x_2 + \lambda_f(n_f - g_{fn}(x)))n_l + \rho_u n_l - \rho_d n_h + \rho_d x_2 \\ &\geq (\lambda_d x_2 + \lambda_f(n_f - g_{fn}(x)))n_l \quad (\text{as } \rho_u n_l = \rho_d n_h) \\ &\geq \min\{\lambda_d x_2 n_l, \lambda_f(n_f - g_{fn}(x))n_l\}. \end{aligned}$$

   As either $x_2 > 0$ or $x_2 = 0$, we have $n_f - g_{fn}(x) = n_v + n_f - n_a > 0$, and $\mathbf{n} \cdot F(x) > 0$.

4. ($x_2 = n_h$ and $0 \leq x_1 \leq n_l$) $\mathbf{n} = (0, 1)$ is an outward normal vector, and

$$\mathbf{n} \cdot F(x) = (\lambda_d x_1 + \lambda_f g_{fn}(x))n_h + \rho_d n_h - \rho_u n_l + \rho_u x_1 \geq \min\{\lambda_d x_1 n_h, \lambda_f g_{fn}(x) n_h\}.$$

   As either $x_1 > 0$ or $x_1 = 0$, we have $g_{fn}(x) = n_a > 0$, and $\mathbf{n} \cdot F(x) > 0$.

5. ($g_{fn}(x) = 0$ and $x_1 > 0$) $\mathbf{n} = (1, -1)$ is an outward normal vector, and $\mathbf{n} \cdot F(x) = F_{lo}(x) - F_{hn}(x) = \lambda_f n_f x_1 > 0$.

6. ($g_{fn}(x) = n_f$ and $x_2 > 0$) $\mathbf{n} = (-1, 1)$ is an outward normal vector, and $\mathbf{n} \cdot F(x) = F_{hn}(x) - F_{lo}(x) = \lambda_f n_f x_2 > 0$.

We are ready to apply the Poincare-Hopf index theorem in Simsek, Ozdaglar, and Acemoglu (2007, p.194). The Euler characteristic of $X(\theta)$ is 1; see their definition on p.193 for the case of non-empty and convex sets. Claim 1 follows immediately from the index theorem, which completes the proof for Part 1 of Proposition 1.

A.3 Proof for Part 2 of Proposition 1

We first reduce the system $P(\theta)$. For any market θ and an initial condition $\mu(0)$, any dynamic solution $\mu : [0, \infty) \to \mathbb{R}^{\mathcal{T}}$ of the system $P(\theta)$ satisfies, for every $t \in [0, \infty)$,

$$\begin{aligned}
\mu_{ho}(t) + \mu_{hn}(t) + \mu_{lo}(t) + \mu_{ln}(t) &= n_v (= 1), \\
\mu_{fn}(t) + \mu_{fo}(t) + \mu_{fe}(t) &= n_f, \quad \text{and} \\
\mu_{ho}(t) + \mu_{lo}(t) + \mu_{fo}(t) + \mu_{fe}(t) &= n_a.
\end{aligned}$$

Without changing the set of dynamic solutions, we can reduce the system $P(\theta)$ for $x(t) \equiv (\mu_{ho}(t), \mu_{hn}(t), \mu_{lo}(t), \mu_{fo}(t))$ by[21]

$$\dot{x} = F(x) \equiv \left(F_{ho}\left(x\right), F_{hn}\left(x\right), F_{lo}\left(x\right), F_{fo}\left(x\right)\right), \tag{13}$$

[21]In the proof of Part 1 of Proposition 1, we reduced $F(x; \theta) = 0$ further as a system of only two equations in Claim 1. The reduction requires $\mu_{hn} + \mu_{ho} = n_h \equiv \frac{\rho_u}{\rho_u + \rho_d}$ and $\mu_{lo} + \mu_{ln} = n_l \equiv \frac{\rho_d}{\rho_u + \rho_d}$, which hold in a steady-state but may not hold on a path of $\mu(t)$ after a perturbation.

where

$$
\begin{aligned}
F_{ho}(x) &\equiv (\lambda_d \mu_{lo} + \lambda_f \mu_{fo} + \lambda_f \mu_{fe}(x))\mu_{hn} - \rho_d \mu_{ho} + \rho_u \mu_{lo}, \\
F_{hn}(x) &\equiv -(\lambda_d \mu_{lo} + \lambda_f \mu_{fo} + \lambda_f \mu_{fe}(x))\mu_{hn} - \rho_d \mu_{hn} + \rho_u \mu_{ln}(x), \\
F_{lo}(x) &\equiv -(\lambda_d \mu_{hn} + \lambda_f \mu_{fn}(x))\mu_{lo} - \rho_u \mu_{lo} + \rho_d \mu_{ho}, \\
F_{fo}(x) &\equiv (\lambda_f \mu_{lo} + \lambda_s \mu_{fe}(x))\,\mu_{fn}(x) - \lambda_f \mu_{hn} \mu_{fo} - \rho_e \mu_{fo}, \quad \text{and} \\
\mu_{ln}(x) &\equiv 1 - \mu_{ho} - \mu_{hn} - \mu_{lo}, \\
\mu_{fe}(x) &\equiv n_a - \mu_{ho} - \mu_{lo} - \mu_{fo}, \\
\mu_{fn}(x) &\equiv n_f - \mu_{fo} - \mu_{fe}(x) = n_f - n_a + \mu_{ho} + \mu_{lo}.
\end{aligned} \tag{14}
$$

The reduction of the system $P(\theta)$ does not change the set of dynamic solutions. If μ is a dynamic (either steady-state or not) solution of $P(\theta)$, then $x \equiv (\mu_{ho}, \mu_{hn}, \mu_{lo}, \mu_{fo})$ solves $F(x;\theta) = 0$; conversely, for any dynamic solution x of $F(x;\theta) = 0$, we can find a dynamic solution μ of $P(\theta)$, from x and the induced μ_{ln}, μ_{fe}, and μ_{fn}. Hence, a dynamic solution μ of $P(\theta)$ is asymptotically stable if and only if $x \equiv (\mu_{ho}, \mu_{hn}, \mu_{lo}, \mu_{fo})$ is asymptotically stable.

A steady-state solution x of $F(x;\theta) = 0$ is asymptotically stable if all eigenvalues of the Jacobian matrix of $F(x;\theta)$ at the steady-state solution x have strictly negative real parts (Hirsch, 2012). The Jacobian matrix is

$$
\nabla F(x) \equiv \left[\frac{\partial F_i(x)}{\partial x_j}\right]_{i,j \in \{ho,hn,lo,fo\}}
$$
$$
= \left[\begin{array}{ccc|c}
-\lambda_f \mu_{hn} - \rho_d & \lambda_v \mu_{lo} + \lambda_f \mu_{fo} + \lambda_f \mu_{fe} & (\lambda_d - \lambda_f)\mu_{hn} + \rho_u & 0 \\
\lambda_f \mu_{hn} - \rho_u & -(\lambda_v \mu_{lo} + \lambda_f \mu_{fo} + \lambda_f \mu_{fe}) - \rho_d - \rho_u & (\lambda_f - \lambda_d)\mu_{hn} - \rho_u & 0 \\
-\lambda_f \mu_{lo} + \rho_d & -\lambda_d \mu_{lo} & -\lambda_f(\mu_{fn} + \mu_{lo}) - \lambda_d \mu_{hn} - \rho_u & 0 \\
\hline
\lambda_f \mu_{lo} + \lambda_s(\mu_{fe} - \mu_{fn}) & -\lambda_f \mu_{fo} & \lambda_f(\mu_{fn} + \mu_{lo}) + \lambda_s(\mu_{fe} - \mu_{fn}) & -\lambda_f \mu_{hn} - \lambda_s \mu_{fn} - \rho_e
\end{array}\right]
$$

where we omit the dependency of μ_{fn} and μ_{fe} on x to simplify the expression.

Due to the block structure, one eigenvalue is $-\lambda_f \mu_{hn} - \lambda_s \mu_{fn} - \rho_e < 0$. The other eigenvalues are the eigenvalues of the sub-matrix with the first three rows and columns. A direct calculation shows that the other three eigenvalues are also strictly negative, which completes the proof.

A.4 Proof of Proposition 2

First, we simplify expositions:

$$\begin{aligned}
g_1 &\equiv g_{fo-hn} = (1/2)(v_{ho} + v_{fn} - v_{fo} - v_{hn}), \\
g_2 &\equiv g_{lo-fn} = (1/2)(v_{fo} + v_{ln} - v_{lo} - v_{fn}), \\
g_3 &\equiv g_{fe-fn} = (1/2)(v_{fo} + v_{fn} - v_{fe} - v_{fn}) = (1/2)(v_{fo} - v_{fe}),
\end{aligned}$$

so that

$$\begin{aligned}
g_{lo-hn} &= (1/2)(v_{ho} + v_{ln} - v_{lo} - v_{hn}) = g_2 + g_1 \quad \text{and} \\
g_{fe-hn} &= (1/2)(v_{ho} + v_{fn} - v_{fe} - v_{hn}) = g_1 + g_3.
\end{aligned}$$

The matrix representations of the value equations (v-hn)-(v-fe) are:

$$\begin{aligned}
\begin{bmatrix} v_{ho} \\ v_{lo} \end{bmatrix} &= \begin{bmatrix} r + \rho_d & -\rho_d \\ -\rho_u & r + \rho_u \end{bmatrix}^{-1} \begin{bmatrix} u_h \\ u_l + \lambda_d \mu_{hn}(g_1 + g_2) + \lambda_f \mu_{fn} g_2 \end{bmatrix}, \\
\begin{bmatrix} v_{hn} \\ v_{ln} \end{bmatrix} &= \begin{bmatrix} r + \rho_d & -\rho_d \\ -\rho_u & r + \rho_u \end{bmatrix}^{-1} \begin{bmatrix} \lambda_d \mu_{lo}(g_1 + g_2) + \lambda_f \mu_{fo} g_1 + \lambda_f \mu_{fe}(g_1 + g_3) \\ 0 \end{bmatrix}, \quad \text{and} \\
\begin{bmatrix} v_{fn} \\ v_{fo} \\ v_{fe} \end{bmatrix} &= \frac{1}{r} \begin{bmatrix} \lambda_f \mu_{lo} g_2 + \lambda_s \mu_{fe} g_3 \\ u_f + \lambda_f \mu_{hn} g_1 - 2\rho_e g_3 \\ u_e + \lambda_f \mu_{hn}(g_1 + g_3) + \lambda_s \mu_{fn} g_3 \end{bmatrix},
\end{aligned} \tag{15}$$

where the inverse matrix is well-defined: i.e., $(r + \rho_d)(r + \rho_u) - \rho_d \rho_u > 0$. As in the case of $n_f = 0$, we compute the gains g_1, g_2, and g_3. Then, the solution v will be uniquely determined by the above matrix equations.

