Liquidity Premia in Dynamic Bargaining Markets

First Version : November 27th 2001

Pierre-Olivier Weill[*]

May 3, 2005

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

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.

In our model, investors cannot trade instantly in multilateral Walrasian market. Instead, trade is bilateral: investors have to search for each others, 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. This search framework applies most directly to over-the-counter markets such as the Treasury market, the corporate-bond market, or markets for financial derivatives. More generally, it applies to trades that are not arranged in a centralized market, such as block trades in the New York Stock Exchange (NYSE) upstairs market. Lastly, the search friction is likely to have an impact on asset prices even in markets where security dealers provide immediacy to outside investors. First, the search friction determines investors' outside option when they trade with dealers. Second, dealers might have to search for end investors in order to unload their inventories, and would charge the associated search cost to their customers. Third, in some markets, such as the corporate-bond market, dealers typically act as a brokers and search for counterparties on

the behalves of their customers.

In the present model, 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 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

$$R_k - R_L = \frac{\bar{\phi}}{\phi_k} (R_M - R_L).$$ (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), ϕ_k denotes the free float of asset k, defined as the portion of the market capitalization available for sale; $\bar{\phi}$ 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 could also be 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,

3

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.

This paper extends the one-asset model of Duffie, Gârleanu and Pedersen [2005] by allowing investors to trade many assets. The present cross-sectional analysis could not have been conducted in the one-asset model, 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 liquidity difference between on-the-run and off-the-run bonds, they provide a two-asset extension of Duffie, Gârleanu and Pedersen [2005]. 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 the analysis on pricing and measurement. In particular, I show that the

FARM (1) holds in an equilibrium with arbitrarily many assets.

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 and Preferences

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]). There are many assets $k \in \{1, \ldots, K\}$ in positive supply. Asset k has a measure $s_k \in (0, 1)$ of shares outstanding, and every share of an asset pays the same dividend rate $\delta > 0$.

The economy is populated by a unit-mass continuum of infinitely-lived and risk-neutral investors who discount the future at the constant rate $r > 0$. An investor enjoys the consumption of a non-storable numéraire good called "cash," with a marginal utility normalized to 1. In order to make side payments, investors are endowed with a technology that instantly produces cash, at unit marginal cost.[1]

An investor has either a high-valuation or a low-valuation for holding assets. When he has a high valuation and holds asset $k \in \{1, \ldots, K\}$, he enjoys the (per unit) utility flow δ. With a low valuation, he enjoys a utility flow $\delta - x$, for some

---

[1]In other words, negative consumption of cash is allowed. Appendix B shows that, equivalently, one could assume that investors borrow and save cash in some "bank account," at the exogenously given interest rate $\bar{r} = r$, and subject to an appropriate Transversality Condition.

holding cost $x > 0$.[2] Investors switch randomly, and pair-wise independently, from a low valuation to a high valuation with intensity[3] γ_u, and from a high valuation to a low valuation with intensity γ_d.

An investor is permitted to hold either zero or one share of some asset,[4] and can choose which asset to hold. We let $s \equiv (s_1, s_2, \ldots, s_K)$ denote the distribution of ownership. We also assume that

$$\sum_{k=1}^{K} s_k \equiv S < \frac{\gamma_u}{\gamma_u + \gamma_d}, \tag{2}$$

which means that the total supply S of assets is less than the steady-state measure of high-valuation investors.[5] Given that investors can hold at most one unit of some asset, equation (2) implies that, in a multilateral Walrasian market, the "marginal investor" has a high valuation. Therefore, in a Walrasian market, all assets have the same equilibrium price δ/r.

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 nonowner n).

---

[2]Duffie, Gârleanu and Pedersen [2005] provide a formal model of the holding cost. They assume that risk-averse investors receive some non-tradable endowment stream which is sometimes highly correlated with the traded asset. In a first-order Taylor expansion of an investor's continuation utility, x represents the cost of holding an asset when it has a high correlation with the endowment.

One could also view the holding cost as the intensity of an investor's need for cash, when he is borrowing constrained and cannot borrow against the full value of his asset holding. Suppose that, if the asset is worth p_k, an investor can only borrow $p_k - h$, for some "haircut" h. If the shadow value of relaxing the borrowing constraint is ϕ, then the holding cost is $x = \phi h$.

[3]For instance, if the investor's valuation is low, the distribution of the next switching time to high is exponential with parameter γ_u. The successive switching times are independent.

[4]Because 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 δ/N.

[5]An application of the Law of Large Numbers implies that the steady-state measure of high-valuation investors is equal to the stationary probability $\gamma_u/(\gamma_d + \gamma_u)$ of being in a state of high valuation.

Hence, the set of investor types is

$$I = \{hn, \ell n, ho1, \ldots, hoK, \ell o1, \ldots, \ell oK\}. \tag{3}$$

For each $i \in I$, μ_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 mass $\bar{\nu}$ 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 ν_{ij} of specialists of type (i, j), subject to the resource constraint $\sum_{j \in I} \nu_{ij} \leq \bar{\nu}$. 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 $\bar{\nu}$ and Λ up and down, respectively, by the same factor has no effect, one can assume without loss of generality that $\bar{\nu} = 1$. Contacts are also pair-wise independent with the investor's valuation 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)

7

and (j, i), for $i \neq j$, occur continually at a total (almost sure) rate of

$$\mu_i \nu_{ij} \Lambda \mu_j \nu_{ji} + \mu_j \nu_{ji} \Lambda \mu_i \nu_{ij} = 2\Lambda \mu_i \nu_{ij} \mu_j \nu_{ji}. \tag{4}$$

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). Specifically, each specialist of the mass $\mu_i \nu_{ij}$ of specialists of type (i, j) makes contacts at rate Λ, 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 ν_{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 λ_{ij} to the search for investors of type j, subject to the ressource 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

$$2\mu_i \mu_j \frac{\lambda_{ij} \lambda_{ji}}{\Lambda}. \tag{5}$$

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.

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,[6] and (b) these two types of investors maintain trading specialists who are searching for each other, that is, if $\lambda_{ij}\lambda_{ji} > 0$.

In equilibrium, the gains from trade will be strictly positive when a low-valuation owner (one of type ℓok) contacts a high-valuation non-owner (of type hn). The ℓok investor will sell her asset to the hn investor, in exchange for some cash payment p_k.[7] The price arises in a simple Nash-bargaining game, as follows. The total surplus of such a transaction is

$$(V_{hok} - V_{hn}) - (V_{\ell ok} - V_{\ell n}) \equiv \Delta V_{hk} - \Delta V_{\ell k}. \tag{6}$$

We study those equilibria in which the ℓok agent receives a fixed fraction $q \in (0,1)$ of the total surplus. This implies that the price of asset k is, in an equilibrium,

$$p_k = q\Delta V_{hk} + (1-q)\Delta V_{\ell k}. \tag{7}$$

The gains from trade can also be positive between a low-valuation owner ℓok and a high-valuation owner hoj. These two investors may swap assets, and one investor may simultaneously transfer cash to the other. The total surplus of a swap between a ℓok agent and a hoj agent is $V_{\ell oj} - V_{\ell ok} + V_{hok} - V_{hoj}$.

