Liquidity Premia in Dynamic Bargaining Markets

First Version: November 27th 2001

Pierre-Olivier Weill*

November 9, 2003

Abstract

This paper develops a search-theoretic model of the cross-sectional distribution of asset returns, abstracting from risk premia and focusing exclusively on liquidity. I derive a float-adjusted return model (FARM), explaining the pricing of liquidity with a simple linear formula: In equilibrium, the liquidity spread of an asset is proportional to the inverse of its free float, the portion of its market capitalization available for sale. This suggests that the free float is an appropriate measure of liquidity, consistent with the linear specifications commonly estimated in the empirical literature. The qualitative predictions of the model corroborate much of the empirical evidence.

Keywords: Liquidity premia, Search
JEL Classification: G12, C78

*Department of Economics, Stanford University. E-mail: poweill@stanford.edu. I am deeply indebted to Darrell Duffie and Tom Sargent, for their supervision, many detailed comments and suggestions. I also would like to thank, for fruitful discussions and comments, Michael Rierson, Fernando Alvarez, Martine Carré, Ken Judd, Narayana Kocherlakota, Guy Laroque, Lars Ljungqvist, Lasse Heje Pedersen, Eva Nagypal, Stijn Van Nieuwerburgh, Dimitri Vayanos, Tan Wang, Randall Wright, participants of seminar at Stanford, the Kellogg School of Management, the Séminaire CREST, the North American Econometric Society Summer 2002 meeting, and the Society for Economic Dynamics 2002 meeting. All errors are mine.
1 Introduction

Why do different assets earn different expected returns? One fundamental reason is that they may bear different risks. Many empirical studies, however, suggest that risk characteristics cannot explain all variation in expected returns. After controlling for risk premia, expected returns appear to be positively related to bid-ask spreads, and negatively related to turnover and dollar trading volume. These patterns suggest that returns are related to liquidity, broadly defined as the ease of buying and selling. Liquidity is reflected in small trading costs, measured for instance by the bid-ask spread, and associated with the opportunity to buy or sell large quantities in a short time, at a similar price. These properties may be proxied by turnover or trading volume.

This paper provides a dynamic asset pricing model in which cross-sectional variation in asset returns is exclusively due to liquidity differences. Its first objective is to explain the pricing of liquidity differences and to suggest an appropriate measure of liquidity. Its second objective is to reproduce some of the qualitative relationships documented by the empirical literature.

The modelling strategy follows Duffie, Gárleanu and Pedersen [2001]. Trade is decentralized: Investors search for each other, meet in pairs, and bargain over prices. In this environment, a more liquid asset is defined as one with smaller trading delays: buyers and sellers of that asset are more likely to be found in a short time. The search assumption captures the fact that, in security markets, it is not always possible to locate instantly a trading counterparty. This applies, for instance, to decentralized security markets such as the NASDAQ and some other over-the-counter markets. More generally, this applies to trades which are not arranged at a centralized marketplace.

In the present model, as opposed to Duffie, Gărleanu and Pedersen [2001], many different assets are traded. Investors allocate their fixed budgets of search efforts to the various assets. They recognize that the value of searching for a particular asset is related to the likelihood of finding a counterparty for that asset in a short time. The first-order condition of the associated search optimization problem is key to the model’s implications, as it reflects

\[\text{For instance, Amihud and Mendelson [1991] describe some trades of off-the-run treasury notes as follows: “Investors who wish to trade a large quantity of notes have to go through considerable search in order to arrange the quantity desired, imposing an additional fixed cost of search.”}\]
how the likelihood of finding an asset is priced in equilibrium. Specifically, in equilibrium, investors are indifferent between searching for alternative traded assets, under natural technical conditions. This indifference property gives rise to a distribution of “liquidity premia.” Namely, an asset that is easier to find is sold at a higher price.

The first contribution of this paper is to derive a float-adjusted return model, or FARM, explaining the pricing of liquidity differences with the following linear formula

\[ R_k - R_L = \frac{\tilde{\phi}}{\phi_k} (R_M - R_L). \]  

(1)

On the left-hand side of (1), \( R_k \) denotes the return of asset \( k \), one of the many assets traded in the steady-state equilibrium, and \( R_L \) denotes the return of an appropriately defined ‘perfectly liquid’ asset. On the right-hand side of (1), \( \phi_k \) denotes the free float of asset \( k \), defined as the portion of the market capitalization available for sale; \( \tilde{\phi} \) denotes the market average free float; and, lastly, \( R_M \) is the float-weighted market return. In words, the FARM (1) states that, in steady state, the liquidity spread \( R_k - R_L \) of asset \( k \) is proportional to the inverse of its (relative) free float. The constant of proportionality is the liquidity spread \( R_M - R_L \) of the float-weighted market return.

Many empirical studies of liquidity spreads estimate linear models. They control for risk with some factor model, and measure an asset liquidity by its bid-ask spread, its trading volume, or its turnover. The FARM (1) suggests that, with such a linear specification, liquidity is best measured by the free float.

In traditional Walrasian asset-pricing models with liquidity effects, such as those of Amihud and Mendelson [1986], Constantinides [1986], Heaton and Lucas [1996], Vayanos [1998], and Huang [2003], assets can be bought and sold instantly, but differ by an exogenously given transaction cost. A more liquid asset is defined as one with a smaller transaction cost. In these models, cross-sectional variation in asset returns is explained by exogenously specified differences in transaction costs. A second contribution of this paper is to explain cross-sectional variation in asset returns without relying on an exogenously specified cross-sectional variation in transaction costs. Although, in the model proposed here, the search technology is the same for
all assets, heterogeneous bid-ask spreads arise endogenously. Cross-sectional variation in asset returns is explained by the distribution of ownership.

The analysis of this paper could not have been conducted in the one-asset model of Duffie, Gârleanu and Pedersen [2001], which examines the impact of liquidity on asset prices only by comparative statics. For instance, in the one-asset model, an increase in the quantity of shareholders results in a positive shift of the supply curve, and thus decreases the price of the asset. In the multiple-assets model, one can keep the total quantity of shareholders constant, and study an equilibrium in which some assets have more shareholders than others. This isolates a liquidity effect: An asset with more shareholders is easier to find, and has a higher price.

Search-theoretic approaches to liquidity have been explored in the monetary literature, following Kiyotaki and Wright [1989]. Most notably, Wallace [2000] focuses on the relative liquidity of intrinsically worthless assets (currency) and assets earning a positive dividend (bonds). The model presented here has no room for currency, and focuses on assets with relatively homogeneous characteristics. This paper is closely related to the independent work of Vayanos and Wang [2003]. In order to study the liquidity difference between on-the-run and off-the-run bonds, they provide a two-asset extension of Duffie, Gârleanu and Pedersen [2001]. They analyze the impact of investor heterogeneity on the concentration of liquidity across markets, and focus most of their analysis on welfare. In the present paper, by contrast, I analyze the impact of asset heterogeneity, and focus most of my analysis on pricing and measurement. In particular, I show that the exact linear formula of FARM (1) holds in an equilibrium with arbitrarily many assets.

The last section of the paper addresses time variation in liquidity. Specifically, I study the dynamic impact of news about “asset fundamentals.” Good news about an asset is represented by a permanent increase in its dividend rate. Investors anticipate the arrival of this news at some exponentially distributed stopping time. When the news is announced, they aggressively search for this asset, causing a temporary increase in its trading volume. As the asset is aggressively purchased, it becomes progressively harder to buy. Once this decrease in liquidity compensates for the increase in dividend rate, trading volume goes back to normal, and the economy slowly approaches its new steady state. The numerical study suggests that time variation in liquidity has a small impact on the level of prices (analogous to Constantinides
[1986]), but may have a temporarily significant impact on capital gains.

2 Trading Many Assets

This section presents the basic model, in which investors cannot buy and sell assets instantly. Rather, they allocate search resources to asset-specific “trading specialists,” who search for counterparties. When two investors meet, they bargain over the terms of trade. (The specialists could bargain on their behalves.)

2.1 The Economic Environment

This subsection describes the model setup.

Information, Preferences and Technology

Time is treated continuously, and runs forever. A probability space $(\Omega, \mathcal{F}, P)$ is fixed, as well as a filtration $\{\mathcal{F}_t, t \geq 0\}$ satisfying the usual conditions (Protter [1990]). The economy is populated by a continuum of risk-neutral and infinitely-lived investors whose total measure is normalized, without loss of generality, to 1. Investors derive utility from two goods: “dividend” and “cash.” Dividends are paid by assets. Cash is used to buy and sell assets. Specifically, investors consume any adapted cumulative dividend and cash processes $\{D(t), C(t), t \geq 0\}$, with net utility

$$E_0 \left( \int_0^{+\infty} e^{-rt} (\theta(t) \, dD(t) + dC(t)) \right),$$

where $r > 0$ is the discount rate and $\{\theta(t), t \geq 0\}$ is some adapted process of marginal utility. An investor’s marginal utility is a two-state Markov

\footnote{Duffie, Gărleanu and Pedersen [2001] provide a model with risk-averse agents where the marginal utility $\theta(t)$ approximates the hedging value of an asset at time $t$. In their economy, each agent owns some non-traded risky asset (for instance, his future labor income), which is correlated with the traded asset. In a first-order Taylor expansion of an agent’s continuation utility, $\theta(t)$ represents the value of hedging the non-traded asset with the traded asset.}
chain: the high-marginal-utility state is normalized to $\theta(t) = 1$ and the low-marginal-utility state is $\theta(t) = 1 - \alpha$, for some $\alpha \in (0,1)$. Investors switch randomly, and pair-wise independently, from a low marginal utility to a high marginal utility with intensity $\gamma_u$, and from a high marginal utility to a low marginal utility with intensity $\gamma_d$. Any two investors have pair-wise independent marginal-utility processes. In order to make side payments, investors are endowed with a technology that instantly produces cash, at unit marginal cost.

Asset Characteristics

The set of asset types is $\{1, \ldots, K\}$. Asset $k$ has a measure $s_k \in (0, 1)$ of shares outstanding, and one share of asset $k$ pays the constant dividend rate $d > 0$ forever. An investor is permitted to hold either zero or one share of some asset, and can choose which asset to hold. This assumption implies that $s_k$ is the fraction of the population holding asset $k$. In order to make the demand side of the economy nontrivial, it is assumed that $S \equiv \sum_{k=1}^{K} s_k < 1$. In other words, some investors do not own an asset and might want to buy one.

Definition 1 (Distribution of Ownership.) A distribution of ownership is some $s = (s_1, \ldots, s_K) \in \mathbb{R}_+^K$, quantities of shareholders of each asset, such that $\sum_{k=1}^{K} s_k = S$.

Investor Types

An investor’s type is made up of her marginal utility (high “h,” or low “l”), and her ownership status, for each asset type $k \in \{1, \ldots, K\}$ (owner “ok”, or

---

3For instance, if $\theta(t) = 1 - \alpha$, the distribution of the next switching time to $\theta(t) = 1$ is exponential with parameter $\gamma_u$. The successive switching times are independent.

4In other words, negative consumption of cash is allowed. Equivalently, one could assume that agents can borrow and save cash in some “bank account,” at the interest rate $\bar{r} = r$.

5Because he has linear utility over dividend, an investor finds it optimal to hold either the minimum quantity of zero share, or the maximum quantity of one share. Normalizing the maximum holding to be one share is without loss of generality, in the following sense: the results would remain unchanged if one assumes a maximum holding of $N$ shares, and redefine the dividend rate to be $d/N$. 

---
nonowner “n”). Hence, the set of investor types is

\[ I = \{ \text{hn, ln, ho1, \ldots, hoK, lo1, \ldots, loK} \}. \]  

For each \( i \in I \), \( \mu_i \) denotes the fraction of investors of type \( i \), and, given the asset fundamentals and the trading environment (to be defined), \( V_i \) denotes the continuation utility of an investor of type \( i \). A precise definition of \( V_i \) is provided in Appendix B.

