

NBER WORKING PAPER SERIES

FINANCIAL-MARKET EQUILIBRIUM WITH FRICTION

Adrian Buss  
Bernard Dumas

Working Paper 19155  
<http://www.nber.org/papers/w19155>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
June 2013

Work on this topic was initiated while Dumas was at the University of Lausanne and Buss was at the University of Frankfurt. Dumas' research has been supported by the Swiss National Center for Competence in Research "FinRisk" and by grant #1112 of the INSEAD research fund. The authors are grateful for useful discussions to Beth Allen, Yakov Amihud, Laurent Barras, Sébastien Bétermier, Bruno Biais, Georgy Chabakauri, Massimiliano Croce, Magnus Dahlquist, Sanjiv Das, Xi Dong, Phil Dybvig, Thierry Foucault, Stefano Giglio, Francisco Gomes, Amit Goyal, Harald Hau, John Heaton, Terrence Hendershott, Julien Hugonnier, Elyès Jouini, Andrew Karolyi, David Lando, John Leahy, Francis Longstaff, Abraham Lioui, Edith Liu, Hong Liu, Frédéric Malherbe, Alexander Michaelides, Pascal Maenhout, Stavros Panageas, Luboš Pástor, Paolo Pasquariello, Patrice Poncet, Tarun Ramadorai, Scott Richard, Barbara Rindi, Leonid Rosu, Olivier Scaillet, Norman Schürhoff, Raman Uppal, Dimitri Vayanos, Grigory Vilkov, Vish Viswanathan, Ingrid Werner, Fernando Zapatero and participants at workshops given at the Amsterdam Duisenberg School of Finance, INSEAD, the CEPR's European Summer Symposium in Financial Markets at Gerzensee, the University of Cyprus, the University of Lausanne, Bocconi University, the European Finance Association meeting, the Duke-UNC Asset pricing workshop, the National Bank of Switzerland, the Adam Smith Asset Pricing workshop at Oxford University, the Yale University General Equilibrium workshop, the Center for Asset Pricing Research/Norwegian Finance Initiative Workshop at BI, the Indian School of Business summer camp, Boston University, Washington University in St Louis, ESSEC Business School, HEC Business School and McGill University. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Adrian Buss and Bernard Dumas. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Financial-market Equilibrium with Friction  
Adrian Buss and Bernard Dumas  
NBER Working Paper No. 19155  
June 2013  
JEL No. G10,G12

**ABSTRACT**

We show that the endogenous stochastic process of the liquidity of securities is as important to investment and valuation as the exogenous stochastic process of their cash flows.

We develop a general-equilibrium model with heterogeneous investors who have an every-day motive to trade and pay transactions fees.

Our model delivers the optimal, market-clearing moves of each investor and the resulting posted and transactions prices. We exhibit the effect of transactions fees on deviations from the consumption CAPM. We compare expected returns on stocks carrying different fees and evaluate the ability of various empirical liquidity measures to act as pricing proxies.

Adrian Buss  
INSEAD  
boulevard de Constance  
77305 Fontainebleau Cedex  
FRANCE  
adrian.buss@insead.edu

Bernard Dumas  
INSEAD  
boulevard de Constance  
77305 Fontainebleau Cedex  
FRANCE  
and NBER  
bernard.dumas@insead.edu

# I Introduction

Following the lead of He and Modest (1995) and Luttmer (1996), we incorporate trading fees in an equilibrium model in which investors optimally and endogenously decide when and how much to trade. Unlike these authors, who established bounds on asset prices, we reach a full characterization.

In the real world, investors do not trade with each other. They trade through intermediaries called brokers and dealers, who incur physical costs, are faced with potentially informed customers and charge a fee that, to an approximation, is proportional to the value of the shares traded. This service charge aims to cover the actual physical cost of trading and the adverse-selection effect plus a profit. Ultimately, however, the end users being the investors, access to a financial market is a service that investors make available to each other. As a way of providing a simple model, we bypass intermediaries and the pricing policy of broker-dealers, and let the investors serve as dealers for, and pay the fees to each other. We just assume that the trading fee is proportional to the value of the shares traded.

We endow investors with an every-day motive for trading, over and above the long-term need to trade for lifetime planning purposes.<sup>1</sup> We assume that investors are long-lived and trade because they have differing risk aversions while they have access only to a menu of linear assets. Whether or not the market without friction would be complete, dynamic completeness is, of course, destroyed by the presence of transactions fees. The imbalance of the portfolios, which investors have to hold because of transactions fees, acts as an inventory cost.

Our goal is to study, in terms both of price and volume, the dynamics of the equilibrium that we can expect to prevail in such a market. When purchasing a security an investor needs not only have in mind the cash flows that the security will pay into the indefinite future, she must also anticipate her desire and ability to resell the security in the marketplace at a later point in time. Given the presence of the fee, an investor may decide not to trade, thereby preventing other investors from trading with her, which is an additional endogenous, stochastic and perhaps quantitatively more important consequence of the fee. We show that the endogenous stochastic process of the liquidity of securities is as important to investment and valuation as is the exogenous stochastic process of their future cash flows.

In the dynamics of our equilibrium, the inventory of securities held by each investor can be viewed as a state variable, a feature that is shared with the inventory-management model of a dealer that has been pioneered by Ho and Stoll (1980, 1983) and which is one of the main pillars of the Microstructure literature. In their work, however, Ho and Stoll focus exclusively on the dealer's problem, taking the arrival of orders to the dealer as an exogenous random process. Here, we fully endogenize each investor's decision to trade and derive the full general equilibrium.<sup>2</sup> Orders do not arrive at random; they implement

---

<sup>1</sup>We discuss trading motives further in Subsection III.A.

<sup>2</sup>Recently, a partial-equilibrium literature has developed, aiming to model the optimal

optimal portfolio adjustments.

Our paper is related to the existing studies of portfolio choice under transactions costs such as Magill and Constantinides (1976), Constantinides (1976a, 1976b, 1986), Davis and Norman (1990), Dumas and Luciano (1991), Edirisinghe, Naik and Uppal (1993), Gennotte and Jung (1994), Shreve and Soner (1994), Cvitanic and Karatzas (1996), Leland (2000), Longstaff (2001), Nazareth (2002), Bouchard (2002), Obizhaeva and Wang (2005), Liu and Lowenstein (2002), Jang *et al.* (2007), Gerhold *et al.* (2011) and Gârleanu and Pedersen (2012) among others. As was noted by Dumas and Luciano, many of these papers suffer from a logical quasi-inconsistency. Not only do they assume an exogenous process for securities returns, as do all portfolio optimization papers, but they do so in a way that is incompatible with the portfolio policy that is produced by the optimization. When transactions costs are linear, the portfolio strategy is of a type that recognizes the existence of a “no-trade” region. Yet, portfolio-choice papers assume that prices continue to be quoted and trades remain available in the marketplace.<sup>3</sup> Obviously, the assumption must be made that some investors, other than the one whose portfolio is being optimized, do not incur costs. In the present paper, we assume that all investors face the trading fee.

The papers of Heaton and Lucas (1996), Vayanos (1998), Vayanos and Vila (1999) and Lo *et al.* (2004) are direct ancestors of the present one in that they have exhibited the equilibrium behavior resulting from a physical, dead-weight cost of transacting.<sup>4</sup> In the neighborhood in which transactions take place, Heaton and Lucas (1996) derive a stationary equilibrium under transactions cost but, in the neighborhood of zero trade, the cost is assumed to be quadratic so that investors trade all the time in small quantities and equilibrium behavior is qualitatively different from the one we produce here. In Vayanos (1998) and Vayanos and Vila (1999), an investor’s only motive to trade is the fact that she has a finite lifetime. Transactions costs induce him to trade twice in her life: when young, she buys some securities that she can resell in order to be able to live during her old age. Here, we introduce a higher-frequency motive to trade. In the paper of Lo *et al.* (2004), costs of trading are fixed costs, all investors have the same negative exponential utility function, individual investors’

---

tactic of an investor who (for unmodelled reasons) needs to trade and determines how to optimally place his orders in a limit-order market. See: Parlour (1998), Foucault (1999), Foucault *et al.* (2005), Goettler *et al.* (2005), Rosu (2009).

<sup>3</sup>Constantinides (1986) in his pioneering paper on portfolio choice under transactions costs attempted to draw some conclusions concerning equilibrium. Assuming that returns were independently, identically distributed (IID) over time, he claimed that the expected return required by an investor to hold a security was affected very little by transactions costs. Liu and Lowenstein (2002), Jang *et al.* (2007) and Delgado *et al.* (2010) have shown that this is generally not true under non IID returns. The possibility of falling in a “no-trade” region is obviously a massive violation of the IID assumption.

<sup>4</sup>In a previous version of our paper, we had assumed that trading entailed physical dead-weight costs proportional to the number of shares traded. Another predecessor is Milne and Neave (2003), which, however, contains few quantitative results. The equilibrium with other costs, such as holding costs and participation costs, has been investigated by Peress (2005), Tuckman and Vila (2010) and Huang and Wang (2010).

endowments provide the motive to trade but the amount of aggregate physical resources available is not stochastic. In our current paper, fees are proportional, preferences are of the power-utility type with heterogeneous risk aversions and aggregate resources are free to follow an arbitrary stochastic process. To our knowledge, ours is the first paper to reach that goal. A form of restricted trading is considered by Longstaff (2009) where a physical asset traded by two logarithmic investors is considered illiquid if, after being bought at time 0, it must be held till some date  $T$  after which it becomes liquid again. The consequences for equilibrium asset prices are drawn in relation to the length  $T$  of the freeze. Here, the trading dates are chosen endogenously by the investors.

Transactions fees are not the only reason for which liquidity and liquidity-risk considerations arise in a financial market. At any given time, an asset is more or less liquid as a function of three conceivable mechanisms and their fluctuating impact, taken in isolation or combined. The first mechanism is the fear of default of the counterparty to the trade. Trade is obviously hampered by the fear that contracts will not be abided by. That is the line of argument behind the concept of “funding liquidity.” The second mechanism is informed trading (asymmetric information) as in the market for “lemons” (Akerlof (1970)).<sup>5</sup> A vast Microstructure literature stemming from Copeland and Galai (1983), Glosten and Milgrom (1985) and Kyle (1985) has shown that informed trading indirectly generates transactions costs.<sup>6</sup> The third mechanism, which we examine here, is the presence of fees charged for transacting, stemming (in an unmodelled way) from order processing costs and inventory holding costs and holding risks, all items which Stoll (2000), as did Demsetz (1968), refers to as “real frictions.”<sup>7</sup> The concept of “market liquidity” captures these last two mechanisms. One can also introduce liquidity considerations in the form of a portfolio constraint. Holmström and Tirole (2001) study a financial-market equilibrium in which investors face an exogenous constraint on borrowing.<sup>8</sup> When they hit their constraint, investors are said to be “liquidity constrained.” Gromb and Vayanos (2002) and Brunnermeier and Pedersen (2008) study situations in which the amount of arbitrage capital is constrained.<sup>9</sup> It would be necessary

<sup>5</sup>Bhattacharya and Spiegel (1998) have shown the way in which the lemon problem can cause markets to close down.

<sup>6</sup>And, in Asset Pricing, an even larger literature stemming from Grossman and Stiglitz (1980) and Hellwig (1980) shows how the risk created by asymmetric information or heterogeneous expectations is priced. See, among many, Kyle (1985), Easley and O’Hara (1987), Admati and Pfleiderer (1988), Easley *et al.* (2002) and O’Hara (2003).

<sup>7</sup>On the various possible determinants of liquidity, see the synthesis paper of Vayanos and Wang (2009). The classic empirical decomposition of stock-market spreads by Stoll (1989) concluded that real frictions represented 53% of the bid-ask spread.

<sup>8</sup>As is apparent below, the cost and the constraint approaches are somewhat similar but are probably not equivalent to each other. As we show, transactions costs or fees give rise to shadow prices of potentially being unable to trade that are specific to each asset and each investor, whereas a constraint gives rise to a dual variable that is specific to each investor only.

<sup>9</sup>Distant antecedents of this idea in the macroeconomic literature can be found in the form of Clower and Bushaw (1954) constraints, which required a household to hold some money balance, as opposed to being able to borrow, when it wanted to consume, as well as the “cash-in-advance” model of Lucas (1982).

to present some microfoundations for the constraint. A constraint on borrowing would best be justified by the risk of default on the loan.<sup>10</sup> Equilibrium with default is an important but separate topic of research.

As far as the solution method is concerned, our analysis is closely related, in ways we explain below, to “the dual method” used by Jouini and Kallal (1995), Cvitanic and Karatzas (1996), Cuoco (1997), Kallsen and Muhle-Karbe (2008) and Deelstra, Pham and Touzi (2002) among others. Furthermore, in computing an equilibrium, one has a choice between a “recursive” method, which solves by backward induction over time, and a “global” method, which solves for all optimality conditions and market-clearing conditions of all states of nature and points in time simultaneously.<sup>11</sup> The global method, often implemented in the form of a homotopy, is limited in terms of the number of periods it can handle. Here, we resort to a recursive technique, which requires the choice of state variables – both exogenous and endogenous – that track the state of the economy. Dumas and Lyasoff (2012) have proposed an efficient method to calculate incomplete-market equilibria recursively with a dual approach, which utilizes state prices as endogenous state variables. We use the same method here with the addition of dual state variables that capture the cost of trading. A crucial advantage of using dual variables as state variables to handle proportional-transactions costs problems is that the variables thus introduced evolve on a fixed domain, namely the interval set by the unit cost of buying and the cost of selling (with opposite signs), whereas primal variables, such as portfolio choices evolve over a domain that has free-floating barriers, to be determined.

Both our paper and Buss *et al.* (2013) derive an equilibrium in a financial market where investors incur a cost when they transact and both use the backward-induction procedure of Dumas and Lyasoff (2012) to solve the model. Technically, the main difference between the two papers is that Buss *et al.* (2013) use a “primal” formulation and we use a “dual” one. In the dual approach, the personal state prices of the investors are among the unknowns and the same system of equations applies in the entire space of values of state variables while the shadow costs of trading are included among the state variables. The primal approach requires the addition of the previous period’s portfolios among the state variables and solves a different system of equations in different regions of the state space, thus introducing some combinatorics, which the dual approach avoids. The economic insights generated by the two papers are also quite different. In our paper, the focus is on the effects of transactions fees and the pricing of liquidity risk. In Buss *et al.* (2013), the transactions cost is a deadweight cost and the focus is on its effect on the cross section of asset returns when there is more than one risky asset and idiosyncratic labor income that is separate from the output process. Finally, the investors in Buss *et al.* (2013) have Epstein-Zin-Weil utility rather than power utility.

Empirical work on equilibria with transactions costs has been couched in

---

<sup>10</sup>In a fascinating empirical rendition of the same idea, Adrian et al. (2012) have estimated a “financial-intermediary stochastic discount factor” and measured, in a CAPM form, the impact on asset prices of the risk of intermediaries becoming constrained.

<sup>11</sup>For an implementation of the global solution, see Herings and Schmedders (2006).

terms of a CAPM that recognizes a number of risk factors. Brennan and Subrahmanyam (1996), Pástor and Stambaugh (2003) and Acharya and Pedersen (2005) have recognized two or more risk factors, one of which is the market return (as in the classic CAPM) or aggregate consumption (as in the consumption-CAPM), and the others are meant to capture stochastic fluctuations in the degree of liquidity of the market, either taken as a whole or individually for each security. Liquidity fluctuations are taken as exogenous and are proxied by fluctuations in volume or in the responsiveness of price to the order flow. The papers cited confirm that there exist in the marketplace significant risk premia related to these factors. Our model also identifies additional risk factors for the investors' willingness to trade, in the form of shadow prices, which are fully endogenized. However, there is one such per investor and they are not directly observable. We use our model to ascertain to what extent proxies used in the empirical literature are to any degree related to these shadow prices.<sup>12 13</sup>

After writing down our model, specifying the solution method (Section II) and illustrating the dynamics of the economy (Section III), we focus our work on two main questions. First, in Section IV we examine the investors' portfolio strategy and ask whether equilibrium securities prices conform to the famous dictum of Amihud and Mendelson (1986a), which says that they are reduced by the present value of transactions costs. In Section V, we examine the behavior of the market over time, asking, for instance, to what degree price changes and transactions volume are related to each other and what effects transactions fees have on the point process of transaction prices. In Section VI, we quantify the additional premia that are created by transactions fees and which are deviations from the consumption-CAPM. These are the drags on expected-return that empiricists would encounter as a result of the presence of transactions fees.

## II Problem statement: the objective of each investor and the definition of equilibrium

We start with a population of two investors  $l = 1, 2$  and a set of exogenous time sequences of individual endowments  $\{e_{l,t} \in \mathbb{R}_{++}; l = 1, 2; t = 0, \dots, T\}$  on a tree or lattice. A given node at time  $t$  is followed by  $K_t$  nodes at time  $t + 1$  at which the endowments are denoted  $\{e_{l,t+1,j}\}_{j=1}^{K_t}$ . The transition probabilities

---

<sup>12</sup>Transactions costs also constitute a “limit to arbitrage” and offer a potential explanation of the observed fact that sometimes securities that are closely related to each other do not trade in the proper price relationship. For these deviations to appear in the first place, however, and subsequently not be obliterated by arbitrage, some category of investors must introduce some form of “demand shock”, that can only result from some departure from von Neumann-Morgenstern utility. Here, we consider only rational behavior so that no opportunities for (costly) arbitrage arise in equilibrium.

<sup>13</sup>Empirical work has also been done by Chordia *et al.* (2008) and others to track the dynamics of liquidity as it moves from one category of assets to another. In the present paper, the menu of assets is too limited to throw any light on the evidence presented by these papers.

are denoted  $\pi_{t,t+1,j}$  ( $\sum_{j=1}^{K_t} \pi_{t,t+1,j} = 1$ ).<sup>14</sup> Notice that the tree accommodates the exogenous state variables only.<sup>15</sup> <sup>16</sup>

In the financial market, there are  $I$  securities, defined by their payoffs  $\{\delta_{t,i}; i = 1, \dots, I; t = 0, \dots, T\}$ .<sup>17</sup> The “posted” prices of the securities, which are not always transactions prices, are denoted:  $\{S_{t,i}; i = 1, \dots, I; t = 0, \dots, T\}$ . The posted price is an effective transaction price if and when a transaction takes place but it is posted all the time by the Walrasian auctioneering computer (which works at no cost). Below, we explain that the real-world interpretation of that posted price is the bid-ask midpoint.

Financial-market transactions entail transactions fees. The fees are calculated on the basis of the transaction price. When an investor sells one unit of security  $i$ , turning it into consumption good, she receives in units of consumption goods the transaction price multiplied by  $1 - \varepsilon_{i,t}$  and, when she buys, she must pay the transaction price times  $1 + \lambda_{i,t}$ . All transactions are cleared and fees paid to a central pot. The fees received by an investor take the form of a transfer, which is taken by her to be a given amount, which enters her budget constraint but plays no role in her first-order conditions. When the buy and sell fees are equal, as is the case in our numerical illustrations below, this assumption is equivalent to investors being compensated (in the Hicksian sense) for the fees they incur on their transactions. As a result, transactions fees generate no income/wealth effect, only substitution effects.

With symbol  $\theta_{l,t,i}$  standing for the number of units of Security  $i$  in the hands of Investor  $l$  *after* all transactions of time  $t$ , Investor  $l$  solves the following problem:

$$\sup_{\{c_l, \theta_l\}} \mathbb{E}_0 \sum_{t=0}^T u_l(c_l, t)$$

subject to:

- terminal conditions:

$$\theta_{l,T,i} = 0,$$

---

<sup>14</sup>Transition probabilities generally depend on the current state but we suppress that subscript.