We guess (and verify later) that, in equilibrium, an ℓok investor will not maintain trading specialists who search for swaps with hoj investors, but only trading specialists who search for an outright sale with hn investors. In other words, the net utility of searching for a swap is, in the equilibrium we analyze, strictly less than the net utility of searching for an outright sale, a condition that

---

[6]An arbitrarily small transaction cost rules out trade when the gain is zero.

[7]A cash payment payments is a lump of consumption good, instantly produced at unit marginal cost.

9

can be written:

$$\mu_{hoj}(V_{\ell oj} - V_{\ell ok} + V_{hok} - V_{hoj}) < \mu_{hn}(\Delta V_{hk} - \Delta V_{\ell k}), \tag{8}$$

for all $(k, j) \in \{1, \ldots, K\}^2$. We verify condition (8) in the proof of Proposition 3.

Under condition (8), an ℓok investor allocates all of her search intensity Λ to the search for hn investors. On the other hand, an hn investor allocates intensities, denoted $\lambda_1, \ldots, \lambda_K$, to simultaneous searches for investors of respective types $\ell o1, \ldots, \ell oK$. These allocations of search intensity are illustrated in Figure 1.

Definition 1 *A search-intensity allocation is some* $\lambda \in \mathbb{R}_+^K$ *with* $\sum_{k=1}^K \lambda_k \leq \Lambda$.
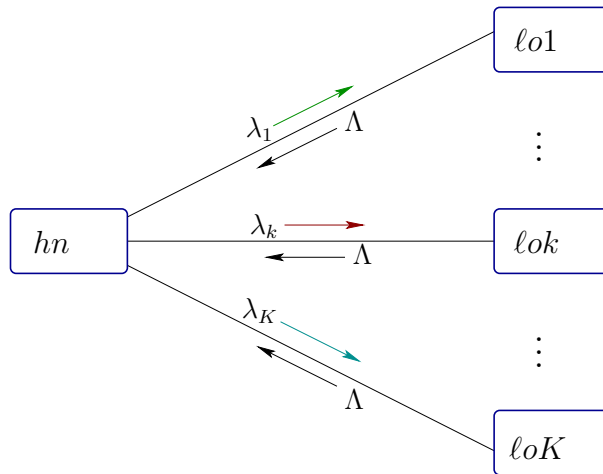
Figure 1: Allocating Search Intensity

An hn investor searches ℓok investors with intensity λ_k, for all $k \in \{1, \ldots, K\}$. An ℓok investor, on the other-hand, searches hn investors with intensity Λ.

Investors' Problem

This paragraph characterizes the equilibrium continuation utilities V_i, $i \in I$. As shown in Appendix B by an optimality verification argument, the Bellman

equation for the continuation utility of a buyer hn is

$$rV_{hn} = \max_{\tilde{\lambda}_1,\ldots,\tilde{\lambda}_K} \left\{ \gamma_d(V_{\ell n} - V_{hn}) + 2\sum_{k=1}^{K} \tilde{\lambda}_k \mu_{\ell ok} \left(V_{hok} - V_{hn} - p_k\right) \right\}, \qquad (9)$$

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 (˜) is used to distinguish the search intensity λ_k that will prevail in equilibrium for all investors of type hn from the intensity $\tilde{\lambda}_k$ that is to be chosen by an individual investor of type hn, taking others' search intensities as given. The Bellman equation (9) breaks up the "flow" continuation utility rV_{hn} into two terms. The first term, $\gamma_d(V_{\ell n} - V_{hn})$, is the expected flow utility of a transition from a high to a low valuation because, with intensity γ_d, a hn investor makes a transition to the ℓn type. The second term is the expected flow utility of searching for alternative assets. Namely, with intensity $2\tilde{\lambda}_k \mu_{\ell ok}$, an hn investor finds asset k, buys it at price p_k and makes a transition to type hok. Similarly, other investors' continuation utilities solve the following system of Bellman equations

$$rV_{hok} = \delta + \gamma_d(V_{\ell ok} - V_{hok}) \qquad (10)$$

$$rV_{\ell ok} = \delta - x + \gamma_u(V_{hok} - V_{\ell ok}) + 2\lambda_k \mu_{hn}(V_{\ell n} - V_{\ell ok} + p_k) \qquad (11)$$

$$rV_{\ell n} = \gamma_u(V_{hn} - V_{\ell n}), \qquad (12)$$

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

Steady-state distribution of Types

We now provide the equations characterizing the steady-state distribution of investors' types. First, of course, all assets are being held and the mass of investors is equal to one:

$$s_k = \mu_{\ell ok} + \mu_{hok} \qquad (13)$$

$$1 = \sum_{k=1}^{K} (\mu_{\ell ok} + \mu_{hok}) + \mu_{hn} + \mu_{\ell n}. \qquad (14)$$

11

Second, in a steady state, the inflow and outflow of investors in each type is zero. For example, for the high-valuation non-owner hn, we have

$$\gamma_u \mu_{\ell n} = \gamma_d \mu_{hn} + 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{\ell o k}. \tag{15}$$

The left-hand side is the flow of ℓn investors who switch from a low valuation to a high valuation, transiting to the hn type. The first term on the right-hand side, $\gamma_d \mu_{hn}$, is the flow of hn investors who switch to a low valuation. The second term is the flow of hn investors who meet sellers of some asset $k \in \{1, \ldots, K\}$ and buy an asset. Similarly, for the low-valuation owner $\ell o k$

$$\gamma_d \mu_{hok} \quad = \quad \gamma_u \mu_{\ell o k} + 2\lambda_k \mu_{hn} \mu_{\ell o k}. \tag{16}$$

for $k \in \{1, \ldots, K\}$. Lastly, similar calculations (see Appendix A) show that the inflow-outflow equations for investors of types ℓn and hok are the same as (15) and (16). We can now define:

Definition 2 *A steady-state symmetric equilibrium is a collection* $p = (p_1, \ldots, p_K)$ *of prices, a collection* $V = (V_{hn}, V_{hok}, V_{\ell o k}, V_{\ell n})_{1 \leq k \leq K}$ *of continuation utilities, a distribution* $\mu = (\mu_{hn}, \mu_{hok}, \mu_{\ell o k}, \mu_{\ell n})_{1 \leq k \leq K}$ *of types, and a search-intensity allocation* $\lambda \gg 0$*, such that*

(i) *Steady-State: Given* λ*,* μ *solves the system (13)-(16).*

(ii) *Optimality: Given* λ*,* μ*, and* p*,* V *and* $(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_K) = \lambda$ *solve the system (9)-(12) of Bellman equations. The no-swap condition (8) holds for all* $(k, j) \in \{1, \ldots, K\}^2$*.*

(iii) *Pricing: the prices satisfy equation (7), for all* $k \in \{1, \ldots, K\}$*.*

Here, symmetry means that all hn investors choose the same search intensity allocation λ. Also, this definition restricts attention to equilibria having two specific properties: there are no swap and all assets are searched, that is $\lambda \gg 0$.[8]

---

[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 could be a coordination failure. For example, there is an equilibrium in which no investor search: namely, given that no other investor searches, not searching is optimal.

