DISCUSSION PAPER SERIES

No. 10737

TRADING FEES AND SLOW-MOVING CAPITAL

Adrian Buss and Bernard J Dumas

FINANCIAL ECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

TRADING FEES AND SLOW-MOVING CAPITAL

Adrian Buss and Bernard J Dumas

Discussion Paper No. 10737
July 2015
Submitted 20 July 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

TRADING FEES AND SLOW-MOVING CAPITAL†

Abstract

In some situations, investment capital seems to move slowly towards profitable trades. We develop a model of a financial market in which capital moves slowly simply because there is a proportional cost to moving capital. We incorporate trading fees in an infinite-horizon dynamic general-equilibrium model in which investors optimally and endogenously decide when and how much to trade. We determine the steady-state equilibrium no-trade zone, study the dynamics of equilibrium trades and prices and compare, for the same shocks, the impulse responses of this model to those of a model in which trading is infrequent because of investor inattention.

JEL Classification: G11 and G12
Keywords: frictions, general equilibrium, slow-moving capital and trading fees

Adrian Buss adrian.buss@insead.edu
INSEAD

Bernard J Dumas bernard.dumas@insead.edu
INSEAD and CEPR

† Previous versions circulated and presented under the titles: "The Equilibrium Dynamics of Liquidity and Illiquid Asset Prices" and "Financial-market Equilibrium with Friction". Buss is with INSEAD (adrian.buss@insead.edu) and Dumas is with INSEAD, NBER and CEPR (bernard.dumas@insead.edu). Work on this topic was initiated while Dumas was at the University of Lausanne and Buss was at the Goethe 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 and comments to Beth Allen, Andrew Ang, Yakov Amihud, Laurent Barras, Sébastien Bétermier, Bruno Biais, John Campbell, Georgy Chabakauri, Massimiliano Croce, Magnus Dahlquist, Sanjiv Das, Xi Dong, Itamar Drechsler, Phil Dybvig, Thierry Foucault, Kenneth French, Xavier Gabaix, Stefano Giglio, Francisco Gomes, Amit Goyal, Harald Hau, John Heaton, Terrence Hendershott, Julien Hugonnier, Ravi Jagannathan, Elyès Jouini, Andrew Karolyi, Felix Kubler, David Lando, John Leahy, Francis Longsta¤, Abraham Lioui, Edith Liu, Hong Liu, Frédéric Malherbe, Ian Martin, Alexander Michaelides, Pascal Maenhout, Stefan Nagel, Stavros Panageas, Lubo.Pástor, Paolo Pasquariello, Patrice Poncet, Tarun Ramadorai, Scott Richard, Barbara Rindi, Jean-Charles Rochet, Leonid Ro¸su, Olivier Scaillet, Norman Schürho¤, Chester Spatt, Raman Uppal, Dimitri Vayanos, Pietro Veronesi, Grigory Vilkov, Vish Viswanathan, Je¤rey Wurgler, 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, McGill University, the NBER Asset Pricing Summer Institute, the University of Zurich and the University of Nantes

Trading Fees and Slow-Moving Capital*

Adrian Buss Bernard Dumas

April 2015

Abstract

In some situations, investment capital seems to move slowly towards profitable trades. We develop a model of a financial market in which capital moves slowly simply because there is a proportional cost to moving capital. We incorporate trading fees in an infinite-horizon dynamic general-equilibrium model in which investors optimally and endogenously decide when and how much to trade. We determine the steady-state equilibrium no-trade zone, study the dynamics of equilibrium trades and prices and compare, for the same shocks, the impulse responses of this model to those of a model in which trading is infrequent because of investor inattention.

*Previous versions circulated and presented under the titles: "The Equilibrium Dynamics of Liquidity and Illiquid Asset Prices" and "Financial-market Equilibrium with Friction". Buss is with INSEAD (adrian.buss@insead.edu) and Dumas is with INSEAD, NBER and CEPR (bernard.dumas@insead.edu). Work on this topic was initiated while Dumas was at the University of Lausanne and Buss was at the Goethe 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 and comments to Beth Allen, Andrew Ang, Yakov Amihud, Laurent Barras, Sébastien Bétermier, Bruno Biais, John Campbell, Georgy Chabakauri, Massimiliano Croce, Magnus Dahlquist, Sanjiv Das, Xi Dong, Itamar Drechsler, Phil Dybvig, Thierry Foucault, Kenneth French, Xavier Gabaix, Stefano Giglio, Francisco Gomes, Amit Goyal, Harald Hau, John Heaton, Terrence Hendershott, Julien Hugonnier, Ravi Jagannathan, Elyès Jouini, Andrew Karolyi, Felix Kubler, David Lando, John Leahy, Francis Longstaff, Abraham Lioui, Edith Liu, Hong Liu, Frédéric Malherbe, Ian Martin, Alexander Michaelides, Pascal Maenhout, Stefan Nagel, Stavros Panageas, Luboš Pástor, Paolo Pasquariello, Patrice Poncet, Tarun Ramadorai, Scott Richard, Barbara Rindi, Jean-Charles Rochet, Leonid Roşu, Olivier Scaillet, Norman Schürhoff, Chester Spatt, Raman Uppal, Dimitri Vayanos, Pietro Veronesi, Grigory Vilkov, Vish Viswanathan, Jeffrey Wurgler, 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, McGill University, the NBER Asset Pricing Summer Institute, the University of Zurich and the University of Nantes.

In his Address as President of the American Finance Association, Darrell Duffie (2010) gives numerous examples – with supporting empirical evidence – of situations in which investment capital does not adjust immediately and seems to move slowly towards profitable trades. When a shock occurs, the price of a security reacts first before the quantities adjust. When they do, the price movement is reversed. Duffie's examples are: additions and deletions from the S&P 500 index, arrival of a new order in the book, natural disasters impacting insurance markets, defaults affecting CDS spreads, and issuance of U.S. Treasury securities affecting yields. Many other authors cited by Duffie exhibit "price-pressure" situations.[1]

One can think of several approaches to the modelling of slow-moving capital. Duffie himself offered as an illustration a model in which the attention of investors is limited, so that they overlook for a while the investment opportunities that appear and keep themselves out of the market. As a result, a "thin subset of investors [has] to absorb the shocks" that occur. In this spirit, Duffie and coauthors (2002, 2007, 2012, 2014) assumed that trading required the physical encounter of one trader with another and that searching for encounters was costly or required some time, whereas, in the Microstructure literature, imperfect, non competitive intermediation or asymmetric information is the reason for the slow movement of capital. We offer a third, more basic microfoundation for sluggish capital. That is, we develop a model of a financial market in which capital moves slowly simply because there is a cost to moving capital. This cost can be interpreted as a trading cost or fee,[2] itself being perhaps a reduced form for a cost of gathering information prior to trading or a cost of facing informed traders.[3]

We incorporate trading fees in a dynamic equilibrium model in which investors optimally and endogenously decide when and how much to trade. In doing this, we follow the lead of He and Modest (1995), Jouini and Kallal (1995) and Luttmer (1996), but, unlike these authors, who established bounds on asset prices, we reach a full characterization of the steady state

[1] Patterns of slow reaction around secondary equity offerings have been documented in several papers, from Scholes (1972) to Kulak (2012). Newman and Rierson (2004) document similar effects around issuances of corporate bonds, including spillover effects to bonds of other firms. Coval and Stafford (2007), Edmans *et al.* (2012) and Jotikasthira *et al.* (2009) provide evidence of "fire sales" following mutual fund redemptions. Mitchell *et al.* (2004) exhibit the price impact, with reversal, of the merger of a company that belongs to an index with one that does not. A large literature shows empirically that prices of assets are affected when financial intermediaries suffer losses that hamper their risk-bearing capacity; see, e.g., Mitchell and Pulvino (2012). Ashcraft and Duffie (2007) describe illiquidity phenomena in the Federal Funds market.

[2] As in Demsetz (1968) and Stoll (2000).

[3] As in Peress (2005) and Glosten and Milgrom (1985), respectively.

when the investors have an infinite horizon. Given the presence of the fee, an investor may decide not to trade, thereby preventing other investors from trading with him, which is an additional endogenous, stochastic and perhaps quantitatively more important consequence of the fee. Liquidity begets liquidity. Conceptually and qualitatively speaking, 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. That is, when purchasing a security an investor needs not only have in mind the cash flows that the security will pay into the indefinite future, he must also anticipate his, and other people's, desire and ability to resell the security in the marketplace at a later point in time.

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. However, the end users being the investors, access to a financial market is ultimately a service that investors make available to each other. As a way of constructing 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-period motive for trading, over and above the long-term need to trade for lifetime planning purposes.[4] Specifically, we assume that investors are long-lived and trade because they receive endowments that are not fully hedgeable as, even without trading fees, the market is incomplete. There are two securities and investors face four states of nature: the aggregate output can go up or down and each investor receives a fraction of it as endowment, with the fraction undergoing a two-state Markov chain. With the addition of proportional transactions costs, investors, as is well-known from the literature on non-equilibrium portfolio choice, tolerate a deviation from their preferred holdings, the zone of tolerated deviation being called the "no-trade region." The imbalance of the portfolios that investors keep when they are in that zone, acts as an inventory cost.

We define a form of Walrasian equilibrium for this market. We invent an algorithm that delivers the exact numerical equilibrium, that synchronizes like clockwork the investors in the implementation of their trades and that allows us to analyze the way in which prices are

[4]We discuss trading motives further in Subsection 3.2.

formed and evolve and in which trades take place. We produce the first representation of an *equilibrium* no-trade region ever displayed. Assuming habit-formation utility and using parameter values that generate realistic asset-pricing moments, we measure the degree to which the aforementioned impediments to trade prevent investors from fully smoothing their consumption and from achieving perfect risk sharing between them. We determine the impact of trading fees on the volume of trading of the securities that are subject to these fees and on the volume of trading of the securities that are not. We compare analytically equilibrium securities prices to the investors' private valuations as well as to the shadow private bid and ask prices of investors; we explain how the gap between them triggers trades. In addition, we ask whether equilibrium securities prices conform to the famous dictum of Amihud and Mendelson (1986), which says that they are reduced by the present value of transactions costs to be paid in the future. Our computations show that the prices of securities are increased only slightly by the presence of fees and we explain precisely why that is the case. Finally, we construct in a proper way the responses of prices to supply shocks and show that *the hysteresis* effect of trading frictions *can explain slow price reversal just as well* as Duffie's model of infrequent trading.

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), Edirsinghe, 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 (2013), Liu and Lowenstein (2002), Jang *et al.* (2007), Gerhold *et al.* (2011) and Gârleanu and Pedersen (2013), 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.[5] Obviously, the assumption must be made that some investors,

[5]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*

other than the one whose portfolio is being optimized, do not incur costs. In the present paper, all investors face the trading fee.

The papers of Heaton and Lucas (1996), Vayanos (1998), Vayanos and Vila (1999) and Lo *et al.* (2004) exhibit the equilibrium behavior resulting from a cost of transacting and are direct ancestors of the present one.[6] Heaton and Lucas (1996) derive a stationary equilibrium under transactions costs 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 his finite lifetime. Transactions costs induce him to trade very little during his life: when young, he buys some securities that he can resell during his old age, in order to be able to consume. Here, we introduce a motive to trade that is operative at every point in time. 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' endowments provide the motive to trade but the amount of aggregate physical resources available is non-stochastic. In our paper, fees are proportional, preferences can be specified at will although we present illustrative results for the power-utility type, and aggregate and individual resources are free to follow an arbitrary stochastic process. To our knowledge, ours is the first paper that characterizes such an equilibrium.

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 fixed date, after which it becomes liquid again. The consequences for equilibrium asset prices are drawn in relation to the length of the freeze. In our model, the trading dates are chosen endogenously by the investors.[7] As we do, Buss *et al.* (2015) derive an equilibrium in a financial market where investors incur a cost of trading. Whereas our paper studies the equilibrium for long-lived agents that trade because of shocks to endowments and is focused on the implications for slow-moving capital, Buss *et al.* (2015) study

al. (2014) have shown that this is generally not true under non IID returns. The possibility of falling in a "no-trade" region is obviously a violation of the IID assumption.

[6] In these papers, the cost is a physical, deadweight cost of transacting. 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).

[7] In Brunnermeier and Pedersen (2008), liquidity is priced and investors, in addition to trading with frictions, face liquidity constraints. However, some investors arrive to the market exogenously.

an economy with two risky assets and investors with different levels of experience, so that they disagree about the growth rate of one of the assets. The focus in their paper is on the interplay between illiquidity and inexperience.

As far as the solution method is concerned, our analysis is closely related, in ways we explain below, to "the dual method" proposed by Bismut (1973), Cvitanic and Karatzas (1992) and 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.[8]

The paper is organized as follows. We write down the model (Section 1), explain how we solve for equilibrium (Section 2) and exhibit the dynamics of the economy (Section 3). Then, in Section 4 we examine the investors' portfolio strategy and equilibrium prices. Finally, in Section 5, we compare the impulse responses produced by our model to those that would result from infrequent trading in the manner of Duffie (2010).

1 The objective of each investor and the definition of equilibrium

The economy is populated with two investors $l = 1, 2$ and a set of exogenous time sequences of individual endowments $\{e_{l,t} \in \mathbb{R}_{++}; l = 1, 2; t = 0, ..., T\}$ on a tree or lattice. Notice that the tree accommodates the exogenous state variables only. In the financial market, there are I securities: $i = 1, .., I$. Some are short lived, make a single payoff, are renewed and are in zero net supply. Others are long-lived and defined by their payoffs $\{\delta_{t,i};\ t = 0, ..., T\}$, which are also exogenous and placed on the tree or lattice.[9]

While the equilibrium construction is based on a finite horizon T, we are able to increase T indefinitely, thereby approaching the behavior of investors whose horizon would be infinite.