<sup>15</sup>As has been noted by Dumas and Lyasoff (2012), because the tree only involves the exogenous endowments, it can be chosen to be *recombining* when the endowments are Markovian, which is a great practical advantage compared to the global-solution approach, which would require a tree in which nodes must be distinguished on the basis of the values of not just the exogenous variables but also the endogenous ones.

<sup>16</sup>It is straightforward to write the equations below for more investors and more complex trees. The implementation of the solution technique is much more computationally intensive with more than two investors while it is not more complicated with a richer tree.

<sup>17</sup>It so happens that, without transactions fees, the market would be dynamically complete. But the derivations and the solution technique depend neither on the number of branches in the tree, nor on the number of securities. We could solve for the equilibrium with transactions fees in a market that would be incomplete to start with.



- a sequence of flow budget constraints:

$$\begin{aligned}
c_{l,t} + \sum_{i=1}^I \max [0, \theta_{l,t,i} - \theta_{l,t-1,i}] \times S_{t,i} \times (1 + \lambda_{i,t}) \\
+ \sum_{i=1}^I \min [0, \theta_{l,t,i} - \theta_{l,t-1,i}] \times S_{t,i} \times (1 - \varepsilon_{i,t}) \\
= e_{l,t} + \sum_{i=1}^I \theta_{l,t-1,i} \delta_{t,i} + \zeta_{l,t}; \forall t
\end{aligned} \tag{1}$$

- and given initial holdings:<sup>18</sup>

$$\theta_{l,-1,i} = \bar{\theta}_{l,i} \quad . \tag{2}$$

In the flow budget constraint, the term  $\sum_{i=1}^I \max [0, \theta_{l,t,i} - \theta_{l,t-1,i}] \times S_{t,i} \times (1 + \lambda_{i,t})$  reflects the net cost of purchases and the term  $\sum_{i=1}^I \min [0, \theta_{l,t,i} - \theta_{l,t-1,i}] \times S_{t,i} \times (1 - \varepsilon_{i,t})$  captures the net cost of sales of securities (i.e., proceeds of sales with a negative sign). And the term  $\zeta_{l,t}$  on the right-hand side stands for the transfer received from the central pot.

The dynamic programming formulation of the investor's problem is:<sup>19</sup>

$$J_l(\{\theta_{l,t-1,i}\}, \cdot, e_{l,t}, t) = \sup_{c_{l,t}, \{\theta_{l,t,i}\}} u_l(c_{l,t}, t) + \mathbb{E}_t J_l(\{\theta_{l,t,i}\}, \cdot, e_{l,t+1}, t+1)$$

subject to the flow budget constraint (1) written at time  $t$  only.

Writing:

$$\widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \triangleq \max [0, \theta_{l,t,i} - \theta_{l,t-1,i}]$$

for purchases of securities and:

$$\widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \triangleq \min [0, \theta_{l,t,i} - \theta_{l,t-1,i}]$$

(a negative number) for sales, so that  $\theta_{l,t,i} = \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i}$ , one can reformulate the same problem to make it more suitable for mathematical programming:

$$\begin{aligned}
J_l(\{\theta_{l,t-1,i}\}, \cdot, e_{l,t}, t) = \sup_{c_{l,t}, \{\widehat{\theta}_{l,t,i}, \widehat{\theta}_{l,t,i}\}} u_l(c_{l,t}, t) \\
+ \mathbb{E}_t J_l\left(\left\{\widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i}\right\}, \cdot, e_{l,t+1}, t+1\right)
\end{aligned} \tag{3}$$

<sup>18</sup>It is assumed that  $\sum_{i=1}^I \bar{\theta}_{l,i} = 0$  or 1 depending on whether the security is assumed to be in zero or positive net supply.

<sup>19</sup>The form  $J_l(\{\theta_{l,t-1,i}\}, \cdot, e_{l,t}, t)$  in which the value function is written refers explicitly only to investor  $l$ 's individual state variables. The complete set of state variables actually used in the backward induction is chosen below.

subject to:

$$\begin{aligned} c_{l,t} + \sum_{i=1}^I \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) S_{t,i} (1 + \lambda_{i,t}) + \sum_{i=1}^I \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) S_{t,i} (1 - \varepsilon_{i,t}) \\ = e_{l,t} + \sum_{i=1}^I \theta_{l,t-1,i} \delta_{t,i} + \zeta_{l,t} \end{aligned} \quad (4)$$

$$\widehat{\theta}_{l,t,i} \leq \theta_{l,t-1,i} \leq \widehat{\theta}_{l,t,i} \quad (5)$$

**Definition 1** An equilibrium is defined as a process for the allocation of consumption  $c_{l,t}$ , a process for securities prices  $\{S_{t,i}\}$ , a set of transfer payments  $\zeta_{l,t}$  and a process for portfolio choices  $\left\{ \widehat{\theta}_{l,t,i}, \widehat{\theta}_{l,t,i} \right\}$  of both investors such that the supremum of (3) is reached for all  $l, i$  and  $t$  and the market-clearing conditions:<sup>20</sup>

$$\sum_{l=1,2} \theta_{l,t,i} = 0 \text{ or } 1; i = 1, \dots, I \quad (6)$$

as well as the central-pot balance conditions for fees:

$$\zeta_{l,t} = \sum_{i=1}^I \left( \widehat{\theta}_{l',t,i} - \theta_{l',t-1,i} \right) S_{t,i} \lambda_{i,t} - \sum_{i=1}^I \left( \widehat{\theta}_{l',t,i} - \theta_{l',t-1,i} \right) S_{t,i} \varepsilon_{i,t}; l' \neq l \quad (7)$$

are also satisfied with probability 1 at all times  $t = 0, \dots, T$ .

In Appendix A, we show, using a shift of equations proposed in the context of incomplete markets by Dumas and Lyasoff (2012), that the equilibrium can be calculated by means of a single backward-induction procedure, for given initial values of some endogenous state variables, which are the dual variables  $\{\phi_{l,t}, R_{l,t,i}\}$  – as opposed to given values of the original state variables, viz., initial positions  $\{\theta_{l,t-1,i}\}$  –, by solving the following equation system written for  $l = 1, 2; j = 1, \dots, K_t; i = 1, \dots, I$ . The shift of equations amounts from the computational standpoint to letting investors at time  $t$  plan their time- $t + 1$  consumption  $c_{l,t+1,j}$  but choose their time- $t$  portfolio  $\theta_{l,t,i}$  (which will finance the time- $t + 1$  consumption).

1. First-order conditions for time  $t + 1$  consumption:<sup>21</sup>

$$u'_l(c_{l,t+1,j}, t + 1) = \phi_{l,t+1,j}$$

<sup>20</sup>One equates  $\sum_{l=1}^2 \theta_{l,t,i}$  to 0 or 1 depending on whether the security is or is not in zero net supply.

<sup>21</sup> $u'_l$  denotes “marginal utility” or the derivative of utility with respect to consumption.

2. The set of time- $t+1$  flow budget constraints for all investors and all states of nature of that time:

$$\begin{aligned}
& c_{l,t+1,j} + \sum_{i=1}^I (\theta_{l,t+1,i,j} - \theta_{l,t,i}) S_{t+1,i,j} R_{l,t+1,i,j} \\
= & e_{l,t+1,j} + \sum_{i=1}^I \theta_{l,t,i} \delta_{t+1,i,j} + \sum_{i=1}^I (\widehat{\theta}_{l,t+1,i,j} - \theta_{l,t,i}) S_{t+1,i,j} \lambda_{i,t+1,j} \\
& - \sum_{i=1}^I (\widehat{\theta}_{l,t+1,i,j} - \theta_{l,t,i}) S_{t+1,i,j} \varepsilon_{i,t+1,j}
\end{aligned}$$

3. The third subset of equations says that, when they trade, all investors must agree on the prices of traded securities and, more generally, they must agree on the “posted prices” inclusive of the shadow prices  $R$  that make units of paper securities more or less valuable than units of consumption. Because these equations, which, for given values of  $R_{l,t+1,i,j}$ , are linear in the unknown state prices  $\phi_{l,t+1,j}$ , restrict these to lie in a subspace, we call them the “kernel conditions:”

$$\begin{aligned}
& \frac{1}{R_{1,t,i} \times \phi_{1,t}} \sum_{j=1}^{K_t} \pi_{t,t+1,j} \times \phi_{1,t+1,j} \times (\delta_{t+1,i,j} + R_{1,t+1,i,j} \times S_{t+1,i,j}) \quad (8) \\
= & \frac{1}{R_{2,t,i} \times \phi_{2,t}} \sum_{j=1}^{K_t} \pi_{t,t+1,j} \times \phi_{2,t+1,j} \times (\delta_{t+1,i,j} + R_{2,t+1,i,j} \times S_{t+1,i,j})
\end{aligned}$$

4. Definitions:

$$\theta_{l,t+1,i,j} = \widehat{\theta}_{l,t+1,i,j} + \widehat{\theta}_{l,t+1,i,j} - \theta_{l,t,i}$$

5. Complementary-slackness conditions:

$$\begin{aligned}
(-R_{l,t+1,i,j} + 1 + \lambda_{i,t+1,j}) \times (\widehat{\theta}_{l,t+1,i,j} - \theta_{l,t,i}) &= 0 \\
(R_{l,t+1,i,j} - (1 - \varepsilon_{i,t+1,j})) \times (\theta_{l,t,i} - \widehat{\theta}_{l,t+1,i,j}) &= 0
\end{aligned}$$

6. Market-clearing restrictions:

$$\sum_{l=1,2} \theta_{l,t,i} = 0 \text{ or } 1$$

7. Inequalities:

$$\widehat{\theta}_{l,t+1,i,j} \leq \theta_{l,t,i} \leq \widehat{\theta}_{l,t+1,i,j}; 1 - \varepsilon_{i,t+1,j} \leq R_{l,t+1,i,j} \leq 1 + \lambda_{i,t+1,j};$$

This is a system of  $2K_j + 2K_j + I + 2K_jI + 2K_jI + 2K_jI + I$  equations (not counting the inequalities) with  $2K_j + 2K_j + 2K_jI + 2I + 2K_jI + 2K_jI$  unknowns

$$\left\{ c_{l,t+1,j}, \phi_{l,t+1,j}, R_{l,t+1,i,j}, \theta_{l,t,i}, \widehat{\theta}_{l,t+1,i,j}, \widehat{\theta}_{l,t+1,i,j}; l = 1, 2; j = 1, \dots, K_j \right\}^{22}$$

We solve the system by means of the Interior-Point algorithm, in a simplified version of the implementation of Armand *et al.* (2008).<sup>23</sup>

Besides the exogenous endowments  $e_{l,t+1,j}$  and payouts  $\delta_{t+1,i,j}$ , the “givens” are the time- $t$  investor-specific shadow prices of consumption  $\{\phi_{l,t}; l = 1, 2\}$  and of paper securities  $\{R_{l,t,i}; l = 1, 2; i = 1, \dots, I\}$ , which must henceforth be treated as state variables and which we refer to as “endogenous state variables.” Actually, given the nature of the equations, the latter variables can be reduced to state variables:  $\frac{R_{2,t,i}}{R_{1,t,i}}$  and  $\frac{\phi_{1,t}}{\phi_{1,t} + \phi_{2,t}}$  all of which are naturally bounded a priori:  $\frac{1 - \varepsilon_{i,t}}{1 + \lambda_{i,t}} \leq \frac{R_{2,t,i}}{R_{1,t,i}} \leq \frac{1 + \lambda_{i,t}}{1 - \varepsilon_{i,t}}$  and  $0 \leq \frac{\phi_{1,t}}{\phi_{1,t} + \phi_{2,t}} \leq 1$ .<sup>24</sup>

In addition, the given securities’ price functions  $S_{t+1,i,j}$  are obtained by backward induction (see, in Appendix A, the third equation in System (18)):

$$\begin{aligned} S_{t,i} &= \frac{1}{R_{l,t,i}\phi_{l,t}} \sum_{j=1}^{K_t} \pi_{t,t+1,j} \phi_{l,t+1,j} \times (\delta_{t+1,i,j} + R_{l,t+1,i,j} \times S_{t+1,i,j}); \\ S_{T,i} &= 0 \end{aligned} \tag{9}$$

and the given future position functions  $\theta_{l,t+1,i,j}$  (satisfying  $\sum_{l=1,2} \theta_{l,t+1,i,j} = 0$  or  $1; i = 1, \dots, I$ ) are also obtained by an obvious backward induction of  $\theta_{l,t,i}$ , the previous solution of the above system, with terminal conditions  $\theta_{l,T,i} = 0$ . All the functions carried backward are interpolated by means of quadratic interpolation based on the modified Shepard method.

Moving back through time till  $t = 0$ , the last portfolio holdings we calculate are  $\theta_{l,0,i}$ . These are the post-trade portfolios held by the investors as they *exit* time 0. We need to translate these into *entering*, or pre-trade, portfolios holdings so that we can meet the initial conditions (2). The way to do that is explained in Appendix B.

We demonstrate a property of scale invariance, which will save on the total amount of computation: all the nodes of a given point in time, which differ only

<sup>22</sup>The size of the system is reduced when some securities do not carry transactions fees.

<sup>23</sup>The Interior-Point method, which involves relaxed Karush-Kuhn-Tucker complementary-slackness conditions, turns inequality constraints into equations. It is more compatible with Newton solvers than the alternative method proposed earlier by Garcia and Zangwill (1981), which involves discontinuous functions such as  $\max[\cdot, \cdot]$ .

<sup>24</sup>The two variables  $\phi_{1,t}$  and  $\phi_{2,t}$  are one-to-one related to the consumption shares of the two investors, so that consumption scales are actually used as state variables. Consumption shares of the two investors add up to 1 because the transactions fees are paid in a reciprocal fashion.

by their value of the exogenous variable, are isomorphic to each other, where the isomorphy simply means that we can factor out the total endowment. In this way, we do not need to perform a new calculation for each node of a given point in time; one suffices. This property, which we prove in Appendix C, holds even though investors have different risk aversions.<sup>25</sup>

### III The dynamics of the economic system

#### A The motives to trade

In the setup that we use for illustration, we consider two investors who have isoelastic utility and have different coefficients of relative risk aversion. As for securities, the subscript  $i = 1$  refers to a short-lived riskless security in zero net supply and the subscript  $i = 2$  refers to “equity” *in positive supply*. We call “equity” a long-lived claim that pays out  $\delta$ . Although the method we just described allows for any stochastic process of payouts and endowments, in our illustration – to be described further in Subsection B–, we assume zero endowments so that the single exogenous process is the payout process, which is represented by a binomial tree, with constant geometric increments mimicking a geometric Brownian motion.

As has been noted in the introduction, investors in our model trade because they have differing risk aversions while they have access only to a menu of linear assets. Frictions will impede that motive somewhat, with the consequence that the improperly balanced portfolios investors have to hold because of transactions fees act as an inventory cost similar to the cost incurred in inventory-management model of the Ho-and-Stoll (1980, 1983) variety. It is clearest to demonstrate the motive in the absence of frictions, where it is given full swing.

Figure 1 displays the equilibrium wealth-sharing rule in the absence of friction. The wealth sharing rule is the function indicating the wealth of each investor as a function of total wealth. Here, total wealth is equal to the stock price. For any given value of total wealth, the slopes of the curves indicate the number of units held in the equity and the intercepts of the curves indicate the amount lent or borrowed. The risk aversion of Investor 1 is set lower than that of Investor 2, so that Investor 1 is a natural borrower, as far as the riskless short-term security is concerned. That is reflected in the intercepts of the straight lines. They both hold the equity long.

The difference in risk aversions causes the wealth function of Investor 1 to be convex and that of Investor 2 to be concave. Consider now a move in the equity price from  $S_{t-1,2}$  to  $S_{t,2}$ . At time  $t$ , Investor 1 borrows additional money at the riskless rate and increases her holding of the equity, while Investor 2 does the opposite. Now, an up move in the equity price only takes place upon an up move

---

<sup>25</sup> Remarkably, the property is valid when  $\frac{R_{2,t,i}}{R_{1,t,i}}$  and  $\frac{\phi_{1,t}}{\phi_{1,t} + \phi_{2,t}}$  are used as endogenous state variables of the backward recursion. With different risk aversions across investors, *it would not have held* if, as in the primal approach, the endogenous state variables had been  $\{\theta_{1,t-1,i}\}$ , the pre-trade portfolios held when entering each point in time  $t$ .

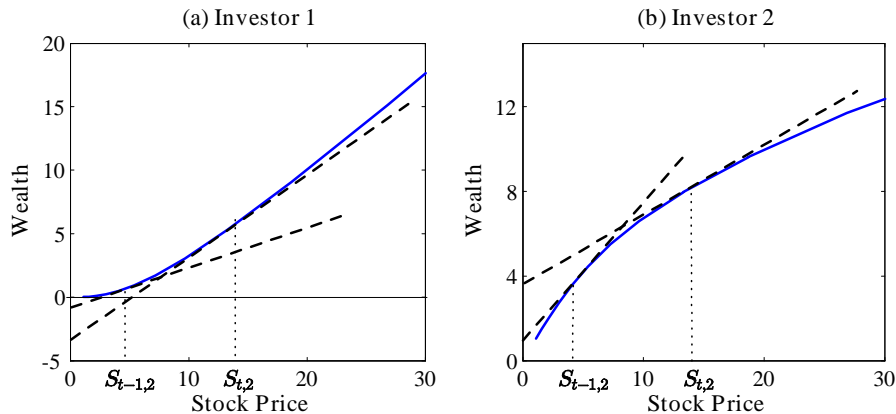


Figure 1: **Wealth sharing rule:** The wealth of each investor as a function of the value of the stock market, which is also total wealth. All parameters are set at their benchmark values indicated in Table 1.

in the payouts’ binomial tree. Therefore, the trading principle is straightforward: the less risk averse investor buys (and borrows) on an up move.

In the presence of frictions, the “liquidity motive” is a second motive to trade. In configurations in which investors receive and anticipate to receive endowments, they trade or hedge a great deal of their amounts at time 0. Thereafter, they trade again when they actually receive the endowment when its amount is above or below the amount they have been able to hedge. The second motive is weaker than the first one. We remove it from the analysis in this paper by assuming zero ongoing endowments of goods.

We leave for later work two other motives for trading that are obviously present in the real world such as the enhanced liquidity-trading motive arising from missing securities and the speculative motive arising from informed trading, private signals or differences of opinion.

## B Calibration

In order to capture some properties of real-world equity, we choose a process for the total payout that reflects the actual behavior of payouts.

Payouts include dividends plus share repurchases minus share issues. The following set of papers document dividend dynamics. Campbell and Shiller (1988) report for periods up to 1986 dividend growth rates of around 4%. Brennan and Xia (2001) calibrated a process for dividends and consumption to match the moments of the joint distribution of these two variables obtained from the Shiller (1989) data. They settled on an average growth rate of dividends and consumption equal to 1.55%/year and 1.69% respectively with standard deviations equal to 12.9%/year and 3.44%/year respectively. Lettau and Ludvigson (2005)

write: “An inspection of the dividend data from the CRSP value-weighted index [...] reveals that [...] the average annual growth rate of dividends has not declined precipitously over the period since 1978, or over the full sample. The average annual growth rate of real, per capita dividends is in fact higher, 5.6%, from 1978 through 1999, than the growth rate for the period 1948 to 1978. The annual growth rate for the whole sample (1948-2001) is 4.2%.” Volatility is reported to be 12.24%. Recently, van Binsbergen and Koijen (2010) estimate a growth rate of 5.89%.

Larrain and Yogo (2008) have estimated net payout dynamics for listed securities in the United States from flow-of-funds data 1926-2004. They found a mean payout growth of 3.83%/year with a standard deviation of 38.37%/year.