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

$$\mu_{\ell ok}(V_{hok} - V_{hn} - p_k) = \mu_{\ell oj}(V_{hoj} - V_{hn} - p_j) \tag{17}$$

$$\iff \quad \mu_{\ell ok}(1-q)(\Delta V_{hk} - \Delta V_{\ell k}) = \mu_{\ell oj}(1-q)(\Delta V_{hj} - \Delta V_{\ell j}), \tag{18}$$

for all $(k, j) \in \{1, \ldots, K\}^2$, and where (18) follows from substituting (7) into (17). This first-order condition reflects "search indifference," meaning that the marginal utility of spending an additional unit of search intensity on a given asset is equated across assets. This marginal utility is decomposed as follows: Conditional on establishing a contact, a seller of asset k is found with probability $\mu_{\ell ok}$. Then, the buyer receives a fraction $1-q$ of the transaction surplus $\Delta V_{hk} - \Delta V_{\ell k}$.

The total transaction surplus may be interpreted as a bid-ask spread, in the following sense. Suppose that the seller's bargaining power is a random variable with support $[0,1]$ and mean q, independently distributed across encounters. Then, the maximum buying price (the ask) is ΔV_{hk} and the minimum selling price price (the bid) is $\Delta V_{\ell k}$. The average price of asset k is $p_k = q\Delta V_{hk} + (1-q)\Delta V_{\ell k}$. Following this interpretation, condition (18) means that an asset that is easier to find (with a larger $\mu_{\ell ok}$) 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 (18).

First, the study of (13)-(16) presented in Appendix A shows the following proposition.

Proposition 1 *Given a search intensity allocation λ, the system (13)-(16) has a unique solution $\mu = (\mu_{hn}, \mu_{hok}, \mu_{\ell n}, \mu_{\ell ok})_{1 \le k \le K} \in [0,1]^{2K+2}$.*

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

$$W_k \equiv \mu_{\ell ok}(1 - q)(\Delta V_{hk} - \Delta V_{\ell k}), \tag{19}$$

for all $k \in \{1, \ldots, K\}$. Clearly, the "search indifference" marginal conditions (18) can be written as

$$W_k = W, \tag{20}$$

for all $k \in \{1, \ldots, K\}$, and for some positive constant W to be determined. Substituting (20) into equation (9), combining the Bellman equations (9) through (11), and using the pricing equation (7) one finds that

$$rW_k = (1 - q)\mu_{\ell ok}x - (\gamma_u + \gamma_d + 2\lambda_k q\mu_{hn})W_k - 2\Lambda(1 - q)\mu_{\ell ok}W, \tag{21}$$

for all $k \in \{1, \ldots, K\}$. Replacing $\mu_{hok} = s_k - \mu_{\ell ok}$ in equation (16), we find that

$$2\lambda_k\mu_{hn} = \frac{\gamma_d s_k}{\mu_{\ell ok}} - (\gamma_d + \gamma_u). \tag{22}$$

Substituting (22) into (21), using (20) and rearranging gives

$$\frac{\gamma_d s_k q}{(1 - q)x} \frac{1}{\mu_{\ell ok}^2} + \frac{r + (1 - q)(\gamma_d + \gamma_u)}{(1 - q)x} \frac{1}{\mu_{\ell ok}} + \frac{2\Lambda}{x} = \frac{1}{W}. \tag{23}$$

This quadratic equation allows one to write $\mu_{\ell ok} = m_k(W)$, for some $W < x/(2\Lambda)$, and for some continuous and increasing function $m_k(\cdot)$.

Now, the steady-state measure of high-valuation investors is equal to the stationary probability of being in a state of high valuation[9]

$$\mu_{hn} + \sum_{k=1}^{K} \mu_{hok} = \frac{\gamma_u}{\gamma_u + \gamma_d}. \tag{24}$$

Combining (24) with (13) shows

$$\mu_{hn} = \frac{\gamma_u}{\gamma_u + \gamma_d} - S + \sum_{k=1}^{K} \mu_{\ell ok}. \tag{25}$$

---

[9]This can also be shown by summing equation (16) over $k \in \{1, \ldots, K\}$, subtracting equation (15) and using (14).

Substituting (25) into (22) gives

$$2\lambda_k \left(\frac{\gamma_u}{\gamma_u + \gamma_d} - S + \sum_{k=1}^{K} m_k(W) \right) = \frac{\gamma_d s_k}{m_k(W)} - (\gamma_d + \gamma_u), \tag{26}$$

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

$$2\Lambda \left(\frac{\gamma_u}{\gamma_u + \gamma_d} - S + \sum_{k=1}^{K} m_k(W) \right) - \sum_{k=1}^{K} \left(\frac{\gamma_d s_k}{m_k(W)} - (\gamma_d + \gamma_u) \right) = 0. \tag{27}$$

The left-hand side of (27) is strictly increasing in W because $m_k(\cdot)$ is strictly increasing for each k. Hence, (27) uniquely characterizes a candidate equilibrium W. Once W is found, the other equilibrium objects are uniquely characterized: the search intensity allocation λ by (26), the distribution μ of types by (13)-(16), the continuation utilities V by (9)-(11), and the prices p by (7). 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 [2005]. Then, one applies the Implicit Function Theorem to equation (27), 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 $\hat{\mu}_{\ell ok} = \hat{\mu}_{\ell o}/K$ and $\hat{\lambda}_k = \Lambda/K$. The equilibrium equations are those of Duffie, Gârleanu and Pedersen [2005], with "λ" there being replaced here by "Λ/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

15

investor to searching for a swap with an *hoj* investor, for all $j \in \{1, \ldots, K\}$. Since the left-hand side of (27) 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 $\hat{s}$, 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, μ, λ) are easily expressed as continuous functions of W and thus as continuous functions of s. The search-indifference marginal conditions (20) are satisfied by construction. The no swap condition (8) as well as all other relevant inequalities hold by continuity. $\blacksquare$

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 of an equilibrium with $\lambda \gg 0$.) *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 with $\lambda \gg 0$ 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 ε is small enough, there is no candidate equilibrium in which $\lambda \gg 0$. If such an equilibrium exists, then both assets would satisfy (23). For Asset 1, since $\mu_{\ell o1} \leq s_1 = \varepsilon$, (23) would imply that the candidate W goes to zero as ε goes to zero. In turn, for Asset 2, (23) would imply that $\mu_{\ell o2}$ goes to zero as ε goes to zero, but then $s_2/\mu_{\ell o2} = (S - \varepsilon)/\mu_{\ell o2}$ goes to infinity as ε goes to zero. Therefore, equation (27) cannot hold. $\blacksquare$

Existence in Proposition 3, and non-existence in Proposition 4, are proved by studying how the left-hand side of (27) 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, (27) 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 the exogenous distributions $s = (s_1, \ldots, s_K)$ of ownership.