[8]Our solution method shares some common features with Buss *et al.* (2015). Both papers 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.* (2015) 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.

[9]As has been noted by Dumas and Lyasoff (2012), because the tree only involves the exogenous variables, it can be chosen to be *recombining* when the endowments are Markovian.

Financial-market transactions entail trading 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, he receives in units of consumption goods the transaction price multiplied by $1 - \varepsilon_{i,t}$ ($0 < \varepsilon_{i,t} < 1$) and, when he buys, he must pay the transaction price times $1 + \lambda_{i,t}$ ($0 < \lambda_{i,t} < 1$). All fees are paid to a central pot. At each point in time, the fees collected at the central pot from one investor are distributed to the other investor in the form of transfers, which are taken by him to be lumpsum (but recurring) amounts.[10] We are interested in exhibiting the distortionary effects of trading fees on trading, not in figuring out who gets rich or poor from them. Tax economists who study the impact of "budget neutral" tax changes will find the concept familiar, as will macroeconomists who differentiate between aggregate capital and individually owned capital, when they handle the recursive, competitive equilibrium of a population of agents.[11] In all cases, the assumption is a way to mimic the pure-competition assumption according to which an agent behaves under the belief that he is very small and has no impact on the behavior of other agents.

We consider a recursive Walrasian market for the securities. Assuming all markets beyond time t are cleared, the auctioneer calls out time-t prices, which we call "posted" prices and denote: $\{S_{t,i}; i = 1, ..., I; t = 0, ..., T\}$. The posted price of a security is an effective transaction price only if and when a transaction takes place but it is posted all the time by the Walrasian auctioneer (which works at no cost). Investors know that fees will be calculated on the basis of that posted price in case of a buy transaction and in case of a sell transaction. They submit to the auctioneer flow quantity schedules. If the flow demanded is positive at the price that is called out, the investor intends to buy; otherwise he intends to sell. The auctioneer determines the intersection between the two schedules, if any. In this way, it clears the market between the two investors, one possible outcome of the clearing being a zero trade.

With symbol $\theta_{l,t,i}$ standing for the number of units of Security i in the hands of Investor

[10] The transfers enter his budget constraint but do not generate an additional term in his first-order conditions. In this way, the trader remains purely competitive in that he only takes into account the cost of his own actions, not the benefits he may receive from the actions of the other trader. We make this assumption purely for the sake of convenience. It will allow us to equate aggregate consumption to aggregate output. Without that, some of the output would be lost to deadweight transactions costs and the sum of the consumption shares would have to bounded away from 100%. It has no material impact on the results. In a previous version of our paper, we had assumed that trading entailed physical deadweight costs.

[11] See Ljunqvist and Sargent (2012), page 233 and page 473. We are referring here to the "K,k" approach.

l *after* all transactions of time t, Investor l solves the following problem:[12]

$$\sup_{\{c_l,\{\theta_{l,i}\}\}} \mathbb{E}_0 \sum_{t=0}^{T} u_l\left(c_{l,t},\cdot,t\right)$$

subject to a sequence of equations, which define the budget set:

- a sequence of flow-of-funds budget constraints for $t = 0, .., T$:

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

  where:

$$\begin{aligned} \zeta_{l,t} \triangleq &\sum_{i=1}^{I} \max\left[0, \theta_{l',t,i} - \theta_{l',t-1,i}\right] \times S_{t,i} \times \lambda_{i,t} \\ -&\sum_{i=1}^{I} \min\left[0, \theta_{l',t,i} - \theta_{l',t-1,i}\right] \times S_{t,i} \times \varepsilon_{i,t}; l' \neq l \end{aligned} \tag{2}$$

- and given initial holdings $\left\{\bar{\theta}_{l,i}\right\}$:[13]

$$\theta_{l,-1,i} = \bar{\theta}_{l,i}. \tag{3}$$

On the left-hand side of the flow budget constraint, the second term reflects the net cost of purchases, the third term 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.

[12] We assume that utility functions are strictly increasing, strictly concave and differentiable to the first order with respect to consumption. Here we write the utility function in the additive form but it may contain other arguments than consumption (hence the $\cdot$ as an argument). Below, we introduce habit formation and, given the recursive technique to be used, it would be easy to handle recursive utility, especially in the isoelastic case.

[13] The initial holdings satisfy the restriction that $\sum_{l=1}^{2} \bar{\theta}_{l,i} = 0$ or 1 depending on whether the security is assumed to be in zero or unit supply.

The recursive dynamic-programming formulation of the investor's problem is:[14]

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

subject to the flow budget constraint (1) for time t only.

A change of notation reformulates the problem in the form of an optimization *under inequality constraints*, which is more suitable for mathematical programming. Writing:

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

for purchases of securities and:

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

(a negative number) for sales, so that $\theta_{l,t,i} \equiv \widehat{\theta}_{l,t,i} + \widehat{\widehat{\theta}}_{l,t,i} - \theta_{l,t-1,i}$, one gets:

$$\begin{aligned} J_l\left(\{\theta_{l,t-1,i}\}, \cdot, e_{l,t}, t\right) &= \sup_{c_{l,t}, \left\{\theta_{l,t,i}, \widehat{\theta}_{l,t,i}, \widehat{\widehat{\theta}}_{l,t,i}\right\}} u_l\left(c_{l,t}, \cdot, t\right) \\ &\quad + \mathbb{E}_t J_l\left(\left\{\widehat{\theta}_{l,t,i} + \widehat{\widehat{\theta}}_{l,t,i} - \theta_{l,t-1,i}\right\}, \cdot, e_{l,t+1}, t+1\right) \end{aligned} \tag{4}$$

subject to:

$$\begin{aligned} c_{l,t} + \sum_{i=1}^{I}\left(\widehat{\theta}_{l,t,i} - \theta_{l,t-1,i}\right) \times S_{t,i} \times (1+\lambda_{i,t}) + \sum_{i=1}^{I}\left(\widehat{\widehat{\theta}}_{l,t,i} - \theta_{l,t-1,i}\right) \times S_{t,i} \times (1-\varepsilon_{i,t}) \\ = e_{l,t} + \sum_{i=1}^{I}\theta_{l,t-1,i}\delta_{t,i} + \zeta_{l,t} \end{aligned} \tag{5}$$

$$\widehat{\widehat{\theta}}_{l,t,i} \leq \theta_{l,t-1,i} \leq \widehat{\theta}_{l,t,i} \tag{6}$$

[14] The form $J_l\left(\{\theta_{l,t-1,i}\}, \cdot, e_{l,t}, t\right)$ 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.

Under standard concavity assumptions on utility functions, the maximization of (4) subject to (5) and (6) is a convex problem. First-order conditions of optimality (including terminal conditions $\theta_{l,T,i} = 0$) are necessary and sufficient for the optimum to be reached.

Definition 1 *An equilibrium is defined as a process for the allocation of consumption $c_{l,t}$, a process for portfolio choices $\left\{\theta_{l,t,i}, \widehat{\theta}_{l,t,i}, \widehat{\widehat{\theta}}_{l,t,i}\right\}$ of both investors and a process for securities prices $\{S_{t,i}\}$ such that the supremum of (4) subject to the budget set is reached for all l, i and t and the market-clearing conditions:*

$$\sum_{l=1,2} \theta_{l,t,i} = \sum_{l=1,2} \bar{\theta}_{l,i}; i = 1, ..., I \tag{7}$$

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

2 The algorithm

The method used to obtain the equilibrium blends in an original fashion the Interior-Point algorithm, which is an optimization technique based on Karush-Kuhn-Tucker (KKT) first-order conditions, with a shift of equations that has been proposed by Dumas and Lyasoff (2012) to facilitate backward induction. To obtain an equilibrium under constraints, one forms a system juxtaposing the first-order conditions of both investors and the market-clearing condition, and one solves.

2.1 A shift of equations

A given node of the tree 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 are denoted $\pi_{t,t+1,j}$ ($\sum_{j=1}^{K_t} \pi_{t,t+1,j} = 1$).[15]

In Appendix A, we show 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 $\left\{\phi_{l,t}, R_{l,t,i}\right\}$ (as opposed to given values of the original state

[15] Transition probabilities and other time-t variables depend on the current state but, for ease of notation only, we suppress the corresponding subscript everywhere.

variables, which were initial positions $\{\theta_{l,t-1,i}\}$), by solving the following equation system written for $l = 1, 2; j = 1, ..., K_t; i = 1, ..., I$. 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:[16]

$$u_l'\left(c_{l,t+1,j}, \cdot, t+1\right) = \phi_{l,t+1,j}$$

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

$$c_{l,t+1,j} + \sum_{i=1}^{I} \left(\theta_{l,t+1,i,j} - \theta_{l,t,i}\right) S_{t+1,i,j} R_{l,t+1,i,j} = e_{l,t+1,j} + \sum_{i=1}^{I} \theta_{l,t,i} \delta_{t+1,i,j} + \zeta_{l,t+1,j}$$

where:

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

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:"

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

[16] u_l' denotes "marginal utility" or the derivative of utility with respect to consumption.

4. Definitions:

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

5. Complementary-slackness conditions:

$$\begin{aligned}
(-R_{l,t+1,i,j} + 1 + \lambda_{i,t+1,j}) \times \left(\widehat{\theta}_{l,t+1,i,j} - \theta_{l,t,i}\right) &= 0 &(9)\\
(R_{l,t+1,i,j} - (1 - \varepsilon_{i,t+1,j})) \times \left(\theta_{l,t,i} - \widehat{\widehat{\theta}}_{l,t+1,i,j}\right) &= 0 &(10)
\end{aligned}$$

6. Market-clearing restrictions:

$$\sum_{l=1,2} \theta_{l,t,i} = \sum_{l=1,2} \bar{\theta}_{l,i}$$

7. Inequalities:

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

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 $\left\{c_{l,t+1,j}, \phi_{l,t+1,j}, R_{l,t+1,i,j}, \theta_{l,t,i}, \widehat{\theta}_{l,t+1,i,j}, \widehat{\widehat{\theta}}_{l,t+1,i,j}; l = 1,2; i = 1, ..., I; j = 1, ..., K_j\right\}$.[17]

Besides the exogenous endowments $e_{l,t+1,j}$ and dividends $\delta_{t+1,i,j}$, the "givens" are the time-t investor-specific shadow prices of consumption $\left\{\phi_{l,t}; l = 1,2\right\}$ and of paper securities $\left\{R_{l,t,i}; l = 1,2; i = 1, ..., I\right\}$, 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 $0 \leq \frac{\phi_{1,t}}{\phi_{1,t}+\phi_{2,t}} \leq 1$.[18]

In addition, the given securities' price functions $S_{t+1,i,j}$ are obtained by backward induc-

[17] The size of the system is reduced when some securities do not carry trading fees.

[18] 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 trading fees are refunded in a lumpsum.

tion (see, in Appendix A, the third equation in System (16)):

$$\begin{aligned} S_{t,i} &= \frac{1}{R_{l,t,i}\phi_{l,t}} \sum_{j=1}^{K_t} \pi_{t,t+1,j}\phi_{l,t+1,j} \times \left(\delta_{t+1,i,j} + R_{l,t+1,i,j} \times S_{t+1,i,j}\right); \\ S_{T,i} &= 0 \end{aligned} \tag{11}$$

and the given future position functions $\theta_{l,t+1,i,j}$ (satisfying $\sum_{l=1,2}\theta_{l,t+1,i,j} = \sum_{l=1,2}\bar{\theta}_{l,i}$; $i = 1,...,I$) are also obtained by an obvious backward induction of 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 (3). The way to do that is explained in Appendix B.

2.2 The interior-point algorithm

At each node and for each value of the endogenous state variables placed on a grid, the system of equations described above can be solved numerically by Newton iterations. However, the iterations can run into indeterminacy because of the KKT complementary-slackness conditions (9, 10), which contain a product of unknowns equated to zero. Indeed, if a Newton step produces, for instance, a value $-R_{l,t+1,i,j} + 1 + \lambda_{i,t+1,j}$ on the boundary where it is equal to zero, then the requirement placed on $\widehat{\theta}_{l,t+1,i,j} - \theta_{l,t,i}$ drops out of the system and indeterminacy follows. The Interior-Point algorithm is a solution to that problem. It amounts to replacing the above equation system by a sequence of equation systems in each of which the KKT conditions are relaxed as follows:

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

where η is a small number, which is made to approach zero as the algorithm progresses. In this way, the indeterminacy is avoided.[19]

In one very convenient implementation, Armand *et al.* (2008) show a way to add to the system one equation that will have the effect of driving η towards zero progressively with each Newton step. We use a simplified version of that implementation.

Parenthetically,[20] the Interior-Point method should be of great interest to microeconomists who study choice problems with inequality constraints. In cases in which limits can be interchanged, the comparative statics of the solution can be obtained by total differentiation of the first-order conditions, for a given value of η, in the same way as is done in Microeconomics textbooks, leading, for instance, to a version of the Slutzky equation. The comparative-statics properties are close to those that would obtain in the original system of first-order conditions with $\eta = 0$.