**Random Matching**

At any point in time, each investor is endowed with a unit mass of “trading specialists” who search for specific trading counterparties, in a sense that is now to be described. A trading specialist of type \((i, j) \in I^2\) works for an investor of type \( i \), and specializes in contacting specialists working for investors of type \( j \). Thus, contacts that could result in a trade occur only between specialists of types \((i, j)\) and \((j, i)\).

An investor of type \( i \) maintains on her “trading staff” a quantity \( \nu_{ij} \) of specialists of type \((i, j)\), subject to the resource constraint \( \sum_{j \in I} \nu_{ij} \leq \tilde{\nu} \). Thus, the mass of specialists of type \((i, j)\) in the entire specialist population is \( \mu_i \nu_{ij} \). A given specialist makes contacts with other specialists, pair-wise independently at Poisson arrival times with intensity \( \Lambda > 0 \). Because scaling \( \tilde{\nu} \) and \( \Lambda \) up and down, respectively, by the same factor has no effect, one can assume without loss of generality that \( \tilde{\nu} = 1 \). Contacts are also pair-wise independent with the liquidity-need processes. Given a contact, because of the random-matching assumption, the probability that the contact is made with a specialist of type \((i, j)\) is \( \mu_i \nu_{ij} \). That is, conditional on making a contact, all trading specialists in the entire specialist population are “equally likely” to be contacted. Adapting the usual random-matching assumption that the Law of Large Numbers applies (see, for instance, Diamond [1982]), contacts between specialists of types \((i, j)\) and \((j, i)\), for \( i \neq j \), occur continually at a total (almost sure) rate of

\[ \mu_i \nu_{ij} \Lambda \mu_j \nu_{ji} + \mu_j \nu_{ji} \Lambda \mu_i \nu_{ij} = 2 \Lambda \mu_i \nu_{ij} \mu_j \nu_{ji}. \]  

The first term on the left-hand side of (4) is the total rate of contacts made by all specialists of type \((i, j)\), and received by specialists of type \((j, i)\). Specifi-
ically, each specialist of the mass $\mu_i \nu_{ij}$ of specialists of type $(i, j)$ makes contacts at rate $\Lambda$, and such contacts are received by some specialist of type $(j, i)$ with probability $\mu_j \nu_{ji}$. Similarly, the second term is the total rate of contact made by specialists of type $(j, i)$ and received by specialists of type $(i, j)$.

For each investor of type $i$, $\lambda_{ij} \equiv \Lambda \nu_{ij}$ is the intensity of contacts with some other specialists, made by the mass $\nu_{ij}$ of specialists of type $(i, j)$. Thus, one can view an investor of type $i$ as endowed with a budget $\Lambda > 0$ of search effort, allocating some intensity $\lambda_{ij}$ to the search for investors of type $j$, subject to the resource constraint $\sum_{j \in I} \lambda_{ij} \leq \Lambda$. With this new notation, adopted for the remainder of the paper, the total (almost sure) rate of contact between investors of types $i$ and $j$ is

$$2\mu_i \mu_j \frac{\lambda_{ij} \lambda_{ji}}{\Lambda}.$$  \hspace{1cm} (5)

An investor maintaining trading specialists can be viewed as an investment firm with separate units that trade specific securities. A typical unit trades securities of a specific industry, such as “telecom” or “entertainment,” or trades securities with a specific payoff structure, such as fixed-income or derivatives. Specialization in trading reflects the costs of collecting and processing information regarding the supply and demand of assets, as well as the fundamentals of the underlying cash flows.

This search-theoretic model abstracts from a decentralized security market, such as the NASDAQ and some other over-the-counter markets. One may argue that, in these markets, search frictions are largely overcome by marketmakers. However, as Duffie, Gârleanu and Pedersen [2001] show, search frictions remain relevant provided investors meet marketmakers sequentially. In one version of their model, two markets coexist: a decentralized search market in which investors trade among each other, and a centralized market in which investors trade with marketmakers. The search frictions determine the reservation values of investors when trading with marketmakers, which in turn affect bid and ask prices.
2.2 Equilibrium

This subsection presents an analysis of the decisions of investors: whether or not to trade in a given encounter, and how to allocate search intensity across types of trading encounters. Then, it describes the dynamics of the distribution of types. Lastly, it defines an equilibrium.

Trade Among Investors

Trade between investors of types $i$ and $j$ occurs at a strictly positive rate if (a) the gain from trade from such a pair is strictly positive, \footnote{An arbitrarily small transaction cost rules out trade when the gain is zero.} and (b) these two types of investors maintain trading specialists who are searching for each other, that is, if $\lambda_{ij}\lambda_{ji} > 0$.

It is anticipated that, in equilibrium, the gains from trade are strictly positive between the following types of investor pairings. First, when a low-marginal-utility owner (one of type $lok$) contacts a high-marginal-utility non-owner (of type $hn$) the $lok$ investor may sell her asset to the $hn$ investor, in exchange for cash. Second, when an $lok$ investor contacts an $hoj$ investor, they may swap assets, and one investor may simultaneously transfer cash to the other. Cash is a lump of consumption good, instantly produced at unit marginal cost. Any ($lok$, $hoj$) pair does not necessarily have a profitable swap, and $lok$ investors do not necessarily search for swaps.

The terms of trade between an $hn$ and an $lok$ investor arise in a simple Nash bargaining game. The total surplus of such a transaction is

$$ (V_{hok} - V_{hn}) - (V_{lok} - V_{ln}) \equiv \Delta V_{hk} - \Delta V_{lk}. \quad (6) $$

This paper studies those equilibria in which the $lok$ agent receives a fixed fraction $q \in (0, 1)$ of the total surplus.\footnote{Allowing $q$ to depend on the pair type would not change the main qualitative results. This point is illustrated by the computations of Section 4.2.} This implies that the price of asset $k$ is, in an equilibrium,

$$ p_k = q\Delta V_{hk} + (1 - q)\Delta V_{lk}. \quad (7) $$

\footnote{An arbitrarily small transaction cost rules out trade when the gain is zero.}

\footnote{Allowing $q$ to depend on the pair type would not change the main qualitative results. This point is illustrated by the computations of Section 4.2.}
Similarly, the total surplus of a swap between a \( \text{lok} \) agent and a \( \text{hoj} \) agent is 
\[ V_{\text{loj}} - V_{\text{lok}} + V_{\text{hok}} - V_{\text{hoj}}. \]
When this surplus is positive, then trade might take place. As before, it is assumed that the \( \text{lok} \) agent receives a fixed fraction \( q \in (0, 1) \) of the total surplus. Simple manipulation shows that the \( \text{lok} \) agent transfers to the \( \text{hoj} \) agent a quantity \( p_j - p_k \) of cash.

It is anticipated that, in equilibrium, \( \text{lok} \) investors do not maintain trading specialists who search for \( \text{hoj} \) investors, but only trading specialists who search for \( \text{hn} \) investors. In other words, the net utility of searching for an asset swap is, in the equilibrium I analyze, strictly less than the net utility of searching for an outright sale, under natural conditions. In other words,

\[
\mu_{\text{hn}}(\Delta V_{\text{hk}} - \Delta V_{\text{lk}}) > \mu_{\text{loj}}(V_{\text{loj}} - V_{\text{lok}} + V_{\text{hok}} - V_{\text{hoj}}),
\]

for all \((k, j) \in \{1, \ldots, K\}^2\). Hence, an \( \text{lok} \) investor allocates all of her search intensity \( \Lambda \) to the search for \( \text{hn} \) investors. On the other hand, an \( \text{hn} \) investor allocates intensities, denoted \( \lambda_1, \ldots, \lambda_K \), to simultaneous searches for investors of respective types \( \text{lo}1, \ldots, \text{lo}K \). These allocations of search intensity are illustrated in Figure 1.

**Definition 2** A search intensity allocation is some \( \lambda \in \mathbb{R}_+^K \) with \( \sum_{k=1}^K \lambda_k = \Lambda \).

**Investors’ Problems and the Distribution of Types**

This paragraph characterizes the equilibrium continuation utilities \( V_i, i \in \{1, \ldots, I\} \). As shown in Appendix [B] by an optimality verification argument, these solve the system of Bellman Equations:

\[
\begin{align*}
    rV_{\text{hn}} &= \max_{\tilde{\lambda}_1, \ldots, \tilde{\lambda}_K} \left\{ \gamma_d(V_{\text{ln}} - V_{\text{hn}}) + 2 \sum_{k=1}^K \tilde{\lambda}_k \mu_{\text{loj}}(V_{\text{hok}} - V_{\text{hn}} - p_k) \right\} \\
    rV_{\text{hok}} &= d + \gamma_d(V_{\text{lok}} - V_{\text{hok}}) \\
    rV_{\text{lok}} &= (1 - \alpha)d + \gamma_u(V_{\text{hok}} - V_{\text{lok}}) + 2\lambda_k \mu_{\text{hn}}(V_{\text{ln}} - V_{\text{lok}} + p_k) \\
    rV_{\text{ln}} &= \gamma_u(V_{\text{hn}} - V_{\text{ln}}),
\end{align*}
\]

for all \( k \in \{1, \ldots, K\} \). The maximization in (9) is subject to \( \sum_{k=1}^K \tilde{\lambda}_k \leq \Lambda \) and \( \tilde{\lambda}_k \geq 0 \), for all \( k \in \{1, \ldots, K\} \). Tilde notation (\( \tilde{\ } \)) is used to distinguish
An $hn$ investor searches $lok$ investors with intensity $\lambda_k$, for all $k \in \{1, \ldots, K\}$. An $lok$ investor, on the other-hand, searches $hn$ investors with intensity $\Lambda$.

Given a search-intensity allocation $\lambda$, the distribution $\mu \equiv (\mu_{hn}, \mu_{ln}, \mu_{lok}, \mu_{hok})$ of types solves

\begin{align*}
0 &= \gamma_u \mu_{ln} - \gamma_d \mu_{hn} - 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{lok} \\
0 &= \gamma_u \mu_{lok} - \gamma_d \mu_{hok} + 2 \lambda_k \mu_{hn} \mu_{lok} \\
0 &= \gamma_d \mu_{hok} - \gamma_u \mu_{lok} - 2 \lambda_k \mu_{hn} \mu_{lok} \\
0 &= \gamma_d \mu_{hn} - \gamma_u \mu_{ln} + 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{lok} \\
s_k &= \mu_{lok} + \mu_{hok} \\
1 &= \sum_{k=1}^{K} (\mu_{lok} + \mu_{hok}) + \mu_{hn} + \mu_{ln},
\end{align*}

Figure 1: Allocating Search Intensity
for \( k \in \{1, \ldots, K\} \). The system (13)-(16) implies that, in a steady state, for each type, the rate of change of the quantity of investors of that type is zero. For instance, in (13), \( \gamma_{u_{hn}} \) is the instantaneous flow of investors of type \( ln \) migrating to the \( hn \) type, \( \gamma_{d_{hn}} \) is the instantaneous flow of investors of type \( hn \) migrating to the \( ln \) type, and \( 2\lambda_{k}\mu_{hn_{k}}\mu_{lok_{k}} \) is the instantaneous flow of investors of type \( hn \) who buy an asset of type \( k \), migrating to the \( hok \) type.

**Steady-State Symmetric Equilibrium**

**Definition 3** A steady-state symmetric equilibrium is a collection \( V = (V_{hn}, V_{hok}, V_{lok}, V_{ln})_{1 \leq k \leq K} \) of continuation utilities, a distribution \( \mu = (\mu_{hn}, \mu_{hok}, \mu_{lok}, \mu_{ln})_{1 \leq k \leq K} \) of types, and a search intensity allocation \( \lambda \gg 0 \), such that

(i) Steady-State: Given \( \lambda \), and \( \mu \), \( V \) and \((\lambda_{1}, \ldots, \lambda_{K}) = \lambda \) solve the system (13)-(18).