3.1   Cross-sectional Prices

The pricing equation (7) can be written

$$p_k = \Delta V_{hk} - (1-q)(\Delta V_{hk} - \Delta V_{\ell k}). \tag{28}$$

The first term on the right-hand side, ΔV_{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 ΔV_{hk} which, when substituted in (28), gives

$$rp_k = \delta - 2\Lambda W - \gamma_d(\Delta V_{hk} - \Delta V_{\ell k}) - r(1-q)(\Delta V_{hk} - \Delta V_{\ell k}). \tag{29}$$

This equation breaks the price of asset k into four components. The first, δ, 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, $\gamma_d(\Delta V_{hk} - \Delta V_{\ell k})$, is the instantaneous cost of switching to the low valuation, 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 δ/r in a multilateral Walrasian market, where all assets can be bought and sold instantly. In the Walrasian case, all discounts in (29) 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 a low valuation 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

$$rp_L \equiv \delta - 2\Lambda W. \tag{30}$$

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.[10] Subtracting equation (30) from equation (28) and rearranging, one finds

$$r(p_L - p_k) = (\gamma_d + r(1-q))(\Delta V_{hk} - \Delta V_{\ell k}). \tag{31}$$

Together with the search-indifference condition (18), (31) implies that

$$p_k \mu_{\ell ok} \left(\frac{\delta}{p_k} - \frac{\delta}{p_L} \right) = \frac{\delta}{rp_L}(\gamma_d + r(1-q))W, \tag{32}$$

---

[10]In the present model, the supply of asset L is zero and (30) 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.

where W is the net utility of searching for alternative assets. The returns of assets k and L are respectively denoted by $R_k \equiv \delta/p_k$ and $R_L \equiv \delta/p_L$. The free float of asset k is $\phi_k \equiv p_k \mu_{\ell ok}$. This is the portion of its market capitalization available for sale. With these notations, (32) can be written

$$\phi_k(R_k - R_L) = M, \tag{33}$$

for some positive constant M which does not depend on k. Equation (33) states that, in equilibrium, the liquidity spread $R_k - R_L$ of asset k is proportional to the inverse of its free float ϕ_k. A convenient expression for the constant of proportionality is obtained by summing equations (33) 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 $k \in \{1, \dots, K\}$,*

$$R_k - R_L = \frac{\bar{\phi}}{\phi_k}(R_M - R_L), \tag{34}$$

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

$$R_M \equiv \sum_{k=1}^{K} \frac{\phi_k}{\sum_{j=1}^{K} \phi_j} R_k. \tag{35}$$

When studying the impact of liquidity on cross-sectional returns, many researchers estimate linear models. For example, if they study stock returns, they would 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]). The FARM (34) suggests that liquidity could also be measured by the free float.

The FARM (34) is a deterministic relationship between returns and float. In a simple stochastic extension of this model, one can let the cumulative dividend rate of asset $k \in \{1, \ldots, K\}$ solve the stochastic differential equation $dD_{kt} = \delta\, dt + \sigma_k dB_{kt}$, where $(\sigma_1, \ldots, \sigma_K) \in \mathbb{R}_+^K$ and $(B_{1t}, \ldots, B_{Kt})$ is a standard Brownian motion. Because our investors are risk-neutral and dividend shocks are independent over time, this stochastic model is essentially equivalent to the present deterministic one. The FARM (34) would also hold, with R_k being the expected return of asset k.[11]

The FARM (34) also holds when holding costs x are heterogenous across assets. It no longer holds, however, when the dividend rates δ are heterogenous: the relationship between return and float would not be (inversely) linear. Numerical calculations (not reported here) suggests that the relationship remains decreasing.

Some evidence on float and liquidity

Some evidence suggest that, in an empirical application, the 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.[12]

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

---

[11]Then, the realized return over a holding period $[T_1, T_2]$ is $R_k(T_2 - T_1) + \sigma_k/p_k \int_{T_1}^{T_2} dB_{kt}$.

[12]Measuring float in practice is a delicate issue because one has to decide which shares are available for investment, and which ones are not. In our model, for example, the only shares of asset k available for investment *at price p_k* are the one held by low-valuation investors. One might argue that the correct definition of float is the total number of share s_k, because they are available for investment at any price greater than ΔV_{hk} (moreover, a larger s_k also results in higher liquidity.) However, no investor in our model is willing to pay such a high price for this asset.

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 [2004] 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 (HSI 33). The monetary authority absorbed about 7.3% of HSI 33 market capitalization and held these stocks for a long time period, resulting in a reduction in the free float of these stocks. The authors show that, relative to some control group with no free-float reduction, the HSI 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.[13]

Here, the cross-sectional variation in asset returns is not explained by an exogenously specified cross-sectional variation in transaction costs, in contrast with the Walrasian models of Amihud and Mendelson [1986], Constantinides [1986], Vayanos [1998], and Huang [2003]. In this 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,

---

[13]The same solution method can be applied to assets which are heterogenous in other dimensions. For example, one could consider cross-sectional variations in dividends, holding costs, and make asset heterogenous in terms of search costs. These extensions of our model are available upon request.

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 (28), written as

$$p_k = \frac{\delta}{r} - \frac{2\Lambda W}{r} - \left(1 + \frac{\gamma_d}{(1-q)r}\right)\frac{W}{\mu_{\ell ok}}. \tag{36}$$

The right-hand side is increasing in $\mu_{\ell ok}$. In other words, an asset that is easier to find (has larger $\mu_{\ell ok}$) is sold at a higher price. The second equation (23) is of the form

$$As_k\frac{1}{\mu_{\ell ok}^2} + B\frac{1}{\mu_{\ell ok}} + C = \frac{1}{W}, \tag{37}$$

for some positive constants A, B, and C, which do not depend on k. The third equation is easily derived from (22), and relates λ_k to the distribution of types and s_k:

$$\frac{\mu_{\ell ok}}{s_k} = \frac{\gamma_d}{\gamma_d + \gamma_u + 2\lambda_k\mu_{hn}}. \tag{38}$$

The quantity $\lambda_k\mu_{hn}$ has several interpretations. First, it represents the demand side of the market. The larger is λ_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 $\mu_{\ell ok}$? Equation (38) 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 τ_h, decomposed as follows. The investor holds the asset k until she switches to a state of low valuation at some time $t+\tau_d$, where τ_d is an exponentially distributed stopping time with parameter γ_d. Then, she either meets a buyer or switches back to a high valuation at some time $t + \tau_d + \min\{\tau_b, \tau_u\}$, where τ_b and τ_u are exponentially distributed stopping

times with respective parameters $2\lambda_k\mu_{hn}$ and γ_u. If she switches back to a high valuation utility, then her mean holding period is some stopping time $\tilde{\tau}_h$. Hence,

$$\tau_h = \tau_d + \mathbb{I}_{\{\tau_u < \tau_b\}}(\tau_u + \tilde{\tau}_h) + \mathbb{I}_{\{\tau_b \leq \tau_u\}}\tau_b = \tau_d + \min\{\tau_u, \tau_b\} + \mathbb{I}_{\{\tau_u < \tau_b\}}\tilde{\tau}_h. \quad (39)$$