Going one step further, Longstaff and Piazzesi (2004) followed by Longstaff (2009) argue that, rather than use dividends or even dividend payouts themselves, which are smoothed over time by corporations, “measures tied to corporate earnings are likely to provide better information about the actual cash flows generated by these firms.” To calibrate the payout process for corporate equity, they use “the real per capita growth rates for corporate profits after tax using annual NIPA data for the 1929-2005 period.” The growth rate is found to be 4.40%/year with a volatility of 21.60%/year. *We adopt this last specification, equating consumption to payout.*

The last parameter of the calibration is the rate of transactions fees. We let transactions fees be levied on trades of equity shares (the “less liquid” asset); none are levied on trades of the riskless asset, which is also, therefore, the “more liquid” asset. In the last section, we introduce a second risky asset so that we can vary the relative degree of liquidity of each of the two risky assets.

While the numbers have been chosen to be realistic, the numerical illustration below cannot easily be seen as being calibrated to a real-world economy since we have two investors, not millions. For these reasons, although our goal is to capture a higher-frequency motive to trade, the amount of trading we are able to generate is not sufficient to match high-frequency data quantitatively. We, therefore, keep a yearly trading interval because we need to cover a sufficient number of years to observe some reasonable amount of trading. Even with this limitation, we are going to document interesting patterns that match real-world data qualitatively.

Finally, at  $t = 0$ , we want to choose initial endowments of securities that are such that the initial conditions have the least possible influence on any of our results and such that other, recurring sources of trading remain dominant factors. For that purpose we put the two investors in a “neutral” portfolio position where they have no immediate need to trade. Specifically, imagining symmetric securities endowments at  $t = -1$  – equal to  $\frac{1}{2}$  for equity shares and 0 for the riskless securities –, we allow them at time  $t = 0$  freely to choose an optimal portfolio, without paying any fee but knowing that, thereafter, they will trade in the market with friction, starting from that initial position. In an attempt further to render the statistical results less dependent on initial conditions, when we display outcomes at a specific point in time, we choose  $t = 25$ . In other

Table 1: **Parameter Values.** This table lists the parameter values used for all the figures in the paper.

Name	Symbol	Value	Range
<i>Parameters for exogenous endowment</i>			
Horizon of the economy	$T$	50 years	
Expected growth rate of payout		4.4%/year	
Time step of the tree		1 year	
Volatility of payout		21.6%/year	
Initial endow. at $t = 0$ (cons. units)		1	
<i>Parameters for the investors</i>			
Investor 1's risk aversion	$\gamma_1$	2	
Investor 2's risk aversion	$\gamma_2$	4	
Investor 1's time preference	$\beta_1$	0.9	[0.9, 0.99]
Investor 2's time preference	$\beta_2$	0.9	
<i>Transactions fees per dollar of equity traded</i>			
When buying and when selling	$\lambda = \varepsilon$	1%	[0%, 3%]

words, we allow for a “burn-in” period equal to 25 time units.<sup>26</sup>

Table 1 shows all the parameter values.

### C Time paths of the economy with friction

We now describe the mechanics of the equilibrium over time and the transactions that take place. In the presence of transactions fees, a key concept, which we further elaborate on in Subsection IV.A is that of “no-trade zone,” which is the area of the space where both investors prefer not to adjust their portfolios. By way of illustration, Figure 2 displays a simulated sample path illustrating how our financial market with transactions fees operates over time. In an attempt to remove the effects of the finite horizon on trade decisions, we only display the first 25 periods of our economy, although it runs for 50 periods.

Panels (a) and (b) show a sample path of: (i) stock holdings as they would be in a zero-transaction fee economy, (ii) the actual stock-holdings with a 1% transaction fee and (iii) the boundaries of the no-trade zone, which fluctuate over time, with transaction dates highlighted by a circle. The boundaries fluctuate very much in tango with the optimal frictionless holdings, allowing a tunnel of deviations on each side. Within that tunnel, the investors' logic is apparent: the actual holdings move up or down whenever they are pushed up or down by the movement of the boundaries, with a view to reduce transactions fees and making sure that there occur as few wasteful round trips as possible. Panel (a) viewed in parallel with Panel (b) illustrates how the two investors are wonderfully synchronized by the market: they are made to trade exactly at the same

<sup>26</sup>If the equilibrium of this economy had been a stationary one, a sufficiently long run-in period would have rendered our results independent of initial conditions. But, with investors of different risk aversions, equilibrium is rarely stationary (see Dumas (1989)).



time exactly opposite amounts. Panels (c) and (d) show the holdings of the riskless security (the “bond”), which basically accommodate the holdings of equity. The optimal holdings under friction are a delayed version of the frictionless holdings. But the length of the delay is not constant.

Panel (e) shows the stock posted price (expressed in units of the consumption good), with transaction dates highlighted by a circle. While the posted price forms a stochastic process with realizations at each point in time, transactions prices materialize as a “point process” with realizations at random times only. The simultaneous observation of Panels (a), (b) and (e) shows the exact way in which the algorithm has synchronized the transactions of the two investors: the first four transactions occur on an up move of the price (Panel (e)), revealing an up move of the payout; on these moves, investor 1 who is less risk averse buys and investor 2 sells (see Subsection III.A above). The next two transactions occur on a down move; investor 1 sells.<sup>27</sup> As Figure 8, Panel (d) further illustrates, and as we explain below, investors trade more often upon an up-move than upon a down-move.

Even though ours is a Walrasian market and neither a limit-order nor a dealer market, we could define a concept of bid and ask as being the prices inclusive of transactions fees at which a person would be willing to buy or sell. More precisely, the bid price of a person could be defined as being equal to the person’s private valuation of payouts (defined below in Subsection IV.B) minus the transactions fees to be paid in case the person buys. When the two valuations differ by more than the sum of the one-way transaction fees for the two investors, a transaction takes place. Equivalently, a trade occurs when the bid price of the dealer (Investor 2) is equal to the ask price of the customer (Investor 1) or vice-versa. Or, the “bid-ask” spread, defined as the higher bid price minus the lower ask price of the two investors, becomes zero at the time of a transaction. That mechanism is displayed in Panel (f).

## IV Equilibrium asset holdings and prices

After solving for the equilibrium process following the backward-induction procedure explained in Section II, we run 50000 simulated paths obtained by walking randomly down the binomial tree of payouts. All quantitative results we display below are statistics computed across simulated paths at date  $t = 25$ . We choose that date in order to avoid the potential effect of the horizon date  $T = 50$ , while allowing a long enough “burn-in” period since the initial date. When more than one date is needed to compute a particular statistic, we use the dates surrounding  $t = 25$  or the sample paths running from  $t = 0$  to  $t = 25$ .

---

<sup>27</sup>The Lee and Ready (1991) algorithm would identify Investor 1 as the “customer”. When, in empirical work, the direction of trade is not observed, they recommend to classify the transaction as a buy (by the customer) if it occurs on an “uptick”.

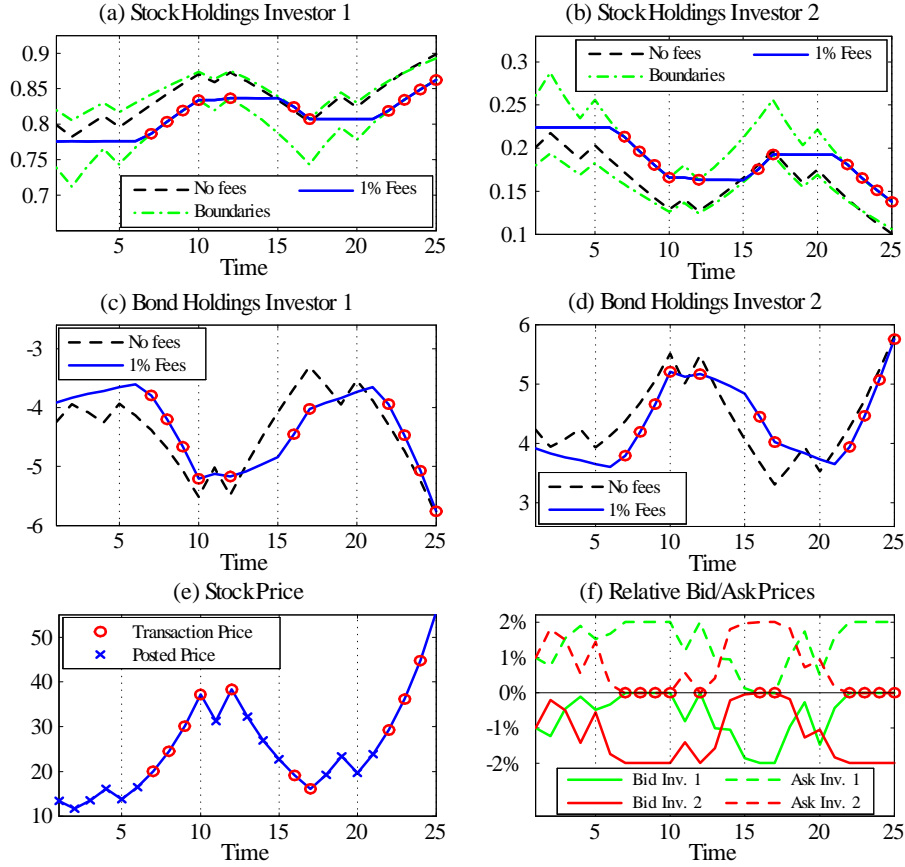


Figure 2: **A sample time path of stock holdings, bond holdings, the stock price and the bid and ask prices of each investor.** Panels (a) and (b) show a sample path of: (i) stock holdings as they would be in a zero-transaction fee economy, (ii) the actual stock-holdings with a 1% transactions fee and (iii) the boundaries of the no-trade zone. Panels (c) and (d) shows the holdings of the riskless security (the “bond”). Panel (e) shows the stock posted price (expressed in units of the consumption good). Panel (f) displays the bid and ask prices of both investors as a percentage difference from the posted price. In all panels, transaction dates are highlighted by a circle. All parameters are set at their benchmark values indicated in Table 1.

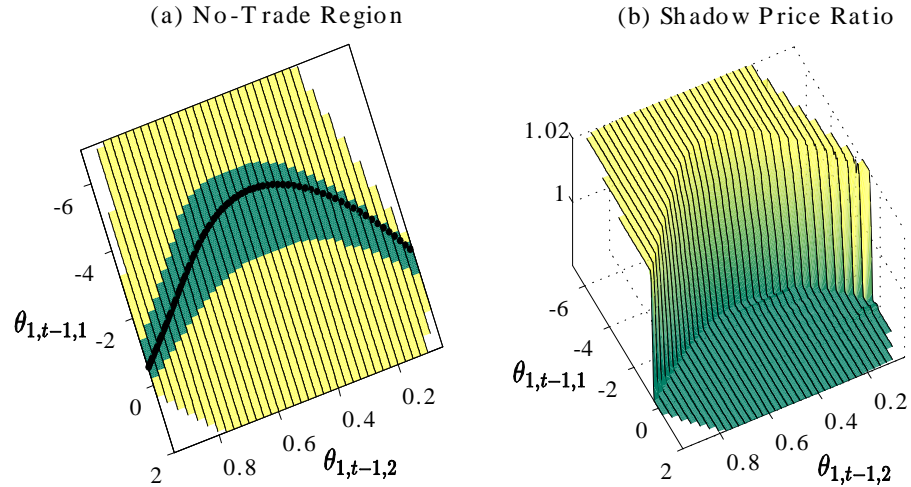


Figure 3: **Equilibrium no-trade region.** Panel (a) shows the no-trade region for different “entering” positions  $\theta_{t-1}$  of the investors at  $t = 25$ . The solid line drawn within the no-trade region shows the sharing rule as it would be in a frictionless market. Transactions fees are equal to  $1\% \times$  value of shares traded, while Panel (b) displays the ratio of shadow prices across the trade and no-trade regions. In both panels, parameters are as in Table 1.

## A Equilibrium asset holdings

It is well-known from the literature on non-equilibrium portfolio choice that proportional transactions costs cause the investors to tolerate a deviation from their preferred holdings. The zone of tolerated deviation is called the “no-trade region.” In previous work, the no-trade region had been derived for a given stochastic process of securities prices. We now obtain the no-trade region in general equilibrium at  $t = 25$ , when two investors make synchronized portfolio decisions and prices are set to clear the market.

### 1 Equilibrium no-trade region

Figure 3, Panel (a) plots the no-trade region at  $t = 25$  for different realized values of the holdings of securities chosen at  $t - 1 = 24$ . To our knowledge, Figure 3, Panel (a) is the first representation of an *equilibrium* no-trade region ever displayed. It is an equilibrium no-trade region in the sense that both investors have been coordinated to trade at the same time in opposite amounts.

The lighter striped grey zone is specifically the trade region while the darker striped zone is the no-trade region. The crescent shape of the no-trade region

is dictated by the shape of the frictionless sharing rule and is the result of the difference in risk aversions between the two investors: the black curve shown inside the zone would be the locus of holdings in a frictionless, complete market, corresponding exactly to the slopes of the wealth-sharing rules of Figure 1. When the holdings with which Investor 1 enters date  $t = 25$  are in one part of the trade region, those of Investor 2 are in the other part and both investors trade to reach opposite edges of the no-trade region; to the contrary, when the holdings upon entering the trading date are within the no-trade regions, the investors do nothing.<sup>28</sup>

Panel (b) illustrates how the shadow prices vary across the trade and no-trade regions: in one trade region, the shadow price of one investor is equal to  $1 + \lambda$  (one plus the buy transaction fee) while the other investor's shadow price is equal to  $1 - \varepsilon$  (one minus the sell transaction fee) and in the other trade region, the opposite is true. The ratio between their two shadow prices is, therefore,  $\frac{1+\lambda}{1-\varepsilon} \cong 1.02$  or  $0.98$ . Within the no-trade region the ratio is between these two numbers, with a discrete-version of the smooth-pasting condition holding on the optimal boundary and causing the shadow-price difference to taper off smoothly. The result is analogous to the no-trade region and the relative price of capital located in two countries in the equilibrium shipping model of Dumas (1992), with the difference that the trades considered are not costly arbitrages between *geographic locations* in which physical resources have different prices but are, instead, costly arbitrages between *people* whose private valuations of paper securities differ.

## 2 Consumption

Figure 4 shows, against the rate of transactions fees, the mean behavior of optimal consumption at  $t = 25$ . As Panels (c) and (d) of the Figure illustrate, on average the presence of transactions fees does not allow the more risk averse investor to smooth her consumption as well as she would without them and it does not allow Investor 1 to provide as much insurance as she would. This is because investors are conflicted between the desire to smooth consumption and the need to smooth trades. Panels (a) and (b) show that Investor 2 receives some compensation for the deteriorated smoothing of her consumption in the form of a higher growth rate of it.

## 3 Holdings

In terms of portfolio positions, the larger volatility of her consumption means that Investor 2 holds more of the stock than she would, and Investor 1 less.<sup>29</sup> Figure 5 shows, against the rate of transactions fees, the mean holdings of the stock and bond with which Investor 1 exits trading period  $t = 25$ . While

<sup>28</sup>The white zone of Figure 3, on both sides of the lighter grey zone, is not admissible; when entering holdings are in that zone, there exists no equilibrium as one investor would, at equilibrium prices, be unable to repay her negative positions to the other investor.

<sup>29</sup>Notice how in Figure 2, Panel (b), the sample path under no friction was closer to the upper boundary of the no trade region than to the lower boundary.

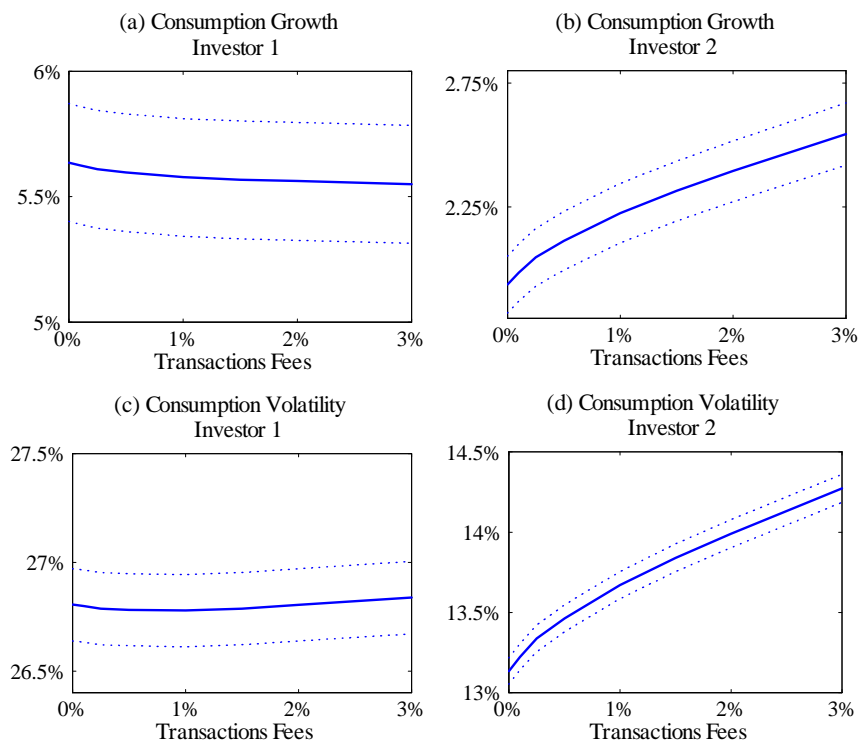


Figure 4: **Optimal consumption behavior** at  $t = 25$ . Average expected growth and volatility (computed from the model) of the two investors for different levels of transactions fees. All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated at  $t = 25$  across 50000 simulated paths. The solid line is the average. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

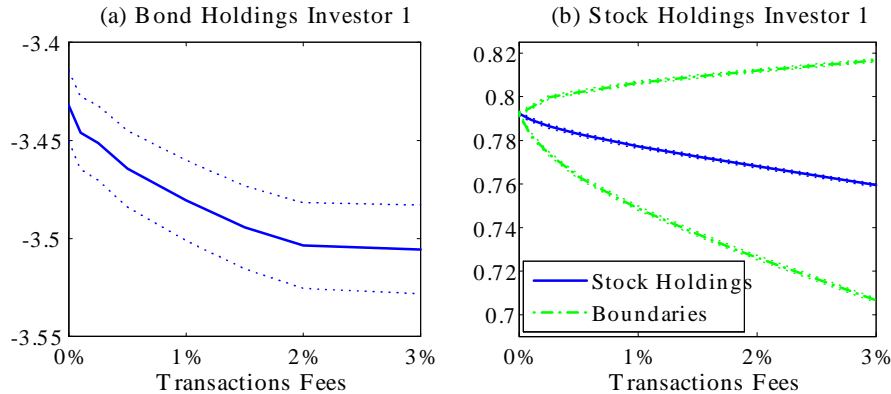


Figure 5: **Optimal “exiting” holdings  $\theta$  of the securities at  $t = 25$ .** Optimal bond and stock holdings of the first investor for different levels of transactions fees. All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated at  $t = 25$  across 50000 simulated paths. The solid line is the average of holdings and, in Panel (b), the semi-dashed lines show the boundaries of the no-trade region.. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

Investor 1, who is less risk averse, is a natural borrower and thus chooses negative positions in the bond, increased transactions fees induce him to carry on with a smaller holding of equity. For that reason alone, she would have to borrow less. However, her wealth is lower for higher transactions fees; the wealth effect dominates and her (negative) holding of bonds also goes down.<sup>30</sup>

#### 4 The clientele effect

Do more patient investors hold less liquid assets as in the “clientele effect” of Amihud and Mendelson (1986)? This is an issue to which most of Longstaff (2009) is devoted since this paper considers two investors with log utility who differ only in their degree of impatience. Longstaff answers the question in the positive.<sup>31</sup> Here, we provide the answer with endogenous trading dates.

We vary the patience parameter of the first investor between 0.9 and 0.99. Figure 6 shows that the clientele effect is negligible.