3 The dynamics of the economic system

3.1 Setup of the illustration

In the setup that we use for illustration, we consider two investors who have the same isoelastic utility function but receive different endowments. There are two securities. The subscript $i = 1$ refers to a short-lived riskless security (the "bond") in zero net supply and the subscript $i = 2$ refers to a "stock" in unit net supply. We call "stock" a long-lived claim that pays out dividend δ_t.[21] Total output (dividend plus endowments) is represented by a binomial tree, with constant geometric increments mimicking a geometric Brownian motion. Although the solution method allows for any stochastic process of dividends and endowments, in our illustration to be described below, we assume that total output minus dividends is distributed as endowments to the two investors, with the shares of endowment following a

[19] This approach, called the "barrier" approach for reasons explained in, e.g., Boyd and Vandenberghe (2004), is more compatible with Newton solvers than the alternative method proposed earlier by Garcia and Zangwill (1981), which involves discontinuous functions such as $\max[\cdot, \cdot]$.

[20] This remark was made by Dimitri Vayanos in a private conversation.

[21] For the case of one stock, 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 by their value of the exogenous variable, are isomorphic to each other, where the isomorphy simply means that we can factor out the total output. In this way, we do not need to perform a new set of calculations for each and every node of a given point in time. We prove this property in Appendix C.

simple two-state Markov chain. For simplicity, we assume that both investors are *ex ante* symmetric and we let the probabilities of transition be identical whether an investor is in the low or the high state, which implies that a person's endowment is positively correlated with total endowment. Each investor l faces four states of nature for the immediate future: high *vs.* low increase in output and high *vs.* low share of endowment. With two securities only, the financial market is incomplete and investors must trade in response to the endowment shocks they receive.

Preferences. Preferences are of the additive external-habit type (in the "Catching-up-with-the-Joneses" form), implemented as *surplus consumption*, similar to Campbell and Cochrane (1999):

$$\mathbb{E}\left[\sum_{t=0}^{T}\beta^t \times \frac{(c_{l,t} - h \times C_{t-1})^{1-\gamma}}{1-\gamma}\right]$$

where C_{t-1} denotes aggregate last period consumption (equal to last period output). Investors have homogeneous preferences, i.e., the same time-preference β, risk appetite γ and habit parameter h. We introduce external habit *solely for the purpose of matching stock return volatility and the equity premium*, our goal being to illustrate the effects of trading fees in the presence of a realistic behavior of stock prices.

Output. Aggregate output O_t follows a simple binomial tree with drift μ_O as well as volatility σ_O and a probability of an up-move (down-move) of 50%.

Financial Assets. In addition to the short-lived security, investors can trade a claim paying a share χ of total output as dividends $\delta_t = \chi \times O_t$. The risky asset is available in unit supply.

Individual Endowments. The remaining part of total output $(1-\chi)\times O_t$ is distributed through endowments. The investors receive individual endowments $e_{l,t}$, e.g., labor income, as a fraction $\upsilon_{l,t}$ of the total endowments $(1-\chi) \times O_t$. Specifically, we assume that the investors' shares of total output follow a simple, symmetric two-state Markov chain, with realizations:

$$\upsilon_{1,t} \in \{0.5 + \alpha; 0.5 - \alpha\}\,; \upsilon_{2,t} = 1 - \upsilon_{1,t}$$

and transition matrix

$$\begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

We choose p in such a way that the share received is persistent. Without persistence, investors would mostly not bother to pay the fee to enter a trade.

3.2 The motives to trade

Investors in our model trade because they receive differing stochastic endowments while the market is incomplete. They are "liquidity" traders. They trade or hedge at time 0 a marketable component of their endowments. Thereafter, they trade again when they actually receive the endowment if its amount is above or below the amount they have previously been able to hedge.[22] In the absence of frictions, the trading motive is completely straightforward: *the investor who receives a high share of endowment uses some of his funds to consume an extra amount and uses the larger part to buy securities.* How much of it he consumes and whether he buys the stock or puts his money into the short-term riskless asset is determined endogenously. Frictions will impede that trading motive somewhat.

3.3 Parameter values

The numerical illustration below is only meant to illustrate in a stylized fashion the workings of the model. It cannot be seen as being calibrated to a real-world economy since we have two investors, not millions, and two securities available to them not tens of thousands. For these reasons, although we incorporate a motive to trade that is present at all times, the volume of trading we are able to generate does not come anywhere close to market data.[23] We, therefore, keep a trading time interval of one year 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.

We set the drift and volatility of output growth equal to their empirical counterparts obtained by Campbell (2003). Specifically, we set $\mu_O = 1.8\%$ and $\sigma_O = 3.2\%$. For the individual endowments, we set $p = 0.85$ and $\alpha = 0.125$, resulting in a growth rate of about 3.85% and a volatility of about 20% for the endowment shocks. Those numbers are comparable to the literature. For example, in Gourinchas and Parker (2002) the total

[22] We leave for future or ongoing research two other motives for trading that are obviously present in the real world such as the sharing of risk between two investors of differing risk aversions and the speculative motive arising from informed trading, private signals or differences of opinion.

[23] Except perhaps in the sense of "net trades" of the market index.

volatility of labor income shocks is about 24%. Carroll (1992,1997) estimates volatilities of 16 to 18% and uses growth rates between 0 and 4%. The resulting average correlation between the two individual endowments is -0.9. The number is excessive but, again, when there are just two traders in a market instead of millions, we have to boost their motive to trade.

We set the rate of time-preference β equal to 0.98 – a common choice in the Finance and Macroeconomics literature. The remaining parameters h, γ and χ are chosen as to closely match the risk-free rate, equity risk premium and stock market volatility that we observe in the data (see Table 2). The resulting values are: $h = 0.2$, $\gamma = 7.5$, $\chi = 0.15$. A value, $\chi = 0.15$, i.e., 85% of total output being distributed as endowments, implies that the average wealth-financial income ratio in our economy is 3.81. Gomes and Michaelides (2005) find median wealth to income ratios of 0.29 for age $20 - 35$, 2.17 for age $35 - 65$ and 7.93 for age greater than 65 based on the Survey of Consumer Finances (2001). The most recent Survey of Consumer Finances (2014) documents similar wealth-to-income ratios ranging from 0.37 (age < 35) to 5.52 (age ≥ 75)

Table 1 collects all the parameter values. The resulting return moments are shown in Table 2 and compared to the data. The table demonstrates that our quantitative experiments are conducted in a realistic financial-market setting.

3.4 Lenghtening the horizon and the steady state

We run the algorithm backward from a fixed horizon date until there is no change in all the functions being carried backward, thereby obtaining the equilibrium *as it would be in an economy where investors are infinitely long-lived.*[24] Besides displaying features that hold for an infinite horizon, we also want to make sure that those features do not depend on initial conditions. For that purpose, we simulate the infinite-horizon economy forward starting from some arbitrary initial conditions and we keep track of the frequency distribution of the consumption shares of the two investors across simulated paths. Figure 1 shows that the steady-state probability distribution of consumption shares obtains after $t = 250$. This many years is a long enough "burn-in" history. In what follows, we discuss only the events that take place well within the steady-state era, namely at $t = 300$.

[24]The criterion for stopping the backward calculation is the mean absolute relative difference from one time step to the next of all iterated functions. We stop when the value of that criterion is below 0.01%.

Name	**Symbol**	**Value**
Parameters for exogenous endowment		
Horizon of the economy	T	∞
Time step of the tree		1
Expected growth rate of output	μ_O	1.8%
Volatility growth rate of output	σ_O	3.2%
Share of dividends	χ	0.15
Parameters for the individual investors		
Time preference	β	0.98
Habit parameter	h	0.2
Risk aversion	γ	7.5
Transition endowment shock	p	0.85
Fraction endowment shock	α	0.125
Trading fees per dollar of equity traded		
When buying and when selling	$\lambda = \varepsilon$	0% to 3%

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

Moment	**Data**	**Model**
Aggregate consumption growth	1.79%	1.82%
Agg. cons. growth volatility	3.22%	3.26%
Risk-free rate	2.02%	2.32%
Equity premium	6.73%	7.47%
Stock return volatility	18.60%	19.65%
Sharpe ratio	0.36	0.38
Price-dividend ratio	23.75	21.01
Volatility of $\log P/D$ ratio	0.32	0.16

Table 2: **Return moments without friction.** The data is based on Campbell (2003) with a sample period spanning from 1891-1998. Consumption growth denotes real per capital consumption growth of non-durables and services for the United States. The stock return data is based on the S&P500 index and the risk-free rate is based on the 6-month U.S. Treasury bill rate.

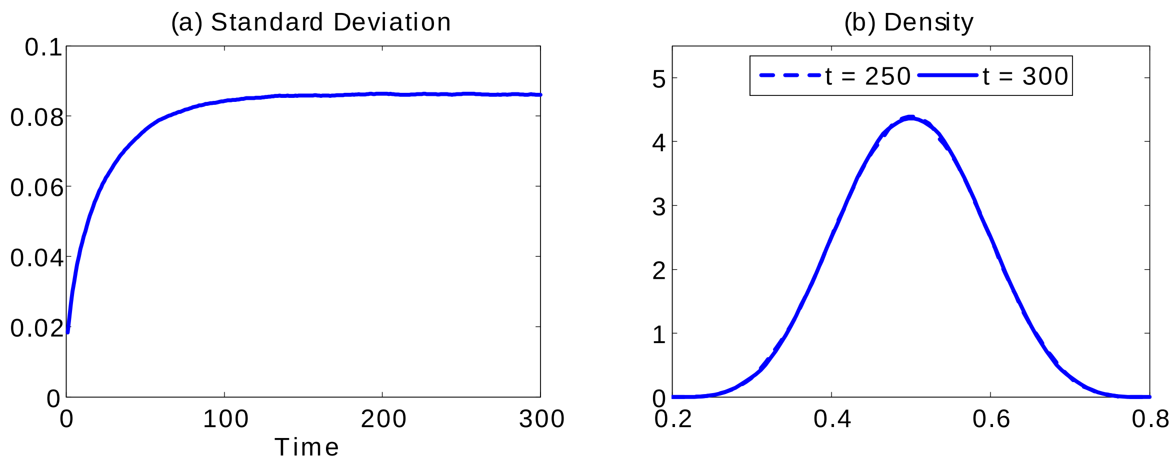

Figure 1: **Steady-state era**: the picture illustrates the existence in our infinite-horizon economy of a time window starting some time before $t = 250$ within which the probability distribution of consumption shares is constant. Panel (a) shows the standard deviation of the consumption share of investor 1 across the 500,000 paths of a simulation. Panel (b) shows the simulated density of the consumption share for two dates. All parameters are set at their benchmark values indicated in Subsection 3.3.

3.5 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 trading fees, a key concept, which we further elaborate on in Subsection 4.1.1 is that of a "no-trade zone," which is the area of the state space where both investors prefer not to adjust their portfolios.

Figure 2 displays a simulated sample path illustrating how our financial market with trading fees operates over time. Panels (a) and (b) show a sample path of: (i) stock holdings (expressed as a fraction of the security available, not as a dollar value) as they would be in a zero-trading fee economy, (ii) the actual stock-holdings with a 2% trading fee and (iii) the boundaries of the optimal 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, except that they allow 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 the amount of trading fees paid 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 algorithm: they are made to trade exactly opposite amounts exactly at the same time. Panels (c) and (d) show the holdings of the riskless, frictionless bond, which basically accommodate the holdings of equity.

The figure also illustrates the degree to which *capital is slow-moving*: the optimal holdings of stocks under friction are a delayed version of the frictionless holdings. But the length of the delay is stochastic. To the opposite, holdings of the riskless asset, which does not entail trading fees, fluctuate more than they would in a frictionless economy. When they receive their endowments, investors use the cost-free, riskless asset as a holding tank and trade it much more than they would if the stock were also cost-free.

Panel (e) shows the stock's posted price, 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 way in which the algorithm has synchronized the trades of the two investors.

Even though ours is a Walrasian market and neither a limit-order nor a dealer market, one can define a virtual concept of bid and ask prices. In Definition 2 below, we define the investors' private valuations of dividends. The *bid price of a person* can be defined as being equal to the person's private valuation of dividends minus the trading fees to be paid in case the person buys, and, similarly, for the ask price. When the two private valuations differ by the sum of the one-way trading fees for the two investors, a transaction takes place. Equivalently, a trade occurs when the bid price of one investor is equal to the ask price of the other investor. Defining the *effective spread* as the difference between the higher and the lower of the two bid prices of the two investors, one could also say that a transaction takes place when the effective spread becomes equal to zero. That mechanism is displayed in Panel (f). The posted price is thus seen as some form of average of the two private valuations, or some form of average of the higher bid and the lower ask prices, or some form of average of the lower bid and the higher ask prices.[25]

[25] For a formal confirmation, see below Proposition 3.