In a steady-state equilibrium, $\tilde{\tau}_h$ and τ_h are identically distributed. Furthermore, all the above stopping times are pairwise independent. Taking expectations of both sides of (39), and using the fact that τ_h and $\tilde{\tau}_h$ are identically distributed, one finds that

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

and therefore

$$E(\tau_h) = \frac{1}{\gamma_d} + \frac{1}{2\lambda_k\mu_{hn}}\left(1 + \frac{\gamma_u}{\gamma_d}\right). \quad (40)$$

This shows that the mean holding period $E(\tau_h)$ is a decreasing function of $\lambda_k\mu_{\ell n}$. Equation (37) has the form

$$F(s_k, \mu_{\ell ok}) = \frac{1}{W}, \quad (41)$$

for some function $F(\cdot, \cdot)$ that is increasing in s_k and decreasing in $\mu_{\ell ok}$. This implies that μ_{ellok} 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 (40), one writes equation (37) as

$$G\left(s_k, \frac{\mu_{\ell ok}}{s_k}\right) = \frac{1}{W}, \quad (42)$$

for some function $G(\cdot, \cdot)$ that is decreasing in s_k and decreasing in $\mu_{\ell ok}/s_k$. This implies that $\mu_{\ell ok}/s_k$ is a decreasing function of s_k. From (38), 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 $2\lambda_k\mu_{hn}\mu_{\ell ok}$, an asset with more shareholders also has a larger trading volume. The above discussion is summarized in

23

Proposition 6 *In equilibrium, $s_k > s_j$ implies that $\mu_{\ell o k} > \mu_{\ell o j}$, $\lambda_k > \lambda_j$, $p_k > p_j$, $R_k < R_j$, and $\Delta V_{hk} - \Delta V_{\ell k} < \Delta V_{hj} - \Delta V_{\ell j}$.*

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.

In contrast with Proposition 6, the one-asset model of Duffie, Gârleanu and Pedersen [2005] implies that an asset with a larger quantity of shareholders has a *lower* price. Indeed, an increase in the quantity of shareholders results in a positive shift of the supply curve, and hence lowers the price of the asset. Similarly, in our model, a larger s_k represents a larger supply. However, a larger s_k also endogenously results in a larger demand, represented by a larger search intensity λ_k. Proposition 6 shows that the "demand shift" dominates, meaning that an asset with larger s_k has a higher price.

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.

3.4   An illustrative numerical example

There is much empirical evidence on the relationship between traditional liquidity proxies and risk-adjusted expected returns. For example, after controlling for risk, expected return appear to be positively related to bid-ask spread, and negatively

Table 1: Parameter Values used in the Numerical Example.

Parameters		Value or Distribution
Contact Intensity	Λ	$62,500$
Intensity of Switch to High	γ_d	1
Intensity of Switch to Low	γ_u	$7/3$
Discount Rate	r	5%
Number of Assets	K	100
Fraction of Shareholders	s_k	$\sim \text{Uniform}([0.0042, 0.0078])$
Holding cost	x	5
Dividend Rate	δ	1
Bargaining Power	q	$1/2$

related to trading volume and turnover. This subsection presents a numerical example suggesting that our model's qualitative predictions are consistent with these empirical evidence.

An equilibrium of the model is computed for a randomly generated economy of $K = 100$ asset types.[14] The ownerships s_k are drawn independently from a uniform distributions on some interval $[\underline{s}, \overline{s}]$. The dividend rate δ is set to 1 and the holding cost x to 5. The bargaining power parameter is set to $1/2$. The equilibrium return $R_k = \delta / p_k$ 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 $1 - \Delta V_{\ell k} / \Delta V_{hk}$. The dollar trading volume is $p_k \lambda_k \mu_{\ell ok} \mu_{hn}$. The turnover is $\lambda_k \mu_{\ell ok} \mu_{hn} / s_k$. The market capitalization is $p_k s_k$. 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, $\Lambda = 62500$ means that an investor establishes a contact with some other investor every 2.5 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 0.6. An investor switches to a low valuation on average once a year, which implies an annual turnover slightly below 100

---

[14]One first solves (27) for W. This can be done quickly since equation (23), characterizing $m_k(W)$, is quadratic. Once W is found, the remaining equilibrium objects are easily computed using the various equations derived during the existence proof.

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 related to returns.
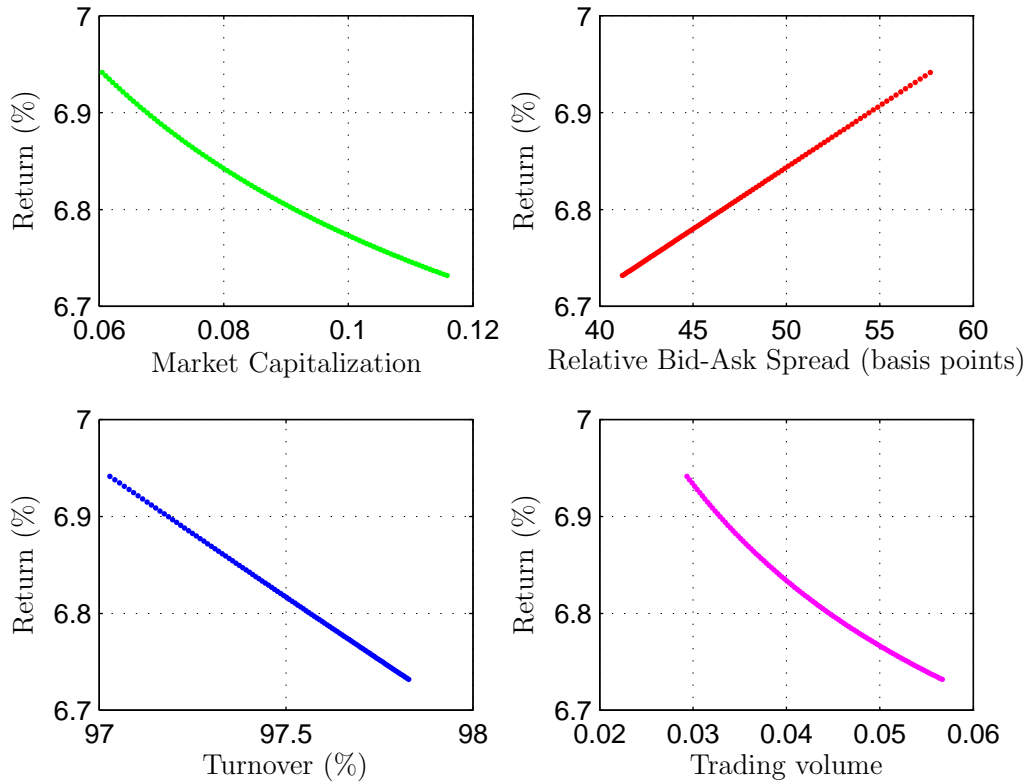

Figure 2: Cross-sectional variation in returns, explained by "liquidity proxies."

4   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.