<sup>30</sup>Here is a second rationale for the result: we explain below (in Subsection V.A) that more transactions occur on an up move of the payoff (and of the stock price) than on a down move. The less risk averse investor knows that, on a down move, he will wish to sell. Since this will be possible relatively infrequently, he chooses to undershoot and, on an average, holds smaller amounts of stock than he would in a frictionless world.

<sup>31</sup>See his Table 2, page 1131.

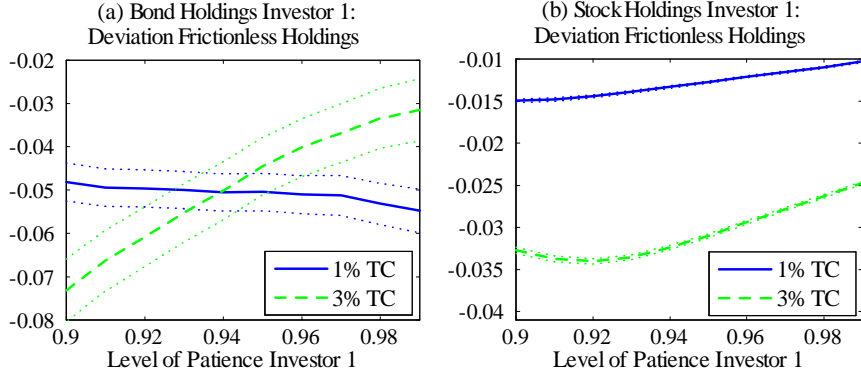


Figure 6: **Clientele effect.** Optimal bond and stock holdings of the first investor at  $t = 25$  for different levels of patience, in the range from 0.9 to 0.99. The figure plots the holdings with friction minus the holdings without friction. All other parameters are as described in Table 1. Transactions fees are, as indicated, equal to 1% or 3% of the value of shares traded. The figure displays averages calculated at  $t = 25$  across 50000 simulated paths. The solid line is the average. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean result.

## B Asset prices

According to Amihud and Mendelson (1986a, Page 228), the price of a security in the presence of transactions costs is equal to the present value of the payouts to be paid on that security minus the present value of transactions costs subsequently to be paid by someone currently holding that security. A similar conclusion was reached by Vayanos and Vila (1999, Page 519, Equation (5.12)).

Our setting and the setting of Amihud and Mendelson are quite different. They consider a large collection of risk-neutral investors each of whom faces different transactions costs and are forced to trade. We consider two investors who are risk averse, face identical trading conditions and trade optimally. Nonetheless, their statement is an appealing conjecture to be investigated using our model.

Recall from Equation (9) that the securities' posted prices  $S_{t,i}$  are:

$$S_{t,i} = \mathbb{E}_t \left[ \frac{\phi_{l,t+1}}{R_{l,t,i}\phi_{l,t}} \times (\delta_{t+1,i} + R_{l,t+1,i} \times S_{t+1,i}) \right];$$

$$S_{T,i} = 0$$

where the terms  $R_{l,t,i}$  ( $1 - \varepsilon_{i,t} \leq R_{l,t,i} \leq 1 + \lambda_{i,t}$ ) capture the effect of current and anticipated trading fees.

We now present two comparisons. First, we compare equilibrium prices to the present value of payouts on security  $i$  calculated at the Investor  $l$ 's equilibrium

state prices *under transactions fees*. We denote this private valuation  $\hat{S}_{l,t,i}$  :

**Definition 2**

$$\hat{S}_{l,t,i} \triangleq \mathbb{E}_t \left[ \frac{\phi_{l,t+1}}{\phi_{l,t}} \times \left( \delta_{t+1,i} + \hat{S}_{t+1,i} \right) \right]; \hat{S}_{T,i} = 0$$

In Appendix D, we show that:

**Proposition 3**

$$R_{l,t,i} \times S_{t,i} = \hat{S}_{l,t,i} \tag{10}$$

which means that the posted prices of securities can at most differ from the private valuation of their payouts as seen by Investor  $l$  by the amount of the transactions fees incurred or imputed by Investor  $l$  at the current date only.<sup>32</sup> The posted price is thus seen as some form of average of the two private valuations.

In Figure 7, Panels (a) and (b) show the bond and stock prices at  $t = 25$  for different levels of transactions fees in the range from 0% to 3%. Panel (c) shows the two investors' present values of payouts  $\hat{S}_{l,t,i}$  for the stock, thus illustrating the decomposition of Equation (10). For instance, for transactions fees of 3%, the difference between the two private valuations is in the range  $[-3\%, +3\%]$  of endowment, where we achieve the boundaries of this range when the system hits the boundaries of the trade region. Within the no-trade region, it is somewhere within the range. While the posted price  $S$  and the present value of payouts  $\hat{S}$  differ from each other at most by one-way transactions fees, it is ambiguous whether they are increased or reduced by the presence of transactions fees.<sup>33</sup>

Second, as a way to understand this ambiguity, we compare equilibrium asset prices that prevail in the presence of transactions fees to those that would prevail in a frictionless economy, based, that is, on state prices that would obtain *under zero transactions fees*. Denoting all quantities in the zero-transactions fees economy with an asterisk \*, and defining:

$$\Delta\phi_{l,t} \triangleq \frac{\phi_{l,t}}{\phi_{l,t-1}} - \frac{\phi_{l,t}^*}{\phi_{l,t-1}^*}$$

we show in Appendix E that:

**Proposition 4**

$$R_{l,t,i} \times S_{t,i} = S_{t,i}^* + \mathbb{E}_t \left[ \sum_{\tau=t+1}^T \frac{\phi_{l,\tau-1}}{\phi_{l,t}} \times \Delta\phi_{l,\tau} \times \left( \delta_{\tau,i} + S_{\tau,i}^* \right) \right] \tag{11}$$

<sup>32</sup>Our proposition is reminiscent of Vayanos (1998) who writes (Page 26): “Second, the effect of transaction costs is smaller than the present value of transaction costs incurred by a sequence of marginal investors.”

<sup>33</sup>Vayanos (1998) has noted that prices can be increased by the presence of transactions costs.



That is, the two asset prices differ by two components: (i) the current shadow price  $R_{l,t,i}$ , acting as a factor, of which we know that it is at most as big as the one-way transactions fees, (ii) the present value of all future price differences arising from the change in state prices and consumption induced by the presence of transactions fees. *The reason for any effect of anticipated transaction fees on prices is not the future fee expense itself. It is, instead, that investors do not hold the optimal frictionless holdings and, therefore, also have consumption schemes that differ from those that would prevail in the absence of transactions fees, as we have seen in Subsection A.2.2. The differences in consumption schemes then influence the future state prices and accordingly the present values of payouts.*

Panel (d) of Figure 7 shows the total difference at  $t = 25$  between the stock price in an economy with transactions fees and the stock price in economies without transactions fees. In addition, it shows the component of that difference that is due to the discounted value of future pricing kernel differences of Investor 1, the remainder being equal to current transactions fees, as per Equation (11).

Equation (11) can be recomposed as follows:

$$\hat{S}_{l,t,i} - S_{t,i}^* = \mathbb{E}_t \left\{ \sum_{\tau=t+1}^T \left[ \frac{\phi_{l,\tau-1}}{\phi_{l,t}} \times \text{cov}_{\tau-1} [\Delta\phi_{l,\tau}, (\delta_\tau + S_\tau^*)] \right. \right. \quad (12)$$

$$\left. \left. + \frac{\phi_{l,\tau-1}}{\phi_{l,t}} \times \mathbb{E}_{\tau-1} \Delta\phi_{l,\tau} \times \mathbb{E}_{\tau-1} (\delta_\tau + S_\tau^*) \right] \right\}$$

The ambiguity in the effect of transactions fees on prices arises from the partially offsetting effects of the two terms of Equation (12). For Investor  $l = 2$ , the term  $\text{cov}_{\tau-1} [\Delta\phi_{l,\tau}, (\delta_\tau + S_\tau^*)]$  is negative because, being more risk averse, she holds more of the risky asset than she would in a frictionless world, as explained in Subsection 3 above; hence, her consumption is more positively correlated with the payoff on the risky asset and her marginal utility is more negatively correlated. However, the term

$$\mathbb{E}_{\tau-1} \Delta\phi_{l,t} \triangleq \beta_l \mathbb{E}_{\tau-1} \left[ \frac{c_{l,t}^{\gamma_l-1}}{c_{l,t-1}^{\gamma_l-1}} \right] - \beta_l \mathbb{E}_{\tau-1} \left[ \frac{(c_{l,t}^*)^{\gamma_l-1}}{(c_{l,t-1}^*)^{\gamma_l-1}} \right]$$

is positive for Investor  $l = 2$  because her consumption is more volatile and marginal utility is a convex function. Of course, the opposite statements hold for the two terms of Investor 1.

Because the affected state prices are applied by investors to all securities, the change in the state prices is also reflected in the one-period bond price even though bonds trading does not incur transactions fees, as is illustrated in Panel (a) of Figure 7.

Finally, Panel (e) shows the effective spread defined as the difference between the highest bid and the lowest ask of the two agents, where the bid price is defined as before as being equal to the person's private valuation of payouts minus the transactions fees to be paid in case the person buys and the ask

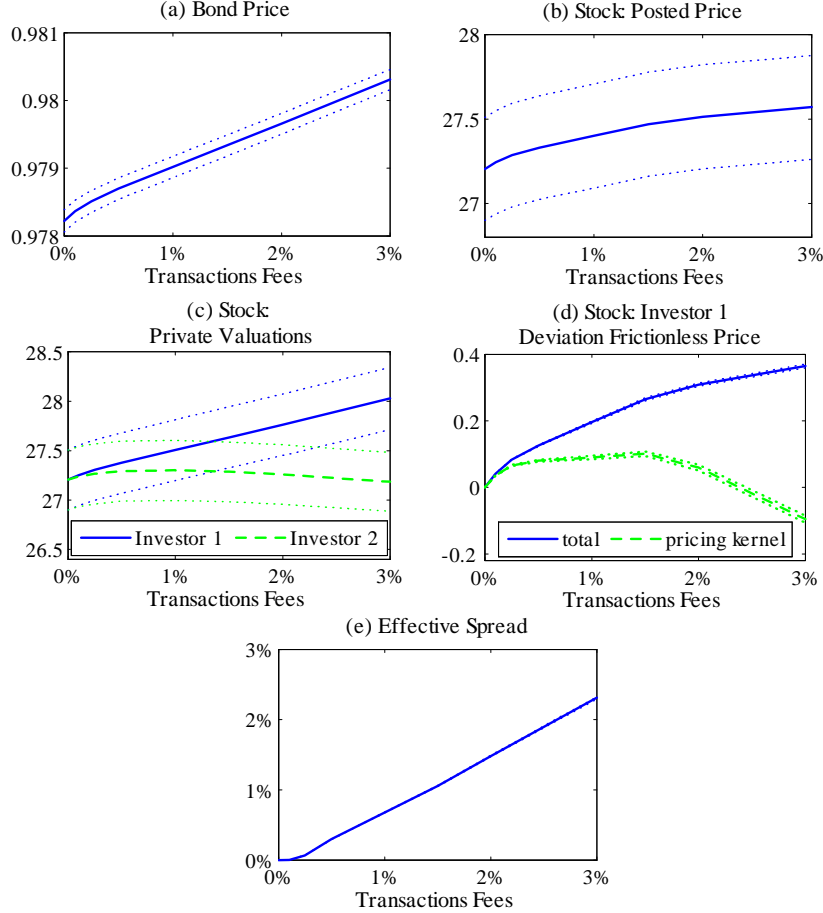


Figure 7: **Asset prices.** Panel (a) shows the bond price at  $t = 25$  for different levels of transactions fees. Panel (b) shows the stock price also at  $t = 25$ . Panel (c) shows the two investors' present values of payouts  $\hat{S}_{t,i,l}$  showing the price comparison as in Equation (10). Panel (d) shows the total difference between the initial stock price in an economy with transactions fees and the stock price in economies without transactions fees. In addition, we show the component of that difference that is due to the discounted value of future pricing kernel differences of Investor  $l = 1$ , as per Equation (11). Panel (e) shows the effective spread. All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated at  $t = 25$  across 50000 simulated paths. The solid line is the average. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

price is defined analogously. As expected, the spread rises with fees as possible deviations between the two agents' private valuations increase. Its mean is approximately equal to, and somewhat below, the actual rate of transactions fee.

## V Time paths of prices and holdings

We now demonstrate some properties of the sample paths. We investigate univariate properties of trades on the one hand and of asset price increments on the other. Then we investigate a bivariate property of trades and price changes.

### A Trades over time

We examine the trading volume and the waiting time between trades induced by transactions fees. The trading volume is defined as the absolute values of changes in  $\theta_2$  (shares of the stock) at time  $t = 25$ . The average trading volume is shown in Figure 8, Panel (a); as one would expect, it decreases with transactions fees. Correspondingly, the average waiting (Panel (b)) between trades rise. We also show in Panel (b) the volatility of the waiting time, which is a measure of the *endogenous liquidity risk* that the investor has to bear because she operates in a market with friction. We examine in section VI below how this risk is priced.

Panel (c) demonstrates that transactions take place more often on an up move than on a down move in the binomial tree of the payout process. For a given price of the stock and given transactions fees, it costs the same to trade on an up or a down move but, since movements up and down of the underlying payoff are of discrete size, if one trades on a down move, one damages one's utility more because of the curvature of the utility function.

The empirical literature has established that trades are autocorrelated and the order flow is predictable (Hasbrouck (1991a, 1991b) and Foster *et al.* (1993)). Looking at the time path in Panels (a) and (b) of Figure 2, we have already pointed out that the investors smooth their trades over time in order to keep transactions fees low. We investigate the matter more systematically in Panel (d). In our Walrasian market, the intertemporal optimization conducted by investors produces a fragmentation of trades over time and positive serial dependence.

### B Prices over time

We are interested in knowing what properties of the sample paths of stock prices that prevail in a frictionless economy are markedly modified by transactions fees. We consider in particular two such properties. One is the length of the time graphs of sample paths and the other is the serial dependence of the paths.

In which way, as one decreases transactions fees, does the point process of transactions prices approach the process that would prevail in the absence of

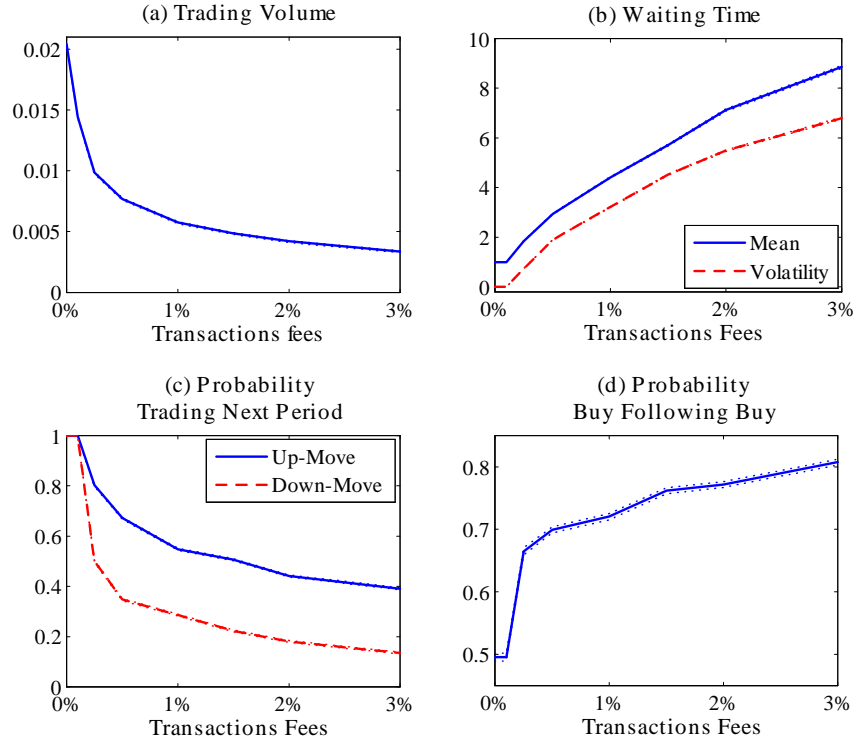


Figure 8: **Patterns of volume and waiting time against transactions fees.** Panels (a) show the mean (across paths at  $t = 25$ ) stock trading volume/year; Panel (b) show the mean expected value of and volatility of the waiting time between trades (measured in years) for different levels of transactions fees. Panel (c) shows the frequency (across paths and dates) of a buy transaction coinciding with an up or down move in price or endowment. Panel (d) shows the frequency (across paths) of a buy transaction following a previous buy transaction. All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated at  $t = 25$  across 50000 simulated paths. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

transactions fees, which in the limit of continuous time would be a continuous-path process? As is well-known, the Brownian motion is characterized by the fact that its total variation, calculated over a finite period of time, is infinite while its quadratic variation is finite.<sup>34</sup> Here, of course, we have represented the random shocks by means of a finite tree so that nothing can become infinite but we can form conjectures about this tendency by shortening the time step.

For zero transactions fees, we calculate average (across paths) total variation and quadratic variation (along each path) over the first 25 periods. Then we generate the same paths of transactions prices with transactions fees equal to 1% and calculate again average (across paths) total variation and quadratic variation. These are plotted against the time step in Figure 9.

Panel (a) of the Figure shows that the total variation of the posted price is finite because this is a finite tree but that, if one took the limit of continuous time, it would be infinite, as is the case for Brownian motions. Similarly, its quadratic variation (Panel (c)) is practically constant. Both variations are practically invariant to transactions fees. The posted price mostly inherits the sample-path properties of a frictionless stock price.

Transactions prices are the result of infrequent sampling of posted prices (with frequency of sampling rising as transactions fees go down), so that one might expect them to inherit the above time-series properties of posted prices. However, that is not right because the sampling (the occurrence of transactions) is not independent of the price movements. Indeed, as Panel (b) shows, in the presence of a 1% transaction fee, the total variation of transactions prices is practically constant and markedly lower than in the absence of friction, and so is the quadratic variation.

The effect of transactions fees on volatility is exhibited in Figure 10, Panel (a). On average, the friction reduces the conditional volatility of the rate of return on the stock.

The serial dependence of price increments plays a crucial role in the empirical Microstructure literature.<sup>35</sup> Figure 10, Panel (b) displays the frequency of an up move in price being followed by an up move in price. Because a move up in the posted price can only be associated with a move up in the payout, and because we have assumed payout moves that are up or down with probabilities 50-50, that frequency is tautologically equal to 1/2 when transactions fees are zero and almost equal to 1/2 when considering the posted price. When consid-

<sup>34</sup>Total variation is the sum of the absolute values of the segments making up a path or connecting the dots, whereas quadratic variation is the sum of their squares.

<sup>35</sup>It serves to decompose real frictions from information frictions. Roll (1984) originally proposed to use the “bid-ask bounce” to measure the effective spread, an approach which was later generalized by Stoll (1989). As Stoll (2000) explains, “Price changes associated with order processing, market power, and inventory are transitory. Prices ‘bounce back’ from the bid to the ask (or from the ask to the bid) to yield a profit to the supplier of immediacy. Price changes associated with adverse information are permanent adjustments in the equilibrium price.” The presumption is that, in response to random customer arrivals (Stoll (2000)), “bid and ask prices are lowered after a dealer purchase in order to induce dealer sales and inhibit additional dealer purchases, and bid and ask prices are raised after a dealer sale in order to induce dealer purchases and inhibit dealer sales.” In our model, there is information coming in (but no information asymmetry).