First, $2rg_3 = r(v_{fo} - v_{fe}) = (u_f - u_e) - 2\rho_e g_3 - \lambda_f \mu_{hn} g_3 - \lambda_s \mu_{fn} g_3$, which implies

$$g_3 = \frac{u_f - u_e}{2r + 2\rho_e + \lambda_f \mu_{hn} + \lambda_s \mu_{fn}} > 0. \tag{16}$$

Next,

$$
\begin{aligned}
2(g_1+g_2) &= v_{ho}+v_{ln}-v_{lo}-v_{hn} = (1,-1)\cdot(v_{ho}-v_{hn}, v_{lo}-v_{ln}) \\
&= \begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} r+\rho_d & -\rho_d \\ -\rho_u & r+\rho_u \end{bmatrix}^{-1} \begin{bmatrix} u_h - \lambda_d\mu_{lo}(g_1+g_2) - \lambda_f\mu_{fo}g_1 - \lambda_f\mu_{fe}(g_1+g_3) \\ u_l + \lambda_d\mu_{hn}(g_1+g_2) + \lambda_f\mu_{fn}g_2 \end{bmatrix}.
\end{aligned}
$$

Since

$$
\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} r+\rho_d & -\rho_d \\ -\rho_u & r+\rho_u \end{bmatrix}^{-1} = \frac{1}{r+\rho_u+\rho_d} \begin{bmatrix} 1 & -1 \end{bmatrix},
$$

we have

$$
\begin{aligned}
(2(r+\rho_u+\rho_d)+\lambda_d(\mu_{lo}+\mu_{hn}))\,(g_1+g_2) + \lambda_f(\mu_{fo}+\mu_{fe})g_1 + \lambda_f\mu_{fn}g_2 \\
= (u_h-u_l) - \lambda_f\mu_{fe}g_3. \qquad (17)
\end{aligned}
$$

On the other hand, by (v-lo), (v-ln), (v-fn), and (v-fo),

$$
\begin{aligned}
2rg_2 =& r(v_{fo}-v_{fn}) - r(v_{lo}-v_{ln}) \\
=& (u_f + \lambda_f\mu_{hn}g_1 - 2\rho_e g_3 - \lambda_f\mu_{lo}g_2 - \lambda_s\mu_{fe}g_3) \\
& - (u_l + \lambda_d\mu_{hn}(g_1+g_2) + \lambda_f\mu_{fn}g_2) + \rho_u(v_{ho}-v_{lo}+v_{ln}-v_{hn}).
\end{aligned}
$$

As $v_{ho}-v_{lo}+v_{ln}-v_{hn} = 2(g_1+g_2)$,

$$
\begin{aligned}
(2\rho_u+\lambda_d\mu_{hn})(g_1+g_2) - \lambda_f\mu_{hn}g_1 + (2r+\lambda_f\mu_{lo}+\lambda_f\mu_{fn})g_2 \\
= (u_f-u_l) - (2\rho_e+\lambda_s\mu_{fe})g_3. \qquad (18)
\end{aligned}
$$

The linear system of equations (17) and (18) is summarized as follows:

$$
\begin{bmatrix} c_1+\lambda_f(\mu_{fo}+\mu_{fe}) & c_1+\lambda_f\mu_{fn} \\ c_2-\lambda_f\mu_{hn} & c_2+2r+\lambda_f(\mu_{lo}+\mu_{fn}) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \end{bmatrix} = \begin{bmatrix} u_h-u_l-\lambda_f\mu_{fe}g_3 \\ u_f-u_l-2\rho_e g_3-\lambda_s\mu_{fe}g_3 \end{bmatrix} \qquad (19)
$$

where $c_1 \equiv 2(r+\rho_u+\rho_d)+\lambda_d(\mu_{lo}+\mu_{hn}) > 0$ and $c_2 \equiv 2\rho_u+\lambda_d\mu_{hn} > 0$.

The determinant of the coefficient matrix is bounded below by

$$2rc_1 + \lambda_f \mu_{fn}(c_1 - c_2) > 4r^2 + \lambda_f \mu_{fn}(2r + 2\rho_d + \lambda_d \mu_{lo}) > 0.$$

Thus, the above linear system has a unique solution (g_1, g_2). This solution, together with g_3, determined the unique solution v of $V(\theta)$. ■

A.5 Conditions for positive trade gains

The gains from trade are all positive if and only if

$$\begin{aligned} g_1 \equiv g_{fo-hn} \geq 0 \iff & (c_2 + 2r + \lambda_f(\mu_{lo} + \mu_{fn}))((u_h - u_l) - \lambda_f \mu_{fe} g_3) \\ & - (c_1 + \lambda_f \mu_{fn})((u_f - u_l) - (2\rho_e + \lambda_s \mu_{fe}) g_3) \geq 0. \quad (20) \\ g_2 \equiv g_{lo-fn} \geq 0 \iff & - (c_2 - \lambda_f \mu_{hn})((u_h - u_l) - \lambda_f \mu_{fe} g_3) \\ & + (c_1 + \lambda_f(\mu_{fo} + \mu_{fe}))((u_f - u_l) - (2\rho_e + \lambda_s \mu_{fe}) g_3) \geq 0, \quad (21) \end{aligned}$$

Note that both expressions depends on the steady-state population measure μ.

A.6 Proof of Lemma 2

For Part 1:

$$\begin{aligned} 2(p_{fo-hn} - p_{fe-hn}) &= v_{fo} - v_{fe} = 2g_{fe-fn} \geq 0, \\ 2(p_{fe-hn} - p_{fe-fn}) &= (v_{ho} - v_{hn}) - (v_{fo} - v_{fn}) = 2g_{fo-hn} \geq 0. \end{aligned}$$

For Part 2:

$$\begin{aligned} 2(p_{fo-hn} - p_{lo-hn}) &= (v_{ho} - v_{hn} + v_{fo} - v_{fn}) - (v_{lo} - v_{ln} + v_{fo} - v_{fn}) \\ &= (v_{ho} - v_{hn}) - (v_{lo} - v_{ln}) = 2g_{lo-hn} \geq 0, \\ 2(p_{lo-hn} - p_{lo-fn}) &= (v_{ho} - v_{hn}) - (v_{fo} - v_{fn}) = 2g_{fo-hn} \geq 0. \end{aligned}$$

A.7 Proof of Lemma 3

First, from (v-hn)-(v-ln),

$$
\begin{aligned}
rW_v \equiv & r(\mu_{ho}v_{ho} + \mu_{hn}v_{hn} + \mu_{lo}v_{lo} + \mu_{ln}v_{ln}) \\
= & \mu_{ho}(u_h + \rho_d(v_{lo} - v_{ho})) + \mu_{hn}(\lambda_d\mu_{lo}g_{lo-hn} + \lambda_f\mu_{fo}g_{fo-hn} + \lambda_f\mu_{fe}g_{fe-hn} + \rho_d\left(v_{ln} - v_{hn}\right)) \\
& + \mu_{lo}(u_l + \lambda_d\mu_{hn}g_{lo-hn} + \lambda_f\mu_{fn}g_{lo-fn} + \rho_u\left(v_{ho} - v_{lo}\right)) + \mu_{ln}\rho_u\left(v_{hn} - v_{ln}\right).
\end{aligned}
$$

We rewrite the above expression in terms of the investors' expected values and their payments to or received from funds. That is, we substitute $g_{fo-hn} = v_{ho} - v_{hn} - p_{fo-hn}$, $g_{fe-hn} = v_{ho} - v_{hn} - p_{fe-hn}$, $g_{lo-fn} = p_{lo-fn} - v_{lo} - v_{ln}$, and $g_{lo-hn} = (1/2)(v_{ho} + v_{ln} - v_{lo} + v_{hn})$. Then, (3) follows from

$$
\begin{aligned}
& rW_v - (\mu_{ho}u_h + \mu_{lo}u_l + \lambda_f\mu_{lo}\mu_{fn}p_{lo-fn} - \lambda_f\mu_{hn}(\mu_{fo}p_{fo-hn} + \mu_{fe}p_{fe-hn})) \\
& = (\rho_u\mu_{lo} - \rho_d\mu_{ho})(v_{ho} - v_{lo}) + (\rho_u\mu_{ln} - \rho_u\mu_{hn})(v_{hn} - v_{ln}) \\
& + \mu_{hn}(\lambda_d\mu_{lo} + \lambda_f\mu_{fo} + \lambda_f\mu_{fe})(v_{ho} - v_{hn}) + \mu_{lo}(\lambda_d\mu_{hn} + \lambda_f\mu_{fn})(v_{ln} - v_{lo}).
\end{aligned}
$$

The combined coefficient of v_{ho} on the right-hand side of the above equation is $-\rho_d\mu_{ho} + \rho_u\mu_{lo} + \mu_{hn}(\lambda_d\mu_{lo} + \lambda_f\mu_{fo} + \lambda_f\mu_{fe})$, which equals the right-hand side of the population equation (μ-ho), so it is zero. We can similarly verify that the combined coefficient of v_i for $i = hn, lo, ln$ are all equal to zero.