(ii) Optimality: Given \( \lambda \), and \( \mu \), \( V \) and \((\lambda_{1}, \ldots, \lambda_{K}) = \lambda \) solve the system (9)-(12) of Bellman equations. The no-swap conditions (8) holds for all \((k, j) \in \{1, \ldots, K\}^{2}\).

Here, symmetry means that all \( hn \) investors choose the same search intensity allocation. Definition 3 restricts attention to those equilibria having two specific properties: there are no swap and all assets are searched, that is \( \lambda \gg 0 \). In particular, since (9) is a linear program, \( \lambda \gg 0 \) implies that \( hn \) investors are indifferent between searching for any two assets. Hence, the first-order condition of the \( hn \) investor’s problem, (9), is

\[
\mu_{lok}(1 - q)(\Delta V_{hk} - \Delta V_{lk}) = \mu_{lok}(1 - q)(\Delta V_{hj} - \Delta V_{lj}),
\]

for all \((k, j) \in \{1, \ldots, K\}^{2}\). This reflects “search indifference,” meaning that the marginal utility of spending an additional unit of search intensity on a given asset is equaled across assets. This marginal utility is decomposed as follows: Conditional on establishing a contact, a seller of asset \( k \) is found with

\[8\] There are other types of equilibria. For instance, because of the random-matching specification, investors of type \( i \) and \( j \) meet only if \( i \) searches for \( j \) and \( j \) searches for \( i \). Hence, there is an equilibrium in which no investor search: namely, given that no other investor searches, not searching is optimal.
probability $\mu_{lok}$. Then, the buyer receives a fraction $1 - q$ of the transaction surplus $\Delta V_{hk} - \Delta V_{lk}$.

The total transaction surplus may be interpreted as a bid-ask spread, in the following sense. Suppose one introduces an “infinitesimal” marketmaker in the steady-state equilibrium. If this marketmaker can make take-it-or-leave-it offers to investors, he would charge $\Delta V_{hk}$ to buyers of asset $k$ (the ask price), and pay $\Delta V_{lk}$ to sellers of asset $k$ (the bid price). In other words, the buyer’s reservation value is the ask price, and the seller’s reservation value is the bid price. Following this interpretation, condition (19) implies that an asset that is easier to find (with a larger $\mu_{lok}$) has a narrower bid-ask spread. This suggests a negative relationship between liquidity and bid-ask spread.

2.3 Existence and Uniqueness

This section provides technical conditions under which an equilibrium exists and is unique. It first analyzes the steady-state distribution of types. Second, in order to prove the existence of an equilibrium, it studies the indifference conditions (19).

Steady-State Distribution of Types

This paragraph studies the system (13)-(18), given a search intensity allocation $\lambda$. Since (17) implies that the sum of (14) and (16) is zero, equation (14) can be eliminated. Similarly, since (18) implies that the sum of (13) through (15) is zero, equation (16) can be eliminated. Hence, (13)-(18) is equivalent to the reduced system

$$0 = \gamma_d s_k - (\gamma_d + \gamma_u)\mu_{lok} - 2\lambda_k \mu_{hn} \mu_{lok}$$  \hspace{2cm} (20)

$$0 = \gamma_u (1 - S) - (\gamma_d + \gamma_u)\mu_{hn} - 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{lok}$$  \hspace{2cm} (21)

$$\mu_{lok} = s_k - \mu_{lok}$$  \hspace{2cm} (22)

$$\mu_{hn} = 1 - S - \mu_{hn}$$  \hspace{2cm} (23)

\footnote{Considering a monopolistic marketmaker that is not “infinitesimal” would change the equilibrium outcome, but would not affect the main results.}
for \( k \in \{1, \ldots, K\} \). The study of (20)-(23) presented in Appendix A shows the following proposition.

**Proposition 1** Given a search intensity allocation \( \lambda \), the system (20)-(23) has a unique solution \( \mu = (\mu_{hn}, \mu_{hok}, \mu_{ln}, \mu_{lok})_{1 \leq k \leq K} \in [0, 1]^{2K+2} \).

**Search-indifference**

The equilibrium equations can be simplified as follows. First, one defines the net utility of searching for asset \( k \),

\[
W_k \equiv \mu_{lok}(1-q)(\Delta V_{hk} - \Delta V_{lk}),
\]

for all \( k \in \{1, \ldots, K\} \). Clearly, the “search indifference” marginal conditions (19) can be written as

\[
W_k = W,
\]

for all \( k \in \{1, \ldots, K\} \), and for some positive constant \( W \) to be determined. Substituting (25) into equation (9), and combining the Bellman equations (9) through (11), one finds that

\[
rW_k = \alpha (1-q) \mu_{lok} d - (\gamma_u + \gamma_d + 2\lambda_k q \mu_{hn}) W_k - 2\Lambda (1-q) \mu_{lok} W.
\]

On the other hand, equation (20) implies that

\[
2\lambda_k \mu_{hn} = \frac{\gamma_d s_k}{\mu_{lok}} - (\gamma_d + \gamma_u).
\]

Substituting (27) into (26), using (25) and rearranging gives

\[
\frac{\gamma_d s_k q}{(1-q)\alpha d \mu_{lok}^2} + \frac{r + (1-q)(\gamma_d + \gamma_u)}{(1-q)\alpha d} \frac{1}{\mu_{lok}} + \frac{2\Lambda}{\alpha d} = \frac{1}{W}.
\]
This quadratic equation allows one to write \( \mu_{lok} = m_k(W) \), for some \( W > 2\Lambda/\alpha d \) and for some continuous and increasing function \( m_k(\cdot) \). Combining (20) and (21), one finds that

\[ \mu_{hn} = y - S + \sum_{k=1}^{K} \mu_{lok}. \]  

(29)

Substituting (29) into (27) gives

\[ 2\lambda_k \left( y - s + \sum_{k=1}^{K} m_k(W) \right) = \left( \frac{\gamma_d s_k}{m_k(W)} - (\gamma_d + \gamma_u) \right), \]  

(30)

which shows that \( \sum_{k=1}^{K} \lambda_k = \Lambda \) only if

\[ 2\Lambda \left( y - s + \sum_{k=1}^{K} m_k(W) \right) - \sum_{k=1}^{K} \left( \frac{\gamma_d s_k}{m_k(W)} - (\gamma_d + \gamma_u) \right) = 0. \]  

(31)

The left-hand side of (31) is strictly increasing in \( W \) because \( m_k(\cdot) \) is strictly increasing for each \( k \). Hence, (31) uniquely characterizes a candidate equilibrium \( W \). Once \( W \) is found, the other equilibrium objects are uniquely characterized: the search intensity allocation \( \lambda \) by (30), the distribution \( m \) of types by (20)-(23), and the continuation utilities \( V \) by (9)-(11). This implies

**Proposition 2 (Uniqueness.)** There is at most one equilibrium.

In order to show existence, one first analyzes the case of identical asset characteristics, for the distribution \( \hat{s} = (S/K, \ldots, S/K) \) of ownership. One shows the existence of a symmetric equilibrium with \( \hat{\lambda}_k = \Lambda/K \), following Duffie, Gârleanu and Pedersen [2001]. Then, one applies the Implicit Function Theorem to equation (31), showing existence in a neighborhood of this symmetric equilibrium.

**Proposition 3 (Existence.)** Let \( \hat{s} = (S/K, \ldots, S/K) \). Then, there is a neighborhood \( N \subset \mathbb{R}_+^K \) of \( \hat{s} \), such that, for all \( s \in N \), there is an equilibrium.
Proof. If the assets have identical characteristics, it is natural to guess that there is a symmetric equilibrium, with \( \hat{\mu}_{lok} = \hat{\mu}_{lo}/K \) and \( \hat{\lambda}_k = \Lambda/K \). The equilibrium equations are those of Due, Gărleanu and Pedersen [2001], with “\( \lambda \)” there being replaced here by “\( \Lambda/K \).” Their results imply that investors’ values are strictly positive, and that there are strictly positive gains from trade between investors of types \( hn \) and \( lok \). Furthermore, since assets have identical characteristics, there is no gain from swapping assets. Thus, \( lok \) investors strictly prefer searching for an outright sale with an \( hn \) investor to searching for a swap with an \( loj \) investor, for all \( j \in \{1, \ldots, K\} \). Since the left-hand side of (31) is strictly increasing in \( W \), the Implicit Function Theorem (see Taylor and Mann [1983], chapter 12) can be applied: This provides a neighborhood \( N \subset \mathbb{R}_+^K \) of \( s \), such that, for all \( s \in N \), there exists a candidate equilibrium \( W = h(s) \), for some continuous function \( h(\cdot) \). The other candidate equilibrium objects \((V, \mu, \lambda)\) are easily expressed as continuous functions of \( W \) and thus as continuous functions of \( s \). The search-indifference conditions (25) are satisfied by construction. All other relevant inequalities hold by continuity. □

The proof shows in particular that, if assets characteristics are sufficiently homogeneous, \( lok \) investors are not searching for swaps. This follows from the fact that the net utility of swapping two assets with nearly identical characteristics is close to zero, and turns out to be strictly less than the net utility of searching for an outright sale.

Does there always exist an equilibrium in which all assets are traded? The following proposition provides a partial answer, in a two-asset economy. Specifically, I show that if the assets have sufficiently different supplies, there cannot be an equilibrium in which both are traded, in the following sense.

**Proposition 4 (Non-Existence.)** Let \( K = 2 \) and \( s = (s_1, S - s_1) \), for some \( S \in (0,1) \) and \( s_1 \in (0, S) \). There is a \( \varepsilon > 0 \) such that, for any \( s_1 < \varepsilon \), an equilibrium cannot exist.

Proof. Let’s consider a two-asset economy with \( s_1 = \varepsilon > 0 \) and \( s_2 = S - \varepsilon > 0 \). One shows that, if \( \varepsilon \) is small enough, there is no candidate equilibrium in which \( \lambda \gg 0 \). If such an equilibrium exists, then both assets would satisfy (28). For Asset 1, since \( \mu_{lo1} \leq s_1 = \varepsilon \), (28) would imply that the candidate \( W \) goes to zero as \( \varepsilon \) goes to zero. In turn, for Asset 2, (28) would imply that \( \mu_{lo2} \) goes to zero as \( \varepsilon \) goes to zero, but then \( s_2/\mu_{lo2} = (1 - \varepsilon)/\mu_{lo2} \) goes to infinity as \( \varepsilon \) goes to zero. Therefore, equation (31) cannot hold. □
Existance in Proposition 3 and non-existence in Proposition 4 are proved by studying how the left-hand side of (31) depends on \( s \). When asset characteristics are sufficiently similar, the equation has a solution. Alternatively, when the quantity of shareholders of an asset is sufficiently small relative to quantities of shareholders of other assets, (31) has no solution.

An equilibrium may fail to exist because, when \( s_1 \) is small, the probability of finding a seller is even smaller. An investor is willing to search for this asset only if she is compensated by a sufficiently low price. If \( s_1 \) is small enough, the appropriate compensation results in a negative price, and thus cannot be the basis of an equilibrium.

### 3 The Pricing of Liquidity Differences

This section analyzes the pricing implications of the model. It first discusses the pricing of liquidity. Then, it derives the float-adjusted return model or FARM, which states that, in equilibrium, the liquidity spread of an asset is proportional to the inverse of its free float. Lastly, it relates the cross-sectional variation in asset returns to \( s = (s_1, \ldots, s_K) \), the exogenous distribution of ownership.