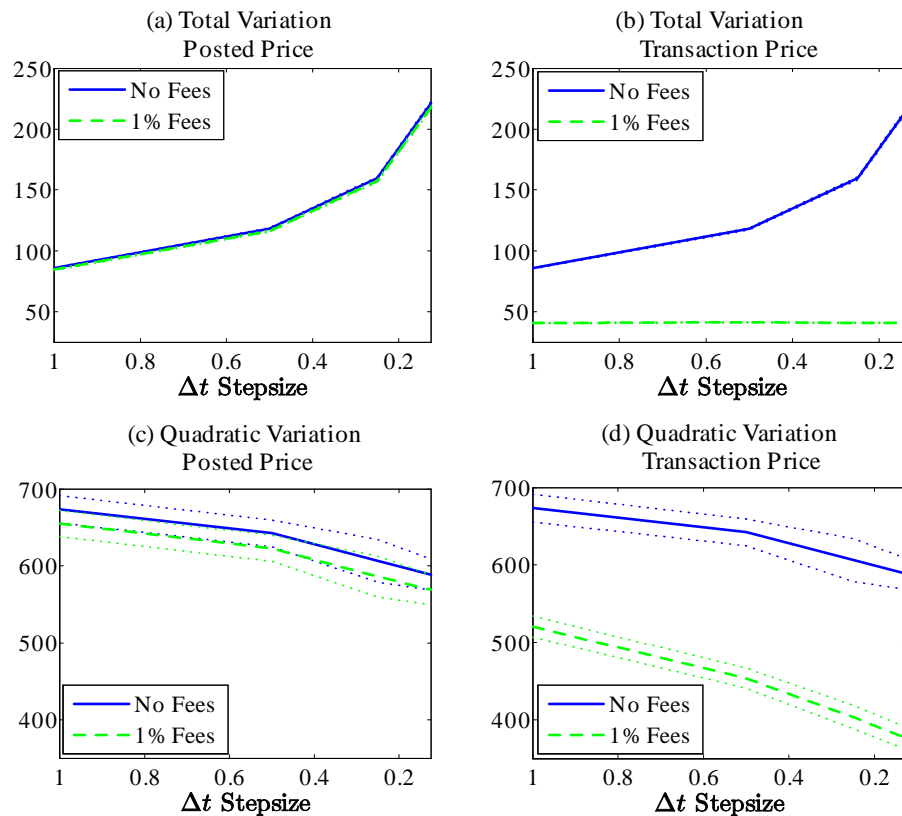


Figure 9: **Total and quadratic variations of stock price depending on transactions fees and time step.** Panels (a) and (b) shows the total variation (defined in footnote 34) of the posted price and the transactions price respectively, up to period 25 for transactions fees of 0% and 1% as a function of the length of the time step. Panels (c) and (d) show the quadratic variation computed the same way. All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated over the interval of time  $t = [0, 25]$  across 50000 simulated paths. The solid line is the average. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

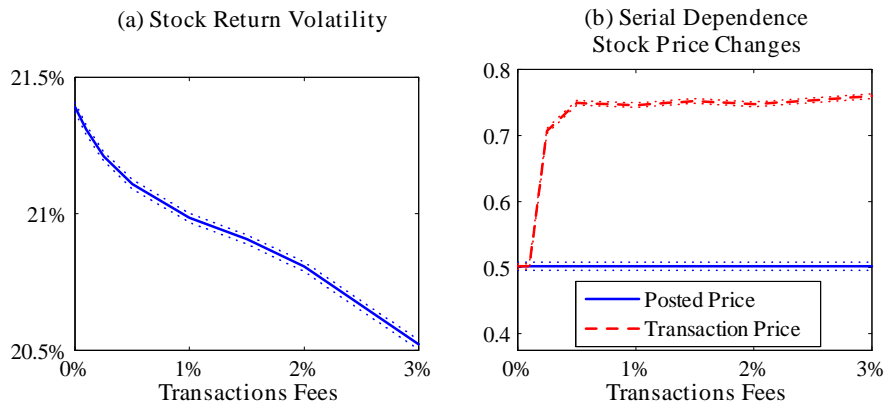


Figure 10: **Volatility and serial dependence of price changes.** Panel (a): the mean (across paths) conditional volatility for an investment made at  $t = 25$ . Panel (b): frequency (across paths and dates from  $t = 0$  to  $t = 25$ ) of a price increase following a previous price increase. All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated across 50000 simulated paths. The solid line is the average. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

ering transactions prices, however, the frequency quickly rises with transactions fees, to above 0.67. Far from displaying a reversion tendency, prices display momentum. Evidently, the absence of a bounce, if observed by an econometrician, should not be regarded as evidence of absence of real frictions. In our model of optimal investor arrival, there can be no buy or sell order coming to the market place unless some information about the fundamental has also arrived. It is not the case that customers act randomly and dealers accommodate them temporarily in an optimal fashion. Everyone here acts optimally.

### C “Price impact”

Finally, we examine a statistic describing the joint dynamic behavior of trades and asset prices. One is tempted to ask: what is the impact of a trade on price? That question can be misconstrued, however, because in our model both quantities are endogenous. Furthermore, even if a trade has an impact on price, the investors behave competitively, take the price as given and, therefore, ignore the price impact. In our model, there is a “price-impact” phenomenon but only because, in equilibrium, when an up move in the payout occurs, the price goes up and simultaneously the less risk averse investor buys, both being endogenous. Both variables are changed by the same exogenous endowment shock. It is not the case that the price goes up “because” of the trade of one investor seen as

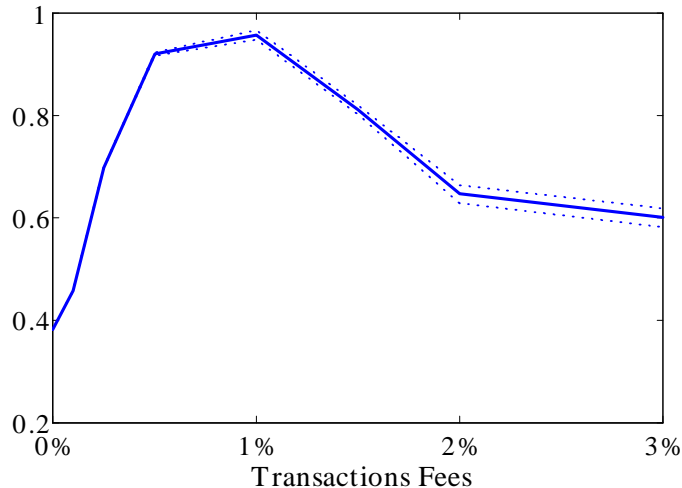


Figure 11: **“Price impact”**: the mean (across paths) of the conditional covariance between next period’s stock return and next period’s unexpected signed trading volume over the variance of the unexpected signed trading volume. All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated at  $t = 25$  across 50000 simulated paths. The solid line is the average, bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

exogenous.

As a way of measuring the price impact that does take place, we take Kyle’s Lambda somewhat out of context and compute the mean (across paths) ratio of the conditional covariance between next period’s posted stock return and next period’s unexpected signed trading dollar volume over the variance of the unexpected signed trading dollar volume.<sup>36 37 38</sup> The result is displayed in Figure 11. The price-impact measure is not zero even at zero transactions fees. It rises steeply as one increases the fee, but then falls again because the increased fee starts causing an increasing number of observations of zero trade, which bring down the covariance.

<sup>36</sup>We use the square root of the signed trading volume as in Hasbrouck (2009), i.e.,  $\text{sign}[\text{signed unexp. volume}(t+1)] \times \sqrt{|\text{abs}(\text{signed unexp. volume}(t+1))|}$ .

<sup>37</sup>The separation between anticipated and unanticipated components (of, say, signed trading volume) is made by means of the conditional expected value provided by the model.

<sup>38</sup>More formal methods to measure price impact are based on reduced forms of theoretical microstructure models. Some are motivated by the desire to capture informed trading (Roll (1984), extended by Glosten and Harris (1988)). Others (Ho and Macris (1984)) are motivated by inventory considerations. Madhavan and Smidt (1991) run a regression which is meant to capture both effects.



## VI The pricing of liquidity and of liquidity risk

Based on a pure portfolio-choice reasoning, Constantinides (1986) argued that transactions costs make little difference to risk premia in the financial market. Liu and Lowenstein (2002) and Delgado *et al.* (2012), still based on portfolio choice alone, challenge that view by pointing out that the conclusion of Constantinides holds only when rates of return are identically, independently distributed (IID) over time. We go one step further than these authors, in that we now get the deviations in a full general-equilibrium model, when endowments are IID but returns themselves are not, and investors must also face the uncertainty about the dates at which they can trade.

### A Deviations from the classic consumption CAPM under transactions fees

In our equilibrium, the capital-asset pricing model is Equation (9) above. It is specific to each investor; we make no attempt at aggregation. The dual variables  $R$ , in addition to the intertemporal marginal rates of substitution  $\phi$ , drive the prices of assets that are subject to transactions fees, as do, in the ‘‘LAPM’’ of Holmström and Tirole (2001), the shadow prices of the liquidity constraints.<sup>39</sup> In effect, there are two distinct pricing kernels: one  $\phi_{l,t+1,j}$  applies to payouts paid in consumption units at time  $t + 1$ , the other  $\phi_{l,t+1,j} \times R_{l,t+1,i,j}$  applies to the time- $t + 1$  posted value of the security. Equivalently, by induction, the present value of all future payouts discounted using the first pricing kernel only, gives the the private valuation  $R_{l,t,i} \times S_{t,i}$ , as we saw in Equation (10).

We show now how, in our setup with fees, various premia arise, relative to the standard consumption-CAPM (CCAPM).

With the usual definition for the gross rate of return on asset  $i$ :<sup>40</sup>

$$r_{t+1,i,j} \triangleq \frac{\delta_{t+1,i,j} + S_{t+1,i,j}}{S_{t,i}}$$

we show in Appendix F that the CCAPM can be rewritten as:

$$\begin{aligned} \mathbb{E}_t[r_{t+1,i}] &= r_{t+1,1} - \text{cov}_t \left( r_{t+1,i}, \frac{\phi_{l,t+1}}{\mathbb{E}_t[\phi_{l,t+1}]} \right) \\ &+ \mathbb{E}_t \left[ (1 - R_{l,t+1,i}) \times \frac{S_{t+1,i}}{S_{t,i}} \right] - (1 - R_{l,t,i}) \times r_{t+1,1} \\ &+ \text{cov}_t \left( (1 - R_{l,t+1,i}) \times \frac{S_{t+1,i}}{S_{t,i}}, \frac{\phi_{l,t+1}}{\mathbb{E}_t[\phi_{l,t+1}]} \right); i \neq 1 \end{aligned} \quad (13)$$

<sup>39</sup>Holmstrom and Tirole (2001) make assumptions such that their liquidity constraint is always binding. Here, the inequality constraints (5) bind whenever it is optimal for them to do so.

<sup>40</sup>Recall that the security numbered  $i = 1$  is the short-term bond, so that  $r_{t+1,1}$  is conditionally riskless at time  $t$ .

Equation (13) opens the door for a decomposition exercise similar to that performed by Acharya and Pedersen (2005). Here, however, the terms have received a formulation that is explicitly related to the optimal decision of investors to trade or not to trade and they have explicit dynamics, which we exhibit and make use of.

The first part of the equation is exactly the CCAPM expression of a frictionless market. The remainder is a deviation from the CCAPM, which we can split into components. Observe that  $1 - R_{l,t+1,i,j}$  is a shadow transaction fee rate applying to asset  $i$  at time  $t$  in state  $j$ , from the point of view of Investor  $l$ , so that  $(1 - R_{l,t+1,i,j}) \times S_{t+1,i,j}$  is a future shadow dollar amount of transaction fee and  $(1 - R_{l,t+1,i,j}) \times \frac{S_{t+1,i,j}}{S_{t,i}}$  is the drag on the asset's rate of return occasioned by the fee, or the "dollar cost per dollar invested" in the words of Acharya and Pedersen (2005).

With that in mind, we can proceed to the following definitions:

**Definition 5**

*CCAPM deviation due to* : (14)

$$\begin{aligned} \text{expected liquidity change} &\triangleq \mathbb{E}_t \left[ (1 - R_{l,t+1,i}) \times \frac{S_{t+1,i}}{S_{t,i}} \right] - (1 - R_{l,t,i}) \times r_{t+1,1} \\ \text{liquidity risk} &\triangleq \text{cov}_t \left( (1 - R_{l,t+1,i}) \times \frac{S_{t+1,i}}{S_{t,i}}, \frac{\phi_{l,t+1}}{\mathbb{E}_t [\phi_{l,t+1}]} \right) \end{aligned}$$

The deviation due to expected liquidity change is equal to expected liquidity minus current liquidity. The deviation due to current liquidity is there because 1 dollar of the asset is potentially purchased today, interest on the fee being included. The deviation due to expected liquidity is there because the asset is potentially sold tomorrow. The deviation in the form of a liquidity risk premium is there because the dollar fee to be paid upon potential resale is uncertain.

Although we refer to the key variables as "liquidity change", notice that the *level* of the liquidity variables  $R$  also play a role in the CCAPM deviation. Indeed, supposing it were known that  $R_{l,t+1,i,j} = R_{l,t,i} \forall j$ , even then the CCAPM deviation would still be equal to:

$$(1 - R_{l,t,i}) \times \left\{ \mathbb{E}_t \left[ \frac{S_{t+1,i}}{S_{t,i}} \right] - r_{t+1,1} + \text{cov}_t \left( \frac{S_{t+1,i}}{S_{t,i}}, \frac{\phi_{l,t+1}}{\mathbb{E}_t [\phi_{l,t+1}]} \right) \right\}$$

which is not equal to zero.<sup>41</sup> Even if the fee rate were known, a stochastic dollar amount of fee would still have to be paid when transacting since the price risk would still be present.

---

<sup>41</sup>Only if the liquidity variable were at the level 1:  $R_{l,t+1,i,j} = R_{l,t,i} = 1$ , would the deviation be zero.

## B A cross-section

We now introduce a cross-section, which we study in this subsection and the next. From now on, there are two stocks, or two trees, in the economy. The sum of their two payouts follows exactly the same process as in the one-stock economy, i.e., with a drift of 4.4% and a volatility of 21.6%. The distribution of the total payouts between the two stocks follows a simple Markov Chain, with the following two states:

- State 1: Stock 1: 80% and Stock 2: 20%,
- State 2: Stock 1: 20% and Stock 2: 80%.

The matrix of probabilities of transition between the two states is:

$$\begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$$

so that the distribution of payouts is persistent. Otherwise the economy is exactly as in Table 1. In the frictionless case, the Markov chain – realizations and transitions –, being totally symmetric, the agents would always (at each state and each time) hold the same amounts of Stock 1 and Stock 2. That implies that the trading volume is always the same. Similarly, the two stocks have on average, i.e. unconditionally, the same expected return and variance but the returns, conditional on each state and time, are not the same as the exact value depends on the current distribution of the payouts (i.e., on the state of the Markov chain).

The two stocks differ by the rate of transactions fees that applies to their trading. Stock 1 is encumbered by an unchanging 1% fee, while we vary the fee levied on Stock 2 from 1% upward.

The mean CCAPM deviations are computed using simulated returns at  $t = 25$  for different levels of transactions fees, in the range from 1% to 3%. Figure 12 displays the result. As could be expected, the absolute value of the mean CCAPM deviation in Panels (a) and (b) is increasing in transactions fees. For high transactions fees of 3%, the total deviation reaches 40bp, which is less than the transactions fees themselves. That deviation is much too small to be able to account for the several percentage points of returns that empirical researchers commonly attribute to liquidity premia.<sup>42</sup> But it does show that trading frictions can play a role when we try to explain empirical deviations from classic asset pricing models.

Given the symmetry we have imposed between the two stocks, they fetch approximately equal expected returns when their transaction fees are equal. For reasons of portfolio diversification, even the stock with an unchanging fee has expected returns that depend slightly on the fee paid on the other one.

---

<sup>42</sup>Furthermore, the terms being of opposite signs for the two investors, their values would be even smaller in any CAPM that would be somehow aggregated across investors.

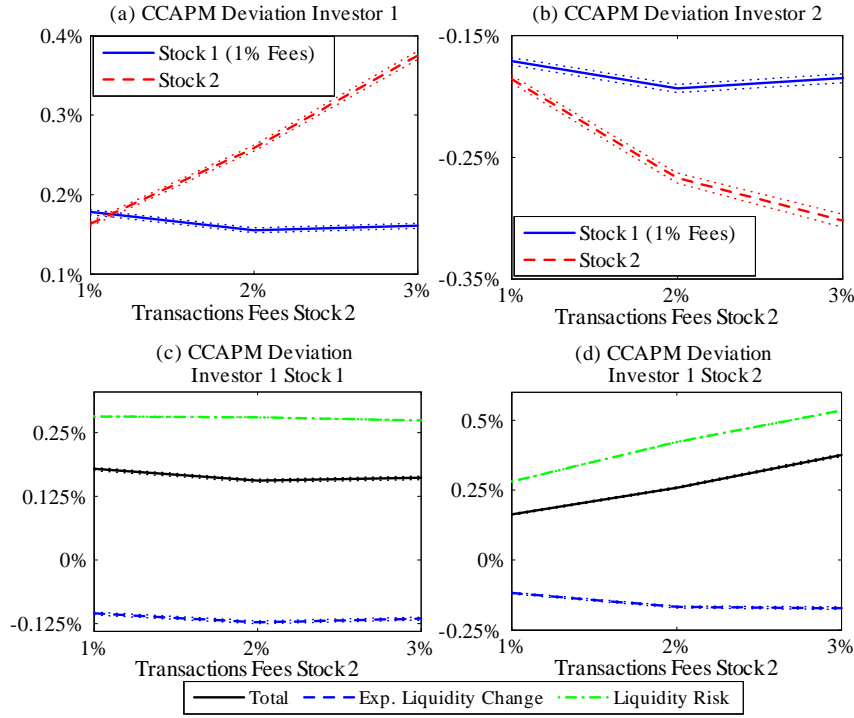


Figure 12: **The Cross-section of unconditional CCAPM deviations.** Panels (a) and (b) show the average across paths of deviations from the classic consumption CAPM (in %/year), as defined in Equation (14), for different levels of transactions fees, in the range from 1% to 3% for Stock 2 while Stock 1 fees remain at 1%. Panels (c) and (d) show the components of the deviations, as defined in Equation (14). All parameters are set at their benchmark values indicated in Table 1. The figure displays averages calculated at  $t = 25$  across 50000 simulated paths. The bolder lines are the averages. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

While Figure 12 displays the unconditional mean of CCAPM deviations,<sup>43</sup> one might also be interested in the unconditional root mean square CCAPM deviation, which would measure to degree to which the classic consumption CAPM is violated in its conditional form. With a fee of 1% on Stock 1 and a fee of 3% on Stock 2, these are somewhat larger than the mean deviations: Investor 1, Stock 1: 0.386%/year; Investor 1 Stock 2: 0.697%; Investor 2, Stock 1: 0.422%; Investor 2, Stock 2: 0.682%.

Panels (c) and (d) display the components of the CCAPM deviation. Because investors, who minimize round trips, tend to trade small amounts repeatedly along the same edge of their trade barrier, current liquidity and expected liquidity tend to be close to each other. Hence, the expected liquidity change is comparatively small unconditionally speaking and the unconditional CCAPM deviation is mostly determined by the liquidity-risk premium.

Two consequences follow from the dominance of the liquidity-risk premium. First, the CCAPM deviation is positive for the first investor, i.e., the less risk-averse investor demands a higher expected return in an economy with transactions fees whereas the more risk-averse Investor 2 demands a lower expected return. The reason is the following. State  $j$  at time  $t+1$  can be an up or a down node; if it is an up node, the marginal utility  $\phi$  of the investors will be lower than its conditional expected value and the less risk-averse investor will, if anything, be buying the security. In that node, therefore, the shadow fee  $1 - R_{l,t+1,i,j}$  is more likely to take the negative value  $-\lambda$ , thus contributing a positive number to the covariance summation of Investor 1. As we have seen, transactions are more likely to occur on an up move so that, for Investor 1, the covariance is on net positive.

