Equilibrium Pricing and Trading Volume under Preference Uncertainty*

Bruno Biais,† Johan Hombert,‡ and Pierre-Olivier Weill§

July 16, 2013

Abstract

Information collection, processing and dissemination in financial institutions is challenging. This can delay the observation by traders of the exact capital charges and constraints of their institution. During this delay, traders face preference uncertainty. In this context, we study optimal trading strategies and equilibrium prices in a continuous centralized market. We focus on liquidity shocks, during which preference uncertainty is likely to matter most. Preference uncertainty generates allocative inefficiency, but need not reduce prices. Traders progressively learning about the preferences of their institution conduct round–trip trades, which generate excess volume relative to the frictionless market. In a cross section of liquidity shocks, the initial price drop is positively correlated with total trading volume. Across traders, the number of round–trips is negatively correlated with trading profits and average inventory.

Keywords: Information Processing, Trading Volume, Liquidity Shock, Preference Uncertainty, Equilibrium Pricing

J.E.L. Codes: D8, G1

---

*Earlier versions of this paper circulated under the titles "Trading and Liquidity with Limited Cognition", "Liquidity Shocks and Order Book Dynamics", and "Pricing and Liquidity with Sticky Trading Plans". Many thanks for insightful comments to the Editor, Dimitri Vayanos, and three referees, as well as Andy Atkeson, Dirk Bergemann, Darrell Duffie, Emmanuel Farhi, Thierry Foucault, Xavier Gabaix, Alfred Galichon, Christian Hellwig, Hugo Hopenhayn, Vivien Lévy–Garboua, Johannes Horner, Boyan Jovanovic, Ricardo Lagos, Albert Menkveld, John Moore, Stew Myers, Henri Pages, Thomas Philippon, Gary Richardson, Jean Charles Rochet, Guillaume Rocheteau, Ioanid Rosu, Larry Samuelson, Tom Sargent, Jean Tirole, Aleh Tsyvinski, Juusso Välimäki, Adrien Verdelhan, and Glen Weyl; and to seminar participants at the Dauphine-NYSE-Euronext Market Microstructure Workshop, the European Summer Symposium in Economic Theory at Gerzensee, École Polytechnique, Stanford Graduate School of Business, New York University, Northwestern University, HEC Montreal, MIT, UCI, CREI/Pompeu Fabra, Bocconi, London School of Economics, University of Zurich, Columbia Economics, New York Fed, NYU Stern, UBC Finance, Wharton Finance, Federal Reserve Bank of Philadelphia Search and Matching Workshop. Paulo Coutinho and Kei Kawakami provided excellent research assistance. Bruno Biais benefitted from the support of the European Research Council, grant 295484 TAP, and Pierre-Olivier Weill from the support of the National Science Foundation, grant SES-0922338.

†Toulouse School of Economics (CNRS-CRM and Federation des Banques Francaise Chair on the Financial Markets and Investment Banking Value Chain), `bruno.biais@univ-tlse1.fr`

‡HEC Paris, `johan.hombert@hec.fr`

§University of California Los Angeles, NBER and CEPR, `poweill@econ.ucla.edu`

1 Introduction

Financial firms' investment choices are constrained by capital requirements and investment guidelines, as well as risk–exposure and position limits. To assess the bindingness and the cost of these constraints, so as to determine the corresponding constrained optimal investment position, each financial firm must collect data from several trading desks and divisions, and then aggregate these data. This is a difficult task, studied theoretically by Vayanos (2003), who analyzes the challenges raised by the aggregation of risky positions within a financial firm subject to communication constraints. These challenges have also been emphasized by several regulators and consultants.[1]

Because data collection and aggregation is challenging, it takes time. For example, Ernst & Young (2012, page 58) finds that "53% of [respondents in its study] aggregate counterparty exposure across business lines by end of day, 27% report it takes two days, and 20% report much longer processes."[2] This delays the incorporation of relevant information into investment decisions, particularly in times of market stress.[3]

From a theoretical perspective, these stylized facts imply that financial firms' traders make decisions under *preference uncertainty*.[4] That is, during the time it takes to reassess financial and regulatory constraints, the traders in charge of implementing the optimal investment policy of a firm are uncertain about the preferences of the latter. The goal of this paper is to examine the consequences of such preference uncertainty for trading strategies, equilibrium pricing and aggregate trading volume.