#### 3.1 Cross-sectional Prices

The pricing equation (7) can be written

\[
p_k = \Delta V_{hk} - (1 - q)(\Delta V_{hk} - \Delta V_{lk}). \tag{32}
\]

The first term on the right-hand side, \( \Delta V_{hk} \), is the reservation value of a \( hn \) investor. The second term is the discount obtained by a \( hn \) investor with bargaining power \( 1 - q \). Subtracting the Bellman equations (9) from (10), gives an expression for the reservation value \( \Delta V_{hk} \) which, when substituted in (32), gives

\[
 rp_k = d - 2\Lambda W - \gamma_d (\Delta V_{hk} - \Delta V_{lk}) - r (1 - q) (\Delta V_{hk} - \Delta V_{lk}). \tag{33}
\]

This equation breaks the price of asset \( k \) into four components. The first, \( d \), is the flow value of dividend payments. The second component, \( 2\Lambda W \), is
the flow value of searching for an asset. An \(hn\) investor obtains this discount because he has the option of not buying asset \(k\) and continuing his search. The third component, \(\gamma_d(\Delta V_{hk} - \Delta V_{lk})\), is the instantaneous cost of switching to the low-marginal-utility state, and not being able to sell the asset instantly. The last component is the bargaining discount.

It is instructive to compare the price \(p_k\) of the asset in this dynamic bargaining market with its price \(p_k^\infty\) in a Walrasian market, where all assets can be bought and sold instantly. When \(S < \gamma_u/(\gamma_d + \gamma_u)\), the marginal investor in a Walrasian market has a high marginal utility, implying that \(p_k^\infty = d/r\). In other words, all discounts in (33) are equal to zero. First, because the net utility of buying an asset is equal to zero, the flow value of searching for an alternative asset is also equal to zero. Second, because the asset can be sold instantly, the cost of switching to the low-marginal-utility state state is equal to zero. Third, because sellers can find alternative buyers instantly, the buyer’s bargaining discount is equal to zero.

### 3.2 A Float Adjusted Return Model

The price \(p_L\) of a hypothetical ‘perfectly liquid’ asset, named asset \(L\), is defined by

\[
r p_L \equiv d - 2\Lambda W. \tag{34}
\]

This price makes an \(hn\) investor indifferent between \((i)\) searching for some asset \(k \in \{1, \ldots, K\}\) and \((ii)\) buying asset \(L\) instantly, stopping search, and having the option to sell the asset instantly at price \(p_L\).

Asset \(L\) is called ‘perfectly liquid’ because it can be bought and sold instantly. Its price includes the discount \(2\Lambda W\) because \(hn\) investors must be compensated from stopping search when they buy it. Since asset \(L\) can be sold instantly at price \(p_L\), however, the cost of switching to the low-marginal-utility state is equal to zero.\footnote{In the present model, the supply of asset \(L\) is zero and \(34\) should be viewed as a definition. One can solve for an equilibrium in which asset \(L\) is in positive supply, and show that its equilibrium price is \(p_L\), under natural technical conditions.} Subtracting equation (34) from equation (32) and rearranging, one finds

\[
r(p_L - p_k) = (\gamma_d + r(1 - q))(\Delta V_{hk} - \Delta V_{lk}). \tag{35}
\]
Together with the search-indifference condition (19), (35) implies that

\[ p_k \mu_{lok} \left( \frac{d}{p_k} - \frac{d}{p_L} \right) = \frac{d}{rp_L} (\gamma_d + r(1 - q))W, \]  

(36)

where \( W \) is the net utility of searching for alternative assets. The returns of assets \( k \) and \( L \) are respectively denoted by \( R_k \equiv d/p_k \) and \( R_L \equiv d/p_L \). The free float of asset \( k \) is \( \phi_k \equiv p_k \mu_{lok} \). This is the portion of its market capitalization available for sale. With these notations, (36) can be written

\[ \phi_k (R_k - R_L) = M, \]  

(37)

for some positive constant \( M \) which does not depend on \( k \). Equation (37) states that, in equilibrium, the liquidity spread \( R_k - R_L \) of asset \( k \) is proportional to the inverse of its free float \( \phi_k \). A convenient expression for \( M \) is obtained by summing equations (37) over \( k \). The result is summarized in

**Proposition 5 (Float Adjusted Return Model, FARM.)** In equilibrium, an asset liquidity spread is proportional to the inverse of its free float. Namely, for all \( k \in \{1, \ldots, K\} \),

\[ R_k - R_L = \frac{\bar{\phi}}{\phi_k} (R_M - R_L), \]  

(38)

where \( R_k \equiv d/p_k \) and \( R_L \equiv d/p_L \) are the returns of assets \( k \) and \( L \), \( \phi_k \equiv p_k \mu_{lok} \) is the free float of asset \( k \), \( \bar{\phi} \equiv 1/K \sum_{k=1}^{K} \phi_k \) is the market average free float, and \( R_M \) is the float-weighted market return

\[ R_M \equiv \sum_{k=1}^{K} \frac{\phi_k}{\sum_{j=1}^{K} \phi_j} R_k. \]  

(39)

When studying the impact of liquidity on cross-sectional returns, many researchers estimate linear models. They first control for risk with a factor model such as Sharpe [1964]'s CAPM, or the three-factor model of Fama and French [1993]. Then, they test the statistical significance of additional independent variables that proxy for liquidity, such as bid-ask spread, trading volume, or turnover (Amihud and Mendelson [1986], Brennan and Subrahmanyam [1996], or Brennan, Chordia and Subrahmanyam [1998]).
FARM (38) suggests that, with the linear specifications that are commonly estimated in the empirical literature, liquidity would be best measured by the free float.

Some evidence suggest that, in an empirical application, free float might be a good measure of liquidity. First, some stock-index producers such as MSCI, NYSE, or Dow Jones have started to publish free-float weighted stock market indexes. They argue that, because indexes aim at being replicated by money managers, stocks included in them should be weighted not only according to their relative size, but also according to their liquidity and their “investability,” measured by their free float.

Second, researchers have documented that recently issued treasury bonds, (or “on-the-run” bonds) are more liquid than older ones (or “off-the-run”) (see, among others, Amihud and Mendelson [1991], Warga [1992], and Krishnamurthy [2003]). A common explanation of the inferior liquidity of older bonds is their lower free float. For instance, Amihud and Mendelson [1991] argue that, because older bonds have been traded for a longer time, a larger part of their supply is “locked away” in investors portfolio (such as insurance companies) who are not standing ready to sell.

Lastly, Chan, Chan and Fong [2002] study the impact of a reduction in free float on asset liquidity. In August 1998, the Hong Kong monetary authority opposed a speculative attack by aggressively buying the 33 stocks of the Hang Seng 33 Index (HSI33). The monetary authority absorbed about 7.3% of the HSI33 market capitalization and held these stocks for a long time period, resulting in a persistent reduction in the free float of these stocks. The authors show that, relative to some control group with no free-float reduction, the HSI33 stocks experienced a decrease in liquidity.

3.3 Explaining Cross-Sectional Returns

In the previous subsection, cross-sectional variation in asset returns was explained by cross-sectional variation in free float, which is an endogenous variable. The present subsection takes a step back and explains the cross-sectional variation in asset returns by an exogenous variable, the distribution $s = (s_1, \ldots, s_K)$ of ownership.

Here, the cross-sectional variation in asset returns is not explained by an exogenously specified cross-sectional variation in transaction costs, in con-
contrast with the Walrasian models of Amihud and Mendelson [1986], Constantinides [1986], Vayanos [1998], and Huang [2003]. In the present model, because of search frictions, investors cannot find buyers and sellers of specific assets instantly, and because investors are impatient, the likelihoods of finding those buyers and sellers in a short time are reflected in prices. One may view cross-sectional variation in the likelihood of finding buyers and sellers as the natural counterpart of cross-sectional variation in transaction costs. This cross-sectional variation is not, however, exogenously specified. Rather, it arises endogenously and is explained by the distribution of ownership.

The following three equations are used. The main equation is the asset pricing equation (32), written as

$$ p_k = \frac{d}{r} - 2 \Lambda W - \left( 1 + \frac{\gamma_d}{(1 - q)r} \right) W \mu_{lok}. $$

(40)

The right-hand side is increasing in $\mu_{lok}$. In other words, an asset that is easier to find (has larger $\mu_{lok}$) is sold at a higher price. The second equation (28) is of the form

$$ A_s k \frac{1}{\mu_{lok}^2} + B \frac{1}{\mu_{lok}} + C = \frac{1}{W}, $$

(41)

for some positive constants $A$, $B$, and $C$, which do not depend on $k$. The third equation is easily derived from (20), and relates $\lambda_k$ to the distribution of types and $s_k$:

$$ \frac{\mu_{lok}}{s_k} = \frac{\gamma_d}{\gamma_d + \gamma_u + 2 \lambda_k \mu_{hn}}. $$

(42)

The quantity $\lambda_k \mu_{hn}$ has several interpretations. First, it represents the demand side of the market. The larger is $\lambda_k$, the more search occurs for asset $k$, and the easier it is to sell this asset. It is natural to ask whether an asset that is easier to sell is also easier to find. That is, can one view $\lambda_k \mu_{hn}$ as an increasing function of $\mu_{lok}$? Equation (42) shows that the answer depends on the quantity $s_k$ of shareholders, and is thus indeterminate at this stage of the analysis.
Second, $\lambda_k \mu_{hn}$ is negatively related to the mean holding period of asset $k$. The holding period of a $hok$ investor is some stopping time $\tau_h$, decomposed as follows. The investor holds the asset $k$ until she switches to a state of low marginal utility at some time $t + \tau_d$, where $\tau_d$ is an exponentially distributed stopping time with parameter $\gamma_d$. Then, she either meets a buyer or switches back to a high marginal utility at some time $t + \tau_d + \min\{\tau_b, \tau_u\}$, where $\tau_b$ and $\tau_u$ are exponentially distributed stopping times with respective parameters $2\lambda_k \mu_{hn}$ and $\gamma_u$. If she switches back to a high marginal utility, then her mean holding period is some stopping time $\bar{\tau}_h$. Hence,

$$\tau_h = \tau_d + \mathbb{1}_{\{\tau_u < \tau_b\}}(\tau_u + \bar{\tau}_h) + \mathbb{1}_{\{\tau_b \leq \tau_u\}}\tau_b = \tau_d + \min\{\tau_u, \tau_b\} + \mathbb{1}_{\{\tau_u < \tau_b\}}\bar{\tau}_h. \quad (43)$$

In a steady-state equilibrium, $\bar{\tau}_h$ and $\tau_h$ are identically distributed. Furthermore, all the above stopping times are pairwise independent. Taking expectations of both sides of (43), and using the fact that $\tau_h$ and $\bar{\tau}_h$ are identically distributed, one finds that

$$E(\tau_h) = \frac{1}{\gamma_d} + \frac{1}{2\lambda_k \mu_{hn}} \left(1 + \frac{\gamma_u}{\gamma_d}\right). \quad (44)$$

Equation (44) has the form

$$F(s_k, \mu_{lok}) = \frac{1}{W}, \quad (45)$$

for some function $F(\cdot, \cdot)$ that is increasing in $s_k$ and decreasing in $\mu_{lok}$. This implies that $\mu_{lok}$ is increasing in $s_k$. In other words, an asset with more shareholders is easier to find, is sold at a higher price, and has a lower return $R_k = d/p_k$. In order to derive a relationship between the quantity $s_k$ of shareholders and the mean holding period (44), one writes equation (41) as

$$G\left(s_k, \frac{\mu_{lok}}{s_k}\right) = \frac{1}{W}, \quad (46)$$

for some function $G(\cdot, \cdot)$ that is decreasing in $s_k$ and decreasing in $\mu_{lok}/s_k$. This implies that $\mu_{lok}/s_k$ is a decreasing function of $s_k$. From (42), it follows
that $\lambda_k \mu_{hn}$ is an increasing function of $s_k$. In other words, an asset with more shareholders has a shorter mean holding period. Lastly, since the total rate of contact between buyers and sellers of asset $k$ is $2\lambda_k \mu_{hn} \mu_{lok}$, an asset with more shareholders also has a larger trading volume. The above discussion is summarized in

Proposition 6 In equilibrium, $s_k > s_j$ implies that $\mu_{lok} > \mu_{loj}$, $\lambda_k > \lambda_j$, $p_k > p_j$, $R_k < R_j$, and $\Delta V_{hk} - \Delta V_{lk} < \Delta V_{hj} - \Delta V_{lj}$.