Secondly, for the same reason, when the fees on the two stocks are different, the less risk-averse investor demands a higher expected return on the stock with the higher transactions fees whereas the more risk-averse Investor 2 demands a lower expected return, which is what we observe in Panels (a) and (b).

An important benefit of our model, in which the liquidity variables are endogenized, is that we can study the variation of each of the terms over time. Whereas Figure 12 has shown that the *unconditional* mean value of the CCAPM deviation is mostly due to the liquidity risk premium, a comparison of the two columns of Table 2 reveals that the conditional expected-liquidity change component is mostly responsible for the fluctuations of the *conditional* CCAPM deviation across paths, the conditional liquidity risk premium being approximately constant.

This theoretical contrast between the conditional and the unconditional properties should provide guidance for empirical researchers currently working on illiquid markets and trying to decide which of the two terms is more important. In a recent contribution, Bongaerts, De Jong and Driessen (2012),

---

<sup>43</sup>The term “unconditional mean” is used here for the first time. It has the exact same meaning as the term “mean (across paths)” that we have used so far. We alter the language slightly at this point in order to conform with the distinction, which is traditional in the Asset-Pricing literature, between tests of the CAPM in its “unconditional” vs. its “conditional” form.

Table 2: **Decomposition of the unconditional variance (across paths) of the conditional CCAPM deviations**, for a 1% fee on Stock 1 and 3% fee on Stock 2, at  $t = 25$ . The complements to 100% are contributed by the unconditional covariance between conditional expected liquidity change and conditional liquidity-risk premium. All parameters are set at their benchmark values indicated in Table 1.

	Proportion of variance due to:	
	Exp. liquidity change	Liquidity risk
Investor 1 Stock 1	91.0%	8.96%
Investor 1 Stock 2	83.4%	15.1%
Investor 2 Stock 1	85.4%	8.25%
Investor 2 Stock 2	88.1%	15.0%

for instance, study very thoroughly the effect of liquidity on corporate-bond expected returns and “find a strong effect of expected liquidity and equity market liquidity risk on expected corporate bond returns, while there is little evidence that corporate bond liquidity risk exposures explain expected corporate bond returns, even during the recent financial crisis.” The model shows that, here especially, empirical conclusions could vary a lot depending on conditioning.

## C Liquidity and empirical asset pricing

Liquidity fluctuations, in addition to current and expected liquidity, have been regarded as a source of risk, and as a risk that receives a price in the market place, which is the liquidity-risk premium. In a number of empirical papers, tests were conducted on a cross-section of monthly portfolio returns, looking at changes in market liquidity as a new risk factor. Brennan and Subrahmanyam (1996) and Brennan *et al.* (2012) base their tests on Kyle’s lambda as a measure of liquidity. Brennan *et al.* (1998) use volume of trading. Acharya and Pedersen (2005) and Bongaerts *et al.* (2012) use ILLIQ of Amihud (2002) as a liquidity measure. Pástor and Stambaugh (2003) use a different measure based on the intensity of return reversals during periods of high vs. low volume.

Our model CAPM (13), however, says that liquidity risk should be captured by the fluctuations in a combination of shadow prices and stock prices, not by the empirical variables used so far. It says further that these shadow liquidity measures should be specific to each security. But the shadow prices are generally not observable and empiricists in a test of the CAPM would replace them with proxies. We now use our cross-section to appraise the theoretical validity of various proxy measures of liquidity, according to our model.

The proxies we consider are the following:<sup>44</sup>

<sup>44</sup>Easley *et al.* (2002) define and utilize the PIN measure. Lesmond *et al.* (1999) suggest a measure (LOT) of transaction costs that does not depend on information about quotes or the order book. LOT uses instead the frequency of zero returns to estimate an implicit trading cost. The frequency of no trade has been proposed by Fong *et al.* (2010) under the acronym “FHT”. We do not reproduce these here.

1. “Amivest” (or LIQ):<sup>45</sup> trading volume (absolute terms) / realized return
2. Effective Spread: same definition as for Figure 7, Panel (e).  
Highest bid – lowest ask of the two agents.
3. Kyle’s Lambda: same definition as for subsection V.C. Ratio of the covariance between next period’s stock return and next period’s unexpected signed trading dollar volume and the variance of the unexpected signed trading dollar volume.<sup>46</sup>
4. Inverse Pastor-Stambaugh measure:<sup>47</sup> defined as the ratio of the following two quantities: trading volume (absolute terms) times sign(last period’s  $(t - 1)$  realized return) over unexpected realized return at time  $t$  (i.e., realized return minus expected return).
5. Volume: simply defined as the absolute value of the difference between the first agent’s stock holdings the period before and time  $t$ .

We use the proxies in such a way as to give each of them the maximum ability to act as a substitute for the correct liquidity measure. Very much in line with Equation (18) of Acharya and Pedersen (2005) or with Equations (7, 8, 9) of Pástor and Stambaugh (2003),<sup>48</sup> we posit a linear relationship between the liquidity change of our model and the proxy:

$$(1 - R_{l,t+1,i,j}) \times \frac{S_{t+1,i,j}}{S_{t,i}} - (1 - R_{l,t,i}) \times r_{t+1,1} = a_{l,i} + b_{l,i} \times \text{liqProxy}_{t+1,i,j} + \text{resid.} \quad (15)$$

and replace in the CCAPM equation the theoretically correct liquidity change with  $a_{l,i} + b_{l,i} \times \text{liqProxy}_{t+1,i,j}$ . Then, at  $t = 25$ , we choose the coefficients  $a_{l,i}$  and  $b_{l,i}$  to minimize the mean (across paths) squared CCAPM error introduced by the proxy being used in lieu of the correct theoretical concept.

Table 3 exhibits the statistics of the error produced by the use of the proxy. It leads to the following conclusions. First, the mean errors are quite small, much smaller in fact than the mean CCAPM deviations created by the friction and displayed in Figure 12, which shows that the proxies are quite adequate for unconditional tests of the CCAPM. Second, the RMSEs are not small compared to the root mean squared CCAPM deviations that are created by the friction and that we mentioned in the previous section. This shows that the proxies are inadequate for conditional tests. Third, no particular proxy seems to perform better than the others. Fourth, the correlations between the theoretically correct

<sup>45</sup>LIQ rather than Amihud’s ILLIQ is used because the trading volume can be zero in our case which would make the liquidity measure equal to infinity.

<sup>46</sup>We remind the reader that the separation between anticipated and unanticipated component is made by means of the conditional expected value provided by the model.

<sup>47</sup>The inverse of the PS measure is used because the trading volume can be zero in our case which would make the liquidity measure equal to infinity.

<sup>48</sup>Pástor and Stambaugh (2003) develop a proxy for market-wide liquidity, not stock specific liquidity.

Table 3: **Quality of CCAPM fit for various proxies.** The table shows the unconditional root mean squared error of the conditional CCAPM and the unconditional CCAPM error, introduced by the use of the proxy. It also shows in the last columns two correlations: “Correlation 1” is the mean conditional correlation and “Correlation 2” is the unconditional correlation between the left-hand side and the right-hand side variable of Equation (15). The term “unconditional” refers to a correlation computed at  $t = 25$  across the 50000 paths of the simulation. Stock 1 carries a fee of 1% and stock 2 a fee of 3%. All parameters are as in Table 1.

<b>CCAPM Deviation Investor 1 Stock 1</b>				
% return/year				
	RMSE	Mean error	Correlation 1	Correlation 2
Amivest	0.3521	0.0007	-0.1152	-0.0015
Effective Spread	0.2311	0.0000	0.1983	0.3000
Kyle’s Lambda	0.3569	0.0017	0.4203	0.0261
Inverse Pastor-Stambaugh	0.3314	-0.0003	-0.0331	-0.0244
Volume	0.3217	0.0000	-0.2265	-0.0926
<b>CCAPM Deviation Investor 1 Stock 2</b>				
	RMSE	Mean error	Correlation 1	Correlation 2
Amivest	0.5950	0.0019	-0.1929	-0.0023
Effective Spread	0.5838	0.0000	0.3130	0.2972
Kyle’s Lambda	0.5958	-0.0017	-0.0015	-0.0029
Inverse Pastor-Stambaugh	0.5586	-0.0001	-0.0579	-0.0753
Volume	0.5798	0.0000	-0.3036	-0.1427

liquidity measure on the left-hand side of Equation (15) and a proxy on the right-hand side do not seem to provide any indication concerning the CCAPM fit one gets with that proxy; these correlations are not sufficient to gauge the quality of a proxy.

We end this section with a methodological remark. In empirical work, the gross rate of return on a security is commonly computed as  $\frac{\delta_{t+1,i,j} + S_{t+1,i,j}}{S_{t,i}}$  between fixed, equally spaced calendar points in time, between which the security is supposed to be held. However, the concept of holding period is quite arbitrary. Absent transactions fees, since investors are ready to trade at any time, the only holding period that would make sense is one approaching zero. In the presence of transactions fees, armed with the current model we have determined the holding period endogenously. In a model with more than two investors, holding periods would generally differ across people. With two investors, who can only trade with each other, the holding periods are identical across investors but, between trades, their desires to trade differ. That desire is reflected in the investor-specific shadow prices, which must be taken into ac-



count if rates of return continue to be based on fixed, equally spaced points in time. If one wanted to test our CCAPM, a better way would be not to use the standard concept of rate of return measured between fixed points in time. Instead, one would use transactions prices only, which do not occur at fixed time intervals, and one would substitute out in the model the values of the posted prices that are “not observed” for lack of transaction.<sup>49</sup> That, however, is not the way empirical tests have been conducted by previous authors. Further work is needed to develop the econometric method.

## VII Conclusion

We have produced a new method to compute financial-market equilibria in the presence of proportional transactions fees. For a given rate of transactions fees, our method delivers the optimal, market-clearing moves of each investor and the resulting posted and transactions prices. In our model, everyone behaves optimally in reaction to the information they receive.

We have concluded that transactions fees have a strong effect on investors’ asset holdings, that deviations in asset prices from a frictionless economy are equal at most to current transactions fees only plus all future state-price differences. We have studied the behavior over time of trades, posted prices and asset holdings.

We have presented a CCAPM model extended for transactions fees, identified the risk factors and displayed their relative sizes and movements over time. We confirmed, however, the view expressed in prior work saying that explicitly observable transactions fees cannot account for the size of what is commonly measured as a liquidity premium. We have commented, in the light of our theoretical model, on the adequacy of extant empirical tests of CCAPMs that include a premium for liquidity risk. Shadow prices that properly capture liquidity are generally not observable but our model validates the variables often used in unconditional tests to proxy for time-varying liquidity, but not so for conditional tests. Further work is needed to develop the econometric method that would be most powerful given the theoretical equilibrium model.

Future theoretical work should aim to model an equilibrium in which trading would not be Walrasian. In it, the rate of transactions fees would not be a given and investors would submit limit and market orders. The behavior of the limit-order book would be obtained. This would be similar to the work of Parlour (1998), Foucault (1999), Foucault *et al.* (2005), Goettler *et al.* (2005) and Rosu (2009), except that trades would arrive at the time and in quantities of the investor’s choice, and would not be driven by an exogenous process.<sup>50</sup>

<sup>49</sup>See the discussion on page 89 of Hasbrouck (2007).

<sup>50</sup>Recently, Kühn and Stroh (2010) have used the dual approach to optimize portfolio choice in a limit-order market and may have shown the way to do that.

# Appendixes

## A Proof of the equation system of Section II.

The Lagrangian for problem (3) is:

$$\begin{aligned}
\mathcal{L}_l(\{\theta_{l,t-1,i}\}, \cdot, e_{l,t}, t) &= \sup_{c_{l,t}, \{\widehat{\theta}_{l,t,i}, \widehat{\theta}_{l,t,i}\}} \inf_{\phi_{l,t}} u_l(c_{l,t}, t) \\
&+ \sum_{j=1}^{K_t} \pi_{t,t+1,j} J_l \left( \left\{ \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right) \\
&\quad + \phi_{l,t} \left[ e_{l,t} + \sum_{i=1}^I \theta_{l,t-1,i} \delta_{t,i} - c_{l,t} + \zeta_{l,t} \right. \\
&\quad \left. - \sum_{i=1}^I \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) S_{t,i} (1 + \lambda_{i,t}) - \sum_{i=1}^I \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) S_{t,i} (1 - \varepsilon_{i,t}) \right] \\
&\quad + \sum_{i=1}^I \left[ \mu_{1,l,t,i} \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) + \mu_{2,l,t,i} \left( \theta_{l,t-1,i} - \widehat{\theta}_{l,t,i} \right) \right]
\end{aligned}$$

where  $\phi_{l,t}$  is obviously the Lagrange multiplier attached to the flow budget constraint (4) and  $\mu_1$  and  $\mu_2$  are the Lagrange multipliers attached to the inequality constraints (5). The Karush-Kuhn-Tucker first-order conditions are:

$$\begin{aligned}
u'_l(c_{l,t}, t) &= \phi_{l,t} \\
e_{l,t} + \sum_{i=1}^I \theta_{l,t-1,i} \delta_{t,i} - c_{l,t} + \zeta_{l,t} \\
- \sum_{i=1}^I \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) S_{t,i} (1 + \lambda_{i,t}) - \sum_{i=1}^I \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) S_{t,i} (1 - \varepsilon_{i,t}) &= 0 \\
\sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left( \left\{ \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right) &= 0 \\
&= \phi_{l,t} \times S_{t,i} \times (1 + \lambda_{i,t}) - \mu_{1,l,t,i} \\
\sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left( \left\{ \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right) &= 0 \\
&= \phi_{l,t} \times S_{t,i} \times (1 - \varepsilon_{i,t}) + \mu_{2,l,t,i} \\
\widehat{\theta}_{l,t,i} \leq \theta_{l,t-1,i} \leq \widehat{\theta}_{l,t,i}; \mu_{1,l,t,i} \geq 0; \mu_{2,l,t,i} \geq 0 \\
\mu_{1,l,t,i} \times \left( \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) = 0; \mu_{2,l,t,i} \times \left( \theta_{l,t-1,i} - \widehat{\theta}_{l,t,i} \right) &= 0
\end{aligned} \tag{16}$$

where the last two equations are referred to as the “complementary-slackness” conditions. Two of the first-order conditions imply that

$$\phi_{l,t} \times S_{t,i} \times (1 + \lambda_{i,t}) - \mu_{1,l,t,i} = \phi_{l,t} \times S_{t,i} \times (1 - \varepsilon_{i,t}) + \mu_{2,l,t,i}$$

Therefore, we can merge two Lagrange multipliers into one,  $R_{l,t,i}$ , defined as:

$$\phi_{l,t} \times R_{l,t,i} \times S_{t,i} \triangleq \phi_{l,t} \times S_{t,i} \times (1 + \lambda_{i,t}) - \mu_{1,l,t,i} = \phi_{l,t} \times S_{t,i} \times (1 - \varepsilon_{i,t}) + \mu_{2,l,t,i}$$

and recognize one first-order condition that replaces two of them:

$$\begin{aligned} \sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left( \left\{ \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right) \\ = \phi_{l,t} \times R_{l,t,i} \times S_{t,i} \end{aligned} \quad (17)$$

In order to eliminate the value function from the first-order conditions, we differentiate the Lagrangian with respect to  $\theta_{l,t-1,i}$  and then make use of (17):

$$\begin{aligned} \frac{\partial J_l}{\partial \theta_{l,t-1,i}} &= \frac{\partial L_l}{\partial \theta_{l,t-1,i}} \\ &= - \sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left( \left\{ \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right) \\ &\quad + \phi_{l,t} [\delta_{t,i} + S_{t,i} \times (1 + \lambda_{i,t}) + S_{t,i} \times (1 - \varepsilon_{i,t})] - \mu_{1,l,t,i} + \mu_{2,l,t,i} \\ &= - \sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left( \left\{ \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right) \\ &\quad + \phi_{l,t} \delta_{t,i} + 2\phi_{l,t} \times R_{l,t,i} \times S_{t,i} \\ &= \phi_{l,t} \times (\delta_{t,i} + R_{l,t,i} \times S_{t,i}) \end{aligned}$$

so that the first-order conditions can also be written:

$$\begin{aligned} u'_l(c_{l,t}, t) &= \phi_{l,t} \\ e_{l,t} + \sum_{i=1}^I \theta_{l,t-1,i} \delta_{t,i} - c_{l,t} - \sum_{i=1}^I \left( \widehat{\theta}_{l,t,i} + \widehat{\theta}_{l,t,i} - 2 \times \theta_{l,t-1,i} \right) \times R_{l,t,i} \times S_{t,i} + \zeta_{l,t} &= 0 \\ \sum_{j=1}^{K_t} \pi_{t,t+1,j} \times \phi_{l,t+1,j} \times (\delta_{t+1,i,j} + R_{l,t+1,i,j} \times S_{t+1,i,j}) &= \phi_{l,t} \times R_{l,t,i} \times S_{t,i} \end{aligned} \quad (18)$$

$$\begin{aligned} \widehat{\theta}_{l,t,i} &\leq \theta_{l,t-1,i} \leq \widehat{\theta}_{l,t,i} \\ 1 - \varepsilon_{i,t} &\leq R_{l,t,i} \leq 1 + \lambda_{i,t}; \\ (-R_{l,t,i} + 1 + \lambda_{i,t}) \times (\widehat{\theta}_{l,t,i} - \theta_{l,t-1,i}) &= 0 \\ (R_{l,t,i} - (1 - \varepsilon_{i,t})) \times (\theta_{l,t-1,i} - \widehat{\theta}_{l,t,i}) &= 0 \end{aligned}$$

As has been noted by Dumas and Lyasoff (2012) in a different context, the system made of (18) and (6) above has a drawback. It must be solved simultaneously (or globally) for all nodes of all times. As written, it cannot be solved recursively in the backward way because the unknowns at time  $t$  include consumptions at time  $t$ ,  $c_{l,t}$ , whereas the third subset of equations in (18) if rewritten as:

$$\begin{aligned} \sum_{j=1}^{K_t} \pi_{t,t+1,j} \times u'_l(c_{l,t+1,j}, t) \times [\delta_{t+1,i,j} + R_{l,t+1,i,j} \times S_{t+1,i,j}] \\ = \phi_{l,t} \times R_{l,t,i} \times S_{t,i}; l = 1, 2 \end{aligned}$$

can be seen to be a restriction on consumptions at time  $t + 1$ , which at time  $t$  would already be solved for.

In order to “synchronize” the solution algorithm of the equations and allow recursivity, we first shift all first-order conditions, except the third one, forward in time and, second, we no longer make explicit use of the investor’s positions  $\theta_{l,t-1,i}$  held when entering time  $t$ , focusing instead on the positions  $\theta_{l,t+1,i,j}$  ( $\sum_{l=1,2} \theta_{l,t+1,i,j} = 0$  or 1) held when exiting time  $t + 1$ , which are carried backward. Regrouping equations in that way, substituting the pot-balance condition (7) and appending market-clearing condition (6) leads to the equation system of Section II.