A The Distribution of Types

This appendix studies the steady-state distribution of types, as well as the dynamics of the distribution of types near that steady-state. For a given search intensity allocation λ, the distribution

$$\mu(t) = (\mu_{hn}(t), \mu_{hok}(t), \mu_{\ell ok}(t), \mu_{\ell n}(t))_{1 \leq k \leq K}$$

of types solves the system of ordinary differential equations (ODE)

$$\dot{\mu}_{hn} = \gamma_u \mu_{\ell n} - \gamma_d \mu_{hn} - 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{\ell ok} \tag{43}$$

$$\dot{\mu}_{hok} = \gamma_u \mu_{\ell ok} - \gamma_d \mu_{hok} + 2 \lambda_k \mu_{hn} \mu_{\ell ok} \tag{44}$$

$$\dot{\mu}_{\ell n} = \gamma_d \mu_{hn} - \gamma_u \mu_{\ell n} + 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{\ell ok} \tag{45}$$

$$\dot{\mu}_{\ell ok} = \gamma_d \mu_{hok} - \gamma_u \mu_{\ell ok} - 2 \lambda_k \mu_{hn} \mu_{\ell ok} \tag{46}$$

$$s_k = \mu_{\ell ok} + \mu_{hok} \tag{47}$$

$$1 = \sum_{k=1}^{K} (\mu_{\ell ok} + \mu_{hok}) + \mu_{hn} + \mu_{\ell n}, \tag{48}$$

where $\dot{\mu} = d\mu(t)/dt$, and time arguments are suppressed. Since equation (47) implies that the sum of (44) and (46) is zero, one can eliminate (44), the ODE for μ_{hok}. Similarly, since equation (48) implies that the sum of equations (43) to (46) is zero, one can eliminate (45), the ODE for $\mu_{\ell n}$, and obtains the equivalent system

$$\dot{\mu}_{\ell ok} = \gamma_d s_k - (\gamma_d + \gamma_u) \mu_{\ell ok} - 2 \lambda_k \mu_{hn} \mu_{\ell ok} \tag{49}$$

$$\dot{\mu}_{hn} = \gamma_u (1 - S) - (\gamma_d + \gamma_u) \mu_{hn} - 2 \sum_{k=1}^{K} \lambda_k \mu_{hn} \mu_{\ell ok} \tag{50}$$

$$\mu_{hok} = s_k - \mu_{\ell ok} \tag{51}$$

$$\mu_{\ell n} = 1 - S - \mu_{hn}, \tag{52}$$

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

Proof of Proposition 1   A steady state is a constant distribution μ of types solving equations (49)-(52). Summing equations (49) over $k \in \{1, \ldots, K\}$, adding equation (50), and imposing the steady-state condition $\dot{\mu} = 0$, one finds

$$\mu_{hn} = \mu_{\ell o} + y - S, \tag{53}$$

where $\mu_{\ell o} \equiv \sum_{k=1}^{K} \mu_{\ell ok}$, and $y \equiv \gamma_u / (\gamma_d + \gamma_u)$. Replacing this last equation in (49) gives

$$\mu_{\ell ok} = \frac{\gamma_d s_k}{\gamma_d + \gamma_u + 2\lambda_k (\mu_{\ell o} + y - S)}. \tag{54}$$

Summing equations (54) over k, one obtains the one equation in one unknown problem

$$\mu_{\ell o} - \sum_{k=1}^{K} \frac{\gamma_d s_k}{\gamma_d + \gamma_u + 2\lambda_k (\mu_{\ell o} + y - S)} = 0. \tag{55}$$

The left-hand side of this equation is increasing in $\mu_{\ell o}$, is negative at $\mu_{\ell o} = 0$, and is positive for $\mu_{\ell o}$ large enough; thus, it has a unique solution. Once the solution $\mu_{\ell o}$ is found, $\mu_{\ell ok}$ is uniquely determined by (54), μ_{hn} by (53), and finally μ_{hok} and $\mu_{\ell n}$ by (51) and (52). 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 (55) is positive when evaluated at $\mu_{\ell o} = S$ and $1 - y$; it is negative when evaluated at $S - y$. Since the left-hand side of (55) is increasing, this shows that

$$S - y < \mu_{\ell o} < \min\{S, 1 - y\}. \tag{56}$$

Next, $s - y < \mu_{\ell o}$ implies that $\mu_{hn} > 0$ and that $\mu_{\ell ok} < s_k$. Finally, $\mu_{\ell o} < 1 - y$ implies that $\mu_{hn} < 1 - S$ and that $0 < \mu_{\ell n} < 1$.

Local Stability This paragraph establishes that, given λ, the steady-state distribution of types is a locally stable point of the following ODE

$$\dot{\mu}_{\ell ok} = \gamma_d s_k - (\gamma_d + \gamma_u)\mu_{\ell ok} - 2\lambda_k \mu_{hn}\mu_{\ell ok} \tag{57}$$

$$\dot{\mu}_{hn} = \gamma_u(1 - S) - (\gamma_d + \gamma_u)\mu_{hn} - 2\sum_{k=1}^{K}\lambda_k \mu_{hn}\mu_{\ell ok}, \tag{58}$$

for all $k \in \{1, \ldots, K\}$. Stacking variables as $(\mu_{\ell o1}, \ldots, \mu_{\ell ok}, \mu_{hn})'$, the Jacobian of the ODE at the steady state is

$$J = -(\gamma_d + \gamma_u)I_{K+1} - D, \tag{59}$$

where $D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$, $D_{11} = \text{diag}\,(2\lambda_k \mu_{hn})$, $D_{12} = \begin{bmatrix} 2\lambda_1 \mu_{\ell o1} \ldots 2\lambda_K \mu_{\ell oK} \end{bmatrix}'$, $D_{21} = \begin{bmatrix} 2\lambda_1 \mu_{hn} \ldots 2\lambda_K \mu_{hn} \end{bmatrix}$, and $D_{22} = \sum_{k=1}^{K} 2\lambda_k \mu_{\ell ok}$.

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 $\xi \in \mathbb{C}$ is denoted y. It solves

$$(\gamma_d + \gamma_u)y_1 + D_{11}y_1 + D_{12}y_2 = -\xi y_1 \tag{60}$$
$$(\gamma_d + \gamma_u)y_2 + D_{12}y_1 + D_{22}y_2 = -\xi y_2. \tag{61}$$

Multiplying equation (60) by e', and subtracting equation (61) gives

$$(\gamma_d + \gamma_u + \xi)\left(e'y_1 - y_2\right) = 0. \tag{62}$$

We distinguishes three cases.
Case 1: $e'y_1 \neq y_2$. From (62), it must be that $\xi = -(\gamma_d + \gamma_u) < 0$.

Case 2: $e'y_1 = y_2 = 0$. Then (60) simplifies to $(\gamma_d + \gamma_u + D_{11})y_1 = -\xi y_1$. Thus, it must be that $\xi = -(\gamma_d + \gamma_u + 2\lambda_k \mu_{hn}) < 0$, for some $k \in \{1, \ldots, K\}$.