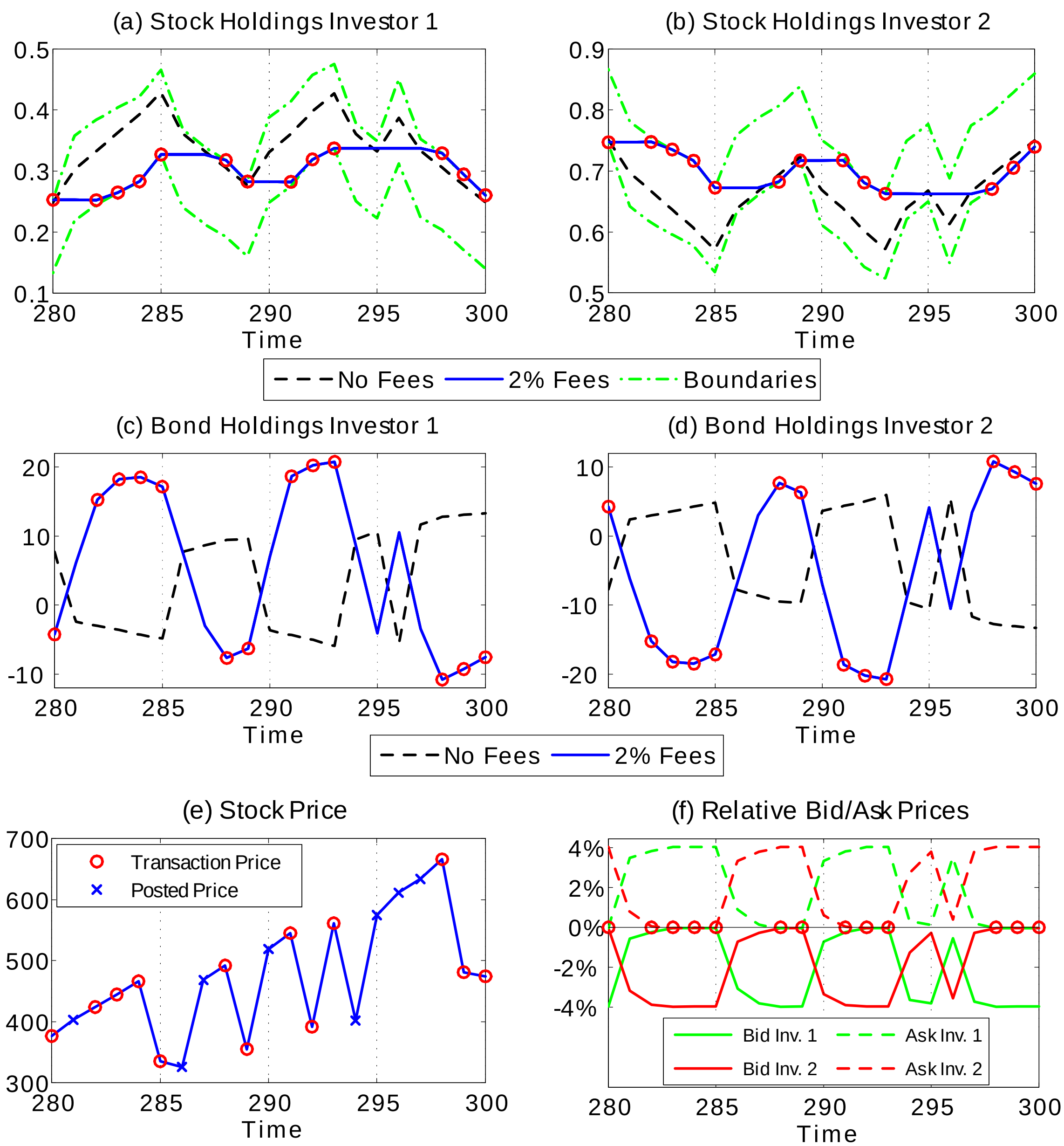

Figure 2: **A sample time path of stock holdings, bond holdings, the stock price and the bid and ask prices of each investor**. Panels (a) and (b) show a sample path of: (i) stock holdings as they would be in a zero-trading fee economy, (ii) the actual stock-holdings with a 2% trading fee and (iii) the boundaries of the no-trade zone. Panels (c) and (d) show the holdings of the riskless security. Panel (e) shows the behavior of the stock price and 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.

4 Equilibrium asset holdings and prices

After solving for the equilibrium process following the backward-induction procedure outlined above, we run $500,000$ simulated paths obtained by walking randomly down the binomial tree of outputs and the Markov chain of endowment shares. All quantitative results we display below are statistics computed across simulated paths at date $t = 300$, at which the system has reached the steady-state era.[26]

4.1 Asset holdings

4.1.1 Equilibrium no-trade region

From the literature on non-equilibrium portfolio choice, it is well-known 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, when two investors make synchronized portfolio decisions and prices are set to clear the market.

Figure 3 plots the no-trade region that applies during the steady-state era. To our knowledge, Figure 3 is the first representation of an *equilibrium* no-trade region ever displayed.[27] 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. Since they trade at the exact same time, the equilibrium no-trade region is *valid for both investors* subject to a relabelling of the axes, as is done in an Edgeworth box.

There is also a central symmetry between the two barriers. Recall that, when one investor receives a low share of the endowment, the other investor receives a high share.

4.1.2 Consumption

Given the exogeneity of the output and the endowments, the probability distribution of aggregate consumption is, of course, unaffected by trading fees. But the conditional joint

[26] Recall that investors are *infinitely long-lived.*

[27] Optimal no-trade regions for given asset price behavior were obtained by Dixit (1991), Fleming et al. (1990) and Svensson (1991).

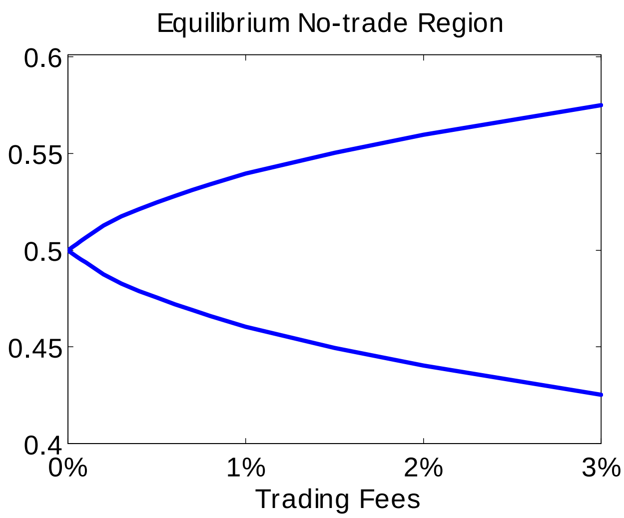

Figure 3: **Equilibrium no-trade region.** The figure displays the average of the positioning of the trade boundaries across $500,000$ paths at $t = 300$. A similar figure could have been obtained from a backward step of the algorithm with the difference that the boundaries would then have been conditional on the current value of the consumption-share state variable. The figure shown here is an average, integrated over the steady-state distribution of consumption shares.

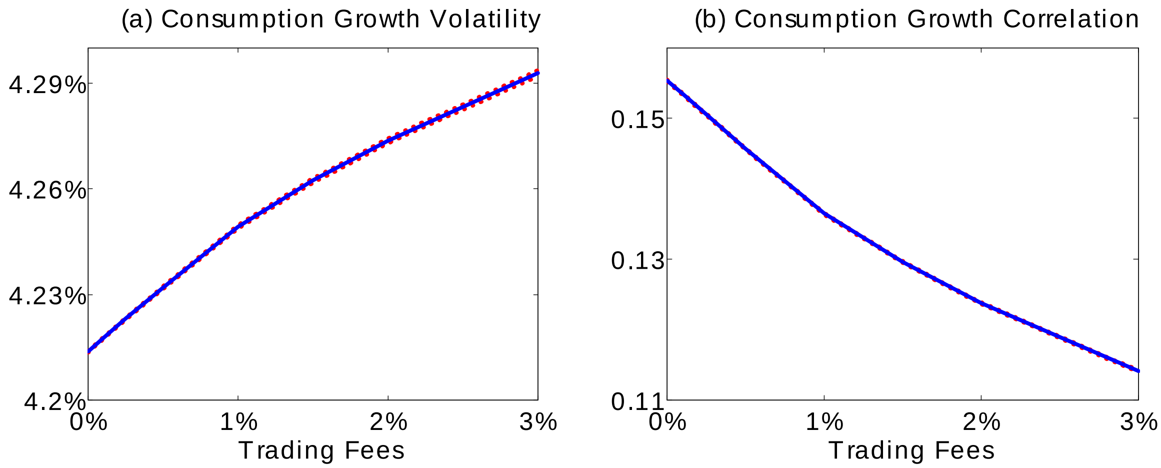

Figure 4: **Optimal consumption behavior**. Average consumption growth conditional volatility and conditional correlation of the two investors for different levels of trading fees. The figure displays averages calculated at $t = 300$ across $500,000$ simulated paths. The solid line is the average. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

distribution of the *individual* consumptions of the two investors reflects asset holdings, which are very much affected, because investors are conflicted between the desire to smooth consumption and the desire to smooth trades. Figure 4, Panel (a) shows, against the rate of trading fees, the average conditional volatility of individual consumption growth at $t = 300$. It illustrates that, on average, the presence of trading fees prevents investors from smoothing their consumption across paths (or states) as effectively as they would in a frictionless economy. Since the aggregate volatility is unchanged, the increased individual volatility must be matched with a reduced correlation of individual consumptions. Panel (b) confirms that idea.[28]

4.2 Trades over time

As we suggested in the introduction and as we demonstrate more amply in Section 5 below, capital flows are slowed down by trading fees. Here, we examine the trading volume and the

[28]Recall that, even at level zero of trading fees, the market is incomplete. That is why the correlation of consumption is never equal to 1.

probability of trading induced by trading fees. The trading volume is defined as the absolute values of changes in θ_2 (shares of the stock) at time $t = 300$. The average volume of stock trading is shown in Figure 5, Panel (a). As one would expect, it decreases with trading fees. For trading volume of the bond, however, Panel (b) exhibits a non monotonic behavior: for sufficiently low fee rates, the trading fees first reduce the trading of the bond against the stock. But, for higher trading fees, the investors are forced to trade bonds as the preferred means of smoothing consumption over time and across states, as was illustrated in Figure 2.

For sufficiently low fees, the probability of trading (Panel (c)) remains at 100%, as it would be in a frictionless word, while the size of the trades gradually drops (Panels (a) and (b)). At some point, the probability of trading decreases. It is a measure of the *endogenous liquidity risk* that the investor has to bear because he operates in a market with friction, in which other traders may not wish to trade and capital moves slowly.

4.3 Asset prices

The work of Amihud and Mendelson (1986, Page 228),[29] is an invitation to compare the prices of assets with and without trading fees. 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 is forced to trade. We consider two investors who are risk averse, face identical trading conditions and trade freely and optimally. Furthermore, the trading fees in our model are refunded.

Recall from Equation (11) that the securities' posted prices $S_{t,i}$ are:

$$\begin{aligned} S_{t,i} &= \mathbb{E}_t\left[\frac{\phi_{l,t+1}}{R_{l,t,i}\phi_{l,t}} \times (\delta_{t+1,i} + R_{l,t+1,i} \times S_{t+1,i})\right]; \\ S_{T,i} &= 0 \end{aligned}$$

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 dividends on security i calculated at Investor l's equilibrium state prices as they are *under trading fees*. We denote this private valuation $\hat{S}_{l,t,i}$:

[29] See also Vayanos and Vila (1999, Page 519, Equation (5.12)).

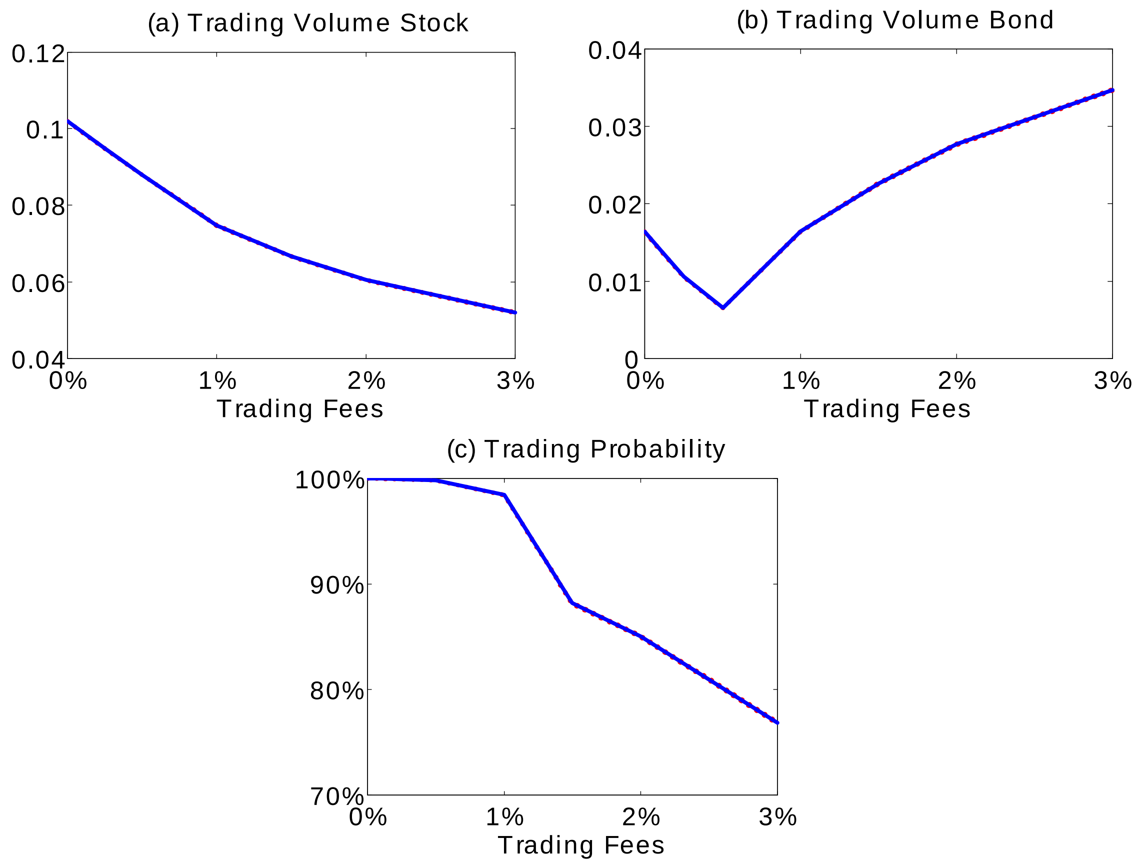

Figure 5: **Patterns of volume and trading frequency against trading fees**. Panel (a) shows, for different trading-fees rates, the mean stock trading volume/year; Panel (b) shows the mean bond trading volume/year. Panel (c) shows the frequency of trading. All parameters are set at their values indicated in Table 1. The figure displays averages calculated at $t = 300$ across $500,000$ simulated paths. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

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

$$R_{l,t,i} \times S_{t,i} = \widehat{S}_{l,t,i} \tag{12}$$

which means that the posted prices of securities can at most differ from the private valuation of their dividends as seen by Investor l by the amount of the potential *one-way trading fee*, actually or not actually incurred by Investor l at the current date only.[30] The posted price is, indeed, some form of average of the two private valuations.