## B Time 0

After solving the equation system of Section II, it remains to solve at time 0 the following equation system ( $t = -1, t + 1 = 0$ ) *from which the kernel conditions only have been removed* (because they were already solved as part of the backward induction):<sup>51</sup>

1. First-order conditions for time 0 consumption:

$$u'_l(c_{l,0}, 0) = \phi_{l,0}$$

2. The set of time-0 flow budget constraints for all investors and all states of nature of that time:

$$\begin{aligned} e_{l,0} + \sum_{i=1}^I \theta_{l,-1,i} \delta_{0,i} - c_{l,0} - \sum_{i=1}^I (\theta_{l,0,i} - \theta_{l,-1,i}) \times R_{l,0,i} \times S_{0,i} \\ + \sum_{i=1}^I (\widehat{\theta}_{l',0,i} - \theta_{l',-1,i}) S_{0,i} \lambda_{i,0} - \sum_{i=1}^I (\widehat{\theta}_{l',0,i} - \theta_{l',-1,i}) S_{0,i} \varepsilon_{i,0} = 0 \end{aligned}$$

3. Definitions:

$$\theta_{l,0,i} = \widehat{\theta}_{l,0,i} + \widehat{\theta}_{l,0,i} - \theta_{l,-1,i}$$

<sup>51</sup>There could be several possible states  $j$  at time 0 but we have removed the subscript  $j$ .

4. Complementary-slackness conditions:

$$\begin{aligned} (-R_{l,0,i} + 1 + \lambda_{i,0}) \times (\widehat{\theta}_{l,0,i} - \theta_{l,-1,i}) &= 0 \\ (R_{l,0,i} - (1 - \varepsilon_{i,0})) \times (\theta_{l,-1,i} - \widehat{\theta}_{l,0,i}) &= 0 \end{aligned}$$

5. Market-clearing restrictions:

$$\sum_{l=1,2} \theta_{l,-1,i} = 0 \text{ or } 1$$

This system can be handled in one of two ways:

1. We can either solve for the unknowns  $\left\{ c_{l,0}, \theta_{l,-1,i}, \widehat{\theta}_{l,0,i}, \widehat{\theta}_{l,0,i}; l = 1, 2; j = 1, \dots, K_t \right\}$  as functions of  $\{\phi_{l,0}\}$  and  $\{R_{l,0,i}\}$ . If we plot  $\theta_{l,-1,i}$  as functions of  $\{\phi_{l,0}\}$  and  $\{R_{l,0,i}\}$ , we have the “Negishi map.”<sup>52</sup> If it is invertible, we can then invert that Negishi map to obtain the values of  $\{\phi_{l,0}\}$  and  $\{R_{l,0,i}\}$  such that  $\theta_{l,-1,i} = \widehat{\theta}_{l,i}$ . If the values  $\widehat{\theta}_{l,i}$  fall outside the image set of the Negishi map, there simply does not exist an equilibrium as one investor would, at equilibrium prices, be unable to repay her debt to the other investor.
2. Or we drop the market-clearing equation also and solve directly this system for the unknowns:  $\left\{ c_{l,0}, \phi_{l,0}, R_{l,0,i}, \widehat{\theta}_{l,0,i}, \widehat{\theta}_{l,0,i}; l = 1, 2; j = 1, \dots, K_t \right\}$  with  $\theta_{l,-1,i}$ , replaced in the system by the given  $\widehat{\theta}_{l,i}$ .

In this paper, the second method has been used.

## C Scale-invariance property

Assuming zero endowments  $e_{l,t,j} = 0$ , we now show that all the nodes of a given point in time, which differ only by their value of the exogenous payout, are isomorphic to each other, where the isomorphy simply means that we can factor out the payout on the stock. We provide the proof for the binomial case  $K_t = 2$  ( $j = 1, 2$ ) with two securities, but the result is valid for any number of states, as long as there is only one payout.<sup>53</sup>

*Time T-1*

Rewriting the investors’ consumptions in terms of consumption shares  $\omega_{l,T,j}$ , given the fact that we have zero transactions fees in the last period  $T$ , using the

<sup>52</sup>For a definition of the “Negishi map” in a market with frictions, see Dumas and Lyasoff (2012).

<sup>53</sup>In the case of Subsection VI.B where we have two stocks, the scale-invariance property only holds with respect to the aggregate payout. States that differ in the relative payout of the two stocks require separate computations.

first-order conditions for consumption, and the system of equations to be solved at time  $T - 1$  is simply:

$$\begin{aligned}
& \sum_{i=1}^2 \theta_{l,T-1,i} \delta_{T,i,j} - \omega_{l,T,j} \times \delta_{T,2,j} = 0 \\
& \beta_1 \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{1,T,j}}{\omega_{1,T-1}} \times \frac{\delta_{T,2,j}}{\delta_{T-1,2}} \right)^{-\gamma_1} \\
& = \beta_2 \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{2,T,j}}{\omega_{2,T-1}} \times \frac{\delta_{T,2,j}}{\delta_{T-1,2}} \right)^{-\gamma_2} \\
& \frac{\beta_1}{R_{1,T-1,i}} \sum_{j=1,2} \pi_{T-1,T,j} \times \left( \frac{\omega_{1,T,j}}{\omega_{1,T-1}} \times \frac{\delta_{T,2,j}}{\delta_{T-1,2}} \right)^{-\gamma_1} \times \delta_{T,i,j} \\
& = \frac{\beta_2}{R_{2,T-1,i}} \sum_{j=1,2} \pi_{T-1,T,j} \times \left( \frac{\omega_{2,T,j}}{\omega_{2,T-1}} \times \frac{\delta_{T,2,j}}{\delta_{T-1,2}} \right)^{-\gamma_2} \times \delta_{T,i,j} \\
& \sum_{l=1,2} \theta_{l,T-1,1} = 0; \quad \sum_{l=1,2} \theta_{l,T-1,2} = 1
\end{aligned}$$

with unknowns  $\{\omega_{l,T,j}; l = 1, 2; j = 1, 2\}$ ,  $\{\theta_{l,T-1,i}; l = 1, 2; i = 1, 2\}$ .

Letting:  $\frac{\delta_{t+1,2,1}}{\delta_{t,2}} = u$  as well as  $\frac{\delta_{t+1,2,2}}{\delta_{t,2}} = d$ , we can solve the flow budget equations:

$$\begin{cases} \theta_{l,T-1,1} + \delta_{T-1,2} \times (\theta_{l,T-1,2} u - \omega_{l,T,1} \times u) = 0 \\ \theta_{l,T-1,1} + \delta_{T-1,2} \times (\theta_{l,T-1,2} d - \omega_{l,T,2} \times d) = 0 \end{cases}$$

The solution for the holdings is:

$$\theta_{l,T-1,1} = u \frac{\delta_{T-1,2}}{d - u} (d \times \omega_{l,T,1} - \omega_{l,T,2} \times d) \quad (19)$$

$$\theta_{l,T-1,2} = \frac{1}{d - u} (-\omega_{l,T,1} \times u + \omega_{l,T,2} \times d) \quad (20)$$

Rewriting the kernel conditions and reducing the system using (19) and (20), we get a system with unknowns  $\{\omega_{l,T,j}; l = 1, 2; j = 1, 2\}$  only:

$$\begin{aligned}
& \beta_1 \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{1,T,j}}{\omega_{1,T-1}} \right)^{-\gamma_1} r_j^{-\gamma_1} = \beta_2 \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{2,T,j}}{\omega_{2,T-1}} \right)^{-\gamma_2} r_j^{-\gamma_2} \\
& \frac{\beta_1}{R_{1,T-1,i}} \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{1,T,j}}{\omega_{1,T-1}} \right)^{-\gamma_1} r_j^{-\gamma_1+1} \\
& = \frac{\beta_2}{R_{2,T-1,i}} \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{2,T,j}}{\omega_{2,T-1}} \right)^{-\gamma_2} r_j^{-\gamma_2+1}
\end{aligned}$$

$$\sum_{l=1,2} \theta_{l,T-1,1} = 0; \quad \sum_{l=1,2} \theta_{l,T-1,2} = 1$$

$$\frac{1}{d-u} (d \times \omega_{1,T,1} \times u - u \times \omega_{1,T,2} \times d) + \frac{1}{d-u} (d \times \omega_{2,T,1} \times u - u \times \omega_{2,T,2} \times d) = 0$$

$$\frac{1}{d-u} (-\omega_{1,T,1} \times u + \omega_{1,T,2} \times d) + \frac{1}{d-u} (-\omega_{2,T,1} \times u + \omega_{2,T,2} \times d) = 1$$

where  $r_j = u$  for  $j = 1$  and  $r_j = d$  for  $j = 2$ .

Importantly this system of equations *does not depend on the current or future levels of payout*, i.e. it is enough to solve the system for one node at time  $T - 1$  as long as  $u$  and  $d$  are not state (node) dependent.

After solving this system, one can compute the implied holdings and asset prices. From (20) we get that the stock holdings are independent of  $T - 1$  payout, while from (19) we know that the bond holdings are scaled by the  $T - 1$  payout:

$$\theta_{l,T-1,1} = \delta_{T-1,2} \times \check{\theta}_{l,T-1,1}, \quad (21)$$

where  $\check{\theta}_{l,T-1,1}$  denotes the normalized bond holdings for  $\delta_{T-1,2} = 1$ . Moreover, we get that the bond price does not depend on  $T - 1$  endowment:

$$S_{T-1,1} = \beta_1 \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{1,T,j}}{\omega_{1,T-1}} \right)^{-\gamma_1} r_j^{-\gamma_1},$$

and that the stock price is scaled by the  $T - 1$  endowment:

$$S_{T-1,2} = \delta_{T-1,2} \times \left[ \frac{\beta_1}{R_{1,T-1,i}} \sum_{j=1}^2 \pi_{T-1,T,j} \left( \frac{\omega_{1,T,j}}{\omega_{1,T-1}} \right)^{-\gamma_1} \frac{r_j^{-\gamma_1+1}}{2} \right]$$

$$\triangleq \delta_{T-1,2} \times \check{S}_{T-1,2}, \quad (22)$$

where  $\check{S}_{T-1,2}$  denotes the normalized price for  $\delta_{T-1,2} = 1$ .

*Time  $t < T-1$*

For time  $t < T-1$  the system of equations is the system of Section II. Rewriting  $c_{l,t+1,j} = \omega_{l,t+1,j} \times \delta_{t+1,2}$ , replacing  $S_{t+1,2}$  and  $\theta_{l,t+1,1}$  with expressions (22) and (21), the flow budget equations are:

$$\left\{ \begin{array}{l} \omega_{l,t+1,1}u + \check{\theta}_{l,t+1,1,1}u + (\theta_{l,t+1,2,1} - \theta_{l,t,2}) \check{S}_{t+1,2,1}uR_{l,t+1,2,1} \\ = \frac{\theta_{l,t,1}}{\delta_{t,2}} + \theta_{l,t,2}u + \left( \widehat{\theta}_{l',t+1,2,1} - \theta_{l',t,2} \right) \check{S}_{t+1,2,1}u\lambda_{2,t+1,1} \\ - \left( \widehat{\theta}_{l',t+1,2,1} - \theta_{l',t,2} \right) \check{S}_{t+1,2,1}u\varepsilon_{2,t+1,1} \\ \omega_{l,t+1,2}d + \check{\theta}_{l,t+1,1,2}d + (\theta_{l,t+1,2,2} - \theta_{l,t,2}) \check{S}_{t+1,2,2}dR_{l,t+1,2,2} \\ = \frac{\theta_{l,t,1}}{\delta_{t,2}} + \theta_{l,t,2}d + \left( \widehat{\theta}_{l',t+1,2,2} - \theta_{l',t,2} \right) \check{S}_{t+1,2,2}d\lambda_{2,t+1,2} \\ - \left( \widehat{\theta}_{l',t+1,2,2} - \theta_{l',t,2} \right) \check{S}_{t+1,2,2}d\varepsilon_{2,t+1,2} \end{array} \right.$$

or:

$$\begin{bmatrix} 1 & u \left( 1 + \check{S}_{t+1,2,1} R_{l,t+1,2,1} \right) \\ 1 & d \left( 1 + \check{S}_{t+1,2,2} R_{l,t+1,2,2} \right) \end{bmatrix} \begin{bmatrix} \frac{\theta_{l,t,1}}{\delta_{t,2}} \\ \theta_{l,t,2} \end{bmatrix} = \begin{bmatrix} u \kappa_{l,t+1,1} \\ d \kappa_{l,t+1,2} \end{bmatrix}$$

where

$$\begin{aligned} \kappa_{l,t+1,j} &= \omega_{l,t+1,j} + \check{\theta}_{l,t+1,1,j} + (\theta_{l,t+1,2,j} - \theta_{l,t,j}) \check{S}_{t+1,2,j} R_{l,t+1,2,j} \\ &\quad - \left( \widehat{\theta}_{l',t+1,2,j} - \theta_{l',t,j} \right) \check{S}_{t+1,2,j} \lambda_{2,t+1,j} + \left( \widehat{\theta}_{l',t+1,2,j} - \theta_{l',t,j} \right) \check{S}_{t+1,2,j} \varepsilon_{2,t+1,j} \end{aligned}$$

Solving for the holdings:

$$\frac{\theta_{l,t,1}}{\delta_{t,2}} = u \frac{1}{d-u} (d \times \kappa_{l,t+1,1} - \kappa_{l,t+1,2} \times d) \quad (23)$$

$$\theta_{l,t,2} = \frac{1}{d-u} (-\kappa_{l,t+1,1} \times u + \kappa_{l,t+1,2} \times d) \quad (24)$$

Rewriting the kernel conditions, we can write the system as:

$$\begin{aligned} \beta_1 \sum_{j=1}^2 \pi_{t,t+1,j} \left( \frac{\omega_{1,t+1,j}}{\omega_{1,t}} \right)^{-\gamma_1} r_j^{-\gamma_1} &= \beta_2 \sum_{j=1}^2 \pi_{t,t+1,j} \left( \frac{\omega_{2,t+1,j}}{\omega_{2,t}} \right)^{-\gamma_2} r_j^{-\gamma_2} \\ \frac{\beta_1}{R_{1,t,2}} \sum_{j=1}^2 \pi_{t,t+1,j} \left( \frac{\omega_{1,t+1,j}}{\omega_{1,t}} \right)^{-\gamma_1} r_j^{-\gamma_1} &\left( R_{1,t+1,2,j} \times \bar{S}_{t+1,2,j} \times r_j + \frac{r_j}{2} \right) \\ = \frac{\beta_1}{R_{2,t,2}} \sum_{j=1}^2 \pi_{t,t+1,j} \left( \frac{\omega_{2,t+1,j}}{\omega_{2,t}} \right)^{-\gamma_2} r_j^{-\gamma_2} &\left( R_{2,t+1,2,j} \times \bar{S}_{t+1,2,j} \times r_j + \frac{r_j}{2} \right) \\ \theta_{l,t+1,2,j} &= \widehat{\theta}_{l,t+1,2,j} + \widehat{\widehat{\theta}}_{l,t+1,2,j} - \theta_{l,t,2} \\ (-R_{l,t+1,2,j} + 1 + \lambda_{2,t+1,j}) \times &\left( \widehat{\theta}_{l,t+1,2,j} - \theta_{l,t,2} \right) = 0 \\ (R_{l,t+1,2,j} - (1 - \varepsilon_{2,t+1,j})) \times &\left( \theta_{l,t,2} - \widehat{\widehat{\theta}}_{l,t+1,2,j} \right) = 0 \\ \sum_{l=1,2} \theta_{l,T-1,1} &= 0; \quad \sum_{l=1,2} \theta_{l,T-1,2} = 1 \end{aligned}$$

with unknowns  $\left\{ \omega_{l,t+1,j}; R_{l,t+1,2,j}; \widehat{\theta}_{l,t+1,2,j}; \widehat{\widehat{\theta}}_{l,t+1,2,j}; l = 1, 2; j = 1, 2 \right\}$ . The holdings implied are given by (23) and (24). One can show that the payout  $\delta_{t+1,2}$  cancels out in the market clearing conditions for the bond. Thus, the full system *does not depend on the level of the payout*  $\delta_{t+1,2}$ , only on  $u$  as well as  $d$ , and therefore we only need to solve the system at one node at time  $t$ .



As backward interpolated values we use the bond price  $S_{t+1,1,j}$  and stock holdings  $\theta_{l,t+1,2,j}$  as well as the normalized stock price  $\hat{S}_{t+1,2,j}$  and normalized bond holdings  $\hat{\theta}_{l,t+1,1,j}$ . After solving the system we can compute the implied time  $t$  holdings and prices. Again, holdings in the bond and the stock price are scaled by  $\delta_{t,2}$ , while the holdings in the stock and the bond price are not scaled. *Using backward induction the scaling invariance holds for any time  $t$ .*

## D Proof of Proposition 3

The proof is by induction.

At date  $t = T - 1$ , the present value of payouts  $\delta$  from the point of view of investor  $l$  is given by:

$$\hat{S}_{T-1,i,l} = \mathbb{E}_{T-1} \left[ \frac{\phi_{l,T}}{\phi_{l,T-1}} \times \delta_{T,i} \right].$$

whereas Equation (9) applied to time  $T - 1$  is:

$$\begin{aligned} R_{l,T-1,i} \times S_{T-1,i} &= \mathbb{E}_{T-1} \left[ \frac{\phi_{l,T}}{\phi_{l,T-1}} \times \delta_{T,i} \right] \\ &= \hat{S}_{T-1,i,l} \end{aligned} \tag{25}$$

At  $t = T - 2$ , the present value of payouts is:

$$\hat{S}_{T-2,i,l} = \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}}{\phi_{l,T-2}} \times \left( \delta_{T-1,i} + \hat{S}_{T-1,i,l} \right) \right]$$

whereas Equation (9) applied to time  $T - 2$  is:

$$\begin{aligned} R_{l,T-2,i} \times S_{T-2,i} &= \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}}{\phi_{l,T-2}} \times \left( \delta_{T-1,i} + R_{l,T-1,i} \times S_{T-1,i} \right) \right] \\ &= \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}}{\phi_{l,T-2}} \times \left( \delta_{T-1,i} + \hat{S}_{T-1,i,l} \right) \right] \\ &= \hat{S}_{T-2,i,l} \end{aligned}$$

where we used equation (25) to replace  $R_{l,T-1,i} \times S_{T-1,i}$ .

By an induction argument one can show the final result (10).

## E Proof of Proposition 4

The proof is by induction.