In words, an asset with more shareholders is easier to find, easier to sell, has a higher price, a lower return, and a narrower bid-ask spread. This implies in turn that it also has a larger trading volume, a larger turnover, and a shorter mean holding period.

This model generates a positive relationship between returns and holding periods with *ex-ante* identical investors, because returns and holding periods are both negatively related to a common exogenous “liquidity” factor, the quantity of shareholders. By contrast Amihud and Mendelson [1986] take the holding period itself to be an exogenous parameter. A positive relationship between returns and holding periods also arises endogenously in general equilibrium models with transaction costs, such as those of Vayanos and Villa [1999] or Huang [2003], but for a different reason. In these models, assets can be bought and sold instantly, and an investor chooses to hold assets with larger transaction costs for a longer period. These assets, in equilibrium, have higher expected returns. In the present model, an asset cannot be bought and sold instantly, and an asset with a higher return is harder to sell, and thus has a longer mean holding period.

4 Cross-Sectional Returns and Liquidity

The objective of this section is to confront the qualitative predictions of the theoretical model with empirical evidence. After a brief review of the empirical literature that relates cross-sectional asset returns to liquidity factors, an equilibrium of the theoretical model is computed. The relationship between returns and liquidity factors is studied numerically for a “random” cross section of 200 assets.
4.1 Empirical Evidence

Amihud and Mendelson [1986] and Amihud and Mendelson [1989] propose an empirical analysis of the “liquidity-premium hypothesis.” They study monthly returns on portfolios of NYSE stocks, over the period 1961-1981. They proxy for liquidity with the relative bid-ask spread, in line with their theoretical model, in which the relative bid-ask spread is an exogenous characteristic of the asset. Controlling for risk premia using a CAPM beta (Sharpe [1964]), and for market-capitalization, they show that there is a significant positive relationship between relative bid-ask spreads and expected returns. Subsequent studies have criticized aspects of their methodology, raising two main methodological questions. The first question is how to proxy for liquidity. Petersen and Fialkowski [1994] argued that the bid-ask spread is a poor measure of trading costs. They study market orders for 144 stocks listed on the NYSE, over the three-months period November 1990 to January 1991, and showed that 50% of transactions do not occur at the quoted bid-ask spread. The second question has been the degree of control for other factors than liquidity, most notably for systematic risk. When controlling for risk premia using CAPM betas, the econometrician is testing jointly the liquidity-premium hypothesis and the CAPM theory. A significant measured liquidity effect may reflect a failure of the CAPM and not necessarily evidence of a liquidity premium.

Alternative Measures of Liquidity

Eleswarapu [1997] studies monthly excess returns on portfolios of NASDAQ stocks, over the period 1973-1990. He controls for risk premia, using a CAPM beta, and for market capitalization. Contrary to the evidence from the earlier cited NYSE study, most trades on NASDAQ occur at the quoted bid-ask spread. Furthermore, the observed variation in bid-ask spread is much larger across NASDAQ stocks than across NYSE stocks. This suggests that a test of NASDAQ data may have more power to reject the null of no liquidity premium. Eleswarapu’s results indicate that liquidity is indeed priced.

Brennan and Subrahmanyam [1996] study monthly excess returns on

\[ \text{This observation is consistent with the model of Duffie, Gárleanu and Pedersen [2001], in which transactions between investors and a monopolistic marketmaker occur at the quoted bid-ask spread, and transactions between pairs of investors occur within the spread.} \]
portfolios of NYSE stocks, over the period 1984-1991. They do not rely on the quoted bid-ask spread, but estimate fixed and proportional trading costs from intraday transactions. In their cross-sectional regressions, the estimated trading cost is positively related to return, after controlling for risk using the three-factor model of Fama and French [1993].

Other authors such as Haugen and Baker [1996], Hu [1997], and Brennan, Chordia and Subrahmanyam [1998], relate liquidity to trading activity. An asset is said to be more liquid if it is traded more frequently and in larger (dollar) quantities. This may reflect the opportunity to conduct a large trade without a large price impact.

Haugen and Baker [1996] study monthly returns on individual stocks listed in the Russell 3000 index, over the period 1979-1993. In their regressions, they use four liquidity factors: market capitalization, market price per share, a ratio of monthly dollar trading volume to market capitalization, as well as the trend of this ratio. In addition, they include factors indicating risk, price level, and growth potential, as well as sector variables and technical factors. The ratio of monthly dollar trading volume to market capitalization, which one may interpret as a measure of turnover, appears to be the most important liquidity factor and is negatively related to expected returns.

Brennan, Chordia and Subrahmanyam [1998] study monthly excess returns on individual stocks traded on NYSE and NASDAQ, over the period 1966-1995. They control for liquidity with the dollar trading volume, for risk with either the three-factor model of Fama and French, or the method of Connor and Korajczyk [1988] that is based on asymptotic principal components. The non-risk factors used by Brennan, Chordia and Subrahmanyam [1998] include the price, a measure of dividend yield, and lagged returns. They find that the dollar trading volume of an asset is negatively related to its expected excess return.

Controlling for Risk

These cited studies use various factor models in order to control for risk premia. As mentioned above, a significant measured liquidity effect may reflect a misspecification of the factor model. The following paragraphs reviews some work that attempts to avoid this criticism.

Amihud and Mendelson [1991] compare expected returns among securities
with similar risk characteristics. They focus on two government securities, treasury bills (shorter maturity), and treasury notes (longer maturity). They match bills and notes with the same time to maturity, cash-flow, and risk. Their sample covers 37 trading days, between April and November 1987, and includes bills and notes with less than 6 months to maturity. Amihud and Mendelson’s presumption is that notes are much less liquid than bills of the same time to maturity: Since notes have been traded for a longer period, part of their supply has been “locked away” in investors portfolios. Amihud and Mendelson find that notes have significantly higher yields to maturity and larger bid-ask spreads, supporting their presumption that notes are less liquid than bills. In a related work, Warga [1992] study the difference in the mean return of on-the-run and off-the-run bonds. He constructs bond portfolios with constant (Macaulay) duration, an alternative to the “constant-time-to-maturity” criterion of Amihud and Mendelson [1991]. For each duration, he forms an on-the-run portfolio, composed of bonds issued in the most recent treasury auction, as well as an off-the-run portfolio, composed of all other bonds. The estimated mean return of the on-the-run portfolios appears to be smaller than that of the off-the-run portfolios, uniformly across durations.

Kadlec and McConnell [1994] study the prices of NASDAQ securities that obtained a NYSE listing during the 1980-1989 period. Since trading costs appear to be smaller on the NYSE than on the NASDAQ (see, among others, Huang and Stoll [1996]), the authors expect after-listing prices to reflect better market liquidity. They measure liquidity with the bid-ask spread, controlling for the increase in shareholders base (Merton [1987]) and for the “good news” associated with a NYSE listing. They find a positive relationship between bid-ask spreads and returns.

4.2 Cross-Sectional Returns: An Example

This subsection presents a numerical example suggesting that the predictions of the theoretical model developed in this paper are qualitatively consistent with much of the evidence from the empirical literature.

An equilibrium of the model is computed for a randomly generated economy of $K = 200$ asset types. The ownerships $s_k$ are drawn independently

One first solves (31) for $W$. This can be done quickly since equation (28), characterizing $m_k(W)$, is quadratic. Once $W$ is found, the remaining equilibrium objects are easily
Table 1: Parameter Values used in the Numerical Example.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value or Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact Intensity</td>
<td>$\Lambda$ 12000</td>
</tr>
<tr>
<td>Intensity of Switch to High</td>
<td>$\gamma_d$ 0.1</td>
</tr>
<tr>
<td>Intensity of Switch to Low</td>
<td>$\gamma_u$ 1</td>
</tr>
<tr>
<td>Discount Rate</td>
<td>$r$ 5%</td>
</tr>
<tr>
<td>Liquidity Shock</td>
<td>$\alpha$ 0.99</td>
</tr>
<tr>
<td>Number of Assets</td>
<td>$K$ 200</td>
</tr>
<tr>
<td>Fraction of Shareholders</td>
<td>$s_k \sim \text{Uniform}([510^{-4}, 1.510^{-3}])$</td>
</tr>
<tr>
<td>Dividend Rate</td>
<td>$d$ 1</td>
</tr>
<tr>
<td>Bargaining Power</td>
<td>$q_k \sim \text{Uniform}([0.45, 0.55])$</td>
</tr>
</tbody>
</table>

from an uniform distribution on some interval $[\underline{s}, \overline{s}]$. The dividend rate $d$ is set to 1. The bargaining powers $q_1, \ldots, q_K$, of sellers of assets $1, \ldots, K$, are drawn independently from an uniform distribution on some interval $[\underline{q}, \overline{q}]$. Varying $q$ across trading pairs is a simple way to check the robustness of the results to the introduction of other forms of asset heterogeneity. The equilibrium return $R_k = d/p_k$ is plotted against various measures of liquidity used in the empirical literature, which have direct counterparts in the theoretical model. The relative bid-ask spread is $1 - \Delta V_{l_k}/\Delta V_{h_k}$. The dollar trading volume is $p_k \lambda_k \mu_{lok} \mu_{hn}$. The turnover is $\lambda_k \mu_{lok} \mu_{hn}/s_k$. The values of the exogenous parameters are as in Table 1.

The unit of time is one year. Assuming that the stock market opens 250 days a year and that there are 10 trading hours per day, $\Lambda = 12,000$ means that an investor establishes a contact every 12.5 minutes, on average. The discount rate $r$ is 5%. Given the chosen uniform distribution for $s_k$, the expected aggregate supply of assets, $E\left(\sum_{k=1}^K s_k\right)$, is 0.2. As in Duffie, Gârleanu and Pedersen [2001], an investor has a low marginal utility, on average, for 1 year out of every 11 years.

Figure 2 displays the results of the computations. Returns and relative bid-ask spreads are positively related. In contrast with the theoretical results of Amihud and Mendelson [1986], the relationship is almost linear and not concave. Consistently with the empirical evidence, returns are negatively related to turnover and trading volume. The holding period is positively computed using the various equations derived during the existence proof.
related to returns.

Figure 2: Cross-sectional variation in returns, explained by “liquidity factors.”

5 The Impact of News on Liquidity

Recent empirical works have shown that time variation in liquidity have a significant impact on asset prices (see, among others, Chordia, Roll and Subrahmanyam [2000], Pastor and Stambaugh [2003], and Acharya and Pedersen [2003]). This section investigates some implications of the search-theoretic
model for time variation in liquidity. Namely, it presents some numerical examples that shed light on the dynamic impact of news on prices and returns. News regarding “fundamentals,” such as dividend rates, cause investors to deviate from their steady-state search allocations. This has an impact on the distribution of investors’ types, and as a consequence the likelihood of finding buyers and sellers of given types. Specifically, when news regarding fundamentals is announced, investors are no longer indifferent between searching for any two assets, and the distribution of investors’ types must adjust in order to eventually restore indifference. Since investors need to contact each other in order to trade, the distribution of investors’ types cannot adjust instantly.