To do so, we focus on situations where the market is hit by an aggregate liquidity shock, in

[1] See Basel Committee on Banking Supervision (2009) and Ernst & Young (2012, page 9): "Many firms face challenges extracting and aggregating appropriate data from multiple siloed systems, which translate into fragmented management information on the degree of risk facing the organization."Ernst & Young (2012, page 20), however, also notes that financial firms use these data to assess their risk appetite and that "close to half [the respondents] (49%) report that stress testing results are significantly incorporated into risk management decision making."

[2] See also Ernst & Young (2012, page 76): "The most prominent challenge is the sheer amount of time it takes to conduct stress testing [via] what is often a manual process of conducting test and gathering results across portfolios and businesses." Similarly, the Institute for International Finance (2011, page 50) mentions, some of the respondents to its study "say that their process lacks the capability to produce near-real-time and real-time reports on exposure and limit usage."

[3] As noted by Mehta et al. (2012, page 7): "Most banks calculate economic capital on a daily (30%) or weekly (40%) basis, actively using it for risk steering and definition of limits in accordance with the risk appetite."See also Mehta et al. (2012, page 5): "Across all banks, the survey found that average Value-at-Risk run time ranges between 2 and 15 hours; in stressed environments, it can take much longer".

[4] By "preference uncertainty" we mean that traders view the utility function of their institution as a random variable, but we don't use the word "uncertainty" in the Knightian sense.

which firms' willingness and ability to hold assets is reduced, due, e.g., to losses (Berndt et al., 2005), increased risk, asset downgrades or index changes (Greenwood, 2005), or margin calls and fund outflows (Coval and Stafford, 2007). As mentioned above, it is in such times of stress that preference uncertainty is likely to be most severe. To cope with the shock, financial firms establish hedges, raise new capital, and adjust positions in several assets and contracts. This process is complex, and involves transactions conducted by different desks in different markets. It takes time to complete it and also, as discussed above, to collect, process and disseminate the corresponding information to all traders in the firm.

To model this situation, we consider an infinite-horizon, continuous-time market for one particular asset. There is a continuum of infinitely lived, risk–neutral and competitive financial firms who derive a non–linear utility flow, denoted by $v(\theta, q)$, from holding q divisible shares of this asset, as in Gârleanu (2009) and Lagos and Rocheteau (2009). At the time of the liquidity shock, the utility flow parameter drops to θ_ℓ for some of the firms, as in Weill (2004, 2007) and Duffie, Gârleanu, and Pedersen (2007). This drop in utility flow reflects the increase in capital charges and the additional regulatory costs of holding the asset induced by the liquidity shock. Then, as time goes by, the firms hit by the shock progressively switch back to a high valuation, $\theta_h > \theta_\ell$. This switch occurs when a firm has successfully established the hedges and adjustments in capital and position necessary to absorb the liquidity shock and correspondingly recover a high valuation for the asset. To model this we assume each firm is associated with a Poisson process and switches back to high-valuation at the first jump in this process. Furthermore, to model preference uncertainty we assume each firm is represented in the market by a trader who observes her firm's current valuation for the asset, θ, at Poisson distributed "updating times." Each firm is thus exposed to two Poisson processes: one jumps with its valuation for the asset, and the other jumps when its trader observes updated information about that valuation. For tractability, we assume that these processes are independent and independent across firms.

In this context, a trader does not continuously observe the utility flow generated for her firm by the position she takes. She, however, designs and implements the trading strategy that is optimal for the firm, given her information. Thus, when a trader observes updated information about the preferences of her firm, she designs a new trading plan, specifying the process of her asset holdings until the next information update, based on rational expectations about future variables and decisions. At each point time, the corresponding demand from a trader is increasing in the probability that her firm has high valuation. Substituting demands in the market clearing condition gives rise to equilibrium prices. We show equilibrium existence and uniqueness. By the law of large numbers, the cross–sectional aggregate distribution of preferences, information sets and demands is deterministic, and so is the equilibrium price.[5]

Unconstrained efficiency would require that low-valuation traders sell to high-valuation traders. However, such asset reallocation is delayed by preference uncertainty. Some traders hold more shares than they would if they knew the exact current status of their firm, while others hold less shares. Thus preference uncertainty generates allocative inefficiency. This does not necessarily translate into lower prices, however. Indeed, preference uncertainty has two effects on asset demand, going in opposite directions. On the one hand, demand increases because traders who currently have low valuation don't know it for sure and believe that they may have a high valuation with positive probability. On the other hand, by the same token, demand decreases because traders who currently have a high valuation believe that they may have a low valuation with positive probability. If the utility function is such that demand is concave in the probability that the firm has high valuation, the former effect dominates the latter, so that preference uncertainty actually increases prices. The opposite holds if asset demand is convex in this probability. We also analyze in closed form a specification where demand is neither globally concave nor convex: in this case, we show that preference uncertainty may increase prices when the liquidity shock hits, but subsequently lowers them as the shock subsides.