Second, we compare equilibrium asset prices that prevail in the presence of trading fees to those that would prevail in a frictionless economy, based, that is, on state prices that *would* obtain under zero trading fees. Denoting all quantities in the zero-trading 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

$$R_{l,t,i} \times S_{t,i} = S^*_{t,i} + \mathbb{E}_t \left[\sum_{\tau=t+1}^{T} \frac{\phi_{l,\tau-1}}{\phi_{l,t}} \times \Delta\phi_{l,\tau} \times \left(\delta_{\tau,i} + S^*_{\tau,i} \right) \right] \tag{13}$$

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 trading fees, (ii) the present value of all future price differences arising from the change in state prices and consumption induced by the presence of trading fees. *The reason for any effect of anticipated trading fees on prices is not the future fee expense itself. It is, instead, that*

[30] 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." Emphasis added.

investors do not hold the optimal, frictionless portfolios and, therefore, also have consumption schemes that differ from those that would prevail in the absence of trading fees, as we have seen in Subsection 4.1.2. The differences in consumption schemes then influence the future state prices and accordingly the present values of dividends. The increased volatility of individual consumption plays a role in setting the price because of the marginal utilities and the reduced correlation of individual consumption also plays a role via the term $\Delta\phi_{l,\tau} \times \left(\delta_{\tau,i} + S^*_{\tau,i}\right)$. Indeed, δ is a fraction of total output and whatever part of total output one group of investors is not consuming because of tradings fees, the other group is consuming.

In Figure 6, Panels (a) and (b) show the bond and stock prices at $t = 300$ for different levels of trading fees in the range from 0% to 3%. They show that, in total (solid curves), trading fees increase securities prices but not by very much.[31] We explain that result by decomposing the price increases.

We have seen in Figure 4 that trading fees have the effect of increasing the volatility of the consumption of both investors and of reducing their correlation. The effect of the increased volatility of consumption resulting from fees can be likened to the effect of increased volatility resulting from more volatile endowment shocks in the absence of trading fees, which would be *a precautionary-savings effect*. It is well known that this effect encourages saving and brings down the rate of interest as long as prudence is positive.[32] In Figure 6, the dashed curves labelled "Prec. Savings Effect" are obtained by slightly modifying the endowment shocks in such a way as to mimic the increased volatility of consumption arising from fees.[33] That increased volatility by itself *fully accounts* for the increase in the price of the bond or the reduction in the rate of interest that we obtain (see Panel (a) and also Panel (a) of Figure 7 below).

The same increased volatility of endowments produces, however, an increase in the price of the stock that is much larger than the total (see Panel (b)). The downward difference between the solid and the dashed curves is accounted for by *the drop in the correlation between individual consumptions*, which is also a drop in the correlation between the aggregate dividend and individual consumption. Wachter (2013) notes: "As is true more generally for

[31] Vayanos (1998) has noted that prices can be increased by the presence of transactions costs. Gârleanu (2009) draws a similar conclusion in a limited-trading context. However, in both of these papers, the rate of interest is exogenous.

[32] See Gollier (2001) and Barro (2009), Equation (5).

[33] Specifically, we increase α a bit from 0.125 to the level required.

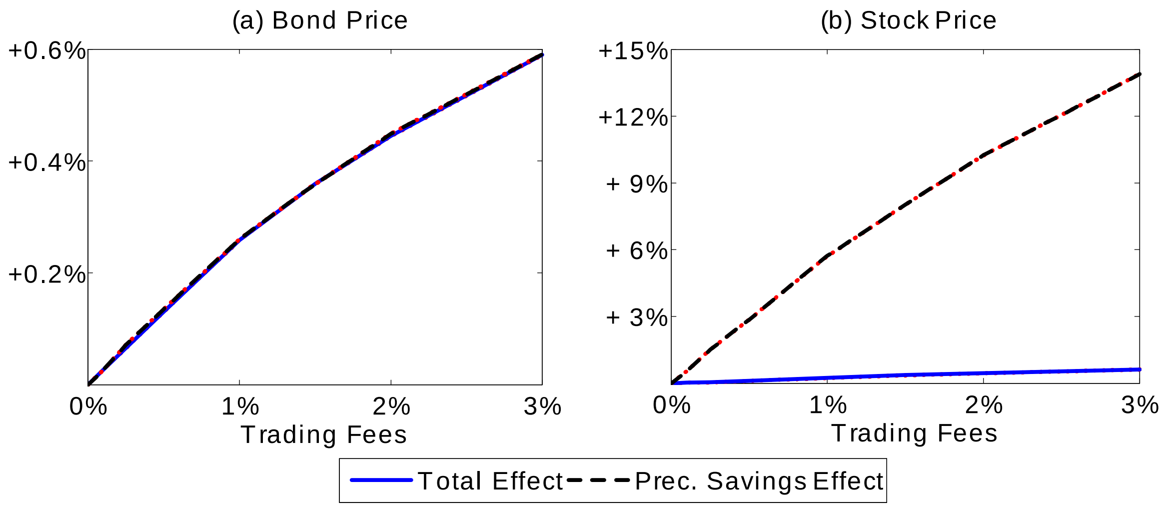

Figure 6: **Asset prices**. Panels (a) and (b) show the bond price and the stock price, for different levels of trading fees relative to their value without fees. The figure displays averages calculated at $t = 300$ across $500,000$ simulated paths. The solid curve is the average of the price. The dashed curve shows the precautionary-savings effect created by endowment shocks that would induce the same consumption volatility as do the trading fees. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

dynamic models of the price–dividend ratio (Campbell and Shiller (1988)), the net effect depends on the interplay of three forces: the effect of the [...] risk on risk premia, on the risk-free rate, and on future cash flows." Here, trading fees have no effect on cash flows but the effect on the rate of interest and on risk premia almost cancel each other, as far as stock prices are concerned.

The same mechanism, as it affects rates of return, is illustrated in Figure 7. Here also we observe the effect of trading fees and compare it to an increase in endowment volatility that artificially mimics the consumption risk added by trading fees. We find that the rate of interest is reduced and the expected stock return left unchanged by trading fees while the latter would be reduced by the endowment risk by exactly the same amount as the reduction of the rate of interest, which is exactly congruent with the earlier finding. The effect on the equity premium follows: it is increased quite markedly by trading fees while it would be left unchanged by endowment risk. The volatility of stock returns is increased somewhat by fees, which is in line with the empirical findings of Hau (2006). It also would be left unchanged by endowment risk. The net effect on the Sharpe ratio is an increase with fees while it would be unchanged with endowment risk.

5 Slow-moving investment capital

In Duffie (2010), capital moves slowly in response to a "supply shock" because some investors are inattentive at the time of the shock. Therefore, when a supply shock occurs the price of a security reacts first before the quantities adjust. When they do, the price movement is reversed. The effect is displayed in the form of impulse-response functions for prices of securities.[34]

[34] The literature on infrequent trading is burgeoning. Bacchetta and van Wincoop (2010) observe that only a small fraction of foreign-currency holdings is actively managed and calibrate a two-country model in which agents make infrequent portfolio decisions. Chien, Cole and Lustig (2012) set up an equilibrium model in which a large mass of investors do not rebalance their portfolio shares in response to aggregate shocks. Hendershott, Li, Menkveld, and Seasholes (2014) expand the Duffie (2010) slow-moving capital model to analyze multiple groups of investors. They quantify the economic effects of limited attention on asset prices by estimating a reduced form version of their model on New York Stock Exchange data. A one standard deviation change in market maker inventories is associated with transitory price movements of 65 basis points at a daily frequency and 159 basis points at a monthly frequency. Rachedi (2014), as Peress (2005) had done, introduces an observation cost in a production economy with heterogeneous agents, incomplete markets and idiosyncratic risk. Bogousslavsky (2015) shows that inattention can explain return autocorrelation patterns

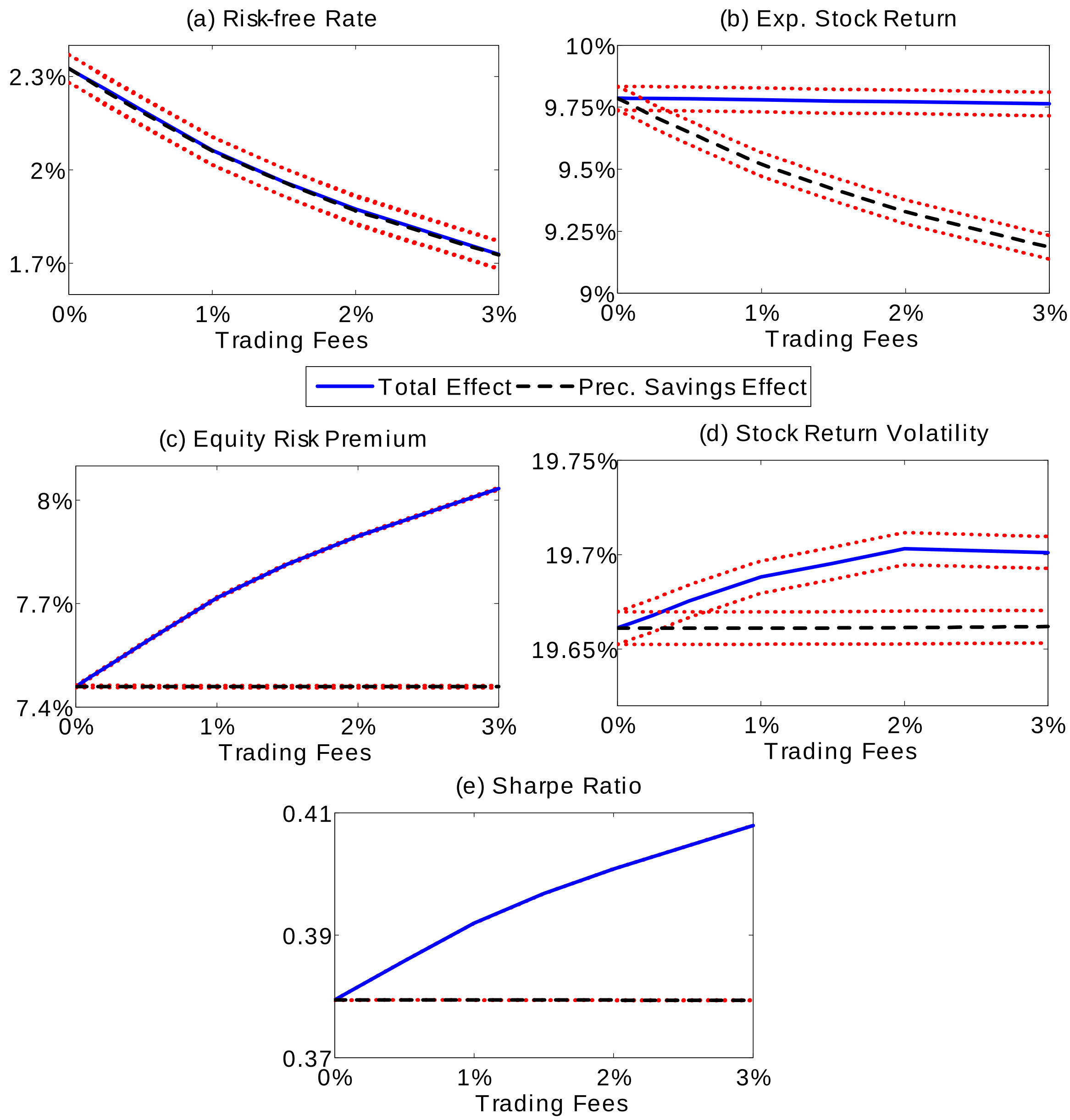

Figure 7: **Asset rates of return**. Panels (a) and (b) show the average bond return and the average conditional expected return on the stock, for different rates of trading fees. Panels (c), (d) and (e) display the average conditional equity premium, conditional volatility and Sharpe ratio of equity, respectively. The figure displays averages calculated at $t = 300$ across $500,000$ simulated paths. The solid curve represent the averages resulting from trading fees. The dashed curve shows the precautionary-savings effect created by endowment shocks that would induce the same consumption volatility as do the trading fees. All curves are bracketed by dotted lines showing the two-sigma confidence intervals for the estimate of the mean.

In the following, we want to evaluate similarly the supply-shock response of our equilibrium with trading fees and compare the two price and equity-premium responses. To perform that experiment, we need to have available a transaction price at all times. For that reason, we expand the model to three investors, *one of which only* (Investor 1) has to pay trading fees while the other two are free to trade and effectively trade all the time. So, there are now three investors receiving endowments. The exogenous process for output is unchanged, i.e., it follows a binomial tree with the same parameters.