Second, we obtain (2) from all population equations (μ-hn)-(μ-fe) such that

$$
\begin{aligned}
& rW - (\mu_{ho}u_h + \mu_{fo}u_f + \mu_{fe}u_e + \mu_{lo}u_l) \\
& = (\rho_u\mu_{lo} - \rho_d\mu_{ho})(v_{ho} - v_{lo}) + (\rho_u\mu_{ln} - \rho_d\mu_{hn})(v_{hn} - v_{ln}) + \rho_e\mu_{fo}(v_{fe} - v_{fo}) \\
& + \mu_{hn}(\lambda_d\mu_{lo} + \lambda_f\mu_{fo} + \lambda_f\mu_{fe})(v_{ho} - v_{hn}) + \mu_{lo}(\lambda_d\mu_{hn} + \lambda_f\mu_{fn})(v_{ln} - v_{lo}) \\
& + ((\lambda_f\mu_{lo} + \lambda_s\mu_{fe})\mu_{fn} - \lambda_f\mu_{hn}\mu_{fo})v_{fo} \\
& + \lambda_f(\mu_{hn}\mu_{fo} + \mu_{hn}\mu_{fe} - \mu_{lo}\mu_{fn})v_{fn} - (\lambda_f\mu_{hn} + \lambda_s\mu_{fn})\mu_{fe}v_{fe}.
\end{aligned}
$$

As before, we can verify that the combined coefficients of v_i for each $i \in \mathcal{T}$ equals the right-hand side of the type's population equation, so it is zero.

Lastly, the expression for W_f follows from $W_f = W - W_v$.

A.8 Proof of Lemma 4

First, consider the path of a lo-type investor in a steady-state equilibrium. This investor can sell its asset upon meeting either a buying investor (hn) or a fund buyer (fn). Each kind of meeting arrives with Poisson rate $\lambda_c\mu_{hn}$ or $\lambda_f\mu_{fn}$. The time until the first meeting of each kind, denoted by τ_{lo-hn} and τ_{lo-fn}, follows the exponential distributions. Thus, the time until selling $\tau_{sc} \equiv \min\{\tau_{lo-hn}, \tau_{lo-fn}\}$ follows an exponential distribution with parameter $\lambda_c\mu_{hn} + \lambda_f\mu_{fn}$:

$$E[\tau_{sc}] = \frac{1}{\lambda_c\mu_{hn} + \lambda_f\mu_{fn}}.$$

Second, consider the path of a fo-type fund in a steady-state equilibrium. The fund sells its asset before receiving a liquidity shock to a buying investor (hn), or receives a liquidity shock and enters the exit phase (after which it can sell to either a buying investor (hn) or a fund buyer (fn)). We denote by τ_{fo} this period for which a fund maintain its type as fo . The time τ_{fo} follows an exponential distribution with parameter $\lambda_f\mu_{hn} + \rho_e$:

$$E[\tau_{fo}] = \frac{1}{\lambda_f\mu_{hn} + \rho_e}.$$

Finally, we evaluate the path of an fe type fund (an outcome of an fo type fund receiving a liquidity shock before meeting a buying investor with probability $\frac{\rho_e}{\lambda_f\mu_{hn}+\rho_e}$). The fe type fund maintains its type until it sells its portfolio asset either to a buying investor (hn) or a fund buyer (fn). Thus, the fund maintains its type for the time period τ_{fe}, which follows an exponential distribution with parameter $\lambda_f\mu_{hn} + \lambda_s\mu_{fn}$:

$$E[\tau_{fe}] = \frac{1}{\lambda_f\mu_{hn} + \lambda_s\mu_{fn}}.$$

As a result, the overall expected time for a fund to sell an asset is:

$$E[\tau_{sf}] = \frac{1}{\lambda_f\mu_{hn} + \rho_e} + \frac{\rho_e}{\lambda_f\mu_{hn} + \rho_e}\left(\frac{1}{\lambda_f\mu_{hn} + \lambda_s\mu_{fn}}\right).$$

A.9 Proof of Proposition 3

A.9.1 Part 1

Let $(\mu(\theta), v(\theta))$ be the unique steady-state solution of population and value for each market θ. We compute the comparative static derivatives with respect to λ_s. It is intuitive that the unique steady-state measure of each investor type $(\mu_i)_{i\in\mathcal{T}_v}$ and the measure μ_{fn} are independent of λ_s. Through a secondary trade, one fund changes its type from fe to fn, replacing another fund of type changed from fn to fo.

To confirm the intuition, from the proof of Proposition 1, take the unique steady-state solution $x(\theta) \equiv (\mu_{lo}(\theta), \mu_{hn}(\theta))$ of $F(x) \equiv (F_{lo}(x), F_{hn}(x)) = 0$, where

$$
\begin{aligned}
F_{lo}(x) &\equiv (\lambda_d x_2 + \lambda_f(n_f - g_{fn}(x)))x_1 + \rho_u x_1 - \rho_d(n_h - x_2), \\
F_{hn}(x) &\equiv (\lambda_d x_1 + \lambda_f g_{fn}(x))x_2 + \rho_d x_2 - \rho_u(n_l - x_1).
\end{aligned}
$$

By Implicit function theorem, $x(\theta)$ is differentiable in λ_s, and

$$
\frac{\partial x(\theta)}{\partial \lambda_s} = -\left[\nabla_x F(x(\theta);\theta)\right]^{-1} \frac{\partial F(x(\theta);\theta)}{\partial \lambda_s}.
$$

We denoted the domain of $F(x)$ by $X(\theta)$ in the proof of Claim 1. The unique solution $x(\theta)$ of $F(x;\theta) = 0$ is an interior point of $X(\theta)$, as shown in Lemma 1 for the case of $n_f > 0$ and in the proof of Part 1 of Proposition 1 for the case of $n_f = 0$. As a result, we have shown in the proof of Claim 1, the Jacobian matrix $\nabla F(x)$ at the unique solution $x(\theta)$ is invertible.

Then,

$$
\frac{\partial F(x(\theta);\theta)}{\partial \lambda_s} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies \frac{\partial x(\theta)}{\partial \lambda_s} = \begin{bmatrix} \partial \mu_{lo}(\theta)/\partial \lambda_s \\ \partial \mu_{hn}(\theta)/\partial \lambda_s \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{22}
$$

For other types, it follows from $\mu_{ho}(\theta) + \mu_{hn}(\theta) = n_h$ and $\mu_{lo}(\theta) + \mu_{ln}(\theta) = n_l$ that $\frac{\partial \mu_{ln}(\theta)}{\partial \lambda_s} = \frac{\partial \mu_{ho}(\theta)}{\partial \lambda_s} = 0$, and from $\mu_{ho}(\theta) + \mu_{lo}(\theta) + (n_f - \mu_{fn}(\theta)) = n_f$ that $\frac{\partial \mu_{fn}(\theta)}{\partial \lambda_s} = 0$. Lastly, from (11) and $\mu_{fo} + \mu_{fe} + \mu_{fn} = n_f$,

$$
\frac{\partial \mu_{fe}(\theta)}{\partial \lambda_s} = \frac{-\rho_e(n_f - \mu_{fn}(\theta))\mu_{fn}(\theta)}{(\rho_e + \lambda_f \mu_{hn}(\theta) + \lambda_s \mu_{fn}(\theta))^2} = -\frac{\partial \mu_{fo}(\theta)}{\partial \lambda_s}. \tag{23}
$$

From the definition of g_3 on p. 31 and Equation 16,

$$v_{fo} - v_{fe} = 2g_{fe-fn} = 2g_3 = \frac{2(u_f - u_e)}{2r + 2\rho_e + \lambda_f \mu_{hn} + \lambda_s \mu_{fn}}.$$

The comparative static derivatives show that μ_{hn} and μ_{fn} are independent of λ_s. Thus, $\frac{\partial(v_{fo}-v_{fe})}{\partial\lambda_s} < 0$ and $\lim_{\lambda_s\to\infty}(v_{fo} - v_{fe}) = 0$.

A.9.2 Part 2

The above comparative static derivatives with respect to λ_s show that $\frac{\partial\mu_{fo}(\theta)}{\partial\lambda_s} > 0$, $\frac{\partial\mu_{fe}(\theta)}{\partial\lambda_s} < 0$, and $\frac{\partial\mu_i(\theta)}{\partial\lambda_s} = 0$, for all $i \neq fo, fe$. Thus,

$$\begin{aligned} r\frac{\partial W(\theta)}{\partial\lambda_s} &= \frac{\partial\mu_{ho}(\theta)}{\partial\lambda_s}u_h + \frac{\partial\mu_{lo}(\theta)}{\partial\lambda_s}u_l + \frac{\partial\mu_{fo}(\theta)}{\partial\lambda_s}u_f + \frac{\partial\mu_{fe}(\theta)}{\partial\lambda_s}u_e \\ &= \frac{\partial\mu_{fo}(\theta)}{\partial\lambda_s}(u_f - u_e) + \frac{\partial(\mu_{fo} + \mu_{fe})(\theta)}{\partial\lambda_s}u_e \\ &= \frac{\partial\mu_{fo}(\theta)}{\partial\lambda_s}(u_f - u_e) - \frac{\partial\mu_{fn}(\theta)}{\partial\lambda_s}u_e = \frac{\partial\mu_{fo}(\theta)}{\partial\lambda_s}(u_f - u_e) > 0. \end{aligned}$$

A.10 Proof of Proposition 4 and Proposition 5

Recall from Claim 1 that the steady-state population is determined by a solution of $F(x; n_f) \equiv (F_{lo}(x; n_f), F_{hn}(x; n_f)) = 0$ where

$$\begin{aligned} F_{lo}(x; n_f) &\equiv (\lambda_d x_2 + \lambda_f(n_f - g_{fn}(x)))x_1 + \rho_u x_1 - \rho_d(n_h - x_2), \\ F_{hn}(x; n_f) &\equiv (\lambda_d x_1 + \lambda_f g_{fn}(x))x_2 + \rho_d x_2 - \rho_u(n_l - x_1), \end{aligned}$$

and $g_{fn}(x) \equiv n_a - n_h + x_2 - x_1$. We extend the system $F(x; n_f) = 0$ such that n_f can be any real number and x can be any real vector of length 2. Each solution $x = (x_1, x_2)$ defines a vector $\mu = (\mu_i)_{i\in\mathcal{T}}$ as $(\mu_{lo}, \mu_{hn}, \mu_{ln}, \mu_{ho}) = (x_1, x_2, n_l - x_1, n_h - x_2)$ and $(\mu_{fn}, \mu_{fo}, \mu_{fe})$ by (10) and (11). According to Claim 1, if $n_f > 0$, a solution exists in certain domain (denoted by $X(\theta)$ in the claim) such that the resulting vector μ is a steady-state population. In general, without any restrictions on n_f, the vector μ may not even be positive.