In this exercise, the main technical difficulty is to construct a numerical approximation of the search-intensity allocation along the equilibrium path. Because an investor’s search-intensity allocation solves a linear program, it typically features jumps from corner to corner. In order to ensure smoothness of this policy function and to apply standard approximation methods, it is convenient to add a small, strictly concave penalty function to the search optimization problem of an $hn$ investor. This provides an approximation of the original optimization problem. One might conjecture that the equilibrium constructed on the basis of this approximation approximates an equilibrium for the actual underlying model. In any case, there are in practice costs to changing the allocation of search efforts. In Appendix C, the numerical method is described in detail.

The numerical exercise is cast in the two-asset economy ($K = 2$) described in Table 2. At $t < 0$, the economy is in a steady-state equilibrium, where agents anticipate a one-time Poisson arrival of a piece of news, with intensity $\eta > 0$. At $t = 0$, a piece of news regarding the assets’ characteristics is announced. Figures 13 through 4 display the “transitional dynamics” of equilibrium quantities, for $t > 0$.

---

13The unit of time for the rates and intensities in Table 2 is one year. In the numerical exercise the economy is close to its new steady state after a few hours of trading. As a result, the time unit in the Figures is in “hours.” Years are converted in hours assuming 250 trading days per year and 10 hours of trading per day.

14Investors anticipate the arrival of the piece of news: hence, the steady state at $t < 0$ (the “initial conditions”) depends on the path of the economy at $t > 0$ (the “transitional dynamics.”) As a result, the following fixed point problem has to be solved. Given some candidate initial conditions, one solves for the transitional dynamics using the numerical
Table 2: Parameter Values used in Studying the Dynamic Impact of News.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact Intensity</td>
<td>Λ</td>
</tr>
<tr>
<td>Intensity of Switch to High</td>
<td>γ_1</td>
</tr>
<tr>
<td>Intensity of Switch to Low</td>
<td>γ_0</td>
</tr>
<tr>
<td>Discount Rate</td>
<td>r</td>
</tr>
<tr>
<td>Liquidity Shock</td>
<td>α</td>
</tr>
<tr>
<td>Number of Assets</td>
<td>K</td>
</tr>
<tr>
<td>Measure of Shareholders</td>
<td>s_k</td>
</tr>
<tr>
<td>Dividend Rate</td>
<td>d_k</td>
</tr>
<tr>
<td>Bargaining Power</td>
<td>q_k</td>
</tr>
<tr>
<td>News Intensity</td>
<td>η</td>
</tr>
</tbody>
</table>

**Permanent Increases in Dividend Rates**

This example describes the effect of an 10% permanent increase in the dividend rate $d_1$, which can be interpreted as a piece of good news regarding the long-run profitability of the firm issuing Asset 1.\footnote{Studying and computing steady-state equilibria in which assets have different dividend rate is a straightforward extension of the method described in Section 2.2.} The results, displayed in Figures 3 and 4, illustrate the endogeneity of time variation in liquidity, driven by investor time-varying search intensity allocation.

The dividend increase temporarily makes Asset 1 more attractive. As a result, investors search for Asset 1 with an intensity close to their full budget Λ. This causes $μ_{1,t1}$ to decrease rapidly and $μ_{1,t2}$ to increase rapidly. In other words, Asset 2 becomes relatively more liquid than Asset 1. This change in relative liquidity compensates for the change in dividend so as to make investors nearly indifferent between searching for both assets. Near indifference is achieved after about an hour of heavy trading, by which time investors’ search intensity allocations have moved close to their new steady-state values.

\footnote{Studying and computing steady-state equilibria in which assets have different dividend rate is a straightforward extension of the method described in Section 2.2.}
In the new steady state, $\mu_{lo1} < \mu_{lo2}$, but $\lambda_1 > \lambda_2$. Because investors seek it more aggressively, Asset 1 is harder to find and easier to sell. Liquidity deteriorates for buyers but improves for seller. The “net” effect may be measured by turnover: Asset 1 is more heavily traded than Asset 2.

At $t = 0$, the prices of both assets jump close to the levels that are their new steady-state values. Because Asset 1 is initially more liquid than in the new steady state, its price is slightly larger than its steady-state value. It subsequently decreases, as liquidity deteriorates. The price impact of time variation in liquidity appears to be small, in line with the results of Constantinides [1986]. On the other hand, the capital gains (time derivatives of prices) are not small, implying that the impact on instantaneous returns is significant in the short run.
Figure 4: The Effects of a Permanent Increase in $d_1$.

6 Conclusion

This paper uses a search-theoretic model to study the impact of heterogeneity in asset liquidity on the cross section and the time series of asset returns. The main result of the paper is the float-adjusted return model, or FARM: in equilibrium, the liquidity spread on an asset is proportional to the inverse of its free float. Hence, the free float is a measure of liquidity that is consistent with the linear specifications used in most empirical studies of the liquidity spread.

Although the search technology is the same for all assets, heterogeneous bid-ask spreads arise endogenously. Cross-sectional variation in returns is explained by cross-sectional variation in share ownership. Theoretical and numerical results shows that the model generates key qualitative facts doc-
The out-of-steady-state dynamics shed light on the short-term impact of news. The price impact is generally small, but the return impact is not negligible at a high frequency. The model suggests that even moderate unexpected news may create temporary but sharp increases in trading volume.

Further work might apply the insights of this model to an empirical study of the cross section of asset returns. It would be helpful to extend the current framework in order to incorporate both risk premia and stochastic variation in aggregate liquidity.
A Dynamics of the Type Distribution

This appendix studies the dynamics of the distribution of types. It first solves for the steady state, and then proves its local stability. For a given search intensity allocation $\lambda$, the distribution $\mu(t) = (\mu_{hn}(t), \mu_{hok}(t), \mu_{lok}(t), \mu_{ln}(t))_{1 \leq k \leq K}$ of types solve

\begin{align*}
\dot{\mu}_{hn} &= \gamma_u \mu_{ln} - \gamma_d \mu_{hn} - 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{lok} \\
\dot{\mu}_{hok} &= \gamma_u \mu_{lok} - \gamma_d \mu_{hok} + 2 \lambda_k \mu_{hn} \mu_{lok} \\
\dot{\mu}_{ln} &= \gamma_d \mu_{hn} - \gamma_u \mu_{ln} + 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{lok} \\
\dot{\mu}_{lok} &= \gamma_d \mu_{hok} - \gamma_u \mu_{lok} - 2 \lambda_k \mu_{hn} \mu_{lok} \\
s_k &= \mu_{lok} + \mu_{hok} \\
1 &= \sum_{k=1}^{K} (\mu_{lok} + \mu_{hok}) + \mu_{hn} + \mu_{ln},
\end{align*}

where $\dot{\mu} = d\mu(t)/dt$, and time arguments are suppressed. Since equation (51) implies that the sum of (48) and (50) is zero, one can eliminate (48), the ODE for $\mu_{hok}$. Similarly, since equation (52) implies that the sum of equations (47) to (50) is zero, one can eliminate (49), the ODE for $\mu_{ln}$, and obtains the equivalent system

\begin{align*}
\dot{\mu}_{lok} &= \gamma_d s_k - (\gamma_d + \gamma_u) \mu_{lok} - 2 \lambda_k \mu_{hn} \mu_{lok} \\
\dot{\mu}_{hn} &= \gamma_u (1 - S) - (\gamma_d + \gamma_u) \mu_{hn} - 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{lok} \\
\mu_{hok} &= s_k - \mu_{lok} \\
\mu_{ln} &= 1 - S - \mu_{hn},
\end{align*}

for $k \in \{1, \ldots, K\}$.

**Steady-State Distribution of Types, Proposition 7**
A steady state solves equations (53)-(56). Summing equations (53) over \( k \), adding equation (54), and imposing the steady-state condition \( \dot{\mu} = 0 \), one finds

\[
\mu_{hn} = \mu_{lo} + y - S, \tag{57}
\]

where \( \mu_{lo} = \sum_{k=1}^{K} \mu_{lok} \), and \( y = \gamma_u / (\gamma_d + \gamma_u) \). Replacing this last equation in (53) gives

\[
\mu_{lok} = \frac{\gamma_d s_k}{\gamma_d + \gamma_u + 2\lambda_k (\mu_{lo} + y - S)}. \tag{58}
\]

Summing equations (58) over \( k \), one obtains the one equation in one unknown problem

\[
\mu_{lo} - \sum_{k=1}^{K} \frac{\gamma_d s_k}{\gamma_d + \gamma_u + 2\lambda_k (\mu_{lo} + y - S)} = 0. \tag{59}
\]

The left-hand side of this equation is increasing in \( \mu_{lo} \), is negative at \( \mu_{lo} = 0 \), and is positive for \( \mu_{lo} \) large enough; thus, it has a unique solution. Once the solution \( \mu_{lo} \) is found, \( \mu_{lok} \) is uniquely determined by (58), \( \mu_{hn} \) by (57), and finally \( \mu_{lok} \) and \( \mu_{ln} \) by (55) and (56). This characterizes a unique candidate steady state. Since the steady-state fractions sum to one by construction, one only needs to show that they are positive as follows: The left-hand side of (59) is positive when evaluated at \( \mu_{lo} = S \) and \( 1 - y \); it is negative when evaluated at \( S - y \). Since the left hand side of (59) is increasing, this shows that

\[
S - y < \mu_{lo} < \min\{S, 1 - y\}. \tag{60}
\]

Next, \( s - y < \mu_{lo} \) implies that \( \mu_{hn} > 0 \) and that \( \mu_{lok} < s_k \). Finally, \( \mu_{lo} < 1 - y \) implies that \( \mu_{hn} < 1 - S \) and that \( 0 < \mu_{ln} < 1 \).

**Local Stability**

This paragraph establishes that, given \( \lambda \), the steady-state distribution of types is a locally stable point of the following ODE

\[
\dot{\mu}_{lok} = \gamma_d s_k - (\gamma_d + \gamma_u)\mu_{lok} - 2\lambda_k\mu_{hn}\mu_{lok}, \tag{61}
\]

\[
\dot{\mu}_{hn} = \gamma_u (1 - S) - (\gamma_d + \gamma_u)\mu_{hn} - 2\sum_{k=1}^{K} \lambda_k \mu_{hn}\mu_{lok}, \tag{62}
\]

35
for all $k \in \{1, \ldots, K\}$. Stacking variables as $(\mu_{lo1}, \ldots, \mu_{loK}, \mu_{hn})'$, the Jacobian of the ODE at the steady state is

$$J = - (\gamma_d + \gamma_u) I_{K+1} - D,$$

(63)

where $D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$, $D_{11} = \text{diag}(2\lambda_k \mu_{hn})$, $D_{12} = [2\lambda_1 \mu_{lo1} \ldots 2\lambda_K \mu_{loK}]'$, $D_{21} = [2\lambda_1 \mu_{hn} \ldots 2\lambda_K \mu_{hn}]$, and $D_{22} = \sum_{k=1}^K 2\lambda_k \mu_{lok}$.

**Lemma 1 (Local Stability.)** The eigenvalues of $J$ have strictly negative real parts.

**Proof.** The vector $(1, \ldots, 1)'$ is denoted by $e$. One has, by construction, $e'D_{11} = D_{21}$ and $e'D_{12} = D_{22}$. An eigenvector of $J$ associated with the eigenvalue $\nu \in \mathbb{C}$ is denoted $x$. It solves

$$(\gamma_d + \gamma_u)x_1 + D_{11}x_1 + D_{12}x_2 = -\nu x_1$$

(64)

$$(\gamma_d + \gamma_u)x_2 + D_{12}x_1 + D_{22}x_2 = -\nu x_2.$$  

(65)

Multiplying equation (64) by $e'$, and subtracting equation (65) gives

$$(\gamma_d + \gamma_u + \nu)(e' x_1 - x_2) = 0.$$  

(66)

Three cases can be distinguished.