With known preferences, each trader observes the valuation of her firm, hence the cross–sectional variance of valuation parameters is $\mathbb{V}[\theta]$. With preference uncertainty, traders don't observe the valuation of their firms. Instead they form expectations about it, conditional on their information $\mathcal{F}$. Hence the cross–sectional variance of valuations across traders is $\mathbb{V}[\mathbb{E}(\theta \,|\, \mathcal{F})]$. Since $\mathbb{V}[\mathbb{E}(\theta \,|\, \mathcal{F})] < \mathbb{V}[\theta]$, in a static model, preference uncertainty would lead to lower trading volume, due to smaller dispersion of valuations. In our dynamic model, however, the opposite occurs, because trades arise due to *changes* in expected valuations. Such changes occur more often with preference uncertainty than with known preferences, which tends to increase trading volume. More precisely, the economic mechanism underlying trades with preference uncertainty is the following: When traders observe their firm still has low valuation, they sell a block of shares. Then, until the next updating time, they remain uncertain about the exact valuation of their firm. They anticipate, however, that it is more and more likely

[5]In Biais, Hombert, and Weill (2012a), we analyzed an extension of our framework where the market is subject to recurring aggregate liquidity shocks, occurring at Poisson arrival times. In Appendix B.5, we consider the case when the number of traders is finite and the law of large numbers no longer applies. While, in both extensions, the price becomes stochastic, the qualitative features of our equilibrium are upheld.

that their firm has emerged from the shock. Correspondingly, under natural conditions, they gradually buy back shares, which they may well sell back at their next updating time if they learn their firm still has low valuation. This generates round trips, and larger trading volume than when preferences are known. In some sense, preference uncertainty implies a tâtonnement process, by which the allocation of the asset progressively converges towards the efficient allocation. Successive corrections in this tâtonnement process generate excess volume relative to the known preference case. When the frequency of information updates increases, the size of the round trip trades decreases, but their number becomes larger. We show that, as the frequency of information updates goes to infinity and preference uncertainty vanishes, the two effects balance out exactly, so that the excess volume converges to some non-zero limit.

The condition under which preference uncertainty raises prices is related to the condition under which it generates excess volume (more precisely the latter condition is necessary for the former.) This is because the force that leads to increased prices is the demand coming from traders who think their firm may have switched to high valuation, while in fact it still has low valuation. It is precisely the same force that leads to the build–up of inventories, that are then unwound via a block sale when the trader observes her firm still has low valuation. And it is this inventory turnover which generates the round-trip trades that are at the origin of excess volume.

A natural measure of the magnitude of the liquidity shock is the fraction of traders initially hit. As this fraction increases, the initial price drop generated by the shock increases, and so does the total trading volume following the shock. Thus, one empirical implication of our analysis is that, in a cross–section of liquidity shocks, the magnitude of the initial price drop should be positively correlated with the total trading volume following the shock. Our theoretical analysis also generates empirical implications for the cross–section of traders in a given liquidity shock episode. In equilibrium, a trader whose institution recovers rapidly holds large inventories, and makes only a few round trips. In contrast, a trader whose institution recovers late in the cycle engages in many successive round trips. Correspondingly she holds inventory during short periods of time. Furthermore, the traders whose institutions recover late earn lower trading profits than those whose institutions recover early, since the latter buy early (at low prices), while the former also buy late (at higher prices.) Thus, our theoretical analysis implies negative correlation, among traders, between the number of round–trips and trading profits, and also positive correlation between traders' average inventories and trading profits.

The latter correlation can be interpreted in terms of reward to liquidity supply: Traders with low valuation sell the asset, thus demanding liquidity. Traders accommodating this demand supply liquidity by buying the asset and holding inventories. Our implication reflects that such liquidity supply is profitable.