The proof consists of three steps. First, for $n_f = 0$, we find a population measure $\hat{\mu}$ such that $\hat{x} \equiv (\hat{\mu}_{lo}, \hat{\mu}_{hn})$ solves $F(x; n_f) = 0$. Second, by Implicit Function Theorem, we differentiate a solution function $x(n_f)$ defined in the neighborhood of $n_f = 0$ and $x = \hat{x}$, and obtain the comparative static derivative $\mu_i' \equiv \frac{\partial \mu_i}{\partial n_f}\Big|_{n_f=0}$ for each $i \in \mathcal{T}$. Last, we prove the following claim:

Claim 2. *There exist $\beta_1 > 0$ and β_2, each being independent of λ_s, such that*

$$\frac{\partial v_{fn}}{\partial n_f}\Big|_{n_f=0} = \beta_1 \lambda_s + \beta_2.$$

Then, Proposition 4 and Proposition 5 follow immediately.

A.10.1 A benchmark model ($n_f = 0$)

We set $\hat{\mu}_i = 0$ for every fund type $i \in \mathcal{T}_f$, and impose

$$\hat{\mu}_{ho} = n_h - \hat{\mu}_{hn}, \quad \hat{\mu}_{lo} = n_a - \hat{\mu}_{ho} = n_a - n_h + \hat{\mu}_{hn}, \quad \text{and}$$
$$\hat{\mu}_{ln} = n_l - \hat{\mu}_{lo} = n_v - n_a - \hat{\mu}_{hn}.$$

By substituting the above expressions of $\hat{\mu}_{lo}$ and $\hat{\mu}_{ln}$ in

$$\lambda_d \hat{\mu}_{lo} \hat{\mu}_{hn} + \rho_d \hat{\mu}_{hn} - \rho_u \hat{\mu}_{ln} = 0, \quad (\mu\text{-hn})$$

we obtain

$$\hat{\mu}_{hn} = \frac{1}{2}\left(\sqrt{(R + n_a - n_h)^2 + 4R \cdot n_h (1 - n_a)} - (R + n_a - n_h)\right), \tag{24}$$

where $R \equiv \frac{\rho_u + \rho_d}{\lambda_d}$. It is clear that $\hat{x} = (\hat{\mu}_{lo}, \hat{\mu}_{hn})$ solves the system $F(x; n_f) = 0$.

A.10.2 Comparative static derivatives of μ with respect to n_f

We apply Implicit Function Theorem. $F(x; n_f)$ is an infinitely differentiable function of $x \in \mathbb{R}^2$ and $n_f \in \mathbb{R}$, and the Jacobian matrix $\nabla_x F(\hat{x}; 0)$ is invertible (see Equation 12).

As such, there is a differentiable function $x(n_f)$ defined in a neighborhood of $n_f = 0$ and $x = \hat{x}$ such that $F(x(n_f); n_f) = 0$. It is important to note that the derivative of $x(n_f)$ at any $n_f > 0$ is independent of the choice of the function $x(n_f)$; Claim 1 ensures that any choice of function $x(n_f)$ gives the same value of x for each $n_f > 0$.

As explained above, the function $x(n_f)$, together with (10) and (11), defines $\mu(n_f) = (\mu_i(n_f))_{i\in\mathcal{T}}$, which is also differentiable. Let $\mu_i' \equiv \frac{\partial \mu}{\partial n_f}\Big|_{n_f=0}$ for each $i \in \mathcal{T}$. Then,

$$\begin{aligned}\begin{bmatrix}\mu_{lo}'\\ \mu_{hn}'\end{bmatrix} &= -\left[\nabla_x F(\hat{x};0)\right]^{-1}\frac{\partial F(\hat{x};0)}{\partial n_f}\\ &= -\begin{bmatrix}\lambda_d\hat{\mu}_{hn} + \lambda_f\hat{\mu}_{lo} + \rho_u & (\lambda_d - \lambda_f)\hat{\mu}_{lo} + \rho_d\\ (\lambda_d - \lambda_f)\hat{\mu}_{hn} + \rho_u & \lambda_d\hat{\mu}_{lo} + \lambda_f\hat{\mu}_{hn} + \rho_d\end{bmatrix}^{-1}\begin{bmatrix}\lambda_f\hat{\mu}_{lo}\\ 0\end{bmatrix}.\end{aligned} \tag{25}$$

Also, $\mu_{ho}' = -\mu_{hn}'$ and $\mu_{ln}' = -\mu_{lo}'$.

From (10),

$$\begin{aligned}\mu_{fn}' &= 1 - \mu_{hn}' + \mu_{lo}'\\ &= 1 - \frac{1}{det(\nabla_x F(\hat{x};0))}\begin{bmatrix}1 & -1\end{bmatrix}\begin{bmatrix}\lambda_d\hat{\mu}_{lo} + \lambda_f\hat{\mu}_{hn} + \rho_d & *\\ -(\lambda_d - \lambda_f)\hat{\mu}_{hn} - \rho_u & *\end{bmatrix}\begin{bmatrix}\lambda_f\hat{\mu}_{lo}\\ 0\end{bmatrix}\\ &= 1 - \frac{\lambda_f\hat{\mu}_{lo}\left(\lambda_d\hat{\mu}_{lo} + \lambda_d\hat{\mu}_{hn} + \rho_d + \rho_u\right)}{det(\nabla_x F(\hat{x};0))},\end{aligned}$$

where

$$\begin{aligned}det(\nabla_x F(\hat{x};0)) &= \left(\lambda_d\hat{\mu}_{hn} + \lambda_f\hat{\mu}_{lo} + \rho_u\right)\left(\lambda_d\hat{\mu}_{lo} + \lambda_f\hat{\mu}_{hn} + \rho_d\right)\\ &\quad - \left((\lambda_d - \lambda_f)\hat{\mu}_{hn} + \rho_u\right)\left((\lambda_d - \lambda_f)\hat{\mu}_{lo} + \rho_d\right)\\ &= \left(\lambda_d\hat{\mu}_{hn} + \rho_u\right)\left(\lambda_f\hat{\mu}_{hn} + \lambda_f\hat{\mu}_{lo}\right) + \left(\lambda_f\hat{\mu}_{hn} + \lambda_f\hat{\mu}_{lo}\right)\left(\lambda_d\hat{\mu}_{lo} + \rho_d\right)\\ &= \left(\lambda_f\hat{\mu}_{hn} + \lambda_f\hat{\mu}_{lo}\right)\left(\lambda_d\hat{\mu}_{lo} + \lambda_d\hat{\mu}_{hn} + \rho_d + \rho_u\right).\end{aligned}$$

It follows that

$$\mu_{fn}' = \frac{\hat{\mu}_{hn}}{\hat{\mu}_{hn} + \hat{\mu}_{lo}} > 0. \tag{26}$$

Next, from (11),

$$\mu_{fo} = \frac{(\lambda_f \mu_{hn} + \lambda_s \mu_{fn})(n_f - \mu_{fn})}{\lambda_f \mu_{hn} + \lambda_s \mu_{fn} + \rho_e}.$$

Thus,

$$\begin{aligned} \mu'_{fo} &= \frac{(1-\mu'_{fn})(\lambda_f \hat{\mu}_{hn})}{\rho_e + \lambda_f \hat{\mu}_{hn}} = \frac{\lambda_f \hat{\mu}_{hn} \hat{\mu}_{lo}}{(\rho_e + \lambda_f \hat{\mu}_{hn})(\hat{\mu}_{hn} + \hat{\mu}_{lo})} > 0, \\ \mu'_{fe} &= 1 - \mu'_{fn} - \mu'_{fo} = \frac{\rho_e \hat{\mu}_{lo}}{(\rho_e + \lambda_f \hat{\mu}_{hn})(\hat{\mu}_{hn} + \hat{\mu}_{lo})} > 0. \end{aligned} \tag{27}$$

A.10.3 Proof of Claim 2

From (15),

$$r v_{fn} = \lambda_f \mu_{lo} g_2 + \lambda_s \mu_{fe} g_3,$$

where g_2 and g_3 are determined by (16) and (19), respectively.

Let $v'_{fn} \equiv \left.\frac{\partial v_{fn}}{\partial n_f}\right|_{n_f=0}$ and $g'_m \equiv \left.\frac{\partial g_m}{\partial n_f}\right|_{n_f=0}$ for $m = 2, 3$. Then

$$r v'_{fn} = \lambda_f \left(\hat{g}_2 \mu'_{lo} + \hat{\mu}_{lo} g'_2\right) + \lambda_s \left(\hat{g}_3 \mu'_{fe} + \hat{\mu}_{fe} g'_3\right).$$

We find the value of each variable on the right-hand side of the above equation. For certain variables that we will use later, we remark whether the values are strictly positive and/or independent of λ_s.

We have observed the following properties:

1. (from (24)) the population $\hat{\mu} = (\mu_i)_{i \in \mathcal{T}}$ is strictly positive for corporate types, zero for fund types, and independent of λ_s,

2. (from (25), (26), and (27)) the derivative μ' is independent of λ_s,

3. (from (16)) As $g_3 = \frac{u_f - u_e}{2r + 2\rho_e + \lambda_f \mu_{hn} + \lambda_s \mu_{fn}}$, we have $\hat{g}_3 = \frac{u_f - u_e}{2r + 2\rho_e + \lambda_f \hat{\mu}_{hn}} > 0$ and $g'_3 = -\frac{(\lambda_f \mu'_{hn} + \lambda_s \mu'_{fn}) \hat{g}_3}{2r + 2\rho_e + \lambda_f \hat{\mu}_{hn}}$, which are independent of λ_s.

It remains to find the values of $\hat{g}_2$ and g'_2.