5.1 Setup and impulse-response functions

To preserve the symmetry between investors 1 and 2 exactly, we consider a Markov chain of endowments that is the compound of two 2×2 Markov chains. In a first chain, Investor 3 gets a share of either 41.68% or 24.98%. In a second chain, the remainder is distributed to the other two investors with each one getting either 39% or 61% out of what is left. That is, on average each investor gets 1/3 of the labor income. Both separate Markov chains have a persistence of 0.85, i.e., $Prob(Future\ state\ i|State\ i) = 0.85$.

Given these parameter values, for all three investors the expected labor income growth rate is 3.86% with a volatility of 19.74% (compared to 3.85% and 19.69% in the two-investor setup).[35] So, the three-investor setup closely resembles the two-investor setup. However, while the moments of labor income growth of all three investors are exactly the same, due to the two-step procedure, their realizations are slightly different, so that only investors 1 and 2 are perfectly symmetric. With three investors, the economy takes a bit longer to reach the steady state. Our computations show that the consumption share distribution is unchanged after $t = 400$, so that our analysis focuses on the steady-state era, starting from that point in time. Specifically, the steady state consumption shares are 0.338 for both investors 1 and 2 and 0.324 for Investor 3.

The habit parameter is changed in such a way that the average surplus consumption ratio is unchanged, resulting in $h = 0.1333$. Everything else is unchanged relative to the parameter values of Table 1.

In most published work, an impulse-response function is defined as the path followed

for intraday returns. Gabaix *et al.* (2007) and Andersen *et al.* (2014) show evidence, and develop a model of, household inattention in the mortgage market.

[35]In simulations, the AR(1) coefficient of the labor income is 0.8026 versus 0.8070 in the two-group case.

by an endogenous variable after a shock occurs at a specific time, followed by a complete absence of shocks. After that point in time, the exogenous variables of the economy remain at the same level as they are after the shock. The economy becomes deterministic. That path is not representative of what will be seen if one observes the economy as an ongoing entity. In the context of the stochastic equilibrium, that path has zero probability of occurring.

A different definition of an impulse-response function is called for, to reflect the concept of a "supply shock" occurring along the way. We generate 500,000 paths of a simulation, at each point in time drawing three $[0, 1]$ uniform random numbers – and transforming them through cumulative probability distributions as needed – to determine (i) whether total output goes up or down, (ii) whether Investor 3 gets a high or a low share (first Markov chain) and (iii) whether Investor 1 or Investor 2 gets a high or a low endowment share. Please, observe that the draws from the uniform, as opposed to the output and endowment realizations themselves, make up *purely transient processes*. Then we segregate the 500,000 paths into two subsets depending on whether at time $t = 450$, the third draw from the uniform distribution (setting the endowment share between investors 1 and 2) is above or below 0.5. We refer to that difference as "the impulse." We compute the average of each of these two subsets of paths and take the difference between them. This is the difference between two sets of paths both of which are expected conditional on two levels of the impulse. They represent the effect of the impulse that an observer would actually witness on average. Empirical event studies à la Fama, Fisher, Jensen and Roll (1969) plot an average path for cumulative abnormal return (CAR) that is defined exactly the same way.

5.2 Response to an endowment shock, for the case of trading fees

The result is shown as the dashed line of Figure 8 for trading fees of 2%. By way of benchmark, the price response in the case of zero trading fees is also shown (as the solid line) and is perfectly flat because the three investors are free to trade and are able to stabilize the price, especially since the two investors receiving the impulse at $t = 450$ (investors 1 and 2) are perfectly symmetric with each other and can, therefore, wipe clean their endowment differences. This is true irrespective of the fact that the endowment impulse of time $t = 450$ is followed by other shocks. In the case of trading fees, relative to the frictionless price, the stock price is depressed by as much as 90 basis points when the fee-paying investor receives

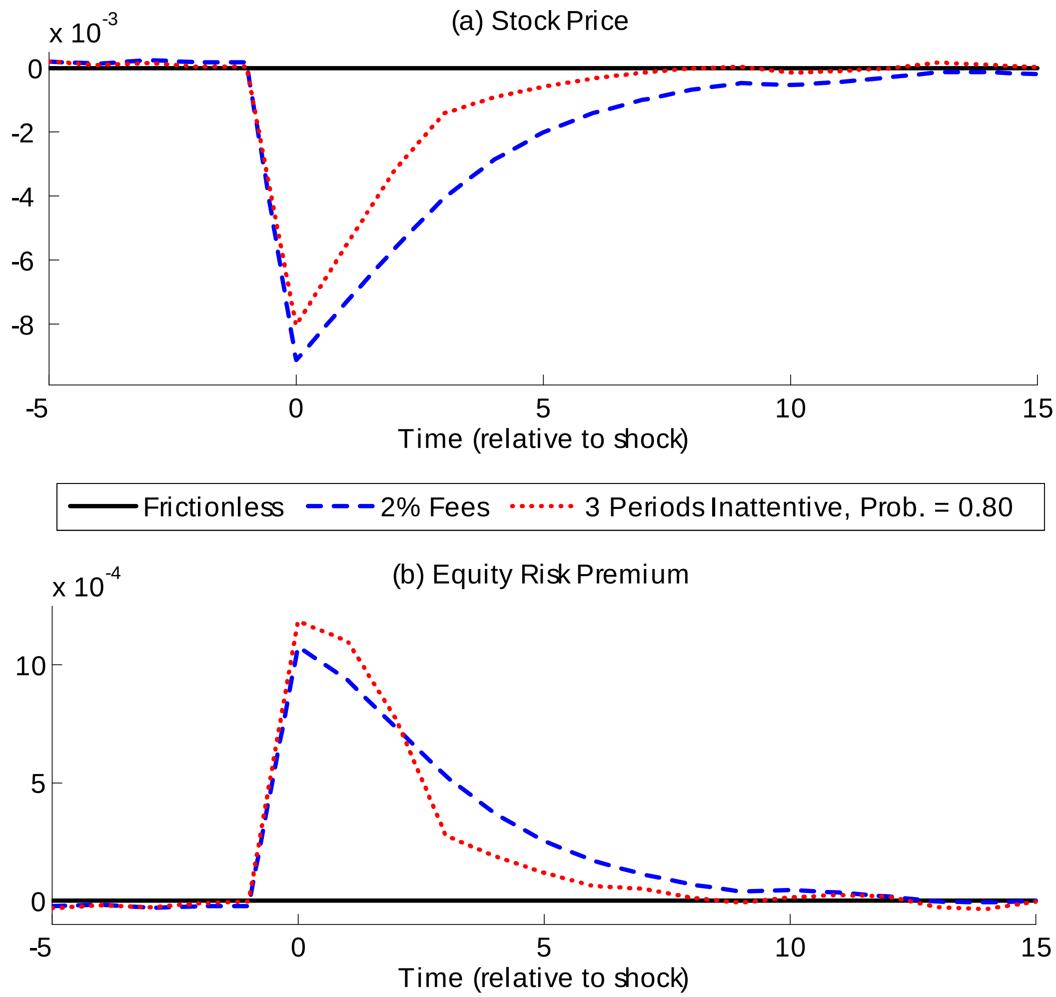

Figure 8: **Impulse-response functions (our definition) following an endowment shock to investors 1 and 2**. This is the difference between two sets of paths both of which are expected conditional on two levels of the impulse. In both panels, the solid line shows the response in the absence of friction or inattention. The dashed line shows the response when Investor 1 pays a trading fee of 2% and the dotted line shows the response when Investor 1 becomes inattentive with probability 80% and, when he is, remains inattentive for three periods.

a positive impulse. In the absence of a trading fee, he would have invested at least some of his positive endowment shock into equity shares. Since he does so in a smaller amount or less often, the average price is lower. At the same time, the equity premium is higher.

After the impulse, Figure 8 shows that the effect gradually disappears: *there is a reversal.* It is a common belief in the profession that trading fees could not produce a reversal, because fee-paying investors react instantaneously, albeit in smaller quantities. Duffie (2010) writes:

> "At the time of a supply or demand shock, the entire population of investors would stand ready to absorb the quantity of the asset supplied or demanded, with an excess price concession relative to a neoclassical model that is bounded by marginal trading costs. After the associated price shock, price reversals would not be required to clear the market."

It is true that, when the shock hits, all investors adjust immediately, and less so than they would in the absence of fees. Then, with the common concept of impulse response criticized above, since there is no more shock after the one shock of the impulse, there is no need for further adjustment and there is no reversal.

However, on an equilibrium path *with ongoing shocks*, the investors will react later on. That is true *because of hysteresis.*[36] Indeed, the impulse has moved the fee-paying investors closer to a trade boundary. When new shocks arrive, these investors will act, more so than they would have acted in the absence of the impulse.

We encounter here a great example of the difference the definition of the impulse-response function makes to the economist's understanding. With our definition, which is right because it compares expected paths, there is a reversal. With the standard concept, which is incorrect since it exhibits a zero-probability path, there would be no reversal.

5.3 Response to an endowment shock, for the case of inattention

Consider now the case in which Investor 1 may randomly become inattentive for n periods. Every investor that is attentive solves a full intertemporal optimization problem, including in his calculation the anticipation of being inattentive later. The other two investors always

[36] Recall that we have segregated the paths based on the draw from a uniform distribution, which is not persistent.

trade and provide us with a stock price. As in Duffie (2010), we have to choose the probability of becoming inattentive and the length n of periods of inattention. In Figure 8, the probability of the investor becoming inattentive is 0.8 and he becomes inattentive for 3 periods of time.

We solve the model recursively, using as endogenous state variables the consumption shares and the last period's stock holdings of the potentially inattentive investor. That's three state variables in total. For each combination of the consumption share variables, we solve a system of equations similar to the one with trading fees with the small difference that, in case the potentially inattentive investor is inattentive, then he does not agree with the other investors on the price of the stock, since he cannot trade the stock. This means that there is one equation fewer (one kernel condition), but also one choice variable fewer as the investor cannot decide on his stock holdings, which remain equal to the holdings in the period before. We keep track of his "private valuation," which we carry backwards until the investor is attentive again. When that happens, all investors agree on the price of the stock.

We draw the impulse response function the same way as in the previous subsection. Figure 8 displays the impulse response under limited attention as the dotted line, alongside the previous one. Here again, the reaction of the price to the impulse is immediate and approximately equal to 80 basis points. And we observe a reversal as in Duffie (2010).[37] It is clear that the time path is extremely similar whether we consider the trading-fee model or the limited-attention model. As far as responses to endowment shocks are concerned, they are empirically indistinguishable.

6 Conclusion

In some situations, investment capital seems to move slowly towards profitable trades. In this paper, we develop a general-equilibrium model of a financial market in which capital moves slowly simply because of trading fees.

We define a form of Walrasian equilibrium for this market. We invent an algorithm that delivers the exact numerical equilibrium, that synchronizes like clockwork the investors in the implementation of their trades and that allows us to analyze the way in which prices are

[37]Duffie (2010) assumed myopic investors. In our model of inattention, investors optimize their decisions intertemporally when they are attentive, and rationally anticipate the events in the market beyond one period. In particular, they anticipate becoming inattentive again.

formed and evolve and in which trades take place. We produce the first representation of an *equilibrium* no-trade region ever displayed.

Using parameter values that generate realistic asset price moments, we measure the degree to which the aforementioned impediments to trade prevent investors from fully smoothing their consumption and from achieving perfect risk sharing between them. Because of the fees, the volatility of individual consumption is increased and the correlation between individual consumptions is markedly reduced.

We determine the impact of trading fees on the volume of trading of the securities that are subject to these fees (the stock) and on the volume of trading of the securities that are not (the short-term bond). The trading volume of the former naturally drop. But, for low trading fees, the trading volume of the latter can also drop because the trades are exchanges of the two security types. As trading fees increase further, however, the trading volume of bonds increase because investors prefer to buy bonds as a holding tank for the stochastic endowments they receive.

We compare analytically equilibrium securities prices to the investors' private valuations and to the shadow private bid and ask prices of investors; we explain how the gap between them triggers trades. In addition, we ask whether equilibrium securities prices conform to the famous dictum of Amihud and Mendelson (1986), which says that they are reduced by the present value of transactions costs to be paid in the future. We show that in our setup the prices of securities are actually increased – but only slightly – by the presence of fees. The increase is related to the increased volatility of individual consumption, which we compare to an increase in endowment risk, and which produces a drop in the rate of interest because of precautionary saving. For risky securities, the reduced correlation of individual consumption largely offsets the increase in their volatilities, leaving only a moderate rise in price.

As for rates of return, we show that the expected rate of return on the stock is left practically unchanged, that the equity premium is increased, the volatility is increased, in line with empirical evidence, and the Sharpe ratio of the market is increased.

Finally, we compare the impulse responses of this model to those of a model in which trading is infrequent because of investor inattention. Contrary to what has been asserted by some authors, limited-attention models and trading-fees models produce very similar market responses and are, so far, equally good contenders as representations of slow-moving capital.

Most of the propositions summarized above are empirically refutable, in the manner of

Hau (2006), who showed already that volatility does increase with trading fees.

The next step in this line of research is to obtain endogenously the behavior over time of the various components of a CAPM that incorporates frictions, in the manner of Pastor and Stambaugh (2003) and Acharya and Pedersen (2005), and to let the theory prescribe how such a CAPM should be tested empirically. That task is undertaken in Buss and Dumas (2015).

Appendixes

A Proof of the equation system of Section 2.