**Case 1:** $e' x_1 \neq x_2$. From (66), it must be that $\nu = - (\gamma_d + \gamma_u) < 0$.

**Case 2:** $e' x_1 = x_2 = 0$. Then (64) simplifies to $(\gamma_d + \gamma_u + D_{11}) x_1 = -\nu x_1$. Thus, it must be that $\nu = - (\gamma_d + \gamma_u + 2\lambda_k \mu_{hn}) < 0$, for some $k \in \{1, \ldots, K\}$.

**Case 3:** $e' x_1 = x_2 \neq 0$. Without loss of generality, one can assume that $x_2 = 1$. Using (64) one solves explicitly for $x_{1k}$,

$$x_{1k} = \frac{2\lambda_k \mu_{lok}}{\nu + \gamma_d + \gamma_u + 2\lambda_k \mu_{hn}}.$$  

(67)

Since the $x_{1k}$ sum to one, it must be that $\text{Re} \left( \sum_{k=1}^K x_{1k} \right) = 1$. Thus, there is one $k \in \{1, \ldots, K\}$ such that $\text{Re}(x_{1k}) > 0$, which is equivalent to $\text{Re}(1/x_{1k}) > 0$ and, from equation (67), to $\text{Re}(\nu) < - (\gamma_d + \gamma_u + 2\lambda_k \mu_{hn}) < 0$. $\blacksquare$
B Formulating and Solving the Investor’s Problem

This Appendix defines the stochastic control problem faced by an individual investor in a candidate steady-state equilibrium, and verifies that the Bellman equations (9)-(12) are sufficient for optimality.

A Candidate Steady-State Equilibrium

A candidate steady-state equilibrium is described as follows. There is a fraction \( \mu_i \) of investors of type \( i \in I \), who apply the search intensity allocations \( (\ell(i,j))_{j \in I} \) and who consume at the constant rate \( c(i) \). The marginal utility process of a type-\( i \) investor switches with intensity \( \ell(i,m) > 0 \). One defines \( H \equiv I \cup \{m\} \).

The terms of trade are described by a “transition function” \( \sigma : I \times H \to I \), and a “payment function” \( p : I \times H \to \mathbb{R}_+ \). Specifically, when an investor of type \( i \in I \) meets an investor of type \( j \in I \), the investor of type \( i \) evolves to type \( \sigma(i,j) \) and makes a payment of \( p(i,j) \) to the investor of type \( j \). When her marginal utility switches, an investor of type \( i \) evolves to one of type \( \sigma(i,m) \) and makes the payment \( p(i,m) = 0 \).

The Investor’s Problem

Given an individual investor with initial type \( i_0 \), one fixes a measurable space \( (\Omega, \mathcal{F}) \) and a measurable counting process \( N_t = (N_t(j))_{j \in H} \) where \( N_t(j) \in \mathbb{N} \) counts the numbers of encounters with some investor of type \( j \in H \) in the time interval \([0,t]\). The process \( N_t \) is associated with a sequence of encounter times \( 0 = T_0 < T_1 < T_2 < \cdots < T_n < \cdots \) and a sequence of \( H \)-valued random variables \( j_1, j_2, \ldots, j_n, \ldots \) such that, at \( t = T_n \), the investor encounters an investor of type \( j_n \). Moreover, the process \( N_t \) generates a \( I \)-valued type process \( X_t \) as follows. First, \( X_0 = i_0 \). Second,

\[
X_t = X_{T_n},
\]

for all \( t \in [T_n, T_{n+1}) \). And third,

\[
X_{T_{n+1}} = \sigma(X_{T_n}, j_{n+1}).
\]

Lastly, the internal history of (filtration generated by) the process \( N_t \) is denoted by \( \{\mathcal{F}_t^N, t \geq 0\} \).
Definition 4 (Admissible Search Intensity Allocation.) An admissible search intensity allocation is a \((\mathcal{F}_t^N)\)-adapted process \(\lambda_t\). It is \(\mathbb{R}_+^I\)-valued, left-continuous with right limit (LCRL), and such that \(\sum_{j \in I} \lambda_t(j) \leq \Lambda\). The set of admissible search intensity allocations is denoted by \(\mathcal{L}\).

An admissible search intensity allocation \(\lambda \in \mathcal{L}\) is associated with a probability \(P_\lambda\) on \((\Omega, \mathcal{F})\) such that \(N_t\) admits the \((P_\lambda - \mathcal{F}_t^N)\) intensity

\[
\eta_t^\lambda(j) = 2\mu_j \lambda_t(j) \frac{\ell(j, X_{t^-})}{\Lambda},
\]

for \(j \in I\), and

\[
\eta_t^\lambda(m) = \ell(X_{t^-}, m).
\]

The search intensity allocation describes how an investor uses her search effort over time. In particular, it incorporates the decision of accepting or rejecting a trade. Namely, any investor of type \(i \in I\) can choose not to trade with investors of type \(j \in I\) by setting \(\lambda_t(j) = 0\). The consumption process associated with the admissible search intensity allocation \(\lambda\) is defined by the stochastic differential equation

\[
dC_t^\lambda = c(X_t) \, dt - \sum_{j \in H} p(X_{t^-}, j) \, dN_t(j),
\]

Definition 5 (Investor’s Problem.) The lifetime utility of an investor with initial type \(i_0\) applying some admissible search intensity allocation \(\lambda \in \mathcal{L}\) is

\[
v_{i_0}(\lambda) = E_{P_\lambda} \left( \int_0^{+\infty} e^{-rt} \, dC_t^\lambda \right).
\]

The investor’s problem is to attain the optimal lifetime utility

\[
V_{i_0} = \sup_{\lambda \in \mathcal{L}} v_{i_0}(\lambda).
\]

Dynamic Programming

An admissible feedback is a function \(\theta : I^2 \to \mathbb{R}_+\), such that, for all \(i \in I\), \(\sum_{j \in I} \theta(i, j) \leq \Lambda\). For each \((i, j) \in I \times H\), one lets \(\eta^\theta(i, j) = 2\mu_j \theta(i, j) \ell(j, i) / \Lambda\),
and \( \eta^\theta(i, s) = \ell(i, s) \). The set of admissible feedbacks is denoted \( T \). With this notation, the system (9)-(12) of Bellman equations is

\[
\rho_j(i) = \max_{\theta \in T} \left\{ c(i) + \sum_{j \in H} \eta^\theta(i, j)(J(\sigma(i, j)) - J(i) - p(i, j)) \right\},
\]

for all \( i \in I \). The Bellman equation (75) is solved in the text. It is shown that the maximum is achieved for some \( \theta^* \). This feedback is associated with the admissible search intensity allocation

\[
\lambda_t^*(j) = \theta^*(X_{t^-}, j).
\]

for all \( j \in J \).

Proposition 7 (Sufficiency of the Bellman Equations.) The suppremum utility \( V_{i_0} \) is bounded above by \( J(i_0) \), and this upper bound is achieved by the admissible control \( \lambda^* \).

Proof. The following adapts the proof of Theorem VII.T1 from Brémaud [1981]. One consider some admissible search intensity allocation \( \lambda \in \mathcal{L} \), with its associated intensity process \( \eta^\lambda \) as in (70)-(71). One can write

\[
J(X_t)e^{-rt} = J(X_0) + \sum_{0 < T_n \leq t} (J(X_{T_n})e^{-rT_n} - J(X_{T_n-1})e^{-rT_{n-1}})
+ J(X_{T_n})(e^{-rt} - e^{-r\tau_t}),
\]

where \( \tau_t = \sup\{T_n, n \geq 0 : T_n \leq t\} \). Equation (77) can be manipulated as follows:

\[
J(X_t)e^{-rt} = J(X_0) + \sum_{0 < T_n \leq t} J(X_{T_n-1})(e^{-rT_n} - e^{-rT_{n-1}}) + J(X_{T_n})(e^{-rt} - e^{r\tau_t})
+ \int_0^t \sum_{j \in H} e^{-rT_n}(J(X_{T_n}) - J(X_{T_{n-1}}))
+ \int_0^t \sum_{j \in H} (J(\sigma(X_{s^-}, j)) - J(X_{s^-}))e^{-rs}dN_s(j)
\]

39
\[
\begin{align*}
J(X_0) + \int_0^t \left( -r J(X_s) + \sum_{j \in H} \eta^\lambda_s(j) (J(\sigma(X_{s^-}, j)) - J(X_{s^-})) \right) e^{-rs} ds \\
+ \int_0^t \sum_{j \in H} (J(\sigma(X_{s^-}, j)) - J(X_{s^-})) e^{-rs} (dN_s(j) - \eta^\lambda_s(j) ds).
\end{align*}
\]

Adding \( \int_0^t e^{-rs} dC^\lambda_s \) to both sides gives

\[
\int_0^t e^{-rs} dC^\lambda_s + J(X_t) e^{-rt} = \tag{78}
\]

\[
J(X_0) + \int_0^t \left( -r J(X_s) + c(X_s) \right) e^{-rs} ds \\
+ \int_0^t \sum_{j \in H} \eta^\lambda_s(j) (J(\sigma(X_{s^-}, j)) - J(X_{s^-}) - p(X_{s^-}, j)) e^{-rs} ds \\
+ \sum_{j \in H} \int_0^t (J(\sigma(X_{s^-}, j)) - J(X_{s^-}) - p(X_{s^-}, j)) e^{-rs} (dN_s(j) - \eta^\lambda_s(j) ds).
\]

Since \( (J(\sigma(X_{t^-}, j) - J(X_{t^-}) - p(X_{t^-}, j)) e^{-rt} \) is a bounded \( \mathcal{F}_t \)-predictable process, it follows by theorem II, T8 in Brémaud [1981] that the last term on the right-hand side of (78) is a martingale. Taking expectations on both sides, and using the Bellman equation (75), one finds

\[
E_{P,N} \left( \int_0^t e^{-rs} dC^\lambda_s + J(X_t) e^{-rt} \right) \leq J(X_0), \tag{79}
\]

with equality for \( \lambda = \lambda^* \), the admissible search intensity allocation associated with the \( \theta^* \) that solves (75). Letting \( t \) go to infinity proves that \( v_{io}(\lambda) \leq J(X_0) \), with equality if \( \lambda = \lambda^* \). \( \blacksquare \)

### C Numerical Method

This Appendix describes the numerical method used to compute transitional dynamics, when studying the impact of pieces of news. In what follows, the assets’ fundamental characteristics are denoted by \((s, d)\), where \( d = (d_1, \ldots, d_K) \in \mathbb{R}_+^K \) is a collection of strictly positive dividend rates. A steady-state equilibrium in which all assets are searched is described by the following objects: a collection \( V^* \) of continuation utilities, a distribution \( \mu^* \) of types, a search intensity allocation

\[
\begin{align*}
&\text{C Numerical Method} \\
&\text{This Appendix describes the numerical method used to compute transitional dynamics, when studying the impact of pieces of news. In what follows, the assets’ fundamental characteristics are denoted by \((s, d)\), where } d = (d_1, \ldots, d_K) \in \mathbb{R}_+^K \text{ is a collection of strictly positive dividend rates. A steady-state equilibrium in which all assets are searched is described by the following objects: a collection } V^* \text{ of continuation utilities, a distribution } \mu^* \text{ of types, a search intensity allocation}
\end{align*}
\]
\(\lambda^*,\) and some net utility \(W^*\) of searching for assets.