To state how g_2 is determined by (19), let $c_1 \equiv 2(r+\rho_u+\rho_d)+\lambda_d(\mu_{lo}+\mu_{hn})$, $c_2 \equiv 2\rho_u+\lambda_d\mu_{hn}$, and

$$D \equiv \begin{bmatrix} c_1+\lambda_f(\mu_{fo}+\mu_{fe}) & c_1+\lambda_f\mu_{fn} \\ c_2-\lambda_f\mu_{hn} & c_2+2r+\lambda_f(\mu_{lo}+\mu_{fn}) \end{bmatrix}.$$

Also, let $\alpha_1 \equiv \frac{-D_{21}}{det(D)}$ and $\alpha_2 \equiv \frac{D_{11}}{det(D)}$. Then,

$$g_2 = \alpha_1\left(u_h-u_l-\lambda_f\mu_{fe}g_3\right)+\alpha_2\left(u_f-u_l-2\rho_e g_3-\lambda_s\mu_{fe}g_3\right).$$

Note that, when $n_f=0$,

$$\begin{aligned}
&\hat{c}_1 = 2(r+\rho_u+\rho_d)+\lambda_d(\hat{\mu}_{lo}+\hat{\mu}_{hn})>0,\\
&\hat{c}_2 = 2\rho_u+\lambda_d\hat{\mu}_{hn}>0,\\
&\hat{D} = \begin{bmatrix} \hat{c}_1 & \hat{c}_1 \\ \hat{c}_2-\lambda_f\hat{\mu}_{hn} & \hat{c}_2+2r+\lambda_f\hat{\mu}_{lo} \end{bmatrix} \quad \text{(with a strictly positive determinant)},\\
&\hat{\alpha}_1 = \frac{-\hat{c}_2+\lambda_f\hat{\mu}_{hn}}{det(\hat{D})} \text{ (the exact value is unnecessary for our proof)}, \quad \text{and}\\
&\hat{\alpha}_2 = \frac{\hat{c}_1}{det(\hat{D})} = \frac{1}{2r+\lambda_f(\hat{\mu}_{lo}+\hat{\mu}_{hn})}>0,
\end{aligned} \tag{28}$$

which are all independent of λ_s. It follows that

$$\hat{g}_2 = \hat{\alpha}_1(u_h-u_l)+\hat{\alpha}_2(u_f-u_l-2\rho_e\hat{g}_3)$$

is independent of λ_s.

Last, let c_1', c_2', α_1', and α_2' be the corresponding variables' derivatives: e.g., $c_1' \equiv \left.\frac{\partial c_1}{\partial n_f}\right|_{n_f=0}$. The derivatives are all independent of λ_s, because $\hat{\mu}$ and μ' are independent of λ_s. Therefore,

$$\begin{aligned}
g_2' =&\alpha_1'(u_h-u_l-\lambda_f\hat{\mu}_{fe}\hat{g}_3)-\hat{\alpha}_1\lambda_f(\mu_{fe}'\hat{g}_3+\hat{\mu}_{fe}g_3')\\
&+\alpha_2'\left(u_f-u_l-2\rho_e\hat{g}_3-\lambda_s\hat{\mu}_{fe}\hat{g}_3\right)-\hat{\alpha}_2\left(2\rho_e g_3'+\lambda_s\mu_{fe}'\hat{g}_3+\lambda_s\hat{\mu}_{fe}g_3'\right)\\
=&\alpha_1'(u_h-u_l)-\hat{\alpha}_1\lambda_f\mu_{fe}'\hat{g}_3+\alpha_2'\left(u_f-u_l-2\rho_e\hat{g}_3\right)-\hat{\alpha}_2\left(2\rho_e g_3'+\lambda_s\mu_{fe}'\hat{g}_3\right). \quad (\text{as } \hat{\mu}_{fe}=0)
\end{aligned}$$

Only the last term $-\hat{\alpha}_2\left(2\rho_e g_3' + \lambda_s\mu'_{fe}\hat{g}_3\right)$ is (affinely) dependent on λ_s, through $-\hat{\alpha}_2\mu'_{fe}\hat{g}_3\lambda_s$ and $g_3' = -\frac{(\lambda_f\mu'_{hn}+\lambda_s\mu'_{fn})\hat{g}_3}{2r+2\rho_e+\lambda_f\hat{\mu}_{hn}}$. As such, $g_2' = \gamma_1\lambda_s + \gamma_2$, for $\gamma_1 = \hat{\alpha}_2\hat{g}_3\left(\frac{2\rho_e\mu'_{fn}}{2r+2\rho_e+\lambda_f\hat{\mu}_{hn}} - \mu'_{fe}\right)$ and some γ_2 which aggregates all remaining terms. Both γ_1 and γ_2 are independent of λ_s. Finally,

$$
\begin{aligned}
rv'_{fn} &= \lambda_f\left(\hat{g}_2\mu'_{lo} + \hat{\mu}_{lo}g_2'\right) + \lambda_s\left(\hat{g}_3\mu'_{fe} + \hat{\mu}_{fe}g_3'\right) \\
&= \lambda_f\left(\hat{g}_2\mu'_{lo} + \hat{\mu}_{lo}(\gamma_1\lambda_s + \gamma_2)\right) + \lambda_s\hat{g}_3\mu'_{fe} \quad (\text{as } \hat{\mu}_{fe} = 0) \\
&= (\lambda_f\hat{\mu}_{lo}\gamma_1 + \hat{g}_3\mu'_{fe})\lambda_s + (\lambda_f\hat{g}_2\mu'_{lo} + \lambda_f\hat{\mu}_{lo}\gamma_2),
\end{aligned}
$$

where the coefficient of λ_s and the last term are both independent of λ_s.

It remains to show that the coefficient of λ_s is strictly positive:

$$
\begin{aligned}
\lambda_f\hat{\mu}_{lo}\gamma_1 + \hat{g}_3\mu'_{fe} &= \lambda_f\hat{\mu}_{lo}\hat{\alpha}_2\hat{g}_3\left(\frac{2\rho_e\mu'_{fn}}{2r+2\rho_e+\lambda_f\hat{\mu}_{hn}} - \mu'_{fe}\right) + \hat{g}_3\mu'_{fe} \\
&> -\lambda_f\hat{\mu}_{lo}\hat{\alpha}_2\hat{g}_3\mu'_{fe} + \hat{g}_3\mu'_{fe} \quad (\text{as } \hat{\mu}_{lo}, \hat{\alpha}_2, \hat{g}_3, \mu'_{fn}, \hat{\mu}_{hn} \text{ are strictly positive}) \\
&= \mu'_{fe}\hat{g}_3\left(1 - \lambda_f\hat{\alpha}_2\hat{\mu}_{lo}\right) \\
&= \mu'_{fe}\hat{g}_3\left(1 - \frac{\lambda_f\hat{\mu}_{lo}}{2r + \lambda_f(\hat{\mu}_{lo} + \hat{\mu}_{hn})}\right) \quad (\text{from (28)}) \\
&> 0. \quad (\text{as } \mu'_{fe} \text{ and } \hat{g}_3 \text{ are strictly positive})
\end{aligned}
$$

A.11 Proof for Part 1 of Proposition 6

For any regular environment $\theta \equiv (n, r, u, \rho, \lambda)$, we consider a sequence $\theta^\kappa \equiv (n, r, u, \rho, \kappa\lambda)$ with $\kappa \to \infty$. Let μ^κ be the unique steady-state solution of $P(\theta^\kappa)$ and v^κ be the unique solution of $V(\theta^\kappa)$ with $\mu(t)$ being replaced by μ^κ.

In solving $P(\theta^\kappa)$, it is more convenient to take $z \equiv 1/\kappa$ and define another market $\psi^z \equiv (n, r, u, z\rho, \lambda)$. (i.e., low type-change rates, instead of high search rates) and solve $P(\psi^z)$. It is easy to verify that the unique steady-state solution μ^κ of $P(\theta^\kappa)$ also uniquely solves $P(\psi^z)$. Last, define $\psi^0 \equiv (n, r, u, 0, \lambda)$.

Lemma 5. $\mu^0 \in \mathbb{R}^{\mathcal{T}}$ *is a steady-state solution of* $P(\psi^0)$ *if and only if*

1. *(when $n_f + n_h > n_a$)* $\mu^0_{ho} = \min\{n_a, n_h\}$, $\mu^0_{hn} = n_h - \mu^0_{ho}$, $\mu^0_{lo} = 0$, $\mu^0_{ln} = n_l$, $\mu^0_{fo} = \max\{0, n_a - n_h\}$, $\mu^0_{fe} = 0$, *and* $\mu^0_{fn} = n_f - \mu^0_{fo} - \mu^0_{fe}$.

2. *(when $n_f + n_h < n_a$)* $\mu^0_{ho} = n_h$, $\mu^0_{hn} = 0$, $\mu^0_{lo} = n_a - n_h - n_f$, $\mu^0_{ln} = n_l - \mu^0_{lo}$, $\mu^0_{fn} = 0$, *and* $\mu^0_{fo} + \mu^0_{fe} = n_f$.

The problem $P(\psi^0)$ has multiple steady-state solutions in Case 2 ($n_f + n_h < n_a$), where many possible combinations of (μ_{fo}, μ_{fe}) satisfy $\mu_{fo} + \mu_{fe} = n_f$.

(Proof) The problem $P(\psi^0)$ consists of

$$(\lambda_d \mu_{lo} + \lambda_f \mu_{fo} + \lambda_f \mu_{fe})\mu_{hn} = 0, \qquad \text{(from (μ-ho))}$$

$$(\lambda_d \mu_{hn} + \lambda_f \mu_{fn})\mu_{lo} = 0, \qquad \text{(from (μ-ln))}$$

$$\lambda_f (\mu_{lo}\mu_{fn} - \mu_{hn}\mu_{fo}) + \lambda_s \mu_{fn}\mu_{fe} = 0, \qquad \text{(from (μ-fo))}$$

and the following four conditions that replace (μ-hn), (μ-lo), (μ-fe), and (μ-fn):

$$\mu_{ho} + \mu_{hn} = n_h, \quad \mu_{lo} + \mu_{ln} = n_l, \quad \mu_{ho} + \mu_{lo} + \mu_{fo} + \mu_{fe} = n_a, \quad \text{and} \quad \mu_{fn} + \mu_{fo} + \mu_{fe} = n_f.$$

It follows from $\lambda_d, \lambda_f, \lambda_d > 0$ that

$$\mu_{lo}\mu_{hn} = \mu_{fo}\mu_{hn} = \mu_{fe}\mu_{hn} = \mu_{lo}\mu_{fn} = \mu_{fn}\mu_{fe} = 0.$$