Case 3: $e'y_1 = y_2 \neq 0$. Without loss of generality, one can assume that $y_2 = 1$. Using (60) one solves explicitly for y_{1k},

$$y_{1k} = -\frac{2\lambda_k \mu_{\ell ok}}{\nu + \gamma_d + \gamma_u + 2\lambda_k \mu_{hn}}. \tag{63}$$

Since the y_{1k} sum to one, it must be that $\operatorname{Re}\left(\sum_{k=1}^{K} y_{1k}\right) = 1$. Thus, there is one $k \in \{1, \ldots, K\}$ such that $\operatorname{Re}(y_{1k}) > 0$, which is equivalent to $\operatorname{Re}(1/y_{1k}) > 0$ and, from equation (63), to $\operatorname{Re}(\xi) < -(\gamma_d + \gamma_u + 2\lambda_k \mu_{hn}) < 0$. ∎

B   The Investor's Problem

This Appendix formulates and solves the stochastic-control problem faced by an individual investor in a steady-state equilibrium. We show that our preference specification is essentially equivalent to assuming that: (i) an investor maximizes the present discounted value of consumption, and (ii) is allowed to borrow and lend at a risk-free rate $\bar{r}$ that is equal to his subjective discount rate r, subject to an appropriate Transversality Condition. We also verify that the Bellman equations (9)-(12) are sufficient for optimality.

B.1   The Investor's Problem

Information   Time is continuous and runs forever, and some measurable space $(\Omega, \mathcal{G})$ is fixed. An investor can be of either one of finitely many types $i \in I$, with an initial type $j_0 \in I$. We fix some counting process $N_t \in \mathbb{N}^I$, where $N_t(i)$ counts the number of transition to type i during the time interval $[0, t]$. The sequence of transition times is denoted by $0 < T_0 < T_1 < \ldots$, and the sequence of successive types by $j_0, j_1, \ldots$ Hence, the process N_t generates a type process i_t such that, for all $n \in \mathbb{N}$ and $t \in [T_n, T_n + 1)$, $i_t = j_n$. We let $\{\mathcal{F}_t , \ t \geq 0\}$ be the internal history (filtration generated by) the process N.

Admissible control   A control u is made up of a $\mathcal{F}_t$-predictable $\mathbb{R}^I$-valued transition intensity process π, and of a $\mathcal{F}_t$-adapted, $\mathbb{R}_+$-valued consumption process c. Given π, we fix a probability P_π such that the counting process N admits P_π-intensity π.

A transition intensity process is admissible if, at each time, π_t belongs to some state-dependent compact set $\Pi(i_{t-}) \in \mathbb{R}^I$. The set $\Pi(i)$ embeds all feasible transitions of an investor in state i. For example, if $i = hn$, $\pi(\ell n) = \gamma_d$ for all $\pi \in \Pi(i)$. Similarly, $\pi(\ell ok) = 0$, because an investor can never transit directly from i to j. Also, $\pi(hok) = 2\lambda_k \mu_{\ell ok}$. Lastly, the set $\Pi(i)$ also incorporates the search-intensity budget constraint. For example, for a buyer hn, we must have $\sum_{k=1}^{K} \pi(hok)/(2\mu_{\ell ok}) \leq \Lambda$.

When an investor makes a transition from state i to j, he receives a payment $Q(i,j)$. For example, if $i = hn$ and $j = hok$, then $Q(i,j) = -p_k$. Conversely, if $i = \ell ok$ and $j = \ell n$, then $Q(i,j) = p_k$. Hence, the investor's wealth process solves the Stochastic Differential Equation (SDE)

$$dW_t = (rW_t - c_t)\, dt + D(i_t)\, dt + \sum_{j \in I} Q(i_{t-}, j)\, dN_t(j), \tag{64}$$

where $D(i_t)$ is the flow utility enjoyed by the investor from his asset holding. Namely, $D(\ell n) = D(hn) = 0$, $D(hok) = \delta$, and $D(\ell ok) = \delta - x$, for all $k \in \{1, \ldots, K\}$. We let an investor's initial wealth be W_0. A control $u = (\pi, c)$ is admissible if π is admissible and if the wealth process satisfies the Transversality Condition

$$\lim_{T \to \infty} E_{P_\pi} \left(W_T e^{-rT} \right) = 0. \tag{65}$$

The set of admissible control is denoted by $\mathcal{U}$.

Lifetime utility The lifetime utility of an investor applying control $u \in \mathcal{U}$ is

$$E_{P_\pi} \left[\int_0^\infty e^{-rt} c_t\, dt \right], \tag{66}$$

Substituting the budget constraint (64) into the investor's objective, one finds that

$$\int_0^T e^{-rt} c_t\, dt + e^{-rT} W_T = W_0 + \int_0^T e^{-rt} \left(D(i_t)\, dt + \sum_{j \in I} Q(i_{t-}, j)\, dN_t(j) \right).$$

Taking expectations on both sides and using the transversality condition (65), one finds that the lifetime utility of an investor is equal to $W_0 + V(\pi)$ where

$$V(\pi) \equiv E_{P_\pi} \left[\int_0^\infty e^{-rt} \left(D(i_t) \, dt + \sum_{j \in I} Q(i_{t-}, j) \, dN_t(j) \right) \right].$$

Because of linear utility, and because the interest rate is equal to the subjective discount rate, an investors is indifferent regarding the timing of his consumption, as long as the associated wealth process satisfies the Transversality Condition. The investor's problem is

$$\sup_{u \in \mathcal{U}} \{W_0 + V(\pi)\}.$$

B.2 Hamilton-Jacobi-Bellman Equation

We find an optimal control by seeking a collection of continuation utilities $J(i)$, $i \in I$, solving the system of Hamilton-Jacobi-Bellman (HJB) equations

$$rJ(i) = \max_{\pi_i \in \Pi(i)} \left\{ D(i) + \sum_{j \in I} \pi_i(j) \left(J(j) - J(i) + Q(i,j) \right) \right\}. \tag{67}$$

Up to some relabelling, this system of HJB is that solved in the text. It is shown that the maximum is achieved by some continuation utility $J^*(i)$ and transition intensity π_i^*.