**Dynamic of the Distribution of Type**

Appendix shows that the distribution \(\mu\) of types is a solution of the system of ODEs:

\[
\begin{align*}
\dot{\mu}_{lok} &= \gamma_d s_k - (\gamma_d + \gamma_u)\mu_{lok} - 2\lambda_k\mu_{hn}\mu_{lok} \\
\mu_{hn} &= \gamma_u(1 - S) - (\gamma_d + \gamma_u)\mu_{hn} - 2\sum_{k=1}^{K}\lambda_k\mu_{hn}\mu_{lok} \\
\mu_{hok} &= s_k - \mu_{lok} \\
\mu_{ln} &= 1 - S - \mu_{hn},
\end{align*}
\]

for some initial conditions \((\mu_{lok}(0), \mu_{hn}(0), \mu_{hok}(0), \mu_{ln}(0))\) \(1 \leq k \leq 0\) in \([0, 1]^{2K+2}\) such that, for all \(k\), \(\mu_{lok} + \mu_{hok} = s_k\) and \(\mu_{hn} + \mu_{ln} + \sum_{k=1}^{K}\mu_{lok} + \mu_{hok} = 1\). In what follows, \(\mu_t = (\mu_{lok}(t), \mu_{hn}(t))\) \(1 \leq k \leq K\) denotes the (reduced) distribution of types. The other fractions \(\mu_{hok}\) and \(\mu_{ln}\) are given by, respectively, equations \((82)\) and \((83)\).

**Penalized Bellman Equations**

Because the search intensity allocation solves a linear program, it typically features jumps from corner to corner. A standard linear approximation relying on the differentiability of investors’ policy function (see Judd [1999]) cannot be used. To ensure smoothness of the policy function, a barrier function is added to the objective of \(hn\) investors. Specifically, the system of Bellman equations is written

\[
\begin{align*}
\dot{V}_{hn} &= \gamma_d(V_{ln} - V_{hn}) + 2\Delta W + \dot{V}_{hn} \\
\dot{V}_{hok} &= d_k + \gamma_d(V_{lok} - V_{hok}) + \dot{V}_{hok} \\
\dot{V}_{ln} &= \gamma_u(V_{hn} - V_{ln}) + \dot{V}_{ln} \\
\dot{V}_{lok} &= (1 - \alpha)d_k + \gamma_u(V_{hok} - V_{lok}) \\
&\quad + 2\lambda_k\mu_{hn}g(\Delta V_{hk} - \Delta V_{lk}) + \dot{V}_{lok},
\end{align*}
\]

where the value \(V_i\) of each type \(i\) and the value \(W\) of searching for assets are implicitly a function of time \((t)\), and \(\dot{V}_i\) denote the derivative of \(V_i(t)\) with respect
to time. The value $W$ of searching for assets is

$$W = \max_{\lambda_1, \ldots, \lambda_K} \frac{1}{\Lambda} \left( \sum_{k=1}^{K} \tilde{\lambda}_k W_k + \varepsilon W^* \lambda_k^* \log \left( \frac{\tilde{\lambda}_k}{\lambda_k^*} \right) \right),$$

(88)

where $\varepsilon$ is a strictly positive constant and $W_k \equiv \mu_{lok}(1 - q)(\Delta V_{hk} - \Delta V_{lk}) > 0$ denotes the net utility of searching for asset $k$. The maximization in (88) is subject to $\sum_{k=1}^{K} \lambda_k \leq \Lambda$ and $\lambda_k \geq 0$ for all $k \in \{1, \ldots, K\}$. Penalizing the linear program in (88) amounts to assuming that it is costly to deviate from the steady-state search intensity allocation. The chosen specification (what is called a “barrier function”) guarantees that a solution is unique and interior.

With the penalized Bellman equations, a steady state is defined, as before, as a collection $(V, \mu, \lambda) \in \mathbb{R}^{5K+4}$ solving equations (80)-(83) and (84)-(88), in which the time derivatives $\dot{V}_i$ and $\dot{\mu}_i$ are set to zero. Clearly, the equilibrium $(V^*, \mu^*, \lambda^*)$ of the economy without penalization ($\varepsilon = 0$) is a steady state of the economy with penalization ($\varepsilon > 0$).

**Approximating the Search Intensity Allocation**

Simple computations show that the unique solution of (88) is

$$\lambda_k (\nu (\varepsilon), \varepsilon) = \frac{\varepsilon W^* \lambda_k^*}{\nu (\varepsilon) - W_k}$$

(89)

$$\sum_{k=1}^{K} \lambda_k (\nu (\varepsilon), \varepsilon) = \Lambda$$

(90)

$$\nu (\varepsilon) > \max_{1 \leq k \leq K} \{W_k\},$$

(91)

where $\nu (\varepsilon)$ is the Lagrange multiplier of the constraint $\sum_{k=1}^{K} \lambda_k \leq \Lambda$, and the dependence of the solution $\lambda_k$ on $\nu (\varepsilon)$ and $\varepsilon$ is made explicit in the notation.

**Lemma 2** Let $M = \arg \max_{1 \leq j \leq K} \{W_j\}$. If $k \notin M$, then, as $\varepsilon \to 0$,

$$\lambda_k (\nu (\varepsilon), \varepsilon) \to 0.$$ 

(92)

If, on the other hand, $k \in M$, then, as $\varepsilon \to 0$,

$$\lambda_k (\nu (\varepsilon), \varepsilon) \to A \frac{\lambda_k^*}{\sum_{j \in M} \lambda_j^*}.$$ 

(93)
Proof. If $k \notin M$, then, from (91), $\nu(\varepsilon) - W_j$ is bounded away from zero. This clearly implies (92). If, on the other hand, $k \in M$, then, letting $W_M = \max_{1 \leq j \leq M} \{W_j\}$, one writes (90) as

$$\sum_{k \in M} \varepsilon W^* \Lambda^*_k + h(\varepsilon) = \Lambda,$$

where $h(\varepsilon)$ goes to zero as $\varepsilon$ goes to zero. This implies (93). \(\blacksquare\)

In words, when $\varepsilon > 0$ is small, the solution of (89)--(91) approximates a solution of the the (linear) program (88) with $\varepsilon = 0$. This property does not guarantee, however, that the equilibrium constructed on the basis of this approximation approximates an equilibrium in which an investor solves the (linear) program (88) with $\varepsilon = 0$. I conjecture as much, however, and proceed.

**Perfect Foresight Dynamics**

The “natural state” of the dynamic system is made up of the dividend rates $d_t = (d_1(t), \ldots, d_K(t))$, the (reduced) distribution $\mu_t = (\mu_{lok}(t), \mu_{hn}(t))_{1 \leq k \leq K}$ of types, and the net utility of searching for each asset $w_t = (W_1(t), \ldots, W_K(t))$. The dividend rates $d_t \in \mathbb{R}^K$ and the distribution of types $\mu_t \in \mathbb{R}^{K+1}$ are the predetermined variables and the net utilities of searching for assets $w_t \in \mathbb{R}^K$ are the non-predetermined variables. The state is denoted by $y_t \equiv (d_t', \mu_t', w_t') \in (3K + 1) \times 1$. Equations (89)--(91) describe an investor’s search intensity allocation $\lambda$ in terms of a smooth function $L(\cdot)$, with $\lambda = L(W)$. Hence, the dynamics of the state are described by the system

\[
\begin{align*}
\dot{d}_t &= G \cdot (d_t - d^*) \quad (95) \\
\dot{\mu}_t &= H(\mu_t, L(w_t)) \quad (96) \\
\dot{w}_t &= R(d_t, \mu_t, w_t, L(w_t)). \quad (97)
\end{align*}
\]

The dynamic of the dividend rate (95) is assumed to be autonomous and linear for convenience. Equation (96) represents the ODE (80)--(81) for the distribution of types, and equation (97) follows from simple manipulation of the system (84)--(87) of penalized Bellman equations.

**Linearized Dynamics**
This paragraph checks the local uniqueness of the perfect-foresight dynamics by linearizing the system (95)-(97) in a neighborhood of its steady state. The linearized dynamics are

\[
\begin{bmatrix}
\frac{d_t}{dt} \\
\bar{\mu}_t \\
\bar{w}_t
\end{bmatrix} = \begin{bmatrix}
G & 0 & 0 \\
0 & J_H^H & J_L^H J_w^L \\
J_d^R & J_{\mu}^R & J_w^R + J_L^R J_w^L
\end{bmatrix} \begin{bmatrix}
d_t - d^* \\
\mu_t - \mu^* \\
w_t - W^*
\end{bmatrix},
\]  

(98)

where \(J_f^y\) denotes the Jacobian of some function \(f : \mathbb{R}^M \to \mathbb{R}^N\). In order to check the local determinacy of the perfect-foresight equilibrium, one uses the eigenvalue decomposition that is standard in linear rational expectations models. (see Buiter [1984] for the continuous-time version.) In all numerical examples considered, \(J\) has as many eigenvalues with positive real part as non-predetermined variables, ensuring local determinacy.

The linearization also provides an approximation of the perfect-foresight equilibrium path. The following paragraph proposes instead computations based on a reverse-shooting method. As the figures make clear, reverse shooting with a barrier function provides a smooth approximation of a “bang-bang” policy function. A linearization, on the other hand, could not capture this “bang-bang” feature.

**Computing Perfect Foresight Equilibrium**

The reverse-shooting method follows Judd [1999]. The perfect-foresight equilibrium path solves the system (95)-(97) of ordinary differential equations, denoted \(\dot{y} = g(y)\). The steady state is denoted by \(y^*\) and the initial condition by \(y_0 = [d_0^0 \ \mu_0^0 \ \bar{w}_0^0]^T\). The time horizon of the computation is \(T\). Given a candidate terminal value \(y_T\), the ODE \(\dot{y} = g(y)\) is solved backward, using a fourth-order Runge-Kutta method. This computation delivers candidate-initial conditions \(\tilde{d}_0(y_T), \tilde{\mu}_0(y_T),\) and \(\tilde{w}_0(y_T)\). The reverse-shooting method is to solve the problem

\[
\min_{y_T} \| \tilde{d}_0(y_T) - d_0 \|^2 + \| \tilde{\mu}_0(y_T) - \mu_0 \|^2,
\]

(99)

subject to \(\| y_T - y^* \| < \eta\), where \(\eta\) is a small positive number.

It is convenient to solve the program (99) using a continuation method. Namely, one solves successive programs indexed by a decreasing sequence of penalty \(\varepsilon_1 > \varepsilon_2 \cdots > \varepsilon_N = \varepsilon\). The solution of the \(n\)-th optimization program (99) is used as the initial condition of the \(n+1\)-st program.

For the study of an unexpected permanent increase of the dividend rate, \(\varepsilon\) is set to 0.001. For the study of a temporary increase, it is set to 0.003.
References


Kadlec, Gregory B., and John J. McConnell, The Effect of Market Segmen-
tation and Illiquidity on Asset Price: Evidence from Exchange listing, 

Kiyotaki, Nobuhiro, and Randall Wright, On Money as a Medium of Ex-

Krishnamurthy, Arvind, The Bond/Old-Bond Spread, *Journal of Financial 

Merton, Robert C., Presidential Address: A Simple Model of Capital Market 
Equilibrium with Incomplete Information, *Journal of Finance*, 1987, 42, 
483–510.

Pastor, Lubos, and Robert F. Stambaugh, Liquidity risk and expected stock 

Petersen, M, and D. Fialkowski, Posted versus Effective Spreads: Good 

Protter, Philip, *Stochastic Integration and Differential Equations*, New York: 
Springer-Verlag, 1990.

Sharpe, William F., Capital Asset Prices: A Theory of Market Equilibrium 

Taylor, Angus E., and Robert W. Mann, *Advanced Calculus*, New-York: 
Wiley, John and Sons, 1983.

Vayanos, Dimitri, Transaction Costs and Asset Prices: A Dynamic Equilib-

____, and Jean-Luc Villa, Equilibrium Interest Rate and Liquidity Premium 

____, and Tan Wang, “A Theory of On-The-Run and Off-The-Run Bond 

Wallace, Neil, A Model of the Liquidity Yield Structure Based on Asset 