Suppose that $n_f + n_h > n_a$. If $\mu_{lo} > 0$ or $\mu_{fe} > 0$, then $\mu_{hn} = 0$ and $\mu_{fn} = 0$, which results in a contradiction: $\mu_{ho} + (\mu_{fo} + \mu_{fe}) = n_h + n_f > n_a$. As $\mu_{lo} = \mu_{fe} = 0$, either $\mu_{fo} = 0$ or $\mu_{hn} = 0$. As $\mu_{ho} + \mu_{lo} + \mu_{fo} + \mu_{fe} = n_a > 0$, if $\mu_{fo} = 0$, then $\mu_{ho} = n_a$; for otherwise $\mu_{hn} = 0$ implies that $\mu_{ho} = n_h$ and $\mu_{fo} = n_a - n_h$. On the other hand, if $n_f + n_h < n_a$, then $\mu_{lo} > 0$, which implies that $\mu_{hn} = \mu_{fn} = 0$. Thus, $\mu_{ho} = n_h$, $\mu_{fo} + \mu_{fe} = n_f$, and $\mu_{lo} = n_a - n_h - n_f$. ■

The following lemma implies that $\lim_{\kappa\to\infty} \mu^\kappa$ exists in $\mathbb{R}^{\mathcal{T}}_+$:

Lemma 6. *There exists a solution μ^* of $P(\psi^0)$ such that $\mu^* \equiv \lim_{\kappa\to\infty} \mu^\kappa$.*

(Proof) For each $z \equiv 1/\kappa$, let $F(\mu, \psi^z)$ denote the right-hand sides of the population equations (μ-hn)-(μ-fe) for a market $\psi^z \equiv (n, r, u, z\rho, \lambda)$. Define $f(\mu, z) \equiv -\|F(\mu, \psi^z)\|$, where $\|\cdot\|$ denotes the Euclidean norm. It is clear that μ^κ with $\kappa = 1/z$ is the unique maximizer of f with the maximum value equals zero. Let $M(z) \equiv \{\mu^{1/z}\}$.

We similarly define $F(\mu, \psi^0)$ as the right-hand sides of the population equations for the market ψ^0 and $f(\mu, 0)$. Let $M(0)$ be the solution set of $\max_\mu f(\mu, 0)$. According to Lemma 5, the solution set $M(0)$ is singleton if $n_h + n_f > n_a$; for otherwise, $M(0)$ contains multiple solutions, each being different from others only in (μ_{fo}, μ_{fe}) under the constraint $\mu_{fo} + \mu_{fe} = n_f$.

The function f is continuous in μ and z because the equations F are continuous. It follows from Berge's Maximum Theorem that the correspondence $M(\cdot)$ is upper hemicontinuous at $z = 0$:

1. (when $n_h + n_f > n_a$) μ^κ converges to the unique solution of $P(\psi^0)$.
2. (when $n_h + n_f < n_a$) for each type $i \neq fo, fe$, the population μ_i^κ converges to μ_i^0 given in Lemma 5, and $\mu_{fo}^\kappa + \mu_{fe}^\kappa$ converges to n_f.

It remains to show that, when $n_h + n_f > n_a$, the sequence μ_{fe}^κ converges. (The convergence of μ_{fo}^κ follows immediately from $\lim_{\kappa\to\infty}(\mu_{fo}^\kappa + \mu_{fe}^\kappa) = n_f$.)

For every $\kappa > 0$, $\kappa(\lambda_f \mu_{hn}^\kappa + \lambda_s \mu_{fn}^\kappa)\mu_{fe}^\kappa = \rho_e(n_f - \mu_{fn}^\kappa - \mu_{fe}^\kappa)$ (from (μ-fe)), or equivalently that

$$\mu_{fe}^\kappa = \frac{\rho_e(n_f - \mu_{fn}^\kappa)}{\rho_e + \kappa(\lambda_f \mu_{hn}^\kappa + \lambda_s \mu_{fn}^\kappa)}. \tag{29}$$

We find $\lim_{\kappa\to\infty} \kappa(\lambda_f \mu_{hn}^\kappa + \lambda_s \mu_{fn}^\kappa$ from

$$\kappa\mu_{hn}^\kappa(\lambda_d \mu_{lo}^\kappa + \lambda_f \mu_{fo}^\kappa + \lambda_f \mu_{fe}^\kappa) = -\rho_d \mu_{hn}^\kappa + \rho_u \mu_{ln}^\kappa, \qquad \text{(from (μ-hn))}$$
$$\kappa(\lambda_v \mu_{hn}^\kappa + \lambda_f \mu_{fn}^\kappa)\mu_{lo}^\kappa = -\rho_u \mu_{lo}^\kappa + \rho_d \mu_{ho}^\kappa. \qquad \text{(from (μ-lo))}$$

By the convergence of μ_i^κ for $i \neq fo, fe$, and the convergence of $\mu_{fe}^\kappa + \mu_{fo}^\kappa$ to n_f,

$$\lim_{\kappa\to\infty} \kappa\mu_{hn}^\kappa = \frac{\rho_u \mu_{ln}^*}{\lambda_d \mu_{lo}^* + \lambda_f n_f}, \quad \text{and}$$
$$\lim_{\kappa\to\infty} \kappa(\lambda_v \mu_{hn}^\kappa + \lambda_f \mu_{fn}^\kappa) = \frac{\rho_d \mu_{ho}^* - \rho_u \mu_{lo}^*}{\mu_{lo}^*} = \frac{\rho_d n_h - \rho_u \mu_{lo}^*}{\mu_{lo}^*} = \frac{\rho_u \mu_{ln}^*}{\mu_{lo}^*}.$$

It follows that

$$\lim_{\kappa\to\infty} \kappa(\lambda_f \mu_{hn}^\kappa + \lambda_s \mu_{fn}^\kappa) = \frac{\lambda_s \rho_u \mu_{ln}^*}{\lambda_f \mu_{lo}^*} + \left(\lambda_f - \frac{\lambda_d \lambda_s}{\lambda_f}\right) \frac{\rho_u \mu_{ln}^*}{\lambda_d \mu_{lo}^* + \lambda_f n_f}$$
$$= \frac{\rho_u \mu_{ln}^*}{\mu_{lo}^*} \frac{\lambda_f \mu_{lo}^* + \lambda_s n_f}{\lambda_d \mu_{lo}^* + \lambda_f n_f} > 0.$$

Therefore,

$$\mu_{fe}^* \equiv \lim_{\kappa\to\infty} \mu_{fe}^\kappa = \frac{n_f}{1 + \frac{\rho_u \mu_{ln}^*}{\rho_e \mu_{lo}^*} \frac{\lambda_f \mu_{lo}^* + \lambda_s n_f}{\lambda_d \mu_{lo}^* + \lambda_f n_f}}, \quad \text{and} \quad \mu_{fo}^* = n_f - \mu_{fe}^*.$$

■

A.12 Proof for Part 2 of Proposition 6

We divide the proof into two lemmas.

Lemma 7. *For every $i \in \mathcal{T}$, if $\mu_i^* = 0$, then $\mu_i^{**} \equiv \lim_{\kappa\to\infty} \kappa \mu_i^\kappa$ exists in $\mathbb{R}$.*

(Proof) The following table summarizes the population limits for some types from Lemma 5 and Lemma 6:

	A. $n_a < n_h$	B. $n_h < n_a < n_h + n_f$	C. $n_h + n_f < n_a$
$\mu_{ho}^* =$	n_a	n_h	n_h
$\mu_{fo}^* =$	0	$n_a - n_h$	$< n_f$
$\mu_{lo}^* =$	0	0	$n_a - n_f - n_h$
$\mu_{fe}^* =$	0	0	> 0

As μ_{ho}^* and μ_{ln}^* are always strictly positive, we consider other types only:

Suppose $\mu_{hn}^* = 0$ (Cases A, B and C): for any κ,

$$\kappa(\lambda_c \mu_{lo}^\kappa + \lambda_f \mu_{fo}^\kappa + \lambda_f \mu_{fe}^\kappa)\mu_{hn}^\kappa = -\rho_d \mu_{hn}^\kappa + \rho_u \mu_{ln}^\kappa. \quad \text{(from (μ-hn))}$$

By Lemma 1, $\mu_i^\kappa > 0$ for every $i \in \mathcal{T}$,

$$\kappa\mu_{hn}^\kappa = \frac{\rho_u\mu_{ln}^\kappa - \rho_d\mu_{hn}^\kappa}{\lambda_d\mu_{lo}^\kappa + \lambda_f\mu_{fo}^\kappa + \lambda_f\mu_{fe}^\kappa} \tag{30}$$

$$\implies \mu_{hn}^{**} \equiv \lim_{\kappa\to\infty}\kappa\mu_{hn}^\kappa = \frac{\rho_u\mu_{ln}^* - \rho_d\mu_{hn}^*}{\lambda_d\mu_{lo}^* + \lambda_f(\mu_{fo}^* + \mu_{fe}^*)} = \frac{\rho_u\mu_{ln}^*}{\lambda_d\mu_{lo}^* + \lambda_f(\mu_{fo}^* + \mu_{fe}^*)} > 0. \tag{31}$$

Suppose $\mu_{lo}^* = 0$ (Cases A and B): for every κ,

$$(\lambda_d\mu_{hn}^\kappa + \lambda_f\mu_{fn}^\kappa)(\kappa\mu_{lo}^\kappa) = \rho_d\mu_{ho}^\kappa - \rho_u\mu_{lo}^\kappa. \quad \text{(from } (\mu\text{-lo}))$$

It follows that

$$\mu_{lo}^{**} \equiv \lim_{\kappa\to\infty}\kappa\mu_{lo}^\kappa = \frac{\rho_d\mu_{ho}^* - \rho_u\mu_{lo}^*}{\lambda_d\mu_{hn}^* + \lambda_f\mu_{fn}^*} = \frac{\rho_d\mu_{ho}^*}{\lambda_d\mu_{hn}^* + \lambda_f\mu_{fn}^*}.$$

Suppose $\mu_{fe}^* = 0$ (Cases A and B): for every κ,

$$(\lambda_f\mu_{hn}^\kappa + \lambda_s\mu_{fn}^\kappa)(\kappa\mu_{fe}^\kappa) = \rho_e\mu_{fo}^\kappa. \quad \text{(from } (\mu\text{-fe}))$$

It follows that

$$\mu_{fe}^{**} \equiv \lim_{\kappa\to\infty}\kappa\mu_{fe}^\kappa = \frac{\rho_e\mu_{fo}^*}{\lambda_f\mu_{hn}^* + \lambda_s\mu_{fn}^*}.$$