The Lagrangian for problem (4) is:

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

where $\phi_{l,t}$ is the Lagrange multiplier attached to the flow budget constraint (5) and $\mu_{1,l,t,i}$ and $\mu_{2,l,t,i}$ are the Lagrange multipliers attached to the inequality constraints (6). The Karush-Kuhn-Tucker first-order conditions are:

$$u_l'\left(c_{l,t},\cdot,t\right)=\phi_{l,t}$$
$$e_{l,t}+\sum_{i=1}^{I}\theta_{l,t-1,i}\delta_{t,i}-c_{l,t}+\zeta_{l,t}$$
$$-\sum_{i=1}^{I}\left(\widehat{\theta}_{l,t,i}-\theta_{l,t-1,i}\right)S_{t,i}\left(1+\lambda_{i,t}\right)-\sum_{i=1}^{I}\left(\widehat{\widehat{\theta}}_{l,t,i}-\theta_{l,t-1,i}\right)S_{t,i}\left(1-\varepsilon_{i,t}\right)=0$$
$$\sum_{j=1}^{K_t}\pi_{t,t+1,j}\frac{\partial J_{l,t+1,j}}{\partial\theta_{l,t,i}}\left(\left\{\widehat{\theta}_{l,t,i}+\widehat{\widehat{\theta}}_{l,t,i}-\theta_{l,t-1,i}\right\},\cdot,e_{l,t+1,j},t+1\right)\tag{14}$$
$$=\phi_{l,t}\times S_{t,i}\times\left(1+\lambda_{i,t}\right)-\mu_{1,l,t,i}$$

$$\sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left(\left\{ \widehat{\theta}_{l,t,i} + \widehat{\widehat{\theta}}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right)$$

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

$$\widehat{\widehat{\theta}}_{l,t,i} \leq \theta_{l,t-1,i} \leq \widehat{\theta}_{l,t,i}; \mu_{1,l,t,i} \geq 0; \mu_{2,l,t,i} \geq 0$$

$$\mu_{1,l,t,i} \times \left(\widehat{\theta}_{l,t,i} - \theta_{l,t-1,i} \right) = 0; \mu_{2,l,t,i} \times \left(\theta_{l,t-1,i} - \widehat{\widehat{\theta}}_{l,t,i} \right) = 0$$

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:

$$\sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left(\left\{ \widehat{\theta}_{l,t,i} + \widehat{\widehat{\theta}}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right) = \phi_{l,t} \times R_{l,t,i} \times S_{t,i} \tag{15}$$

In order to eliminate the value function from the first-order conditions, we differentiate the Lagrangian with respect to $\theta_{l,t-1,i}$ (invoking the Envelope theorem) and use (15):

$$\frac{\partial J_l}{\partial \theta_{l,t-1,i}} = \frac{\partial L_l}{\partial \theta_{l,t-1,i}} = -\sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left(\left\{ \widehat{\theta}_{l,t,i} + \widehat{\widehat{\theta}}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right)$$

$$+\phi_{l,t} \left[\delta_{t,i} + S_{t,i} \times (1 + \lambda_{i,t}) + S_{t,i} \times (1 - \varepsilon_{i,t}) \right] - \mu_{1,l,t,i} + \mu_{2,l,t,i}$$

$$= -\sum_{j=1}^{K_t} \pi_{t,t+1,j} \frac{\partial J_{l,t+1,j}}{\partial \theta_{l,t,i}} \left(\left\{ \widehat{\theta}_{l,t,i} + \widehat{\widehat{\theta}}_{l,t,i} - \theta_{l,t-1,i} \right\}, \cdot, e_{l,t+1,j}, t+1 \right)$$

$$+\phi_{l,t} \delta_{t,i} + 2\phi_{l,t} \times R_{l,t,i} \times S_{t,i}$$

$$= \phi_{l,t} \times (\delta_{t,i} + R_{l,t,i} \times S_{t,i})$$

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

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

As has been noted by Dumas and Lyasoff (2012) in a different context, the system made of (16) and (7) 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 (16) if rewritten as:

$$
\sum_{j=1}^{K_t}\pi_{t,t+1,j}\times u_l'\left(c_{l,t+1,j},\cdot,t\right)\times\left[\delta_{t+1,i,j} + R_{l,t+1,i,j}\times S_{t+1,i,j}\right] = \phi_{l,t}\times R_{l,t,i}\times S_{t,i}; l=1,2
$$

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 kernel conditions, 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 $\theta_{l,t+1,i,j}$ ($\sum_{l=1,2}\theta_{l,t+1,i,j} = \sum_{l=1,2}\bar{\theta}_{l,i}$) held when exiting time $t+1$, which are carried backward. Regrouping equations in that way, substituting the rewritten definition (2) of the pot ζ and appending market-clearing condition (7) leads to the equation system of Section 2.

B Time 0

After solving the equation system of Section 2 till and including $t = 0$, it remains to solve for $t = -1, t + 1 = 0$ the following equation system, *not containing the kernel conditions, which have already been solved*:[38]

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:

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

3. Definitions:

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

4. Complementary-slackness conditions:

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

5. Market-clearing restrictions:

$$\sum_{l=1,2} \theta_{l,-1,i} = \sum_{l=1,2} \bar{\theta}_{l,i}$$

This system can be handled in one of two ways:

[38] There could be several possible states j at time 0 but we have removed the subscript j.

1. We can either solve for the unknowns $\left\{c_{l,0}, \theta_{l,-1,i}, \widehat{\theta}_{l,0,i}, \widehat{\widehat{\theta}}_{l,0,i}; l = 1, 2; j = 1, ..., K_t\right\}$ as functions of $\left\{\phi_{l,0}\right\}$ and $\{R_{l,0,i}\}$. If we plot $\theta_{l,-1,i}$ as functions of $\left\{\phi_{l,0}\right\}$ and $\{R_{l,0,i}\}$, we have the "Negishi map."[39] If it is invertible, we can then invert that Negishi map to obtain the values of $\left\{\phi_{l,0}\right\}$ and $\{R_{l,0,i}\}$ such that $\theta_{l,-1,i} = \bar{\theta}_{l,i}$. If the values $\bar{\theta}_{l,i}$ 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: $\left\{c_{l,0}, \phi_{l,0}, R_{l,0,i}, \widehat{\theta}_{l,0,i}, \widehat{\widehat{\theta}}_{l,0,i}; l = 1, 2; j = 1, ..., K_t\right\}$ with $\theta_{l,-1,i}$, replaced in the system by the given $\bar{\theta}_{l,i}$.

In this paper, the second method has been used.

C Scale-invariance property

Assuming the setup specified in Subsection 3.1, in which dividends and endowments are proportional to total output O_t, we now show that all the nodes of a given point in time, which differ only by their value of total output, are isomorphic to each other, where the isomorphism simply means that we can factor out the output. We provide the proof for the binomial case $K_t = 2$ $(j = 1, 2)$ with two securities $(i = 1, 2)$, one in zero net supply, the other in positive supply, and with time-additive utility but the result is valid for any number of states and habit formation utility, as long as the assumption just stated holds. The property delivers a computational benefit: at each point in time t and for each of the two values of the endowments shares, we need to solve the system (for each value of the endogenous state variables) at one node only of the binomial tree for output.[40]

Time T-1

Rewriting the investors' consumptions in terms of consumption shares $\omega_{l,T,j}$, and endowment shares of output as $\bar{\omega}_{l,T,j} \triangleq (1 - \chi) \times \upsilon_{l,T,j}$, given the fact that exiting prices are set

[39] For a definition of the "Negishi map" in a market with frictions, see Dumas and Lyasoff (2012).

[40] With habit-formation utility, however, we need to carry out the computation twice depending on which of the two earlier nodes one comes from.

at zero in the last period T, which implies that fees are also equal to zero, and given the first-order conditions for consumption, the system of equations to be solved at time $T-1$ is simply:

$$\sum_{i=1}^{2} \theta_{l,T-1,1} 1 + \sum_{i=1}^{2} \theta_{l,T-1,2} \chi O_{T,j} - (\omega_{l,T,j} - \bar{\omega}_{l,T,j}) \times O_{T,j} = 0$$

$$\begin{aligned} & \beta_1 \sum_{j=1}^{2} \pi_{T-1,T,j} \left(\frac{\omega_{1,T,j}}{\omega_{1,T-1}} \times \frac{O_{T,j}}{O_{T-1}} \right)^{-\gamma_1} \\ = \quad & \beta_2 \sum_{j=1}^{2} \pi_{T-1,T,j} \left(\frac{\omega_{2,T,j}}{\omega_{2,T-1}} \times \frac{O_{T,j}}{O_{T-1}} \right)^{-\gamma_2} \end{aligned} \tag{17}$$

$$\begin{aligned} & \frac{\beta_1}{R_{1,T-1,i}} \sum_{j=1,2} \pi_{T-1,T,j} \times \left(\frac{\omega_{1,T,j}}{\omega_{1,T-1}} \times \frac{O_{T,j}}{O_{T-1}} \right)^{-\gamma_1} \times \chi O_{T,,j} \\ = \quad & \frac{\beta_2}{R_{2,T-1,i}} \sum_{j=1,2} \pi_{T-1,T,j} \times \left(\frac{\omega_{2,T,j}}{\omega_{2,T-1}} \times \frac{O_{T,j}}{O_{T-1}} \right)^{-\gamma_2} \times \chi O_{T,,j} \end{aligned} \tag{18}$$

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

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

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

$$\begin{cases} \theta_{l,T-1,1} + O_{T-1} \times [\theta_{l,T-1,2} \chi u - (\omega_{l,T,1} - \bar{\omega}_{l,T,1}) \times u] = 0 \\ \theta_{l,T-1,1} + O_{T-1} \times [\theta_{l,T-1,2} \chi d - (\omega_{l,T,2} - \bar{\omega}_{l,T,2}) \times d] = 0 \end{cases}$$

The solution for the holdings is:

$$\theta_{l,T-1,1} = u \frac{O_{T-1}}{\chi (d-u)} [d \times (\omega_{l,T,1} - \bar{\omega}_{l,T,1}) - (\omega_{l,T,2} - \bar{\omega}_{l,T,2}) \times d] \tag{19}$$

$$\theta_{l,T-1,2} = \frac{1}{\chi (d-u)} (- (\omega_{l,T,1} - \bar{\omega}_{l,T,1}) \times u + (\omega_{l,T,2} - \bar{\omega}_{l,T,2}) \times d) \tag{20}$$

Rewriting the kernel conditions (17, 18) 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:

$$\beta_1 \sum_{j=1}^{2} \pi_{T-1,T,j} \left(\frac{\omega_{1,T,j}}{\omega_{1,T-1}}\right)^{-\gamma_1} r_j^{-\gamma_1} = \beta_2 \sum_{j=1}^{2} \pi_{T-1,T,j} \left(\frac{\omega_{2,T,j}}{\omega_{2,T-1}}\right)^{-\gamma_2} r_j^{-\gamma_2}$$

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

where $r_j = u$ for $j = 1$ and $r_j = d$ for $j = 2$. The market-clearing conditions:

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

become:

$$\begin{aligned} &\frac{1}{\chi(d-u)} \left[d \times (\omega_{1,T,1} - \bar{\omega}_{1,T,1}) \times u - u \times (\omega_{1,T,2} - \bar{\omega}_{1,T,2}) \times d\right] \\ &+ \frac{1}{\chi(d-u)} \left[d \times (\omega_{2,T,1} - \bar{\omega}_{2,T,1}) \times u - u \times (\omega_{2,T,2} - \bar{\omega}_{2,T,2}) \times d\right] = 0 \end{aligned}$$

$$\begin{aligned} &\frac{1}{\chi(d-u)} \left[-(\omega_{1,T,1} - \bar{\omega}_{1,T,1}) \times u + (\omega_{1,T,2} - \bar{\omega}_{1,T,2}) \times d\right] \\ &+ \frac{1}{\chi(d-u)} \left(-(\omega_{2,T,1} - \bar{\omega}_{2,T,1}) \times u + (\omega_{2,T,2} - \bar{\omega}_{2,T,2}) \times d\right) = 1 \end{aligned}$$

Importantly this system of equations *does not depend on the current or future levels of output*, 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) one gets that the stock holdings are independent of $T-1$ output, while from (19) we know that the bond holdings are scaled by the $T-1$ output:

$$\theta_{l,T-1,1} = O_{T-1} \times \breve{\theta}_{l,T-1,1}, \tag{21}$$

where $\breve{\theta}_{l,T-1,1}$ denotes the normalized bond holdings for $O_{T-1} = 1$. Moreover, 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 the stock price is scaled by the $T-1$ output:

$$\begin{aligned} S_{T-1,2} &= \chi O_{T-1} \times \left[\frac{\beta_1}{R_{1,T-1,i}} \sum_{j=1}^{2} \pi_{T-1,T,j} \left(\frac{\omega_{1,T,j}}{\omega_{1,T-1}} \right)^{-\gamma_1} \frac{r_j^{-\gamma_{1+1}}}{2} \right] \\ &\triangleq O_{T-1} \times \breve{S}_{T-1,2}, \end{aligned} \tag{22}$$