We now verify that the investor's problem is solved by the control $c_t^* = rW_0 + rJ^*(j_0)$, and $\pi_t^* = \pi_{i_t}^*$. We first show

Proposition 7 (Sufficiency) *The suppremum utility* $\sup_{\pi \in \mathcal{U}} W_0 + V(\pi)$ *is bounded above by* $W_0 + J^*(j_0)$*, and this upper bound is achieved by the admissible control* (c^*, π^*)*.*

Proof The proof goes in two steps. First, we show that $V(\pi) \leq J^*(j_0)$, adapting the proof of Theorem VII,T1 in Brémaud [1981]. Then we show that the associated wealth process satisfies the Transversality Condition. For the first part of the proof, let's consider any admissible control u. We can write

33

$$
\begin{aligned}
J^*(i_T)e^{-rT} &= J^*(j_0) + \sum_{0 < T_n \le T} \left(J^*(j_n)e^{-rT_n} - J^*(j_{n-1})e^{-rT_{n-1}} \right) \\
&\quad + J^*(i_{\tau_T})(e^{-rT} - e^{-r\tau_T}),
\end{aligned}
\tag{68}
$$

where $\tau_T = \sup\{T_n, n \ge 0 : T_n \le T\}$. Equation (68) can be manipulated as follows:

$$
\begin{aligned}
J^*(i_T)e^{-rT} &= J^*(j_0) + \sum_{0 < T_n \le T} J^*(j_{n-1})(e^{-rT_n} - e^{-rT_{n-1}}) + J^*(i_{\tau_T})(e^{-rT} - e^{-r\tau_T}) \\
&\quad + \sum_{0 < T_n \le T} e^{-rT_n}\left(J^*(j_n) - J^*(j_{n-1}) \right) \\
&= J^*(j_0) + \int_0^T \frac{d}{dt}\left(J^*(i_t)e^{-rt} \right) dt + \int_0^T \sum_{j \in I} \left(J^*(j) - J^*(i_{t-}) \right) e^{-rt} \, dN_t(j) \\
&= J^*(j_0) + \int_0^T \left(-rJ^*(i_t) + \sum_{j \in I} \pi_t(j)\left(J^*(j) - J^*(i_{t-}) \right) \right) e^{-rt} \, dt \\
&\quad + \int_0^T \sum_{j \in I} e^{-rt}\left(J^*(j) - J^*(i_{t-}) \right)\left(dN_t(j) - \pi_t(j)\,dt \right).
\end{aligned}
$$

Adding $\int_0^T e^{-rt}D(i_t)\,dt + \int_0^T \sum_{j \in I} Q(i_{t-}, j)\,dN_t(j)$ to both sides gives

$$
\begin{aligned}
&\int_0^T e^{-rt}D(i_t)\,dt + \int_0^T \sum_{j \in I} Q(i_{t-}, j)\,dN_t(j) + J^*(i_T)e^{-rT} = \\
&J^*(j_0) + \int_0^T e^{-rt}\left(-rJ^*(i_t) + D(i_t) + \sum_{j \in I} \pi_t(j)\left[J^*(j) - J^*(i_{t-}) + Q(i_{t-}, j) \right] \right) dt \\
&+ \int_0^T \sum_{j \in I}\left(J^*(j) - J^*(i_{t-}) + Q(i_{t-}, j) \right) e^{-rt}\left(dN_t(j) - \pi_t(j)\,dt \right).
\end{aligned}
\tag{69}
$$

Since $\left(J^*(j) - J^*(i_{t-}) + Q(i_{t-}, j) \right)e^{-rt}$ 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 (69) is a martingale. Taking expectations on both sides, and using the HJB equation (67), one finds

$$
E_{P_\pi}\left(\int_0^T e^{-rt}D(i_t)\,dt + \int_0^T \sum_{j \in I} Q(i_{t-}, j)\,dN_t(j) + J^*(i_T)e^{-rT} \right) \le J^*(j_0),
$$

with equality for $\pi = \pi^*$. Letting t go to infinity proves that $V(\pi) \le J(j_0)$, with equality if $\pi = \pi^*$.

The second step of the proof is to show that the associated wealth process satisfies the Transversality Condition (65). Replacing $c_t^* = rW_0 + rJ^*(j_0)$ into the budget constraint (64), multiplying by e^{-rt}, and integrating from $t = 0$ to $t = T$, we find

$$W_T^* e^{-rT} = W_0 e^{-rT} - J^*(j_0)(1 - e^{-rT}) + \int_0^T e^{-rt} D(i_t) \, dt + \int_0^T \sum_{j \in I} Q(i_{t-}, j) \, dN_t(j).$$

To show that the Transversality Condition is satisfied, one takes expectations $E_{P_{\pi^*}}$ on both sides, lets T goes to infinity, and uses the fact that

$$J^*(j_0) = E_{P_{\pi^*}} \left(\int_0^\infty e^{-rt} D(i_t) \, dt + \int_0^\infty \sum_{j \in I} Q(i_{t-}, j) \, dN_t(j) \right).$$

∎

References

Amihud, Yakov, and Haim Mendelson, Asset Pricing and the Bid-Ask Spread, *Journal of Financial Economics*, 1986, *17*, 223–249.

___ , and ___ , Liquidity, Maturity, and the Yield on U.S. Treasury Securities, *Journal of Finance*, 1991, *46*, 479–486.

Brémaud, Pierre, *Point Processes and Queues*, New-York: Springer-Verlag, 1981.

Brennan, Michael J., and Avanidhar Subrahmanyam, Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns, *Journal of Financial Economics*, 1996, *41*, 441–464.

___ , Tarum J. Chordia, and Avanidhar Subrahmanyam, Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns, *Journal of Financial Economics*, 1998, *49*, 345–373.

Chan, Kalok, Yue-Cheong Chan, and Wai-Ming Fong, Free Float and Market Liquidity: A Study of Hong Kong Government Intervention, *The Journal of Financial Research*, 2004, *27*, 179–197.

Constantinides, George M., Capital Market Equilibrium with Transaction Costs, *Journal of Political Economy*, 1986, *94*, 842–862.

Diamond, Peter A., Aggregate Demand Management in Search Equilibrium, *Journal of Political Economy*, 1982, *90*, 881–894.

Duffie, Darrell, Nicolae Gârleanu, and Lasse H. Pedersen, Over-the-counter Markets,, *Econometrica*, 2005, *Forthcoming*.

Fama, Eugene F., and Kenneth R. French, Common Risk Factors in the Returns on Stock and Bonds, *Journal of Financial Economics*, 1993, *33*, 131–156.

Heaton, John, and Deborah J. Lucas, Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing, *Journal of Political Economy*, 1996, *104*, 443–487.

Huang, Ming, Liquidity Shocks and Equilibrium Liquidity Premia, *Journal of Economic Theory*, 2003, *109*, 104–129.

Kiyotaki, Nobuhiro, and Randall Wright, On Money as a Medium of Exchange, *Journal of Political Economy*, 1989, *97*, 927–954.

Krishnamurthy, Arvind, The Bond/Old-Bond Spread, *Journal of Financial Economics*, 2003, *66*, 463–506.

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

Sharpe, William F., Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk, *Journal of Finance*, 1964, pp. 425–442.

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 Equilibrium Model, *Review of Financial Studies*, 1998, *11*, 1–58.

—— , and Jean-Luc Villa, Equilibrium Interest Rate and Liquidity Premium with Transaction Costs, *Economic Theory*, 1999, *13*, 509–539.

—— , and Tan Wang, "A Theory of On-The-Run and Off-The-Run Bond Markets," Working Paper, MIT 2003.

Wallace, Neil, A Model of the Liquidity Yield Structure Based on Asset Indivisibility, *The Journal of Monetary Economics*, 2000, *45*, 55–68.

Warga, Arthur, Bond Returns, Liquidity, and Missing Data, *Journal of Financial and Quantitative Analysis*, 1992, *27*, 605–617.