Suppose $\mu_{fo}^* = 0$ (Case A): for every κ,

$$\kappa(\lambda_f\mu_{lo}^\kappa + \lambda_s\mu_{fe}^\kappa)\mu_{fn}^\kappa = \kappa\lambda_f\mu_{hn}^\kappa\mu_{fo}^\kappa + \rho_e\mu_{fo}^\kappa. \quad \text{(from } (\mu\text{-fo}))$$

It follows from the convergence of $\kappa\mu_{lo}^\kappa$ and $\kappa\mu_{fe}^\kappa$ in Case A that

$$\mu_{fo}^{**} \equiv \lim_{\kappa\to\infty}\kappa\mu_{fo}^\kappa = \lim_{\kappa\to\infty}\frac{(\lambda_f(\kappa\mu_{lo}^\kappa) + \lambda_s(\kappa\mu_{fe}^\kappa))\mu_{fn}^\kappa - \rho_e\mu_{fo}^\kappa}{\lambda_f\mu_{hn}^\kappa} = \frac{(\lambda_f\mu_{lo}^{**} + \lambda_s\mu_{fe}^{**})\mu_{fn}^*}{\lambda_f\mu_{hn}^*}.$$

Finally, suppose $\mu_{fn}^* = 0$ (Case C): for every κ,

$$\mu_{hn}^\kappa\mu_{fo}^\kappa + \mu_{hn}^\kappa\mu_{fe}^\kappa = \mu_{lo}^\kappa\mu_{fn}^\kappa. \quad \text{(from } (\mu\text{-fn}))$$

As $\mu_{lo}^\kappa > 0$ (Lemma 1), we have

$$\kappa\mu_{fn}^\kappa = \frac{\kappa\mu_{hn}^\kappa(\mu_{fo}^\kappa + \mu_{fe}^\kappa)}{\mu_{lo}^\kappa}. \tag{32}$$

It follows from the convergences of $\kappa\mu_{hn}^\kappa$ that

$$\mu_{fn}^{**} \equiv \lim_{\kappa\to\infty} \kappa\mu_{fn}^\kappa = \lim_{\kappa\to\infty} \frac{\kappa\mu_{hn}^\kappa(\mu_{fo}^\kappa + \mu_{fe}^\kappa)}{\mu_{lo}^\kappa} = \frac{\mu_{hn}^{**}(\mu_{fo}^* + \mu_{fe}^*)}{\mu_{lo}^*} > 0. \tag{33}$$

■

Lemma 8. *For any $i \in \mathcal{T}$, if $\mu_i^* > 0$, then $\mu_i^{**} \equiv \lim_{\kappa\to\infty} \kappa(\mu_i^\kappa - \mu_i^*)$ exists in $\mathbb{R}$.*

(Proof)

First, consider Case A ($n_a < n_h$): Only $\mu_{ho}^*, \mu_{fn}^*, \mu_{ln}^*$ are strictly positive. As $\kappa \to \infty$,

$$\kappa(\mu_{ho}^\kappa - \mu_{ho}^*) = \kappa(n_a - \mu_{lo}^\kappa - \mu_{fo}^\kappa - \mu_{fe}^\kappa) - \kappa(n_a - \mu_{lo}^* - \mu_{fo}^* - \mu_{fe}^*) \to -\mu_{lo}^{**} - \mu_{fo}^{**} - \mu_{fe}^{**},$$

where the convergence of $\kappa\mu_{lo}^\kappa$, $\kappa\mu_{fo}^\kappa$, and $\kappa\mu_{fe}^\kappa$ holds by Lemma 7.

We similarly find the convergence speed for μ_{fn}^κ and μ_{ln}^κ:

$$\begin{aligned}
\kappa(\mu_{fn}^\kappa - \mu_{fn}^*) &= \kappa(n_f - \mu_{fo}^\kappa - \mu_{fe}^\kappa) - \kappa(n_f - \mu_{fo}^* - \mu_{fe}^*) \to -\mu_{fo}^{**} - \mu_{fe}^{**}, \quad \text{and} \\
\kappa(\mu_{ln}^\kappa - \mu_{ln}^*) &= \kappa(n_l - \mu_{ln}^\kappa) - \kappa(n_l - \mu_{lo}^*) \to -\mu_{lo}^{**}.
\end{aligned}$$

Next, consider Case B ($n_h < n_a < n_h + n_f$):

Only $\mu_{ho}^*, \mu_{fo}^*, \mu_{ln}^*$, and μ_{fn}^* are strictly positive. As $\kappa \to \infty$,

$$\begin{aligned}
\kappa(\mu_{ho}^\kappa - \mu_{ho}^*) &= \kappa(n_h - \mu_{hn}^\kappa) - \kappa(n_h - \mu_{hn}^*) \to -\mu_{hn}^{**}, \\
\kappa(\mu_{ln}^\kappa - \mu_{ln}^*) &= \kappa(n_l - \mu_{lo}^\kappa) - \kappa(n_l - \mu_{lo}^*) \to -\mu_{lo}^{**}, \\
\kappa(\mu_{fo}^\kappa - \mu_{fo}^*) &= \kappa(\mu_{fo}^\kappa - (n_a - n_h)) = -\kappa\mu_{lo}^\kappa - \kappa\mu_{fe}^\kappa - \kappa(\mu_{ho}^\kappa - n_h) \\
&\to -\mu_{lo}^{**} - \mu_{fe}^{**} + \mu_{hn}^{**}, \quad \text{and} \\
\kappa(\mu_{fn}^\kappa - \mu_{fn}^*) &= \kappa(\mu_{fn}^\kappa - (n_f - n_a + n_h)) = -\kappa\mu_{fe}^\kappa - \kappa(\mu_{fo}^\kappa - (n_a - n_h)) \\
&\to -\mu_{fe}^{**} + (\mu_{lo}^{**} + \mu_{fe}^{**} - \mu_{hn}^{**}).
\end{aligned}$$

Finally, in Case C ($n_h + n_f < n_a$), we have $\mu^*_{ho}, \mu^*_{lo}, \mu^*_{ln}, \mu^*_{fo}$, and μ^*_{fe} that are strictly positive. The proof for the first three types are similar to the previous cases: as $\kappa \to \infty$,

$$\begin{aligned}
\kappa(\mu^\kappa_{ho} - \mu^*_{ho}) &= \kappa(n_h - \mu^\kappa_{hn}) - \kappa(n_h - \mu^*_{hn}) \to -\mu^{**}_{hn}, \\
\kappa(\mu^\kappa_{lo} - \mu^*_{lo}) &= \kappa(\mu^\kappa_{lo} - (n_a - n_h - n_f)) = -\kappa(\mu^\kappa_{fo} + \mu^\kappa_{fe} - n_f) - \kappa(\mu^\kappa_{ho} - n_h) \\
&\to -\mu^{**}_{fn} + \mu^{**}_{hn}, \qquad (34) \\
\kappa(\mu^\kappa_{ln} - \mu^*_{ln}) &= \kappa(n_l - \mu^\kappa_{lo}) - \kappa(n_l - \mu^*_{lo}) \to -\mu^{**}_{lo} = \mu^{**}_{fn} - \mu^{**}_{hn}. \qquad (35)
\end{aligned}$$

It remains to show the convergence speed for μ^κ_{fo} and μ^κ_{fe}. On the one hand, from (μ-fe) and the convergence of μ^κ_{fe}, μ^κ_{fo}, $\kappa\mu^\kappa_{hn}$, and $\kappa\mu^\kappa_{fn}$, we have

$$\kappa(\lambda_f \mu^\kappa_{hn} + \lambda_s \mu^\kappa_{fn})\mu^\kappa_{fe} = \rho_e \mu^\kappa_{fo} \quad \text{and} \quad (\lambda_f \mu^{**}_{hn} + \lambda_s \mu^{**}_{fn})\mu^*_{fe} = \rho_e \mu^*_{fo}.$$

Let $\phi^\kappa \equiv \kappa(\lambda_f \mu^\kappa_{hn} + \lambda_s \mu^\kappa_{fn})$, and $\phi^{**} \equiv \lambda_f \mu^{**}_{hn} + \lambda_s \mu^{**}_{fn}$. Then,

$$\rho_e \kappa(\mu^\kappa_{fo} - \mu^*_{fo}) = \phi^\kappa \kappa \mu^\kappa_{fe} - \phi^{**}\kappa\mu^*_{fe} = \kappa(\phi^\kappa - \phi^{**})\mu^\kappa_{fe} + \phi^{**}\kappa(\mu^\kappa_{fe} - \mu^*_{fe}). \qquad (36)$$

On the other hand, from $\mu^\kappa_{fn} + \mu^\kappa_{fo} + \mu^\kappa_{fe} = n_f$ and $\mu^*_{fo} + \mu^*_{fe} = n_f$, we have

$$\kappa(\mu^\kappa_{fo} - \mu^*_{fo}) + \kappa(\mu^\kappa_{fe} - \mu^*_{fe}) = -\kappa\mu^\kappa_{fn}. \qquad (37)$$

By summarizing (36) and (37), for every κ,

$$\begin{bmatrix} \kappa(\mu^\kappa_{fo} - \mu^*_{fo}) \\ \kappa(\mu^\kappa_{fe} - \mu^*_{fe}) \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \rho_e & -\phi^{**} \end{bmatrix}^{-1} \begin{bmatrix} -\kappa\mu^\kappa_{fn} \\ \kappa(\phi^\kappa - \phi^{**})\mu^\kappa_{fe} \end{bmatrix},$$

where the inverse matrix is well-defined because $\phi^{**} > 0$ (see (31) and (33)). Note that $\kappa\mu^\kappa_{fn}$ and μ^κ_{fe} converge (see (33) and Lemma 6). It remains to prove that

$$\kappa(\phi^\kappa - \phi^{**}) = \lambda_f \kappa(\kappa\mu^\kappa_{hn} - \mu^{**}_{hn}) + \lambda_s \kappa(\kappa\mu^\kappa_{fn} - \mu^{**}_{fn}) \quad \text{converges as } \kappa \to \infty.$$