At date  $t = T - 1$ , the stock price in an economy without transactions fees is given by:

$$S_{T-1}^* = \mathbb{E}_{T-1} \left[ \frac{\phi_{l,T}^*}{\phi_{l,T-1}^*} \delta_T \right]$$

whereas Equation (9) applied to time  $T - 1$  is:

$$R_{l,T-1} \times S_{T-1} = \mathbb{E}_{T-1} \left[ \frac{\phi_{l,T}}{\phi_{l,T-1}} \delta_T \right]$$

which can be rewritten as:

$$\begin{aligned} R_{l,T-1} \times S_{T-1} &= \mathbb{E}_{T-1} \left[ \frac{\phi_{l,T}^*}{\phi_{l,T-1}^*} \delta_T \right] + \mathbb{E}_{T-1} \left[ \left( \frac{\phi_{l,T}}{\phi_{l,T-1}} - \frac{\phi_{l,T}^*}{\phi_{l,T-1}^*} \right) \delta_T \right] \\ &= \mathbb{E}_{T-1} \left[ \frac{\phi_{l,T}^*}{\phi_{l,T-1}^*} \delta_T \right] + \mathbb{E}_{T-1} [\Delta\phi_{l,T} \times \delta_T] \end{aligned}$$

where we defined:

$$\Delta\phi_{l,T} \triangleq \frac{\phi_{l,T}}{\phi_{l,T-1}} - \frac{\phi_{l,T}^*}{\phi_{l,T-1}^*}.$$

We can thus derive the following relation between the stock price in a zero-transactions fees economy  $S_{T-1}^*$  and the stock price in an economy with transactions fees  $S_{T-1}$ :

$$R_{l,T-1} \times S_{T-1} - S_{T-1}^* = \mathbb{E}_{T-1} [\Delta\phi_{l,T} \delta_T] \quad (26)$$

At  $t = T - 2$ , the stock price in an economy without transactions costs is given by:

$$S_{T-2}^* = \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}^*}{\phi_{l,T-2}^*} (\delta_{T-1} + S_{T-1}^*) \right]$$

whereas Equation (9) applied to time  $T - 2$  is:

$$R_{l,T-2} \times S_{T-2} = \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}}{\phi_{l,T-2}} (\delta_{T-1} + R_{l,T-1} \times S_{T-1}) \right]$$

Replacing  $R_{l,T-1} \times S_{T-1}$  with expression (26), this can be rewritten as:

$$\begin{aligned} R_{l,T-2} \times S_{T-2} &= \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}}{\phi_{l,T-2}} (\delta_{T-1} + S_{T-1}^* + \mathbb{E}_{T-1} [\Delta\phi_{l,T} \delta_T]) \right] \\ &= \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}^*}{\phi_{l,T-2}^*} (\delta_{T-1} + S_{T-1}^*) \right] \\ &\quad + \mathbb{E}_{T-2} [\Delta\phi_{l,T-1} (\delta_{T-1} + S_{T-1}^*)] + \mathbb{E}_{T-2} \left[ \frac{\phi_{l,T-1}}{\phi_{l,T-2}} \Delta\phi_{l,T} \delta_T \right] \\ &= S_{T-2}^* \\ &\quad + \mathbb{E}_{T-2} \left[ \Delta\phi_{l,T-1} (\delta_{T-1} + S_{T-1}^*) + \frac{\phi_{l,T-1}}{\phi_{l,T-2}} \Delta\phi_{l,T} \delta_T \right] \end{aligned}$$

We can thus derive the following relation between the stock price in a zero-transactions fees economy  $S_{T-2}^*$  and the stock price in an economy with transactions fees  $S_{T-2}$ :

$$R_{l,T-2} \times S_{T-2} - S_{T-2}^* = \mathbb{E}_{T-2} \left[ \Delta\phi_{l,T-1} (\delta_{T-1} + S_{T-1}^*) + \frac{\phi_{l,T-1}}{\phi_{l,T-2}} \Delta\phi_{l,T} \delta_T \right]$$

By an induction argument one can show the final result (11).

## F Proof of the CCAPM Equation (13)

The derivation is straightforward. We start from:

$$1 = \frac{1}{R_{l,t,i}\phi_{l,t}} \sum_{j=u,d} \pi_{t,t+1,j} \phi_{l,t+1,j} \times \frac{\delta_{t+1,i,j} + R_{l,t+1,i,j} \times S_{t+1,i,j}}{S_{t,i}};$$

$$\frac{1}{r_{t+1,1}} = \frac{1}{\phi_{l,t}} \sum_{j=u,d} \pi_{t,t+1,j} \phi_{l,t+1,j};$$

Then:

$$r_{t+1,1} = \frac{1}{R_{l,t,i}} \sum_{j=u,d} \pi_{t,t+1,j} \frac{\phi_{l,t+1,j}}{\sum_{j=u,d} \pi_{t,t+1,j} \phi_{l,t+1,j}} \times \frac{\delta_{t+1,i,j} + R_{l,t+1,i,j} \times S_{t+1,i,j}}{S_{t,i}}$$

$$R_{l,t,i} \times r_{t+1,1} = \sum_{j=u,d} \pi_{t,t+1,j} \frac{\phi_{l,t+1,j}}{\sum_{j=u,d} \pi_{t,t+1,j} \phi_{l,t+1,j}} \times \frac{\delta_{t+1,i,j} + S_{t+1,i,j}}{S_{t,i}}$$

$$+ \sum_{j=u,d} \pi_{t,t+1,j} \frac{\phi_{l,t+1,j}}{\sum_{j=u,d} \pi_{t,t+1,j} \phi_{l,t+1,j}} \times \frac{(R_{l,t+1,i,j} - 1) \times S_{t+1,i,j}}{S_{t,i}}$$

$$R_{l,t,i} \times r_{t+1,1} = \mathbb{E}_t [r_{t+1,i}] + \text{cov}_t \left( r_{t+1,i}, \frac{\phi_{l,t+1}}{\mathbb{E}_t [\phi_{l,t+1}]} \right)$$

$$+ \mathbb{E}_t \left[ (R_{l,t+1,i} - 1) \times \frac{S_{t+1,i}}{S_{t,i}} \right] + \text{cov}_t \left( (R_{l,t+1,i,j} - 1) \times \frac{S_{t+1,i}}{S_{t,i}}, \frac{\phi_{l,t+1}}{\mathbb{E}_t [\phi_{l,t+1}]} \right)$$

## References

- Acharya, V. and L. H. Pedersen, 2005, "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77, 375-4010.
- Admati, A., and P. Pfleiderer, 1988, "A theory of intraday patterns: Volume and price variability," *Review of Financial Studies*, 1, 3-40.
- Adrian, T., E. Etula and T. Muir, 2012, "Financial Intermediaries and the Cross-Section of Asset Returns," *The Journal of Finance*, forthcoming
- Akerlof, G., 1970, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84, 488-500.
- Amihud, Y. and H. Mendelson, 1986a, "Asset Pricing and the Bid-ask Spread," *Journal of Financial Economics*, 17, 223-249.
- Amihud, Y. and H. Mendelson, 1986b, "Liquidity and Stock Returns," *Financial Analysts Journal*, 42, 43-48.
- Amihud, Y., H. Mendelson and L. Pedersen, 2005, "Liquidity and Asset Prices," *Foundations and Trends in Finance*, 1, 269-364.
- Akerlof, G. A., 1970, "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 84, 488-500.
- Armand, P., J. Benoist and D. Orban, 2008, "Dynamic updates of the barrier parameter in primal-dual methods for nonlinear programming," *Computational Optimization and Applications*, 41, 1-25.
- Bhattacharya, U. and M. Spiegel, 1998, "Anatomy of a Market Failure: NYSE Trading Suspensions (1974-1988)," *Journal of Business & Economic Statistics*, 16, 216-226.
- Bongaerts, D., F. De Jong and J. Driessen, 2012, "An Asset Pricing Approach to Liquidity Effects in Corporate Bond Markets," Available at SSRN: <http://ssrn.com/abstract=1762564> or <http://dx.doi.org/10.2139/ssrn.1762564>
- Bouchard, B., 2002, "Utility Maximization on the Real Line under Proportional Transactions Costs," *Finance and Stochastics*, 6, 495-516.
- Brennan, M. J., T. Chordia, and A. Subrahmanyam, 1998, "Alternative Factor Specifications, Security Characteristics, and the Cross-section of Expected Returns," *Journal of Financial Economics* 49, 345-373.
- Brennan, M. J., T. Chordia, A. Subrahmanyam, and Q. Tong, 2012, "Sell-order liquidity and the cross-section of expected stock returns," *Journal of Financial Economics*, 105, 523-541.
- Brennan, M. J., and A. Subrahmanyam, 1996, "Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns," *Journal of Financial Economics*, 41, 441-464.
- Brennan, M. J., and Y. Xia, 2001, "Stock price volatility and equity premium," *Journal of Monetary Economics*, 47, 249-283.
- Brunnermeier, M. and L. H. Pedersen, 2008, "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201-2238.
- Buss, A., R. Uppal and G. Vilkov, 2011, "Asset Prices in General Equilibrium with Transactions Costs and Recursive Utility," working paper, University of Frankfurt.

- Campbell, J. Y., and R. J. Shiller, 1988, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195-228.
- Chordia, T., R. Roll and A. Subrahmanyam, 2008, "Liquidity and Market Efficiency," *Journal of Financial Economics*, 87.
- Clower, R. W. and D. W. Bushaw, 1954. "Price Determination in a Stock-Flow Economy," *Econometrica*, 22, 328-343.
- Constantinides, G. M., 1976a, "Optimal Portfolio Revision with Proportional Transactions Costs: Extension to HARA Utility Functions and Exogenous Deterministic Income," *Management Science*, 22, 921-923.
- Constantinides, G. M., 1976b, "Stochastic Cash Management with Fixed and Proportional Transactions Costs," *Management Science*, 22, 1320-1331.
- Constantinides, G. M., 1986, "Capital Market Equilibrium with Transactions Costs," *Journal of Political Economy*, 94, 842-862.
- Copeland, T. C., and D. Galai, 1983, "Information effects of the bid-ask spread," *The Journal of Finance*, 38, 1457-1469.
- Cuoco, D., 1997, "Optimal Consumption and Equilibrium Prices with Portfolio Constraints and Stochastic Income," *Journal of Economic Theory*, 72, 33-73.
- Cvitanic J. and I. Karatzas, 1996, "Hedging and Portfolio Optimization under Transactions Costs: a Martingale Approach," *Mathematical Finance*, 6, 113-165.
- Davis, M. H. and A. R. Norman, 1990, "Portfolio Selection with Transaction Costs," *Mathematics of Operations Research*, 15, 676-713.
- Deelstra, G., Pham, H. and N. Touzi, 2001, "Dual Formulation of the Utility Maximization Problem under Transaction Costs," *The Annals of Applied Probability*, 11, 1353-1383.
- de Jong, F. and B. Rindi, 2009, *The Microstructure of Financial Markets*, Cambridge University Press.
- Delgado, F., Dumas, B. and G. Puopolo, 2012, "Hysteresis Bands, Holding Period and Transactions Costs," working paper, INSEAD and Bocconi University.
- Demsetz, H., 1968, "The Cost of Transacting," *Quarterly Journal of Economics*, 82, 33-53.
- Dumas, B., 1989, "Two-person Dynamic Equilibrium in the Capital Market," *Review of Financial Studies*, 2 (1989), 2, 157-188.
- Dumas, B., 1992, "Capital Market Equilibrium and the Real Exchange Rate in a Spatially Separated World," *Review of Financial Studies*, 5, 153-180.
- Dumas, B. and E. Luciano, 1991, "An Exact Solution to a Dynamic Portfolio Choice Problem under Transactions Costs," *The Journal of Finance*, 46, 577-595.
- Dumas, B. and A. Lyasoff, 2012, "Incomplete-Market Equilibria Solved Recursively on an Event Tree," *The Journal of Finance*, LXVII, 1897-1940.
- Easley, D., S. Hvidkjaer and M. O'Hara, 2002, "Is Information Risk A Determinant of Asset Prices?" *The Journal of Finance*, 57, 2185-2221.

- Easley, D. S., and M. O'Hara, 1987, "Price, trade size, and information in securities markets," *Journal of Financial Economics*, 19, 69-90.
- Edirsinghe, C., Naik, V., and R. Uppal, 1993, "Optimal Replication of Options with Transactions Costs and Trading Restrictions," *The Journal of Financial and Quantitative Analysis*, 28, 117-138.
- Fong, K, C. W. Holden and C. A. Trzcinka, 2010, "Can Global Stock Liquidity Be Measured?" working paper, Indiana University.
- Foster, F. D. and S. Viswanathan, 1993, "Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models," *The Journal of Finance*, 1993, 48, 187-211.
- Foucault, T., 1999, "Order Flow Composition and Trading Costs in a Dynamic Limit Order Market," *Journal of Financial Markets*, 99-134.
- Foucault, T., Kadan, O., and E. Kandel, 2005. "Limit Order Book as a Market for Liquidity," *Review of Financial Studies* 18: 1171-1217.
- Garcia, C. and W. Zangwill, 1981, *Pathways to Solutions, Fixed Points and Equilibria*, Prentice-Hall, Englewood Cliffs.
- Gârleanu, N. and L. Pedersen, 2012, "Dynamic Trading with Predictable Returns and Transactions Costs," working paper, Haas School of Business.
- Gennotte, G. and A. Jung, 1994, "Investment Strategies under Transaction Costs: The Finite Horizon Case," *Management Science*, 40, 385-404.
- Gerhold, S., P. Guasoni, J. Muhle-Karbe and W. Schachermayer, 2011, "Transactions Costs, Trading Volume and the Liquidity Premium," working paper, Technische Universität, Wien.
- Glosten, L. R. and P. R. Milgrom, 1985, "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneous'ly Informed Agents," *Journal of Financial Economics*, 71-100.
- Goettler, R. L., Parlour, C. A., and U. Rajan, 2005, "Equilibrium in a Dynamic Limit Order Market," *The Journal of Finance*, 60: 2149-2192.
- Gromb, D. and D. Vayanos, 2002, "Equilibrium and Welfare in Markets with Constrained Arbitrageurs," *Journal of Financial Economics*. 66, 361-407.
- Grossman, S. J. and J. Stiglitz, 1980, "On the Impossibility of Informationally Efficient Markets," *American Economic review*, 70, 303-408.
- Hasbrouck, J., 1991a, "Measuring the information content of stock trades," *The Journal of Finance*, 46, 179-207.
- Hasbrouck, J., 1991b, "The summary informativeness of stock trades: an econometric analysis," *Review of Financial Studies*, 4, 571-595.
- Hasbrouck, J., 2009, "Trading Costs and Returns for US Equities: Estimating Effective Costs from Daily Data," *The Journal of Finance*, 64, 1445-1477.
- He, H. and D. M. Modest, 1995, "Market Frictions and Consumption-Based Asset Pricing," *Journal of Political Economy*, 103, 94-117.
- Heaton, J. and D. Lucas, 1996, Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing, *Journal of Political Economy* 104, 443-487.
- Hellwig, M., 1980, "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, 22, 477 - 498.
- Herings, P. J.-J. and K. Schmedders, 2006, "Computing Equilibria in Finance Economics with Incomplete Markets and Transaction Costs," *Economic*

*Theory*, 27, 493-512.

Ho, T. and R. Macris, 1994, "Dealer Bid-ask Quotes and Transactions Prices: an Empirical Study of Some AMEX Options," *Journal of Finance*, 39, 23-45.

Ho, T. and H. R. Stoll, 1980, "On Dealer Markets under Competition," *Journal of Finance*, 35, 259-67.

Ho, T. and H. R. Stoll, 1983, "The Dynamics of Dealer Markets under Competition," *The Journal of Finance*, 38, 1053-1074.

Holmström, B. and J. Tirole, 2001, "LAPM: A Liquidit-Based Asset Pricing Model," *The Journal of Finance*, LVI, 1837-1867

Huang, J. and J. Wang, 2010, "Market Liquidity, Asset Prices, and Welfare," *Journal of Financial Economics*, 95, 107-127.

Jang, B.-G., H. K. Koo, H. Liu and M. Lowenstein, 2007, "Liquidity Premia and Transactions Costs," *The Journal of Finance*, 62, 2329-2366.

Jouini, E. and H. Kallal, 1995, "Martingales and Arbitrage in Securities Markets with Transactions Costs," *Journal of Economic Theory*, 66, 178-197.

Kallsen, J., J. Muhle-Karbe, 2008, "On Using Shadow Prices in Portfolio Optimization with Transaction Costs," working paper, Christian-Albrechts-Universität zu Kiel.

Kühn C. and M. Stroh, 2010, "Optimal Portfolios of a Small Investor in a Limit-order Market: A Shadow-Price Approach," *Mathematics and Financial Economics*, 3, 45-72.

Kyle, A. S., 1985, "Continuous auctions and insider trading," *Econometrica*, 53, 1315-1336.

Larrain, B. and M. Yogo, 2008, "Does firm value move too much to be justified by subsequent changes in cash flow?" *Journal of Financial Economics*, 87, 200-226.

Lee, C. and M. Ready, 1991, "Inferring Trade Direction from Intra-day data," *The Journal of Finance*, 46, 733-746.

Leland, H. E., 2000, "Optimal Portfolio Implementation with Transactions Costs and Capital Gains Taxes," working paper, University of California, Berkeley.

Lesmond, D., Ogden J., Trzcinka C., 1999, "A new estimate of transaction costs," *Review of Financial Studies*, 12, 1113-1141.

Lettau, M. and S. C. Ludvigson, 2005, "Expected returns and expected dividend growth," *Journal of Financial Economics*, 76, 583-626.

Liu, H. and M. Loewenstein, 2002, "Optimal Portfolio Selection with Transactions Costs nad Finite Horizons," *Review of Financial Studies*, 15, 805-835.

Llorente, G., R. Michaely, G. Saar and J. Wang, 2002, "Dynamic Volume-Return Relation of Individual Stocks," *The Review of Financial Studies*, 15, 1005-1047.

Lo, A. W., H. Mamaysky and J. Wang, 2004, "Asset Prices and Trading Volume under Fixed Transactions Costs," *Journal of Political Economy*, 112, 1054-1090.

Longstaff, F, 2001, "Optimal Portfolio Choice and the Valuation of Illiquid Securities," *Review of Financial Studies*, 14, 407-431.

- Longstaff, F., 2009, "Portfolio Claustrophobia: Asset Pricing in Markets with Illiquid Assets," *The American Economic Review*, 99, 1119-1144.
- Lucas, R. E. Jr., 1982, "Interest rates and currency prices in a two-country world," *Journal of Monetary Economics*, 10, 335-359.
- Luttmer, E. G. J., 1996, "Asset Pricing in Economies with Frictions," *Econometrica*, 64, 1439-1467.
- Madhavan, A. and S. Smidt, 1991, "A Bayesian Model of Intraday Specialist Trading," *Journal of Financial Economics*, 30, 99-134.
- Magill, M. and G. M. Constantinides, 1976, "Portfolio Selection with Transactions Costs," *Journal of Economic Theory*, 13, 245-263.
- Milne, F., E. Neave, 2003, "A General Equilibrium Financial Asset Economy with Transaction Costs and Trading Constraints," working paper Queen's University.
- Nazareth, M., 2002, "Portfolio Selection with Stochastic Transaction Costs," working paper.
- O'Hara, M., 2003, "Presidential Address: Liquidity and Price Discovery," *The Journal of Finance*, 58, 1335-1354.
- Obizhaeva, A. and J. Wang, 2005, "Optimal Trading Strategy and Supply/demand Dynamics," NBER Working Paper 11444.
- Parlour, C. A., 1998, "Price Dynamics in Limit Order Markets," *Review of Financial Studies*, 11, 789-816.
- Pástor, L. and R. F. Stambaugh, 2003, "Liquidity Risk and Expected Stock Returns," *The Journal of Political Economy*, 111, 642-685.
- Peress, J., 2005, "Information vs. Entry Costs: What Explains U.S. Stock Market Evolution," *The Journal of Financial and Quantitative Analysis*, 40, 563-594.
- Roll, R., 1984, "A Simple Implicit Measure of the Effective Bid-ask Spread in an Efficient Market," *The Journal of Finance*, 39, 1127-1139.
- Roşu, I., 2009, "A Dynamic Model of the Limit Order Book," *Review of Financial Studies*, 22, 4601-4641.
- Shiller, R. J., 1989, *Market Volatility*, MIT Press, Cambridge.
- Shreve, S. E. and M. H. M. Soner, 1994, "Optimal Investment and Consumption with Transactions Costs," *Annals of Applied Probability*, 4, 609-692.
- Stoll, H. R., 1989, "Inferring the components of the bid-ask spread: Theory and empirical tests," *The Journal of Finance*, 44, 115-134.
- Stoll, H. R., 2000, "Friction," *The Journal of Finance*, 55, 1479-1514.
- Tuckman, B. and J.-L. Vila, 2010, "Holding Costs And Equilibrium Arbitrage," [bikibook.com](http://bikibook.com).
- Van Binsbergen, J. H. and R. S. J. Koijen, 2010, "Predictive Regressions: A Present-Value Approach," *The Journal of Finance*, 65, 1439-1471.
- Vayanos, D., 1998, "Transaction Costs and Asset Prices: A Dynamic Equilibrium Model," *The Review of Financial Studies*, 11, 1-58.
- Vayanos, D. and J.-L. Vila, 1999, "Equilibrium interest rate and liquidity premium with transaction costs," *Economic Theory*, 13, 509-539.
- Vayanos, D. and J. Wang, 2009, "Liquidity and Asset Prices: A Unified Framework," working paper MIT, CAFR and NBER.