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

Time t<T-1

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

$$\begin{cases} \left(\omega_{l,t+1,1} - \bar{\omega}_{l,t+1,1}\right) u + \breve{\theta}_{l,t+1,1,1} u + \left(\theta_{l,t+1,2,1} - \theta_{l,t,2}\right) \breve{S}_{t+1,2,1} u R_{l,t+1,2,1} \\ \qquad = \frac{\theta_{l,t,1}}{O_t} + \theta_{l,t,2} \chi u + \left(\widehat{\theta}_{l',t+1,2,1} - \theta_{l',t,2}\right) \breve{S}_{t+1,2,1} u \lambda_{2,t+1,1} \\ \qquad\qquad - \left(\widehat{\widehat{\theta}}_{l',t+1,2,1} - \theta_{l',t,2}\right) \breve{S}_{t+1,2,1} u \varepsilon_{2,t+1,1} \\ \left(\omega_{l,t+1,2} - \bar{\omega}_{l,t+1,2}\right) d + \breve{\theta}_{l,t+1,1,2} d + \left(\theta_{l,t+1,2,2} - \theta_{l,t,2}\right) \breve{S}_{t+1,2,2} d R_{l,t+1,2,2} \\ \qquad = \frac{\theta_{l,t,1}}{O_t} + \theta_{l,t,2} d + \left(\widehat{\theta}_{l',t+1,2,2} - \theta_{l',t,2}\right) \breve{S}_{t+1,2,2} d \lambda_{2,t+1,2} \\ \qquad\qquad - \left(\widehat{\widehat{\theta}}_{l',t+1,2,2} - \theta_{l',t,2}\right) \breve{S}_{t+1,2,2} d \varepsilon_{2,t+1,2} \end{cases}$$

or:

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

where

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

Solving for the holdings:

$$\frac{\theta_{l,t,1}}{O_t} = u\frac{1}{d-u}\left(d \times \kappa_{l,t+1,1} - \kappa_{l,t+1,2} \times d\right) \tag{23}$$

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

Rewriting the kernel conditions,the system is:

$$\beta_1 \sum_{j=1}^{2} \pi_{t,t+1,j} \left(\frac{\omega_{1,t+1,j}}{\omega_{1,t}}\right)^{-\gamma_1} r_j^{-\gamma_1} = \beta_2 \sum_{j=1}^{2} \pi_{t,t+1,j} \left(\frac{\omega_{2,t+1,j}}{\omega_{2,t}}\right)^{-\gamma_2} r_j^{-\gamma_2}$$

$$\begin{aligned} & \frac{\beta_1}{R_{1,t,2}} \sum_{j=1}^{2} \pi_{t,t+1,j} \left(\frac{\omega_{1,t+1,j}}{\omega_{1,t}}\right)^{-\gamma_1} r_j^{-\gamma_1} \left(R_{1,t+1,2,j} \times \bar{S}_{t+1,2,j} \times r_j + \frac{r_j}{2}\right) \\ = \; & \frac{\beta_1}{R_{2,t,2}} \sum_{j=1}^{2} \pi_{t,t+1,j} \left(\frac{\omega_{2,t+1,j}}{\omega_{2,t}}\right)^{-\gamma_2} r_j^{-\gamma_2} \left(R_{2,t+1,2,j} \times \bar{S}_{t+1,2,j} \times r_j + \frac{r_j}{2}\right) \end{aligned}$$

$$\theta_{l,t+1,2,j} = \widehat{\theta}_{l,t+1,2,j} + \widehat{\widehat{\theta}}_{l,t+1,2,j} - \theta_{l,t,2}$$

$$\begin{aligned} \left(-R_{l,t+1,2,j} + 1 + \lambda_{2,t+1,j}\right) \times \left(\widehat{\theta}_{l,t+1,2,j} - \theta_{l,t,2}\right) &= 0 \\ \left(R_{l,t+1,2,j} - (1 - \varepsilon_{2,t+1,j})\right) \times \left(\theta_{l,t,2} - \widehat{\widehat{\theta}}_{l,t+1,2,j}\right) &= 0 \end{aligned}$$

$$\sum_{l=1,2} \theta_{l,t,1} = 0; \ \sum_{l=1,2} \theta_{l,t,2} = 1$$

where the unknowns are $\left\{\omega_{l,t+1,j}; R_{l,t+1,2,j}; \widehat{\theta}_{l,t+1,2,j}; \widehat{\widehat{\theta}}_{l,t+1,2,j}; l = 1,2; j = 1,2\right\}$ and where the holdings are given by (23) and (24). The output O_t cancels out in the market clearing conditions for the bond. Thus, the full system *does not depend on the level of the ouput* O_t, only on u as well as d.

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 $\breve{S}_{t+1,2,j}$ and normalized bond holdings $\breve{\theta}_{l,t+1,1,j}$. 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 O_t, while the holdings in the stock and the bond price are not scaled. *Using backward induction, it follows that 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 dividends δ from the point of view of investor l is given by (subscript i omitted):

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

whereas Equation (11) applied to time $T-1$ is:

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

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

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

whereas Equation (11) applied to time $T-2$ is:

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

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

By an induction argument one obtains the final result (12).

E Proof of Proposition 4

The proof is by induction.

At date $t = T-1$, the stock price in an economy without trading fees is given by (subscript i omitted):

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

whereas Equation (11) 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:

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

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-trading fees economy S^*_{T-1} and the stock price in an economy with trading fees S_{T-1}:

$$R_{l,T-1} \times S_{T-1} - S^*_{T-1} = \mathbb{E}_{T-1}\left[\Delta\phi_{l,T}\delta_T\right] \tag{26}$$

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

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

whereas Equation (11) applied to time $T-2$ is:

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

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

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

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

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

By an induction argument one reaches the final result (13).

References

Acharya, V. and L. H. Pedersen, 2005, "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77, 375-4010.

Amihud, Y. and H. Mendelson, 1986, "Asset Pricing and the Bid-ask Spread," *Journal of Financial Economics,* 17, 223-249.

Andersen, S., J. Y. Campbell, K. Meisner-Nielsen, and T. Ramadorai, 2014, "Inattention and Inertia in Household Finance: Evidence from the Danish Mortgage Market," working paper, Copenhagen Business School.

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.

Ashcraft, A. and D. Duffie, 2007, "Systemic illiquidity in the federal funds market," *American Economic Review Papers and Proceedings*, 97, 221–225.

Bacchetta, P. and E. van Wincoop, 2010, "Infrequent Portfolio Decisions: A Solution to the Forward Discount Puzzle," *American Economic Review*, 100, 870–904.

Barro, Robert J., 2009, "Rare Disasters, Asset Prices, and Welfare Costs," *American Economic Review*, 99, 243-64.

Bismut, J. M., 1973, "Conjugate convex functions in optimal stochastic control," *Journal of Mathematical Analysis and Applications*, 44, 384-404.

Bouchard, B., 2002, "Utility Maximization on the Real Line under Proportional Transactions Costs," *Finance and Stochastics*, 6, 495-516.

Bogousslavsky, V., 2015, "Infrequent Rebalancing, Return Autocorrelation, and Seasonality," working paper, Swiss Finance Institute.

Brunnermeier, M. and L. H. Pedersen, 2008, "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22, 2201-2238.

Buss, A. and B. Dumas, 2015, "The Pricing of Liquidity Risk in General Equilibrium," working paper, INSEAD.

Buss, A., R. Uppal and G. Vilkov, 2015, "Where Experience Matters: Asset Allocation and Asset Pricing with Opaque and Illiquid Assets," working paper, INSEAD.

Campbell, J. Y., 2003, "Consumption-based asset pricing, " *Handbook of the Economics of Finance,* 1, Part B, 803–887.

Campbell, J. Y., and J. H. Cochrane, 1999, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205 -251.

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.

Carroll, C. D., 1993, "The Buffer-Stock Theory of Savings: Some Macroeconomic Evidence," *Brookings Papers on Economic Activity 1992:2, Macroeconomis.*

Carroll, C. D., 1997, "Buffer-Stock Saving and the Life-Cycle/Permanent-Income Hypothesis," *The Quarterly Journal of Economics*, CXII, 1-55.

Chien, Y., H. Cole and H. Lustig, 2012, "Is the Volatility of the Market Price of Risk Due to Intermittent Portfolio Rebalancing?", *American Economic Review*, 102, 2859–2896.

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.

Coval, J., and E. Stafford, 2007, "Asset fire sales (and purchases) in equity markets," *Journal of Financial Economics*, 86, 479–512.

Cvitanic J. and I. Karatzas, 1996, "Convex Duality in Constrained Portfolio Optimization," *The Annals of Applied Probability*, 2, 767-818.

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.

Delgado, F., Dumas, B. and G. Puopolo, 2014, "Hysteresis Bands on Returns, Holding Period and Transactions Costs," *Journal of Banking and Finance*, forthcoming.

Demsetz, H., 1968, "The Cost of Transacting," *Quarterly Journal of Economics*, 82, 33-53.

Dixit, A., 1991, "Analytical Approximations in Models of Hysteresis," *Review of Economic Studies*, Vol. 58, No. 1, 141-151.

Duffie, D., 2010, "Presidential Address: Asset Price Dynamics with Slow-Moving Capital," *The Journal of Finance*, 55, 1237-167.

Duffie, D., N. Gârleanu, and L. H. Pedersen, 2002, "Securties lending, shorting, and pricing," *Journal of Financial Economics*, 66, 307–339.

Duffie, D., N. Gârleanu, and L. H. Pedersen, 2007, "Valuation in over-the-counter markets," *Review of Financial Studies*, 20, 1865–1900.

Duffie, D., G. Manso, and S. Malamud, 2014, "Information percolation with segmented markets," *Journal of Economic Theory*, 153, 1–32.

Duffie, D., and B. Strulovici, 2012, "Capital mobility and asset pricing," *Econometrica*, 80, 2469–2509.

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*, 67,1897-1940.

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.

Edmans, A., I. Goldstein and W. Jiang, 2012, "The Real Effects of Financial Markets: The Impact of Prices on Takeovers," *The Journal of Finance*, 67, 933-971.

Fama, E. F., L. Fisher, M. C. Jensen and R. Roll, 1969, "The Adjustment of Stock Prices to New Information," *International Economic Review*, 10, 1-21.

Fleming, W. H., S. G. Grossman, J.-L. Vila and T. Zariphopoulou, 1990, "Optimal Portfolio Rebalancing with Transactions Costs," working paper, Brown University.

Gabaix, X., A. Krishnamurthy, and O. Vigneron, 2007, "Limits of Arbitrage: Theory and Evidence from the Mortgage-Backed Securities Market", *The Journal of Finance*, 62, 557–595.

Garcia, C. and W. Zangwill, 1981, *Pathways to Solutions, Fixed Points and Equilibria*, Prentice-Hall, Englewood Cliffs.

Gârleanu, N., 2009, "Portfolio choice and pricing in illiquid markets," *Journal of Economic Theory*, 144, 532–564

Gârleanu, N. and L. Pedersen, 2013, "Dynamic Trading with Predictable Returns and Transactions Costs," *The Journal of Finance*, 68, 2309–2340.

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.

Gollier, C., 2001, *The Economics of Risk and Time*, The MIT Press.

Gomes, F. and A. Michaelides, 2005, "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence," *The Journal of Finance*, 60, 869–904.

Gourinchas, P.-O., and J. Parker, 2002, "Consumption over the Life-Cycle," *Econometrica*, 70, 47–89.

Hau, H., 2006, "The Role of Transaction Costs for Financial Volatility: Evidence from the Paris Bourse," *Journal of the European Economic Association*, 4, 862-890.

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.

Hendershott, P., S. X. Li, A. Menkveld and M. Seasholes, 2014, "Asset Price Dynamics with Limited Attention," working paper, VU University of Amsterdam.

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.

Jotikasthira, C., C. Lundblad and T. Ramadorai, 2012, "Asset Fire Sales and Purchases and the International Transmission of Funding Shocks," *The Journal of Finance*, 67, 2015–2050.

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.

Kulak, J. P., 2012, *Essays in Financial Economics*, Swiss Finance Institute.

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

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

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., 2009, "Portfolio Claustrophobia: Asset Pricing in Markets with Illiquid Assets," *The American Economic Review*, 99, 1119-1144.

Ljunqvist, L. and T. Sargent, 2012, *Recursive Macroeconomic Theory*, third edition, MIT Press.

Luttmer, E. G. J., 1996, "Asset Pricing in Economies with Frictions," *Econometrica*, 64, 1439-1467.

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.

Mitchell, M., and T. Pulvino, 2012, "Arbitrage crashes and the speed of capital," *Journal of Financial Economics*, 104, 469–490.

Mitchell, M., T. Pulvino, and E. Stafford, 2004, Price pressure around mergers," *The Journal of Finance*, 59, 31–63.

Nazareth, M., 2002, "Portfolio Selection with Stochastic Transaction Costs," working paper.

Newman, Y., and M. Rierson, 2004, "Illiquidity spillovers: Theory and evidence from European telecom bond issuance," Working paper, Graduate School of Business, Stanford University.

Obizhaeva, A. and J. Wang, 2005, "Optimal Trading Strategy and Supply/demand Dynamics," *Journal of Financial Markets*, 16, 1–32

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.

Rachedi, O., 2014, "Asset Pricing with Heterogeneous Inattention," working paper, Universidad Carlos III de Madrid.

Scholes, M. S., 1972, "The market for securities: Substitution versus price pressure and the effects of information on share price," *Journal of Business*, 45, 179–211.

Shreve, S. E. and H. M. Soner, 1994, "Optimal Investment and Consumption with Transactions Costs," *Annals of Applied Probability*, 4, 609-692.

Stoll, H. R., 2000, "Friction," *The Journal of Finance*, 55, 1479-1514.

Svensson, L. E. O., 1991, "Target Zones and Interest Rate Variability," *Journal of International Economics*, 31, 27-54.

Tuckman, B. and J.-L. Vila, 2010, "Holding Costs And Equilibrium Arbitrage," bikibook.com.

Vayanos, D., 1998, "Transaction Costs and Asset Prices: A Dynamic Equilibrium Model," *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, 2012, "Liquidity and Asset Prices under Asymmetric Information and Imperfect Competition," *Review of Financial Studies*, 25, 1339-1365.

Wachter, J. A., 2013, "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *The Journal of Finance*, 68, 987-1035.