From (30), (31), $\mu^*_{hn} = 0$, and $\mu^*_{fo} + \mu^*_{fe} = n_f$,[22] we have

$$\kappa(\kappa\mu^\kappa_{hn} - \mu^{**}_{hn}) = \frac{\rho_u\kappa\mu^\kappa_{ln} - \rho_d\kappa\mu^\kappa_{hn}}{\lambda_d\mu^\kappa_{lo} + \lambda_f(\mu^\kappa_{fo} + \mu^\kappa_{fe})} - \frac{\rho_u(\kappa\mu^*_{ln})}{\lambda_d\mu^*_{lo} + \lambda_f n_f}.$$

To ease expositions, let A^κ and A^* denote the denominators in the above equation. Then,

$$\begin{aligned}
\kappa(\kappa\mu^\kappa_{hn} - \mu^{**}_{hn}) &= \frac{\rho_u\kappa\mu^\kappa_{ln} - \rho_d\kappa\mu^\kappa_{hn}}{A^\kappa} - \frac{\rho_u\kappa\mu^*_{ln}}{A^*} \\
&= \frac{\rho_u\kappa(\mu^\kappa_{ln} - \mu^*_{ln}) - \rho_d\kappa\mu^\kappa_{hn}}{A^\kappa} + \rho_u\mu^*_{ln}\kappa\left(\frac{1}{A^\kappa} - \frac{1}{A^*}\right) \\
&= \frac{\rho_u\kappa(\mu^\kappa_{ln} - \mu^*_{ln}) - \rho_d\kappa\mu^\kappa_{hn}}{A^\kappa} - \rho_u\mu^*_{ln}\frac{\lambda_d\kappa(\mu^\kappa_{lo} - \mu^*_{lo}) - \lambda_f\kappa\mu^\kappa_{fn}}{A^\kappa A^*},
\end{aligned} \tag{38}$$

which converges by (31), (33), (34), and (35).

Then, from (32), (33), and $\mu^*_{fo} + \mu^*_{fe} = n_f$, we have

$$\kappa(\kappa\mu^\kappa_{fn} - \mu^{**}_{fn}) = \frac{\kappa\mu^\kappa_{hn}\kappa(\mu^\kappa_{fo} + \mu^\kappa_{fe})}{\mu^\kappa_{lo}} - \frac{\mu^{**}_{hn}\kappa n_f}{\mu^*_{lo}}.$$

Then,

$$\begin{aligned}
\kappa(\kappa\mu^\kappa_{fn} - \mu^{**}_{fn}) &= \frac{\kappa\mu^\kappa_{hn}\kappa(\mu^\kappa_{fo} + \mu^\kappa_{fe} - n_f)}{\mu^\kappa_{lo}} + \frac{\kappa^2\mu^\kappa_{hn}n_f}{\mu^\kappa_{lo}} - \frac{\kappa\mu^{**}_{hn}n_f}{\mu^*_{lo}} \\
&= -\frac{(\kappa\mu^\kappa_{hn})(\kappa\mu^\kappa_{fn})}{\mu^\kappa_{lo}} + \frac{\kappa(\kappa\mu^\kappa_{hn} - \mu^{**}_{hn})n_f}{\mu^\kappa_{lo}} + \frac{\kappa\mu^{**}_{hn}n_f}{\mu^\kappa_{lo}} - \frac{\kappa\mu^{**}_{hn}n_f}{\mu^*_{lo}} \\
&= -\frac{(\kappa\mu^\kappa_{hn})(\kappa\mu^\kappa_{fn})}{\mu^\kappa_{lo}} + \frac{\kappa(\kappa\mu^\kappa_{hn} - \mu^{**}_{hn})n_f}{\mu^\kappa_{lo}} - \frac{\mu^{**}_{hn}n_f\kappa(\mu^\kappa_{lo} - \mu^*_{lo})}{\mu^\kappa_{lo}\mu^*_{lo}},
\end{aligned}$$

which converges by (31), (33),(34), and (38). ■

A.13 Proof of Proposition 7

By Lemma 5 and Lemma 6, if $n_a < n_h + n_f$, then $\mu^*_{ho} = \min\{n_a, n_h\}$, $\mu^*_{fo} = \max\{0, n_a - n_h\}$, $\mu^*_{fe} = 0$, and $\mu^*_{lo} = 0$. Since μ^* coincides with the efficient asset allocation $\overline{\mu}$, we have

[22] Recall that we are considering Case C ($n_h + n_f < n_a$).

$W^* = \overline{W}$. The independence of W^* on u_f and u_e is trivial as $\mu^*_{fo} = \mu^*_{fe} = 0$. The independent of W^* on λ_d also follows from $\overline{W}$'s independence of any search friction. When $n_h < n_a < n_h + n_f$, we have $\mu_{fo} > 0$, so $W^* = \overline{W}$ is strictly increasing in u_f.

If $n_a > n_h + n_f$, then $rW^* = r\overline{W} - \mu^*_{fe}(u_f - u_e)$. We have $W^* < \overline{W}$ because

$$\mu^*_{fe} = \frac{n_f}{1 + \frac{\rho_u \mu^*_{ln}}{\rho_e \mu^*_{lo}} \frac{\lambda_f \mu^*_{lo} + \lambda_s n_f}{\lambda_d \mu^*_{lo} + \lambda_f n_f}} > 0,$$

The welfare W^* is increasing in u_f and u_e as μ^*_{fo} and μ^*_{fe} are strictly positive. Moreover, μ^*_{fe} is decreasing in λ_s and increasing in λ_d. Thus, the welfare W^* is increasing in λ_s and decreasing in λ_d.

References

Ahern, K. R. (2012): "Bargaining power and industry dependence in mergers," *Journal of Financial Economics*, 103(3), 530–550.

Allen, F., and A. Babus (2009): "Networks in finance," *The Network Challenge*, pp. 367–382.

Almeida, H., M. Campello, and D. Hackbarth (2011): "Liquidity mergers," *Journal of Financial Economics*, 102(3), 526–558.

Arcot, S., Z. Fluck, J.-M. Gaspar, and U. Hege (2015): "Fund managers under pressure: Rationale and determinants of secondary buyouts," *Journal of Financial Economics*, 115(1), 102–135.

Atkeson, A. G., A. L. Eisfeldt, and P.-O. Weill (2015): "Entry and exit in OTC derivatives markets," *Econometrica*, 83(6), 2231–2292.

David, J. (2017): "The aggregate implications of mergers and acquisitions," *Working Paper*.

Degeorge, F., J. Martin, and L. Phalippou (2016): "On secondary buyouts," *Journal of Financial Economics*, 120(1), 124–145.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005): "Over-the-Counter Markets," *Econometrica*, 73(6), 1815–1847.

Eisfeldt, A. L., and A. A. Rampini (2008): "Managerial incentives, capital reallocation, and the business cycle," *Journal of Financial Economics*, 87(1), 177–199.

Farboodi, M., G. Jarosch, and R. Shimer (2017): "The Emergence of Market Structure," Working Paper 23234, National Bureau of Economic Research.

Gofman, M. (2014): "A Network-Based Analysis of Over-the-Counter Markets," *Working Paper*.

Hirsch, M. W. (2012): *Differential topology*, vol. 33. Springer Science & Business Media.

Hirsch, M. W., and S. Smale (1973): *Differential equations, dynamical systems and linear algebra*. Academic Press College Division.

Hochberg, Y. V., A. Ljungqvist, and Y. Lu (2007): "Whom you know matters: Venture capital networks and investment performance," *The Journal of Finance*, 62(1), 251–301.

Hugonnier, J., B. Lester, and P.-O. Weill (2018): "Frictional intermediation in over-the-counter markets," Discussion paper.

——— (2020): "Frictional intermediation in over-the-counter markets," *The Review of Economic Studies*, 87(3), 1432–1469.

Jensen, M. C. (1991): "Eclipse of the Public Corporation," in *In The Law of Mergers, Acquisitions, and Reorganizations*.

Jovanovic, B., and P. L. Rousseau (2002): "The Q-theory of mergers," *American Economic Review*, 92(2), 198–204.

Kaplan, S. N., and P. Stromberg (2009): "Leveraged buyouts and private equity," *Journal of economic perspectives*, 23(1), 121–46.

Neklyudov, A. (2012): "Bid-ask spreads and the over-the-counter interdealer markets: Core and peripheral dealers," Discussion paper, Working Paper HEC Lausanne.

Nosal, E., and G. Rocheteau (2011): *Money, payments, and liquidity*. MIT press.

Nosal, E., Y.-Y. Wong, and R. Wright (2016): "Who Wants to be a Middleman?," Discussion paper, Mimeo, University of Wisconsin Madison.

Pagano, M., and P. Volpin (2012): "Securitization, transparency, and liquidity," *The Review of Financial Studies*, 25(8), 2417–2453.

Phillips, G. M., and A. Zhdanov (2017): "Venture Capital Investments and Merger and Acquisition Activity Around the World," Discussion paper, National Bureau of Economic Research.

Rhodes-Kropf, M., and D. T. Robinson (2008): "The market for mergers and the boundaries of the firm," *The Journal of Finance*, 63(3), 1169–1211.

Rhodes-Kropf, M., D. T. Robinson, and S. Viswanathan (2005): "Valuation waves and merger activity: The empirical evidence," *Journal of Financial Economics*, 77(3), 561–603.

Rubinstein, A., and A. Wolinsky (1987): "Middlemen," *The Quarterly Journal of Economics*, 102(3), 581–593.

Shen, J., B. Wei, and H. Yan (2021): "Financial intermediation chains in an over-the-counter market," *Management Science*, 67(7), 4623–4642.

Simsek, A., A. Ozdaglar, and D. Acemoglu (2007): "Generalized Poincare-Hopf theorem for compact nonsmooth regions," *Mathematics of Operations Research*, 32(1), 193–214.

Trejos, A., and R. Wright (2016): "Search-based models of money and finance: An integrated approach," *Journal of Economic Theory*, 164, 10–31.

Uslu, S. (2019): "Pricing and liquidity in decentralized asset markets," *Econometrica, forthcoming.*

Vayanos, D., and P.-O. Weill (2008): "A search-based theory of the on-the-run phenomenon," *The Journal of Finance*, 63(3), 1361–1398.

Wang, Y. (2012): "Secondary buyouts: Why buy and at what price?," *Journal of Corporate Finance*, 18(5), 1306–1325.

Yang, M., and Y. Zeng (2018): "The Coordination of Intermediation," *Available at SSRN 3265650.*