Our assumption that institutions are unable to collect, process and disseminate all information instantaneously is in the spirit of the rational inattention literature, which emphasizes that economic agents have limited information processing ability (see, e.g., Lynch, 1996, Reis, 2006a, 2006b, Mankiw and Reis, 2002, Gabaix and Laibson, 2002, Alvarez, Lippi, and Paciello, 2011, Alvarez, Guiso, and Lippi, 2010.) Our analysis complements theirs by focusing on a different object: financial institutions and traders during liquidity shocks, and the corresponding equilibrium prices and trading strategies.

Much of our formalism builds on search models of over-the-counter (OTC) markets, such as those of Duffie, Gârleanu, and Pedersen (2005), Weill (2007), Gârleanu (2009), Lagos and Rocheteau (2009), Lagos, Rocheteau, and Weill (2011), and Pagnotta and Philippon (2011). In particular, we follow these models in assuming that investors' valuations change randomly. Also, the friction we consider (infrequent observation of types) is, in some sense, comparable to that they study (infrequent contact with markets). That being said, the market structure we consider (centralized, continuous, limit order markets), is very different from the fragmented dealer market structure they consider. Correspondingly, what happens between times at which our traders observe their valuation differs from what happens in search models of OTC markets between times at which traders contact the market. During this time interval, in our approach, traders engage in active trading strategies, while they must stay put in models of OTC markets. This, in turn, leads to a qualitative difference in implications: In our analysis, the friction leads to excess–volume with respect to the frictionless market. In contrast, in the above models of OTC markets, an increase in the magnitude of the friction reduces trading volume.

One of the major implications of our theoretical analysis is that each trader will generally engage in several consecutive round–trips. The round–trips arising in our centralized limit order market are different, however, from those arising in dealer markets.[6] Dealers aim, after a sequence of round–trip trades, to hold zero net inventory position. And, in turning over their

[6]They also differ from those arising in fragmented OTC markets such as those analyzed by Afonso and Lagos (2011), Atkeson, Eisfeldt, and Weill (2012) and Babus and Kondor (2012). In these models, excess volume arises because all trades are bilateral, which give investors incentives to provide immediacy to each other, buying from those with lower valuation than them, and then selling to those with higher valuation. In our model, excess volume arises even though all trades occur in a centralized market.

position, they earn the realized bid–ask spread. This logic, underlying Grossman and Miller (1988), stands in contrast with the economic mechanism prevailing in our model where there are no designated dealers and round-trips are not motivated by the desire to move back to an ideal zero net position or to earn the bid-ask spread.

Round–trips can also arise when potentially informed investors buy, driving the price up, and then re-sell, before the price eventually reverts (see for instance Allen and Gale, 1992, Brunnermeier, 2005). Our model has different implications in at least two dimensions: First, the investors who engage in round–trips eventually hold the asset. Second, the price continues to go up after the round–trip trade, i.e., there is no overshooting.

Gromb and Vayanos (2002, 2010) and Brunnermeier and Pedersen (2009) also study liquidity shocks in markets with frictions. But the frictions they consider differ from ours. They consider traders' funding constraints, while we consider information processing constraints. In contrast with these papers, in our framework frictions don't necessarily amplify the initial price drop. And they can also increase trading volume.

The consequences of the informational frictions in our analysis vastly differ from those of asymmetric information on common values. When traders observe private information on the common value of the asset, this creates a "speculative" motive for trade. With rational traders, however, this does not increase trading volume, but reduces it. This is because private information induces adverse selection, making traders reluctant to trade, as in Akerlof (1970).[7]

The next section presents our model. Section 3 presents the equilibrum. The implications of our analysis are outlined in Section 4. Section 5 briefly concludes. The main proofs are in the appendix. A supplementary appendix collects the proofs omitted in the paper. It also discusses a model in which institutions choose their information collection effort as well as what happens with a finite number of traders.

[7]See Appendix B.3 for a formal argument. Of course the effect of adverse selection disappears if uninformed traders are noise traders. Noise traders do not optimize so, by assumption, never worry about adverse selection.

2 Model

2.1 Assets and agents

Time is continuous and runs forever. A probability space $(\Omega, \mathcal{F}, P)$ is fixed, as well as an information filtration satisfying the usual conditions (Protter, 1990).[8] There is an asset in positive supply $s > 0$ exchanged in a centralized continuous market. The economy is populated by a $[0, 1]$-continuum of infinitely-lived financial firms (banks, funds, insurers, etc...) discounting the future at the same rate $r > 0$.

Financial firms can either be in a high valuation state, θ_h, or in a low valuation state, θ_ℓ. The firm's utility flow from holding q units of the asset in state $\theta \in \{\theta_\ell, \theta_h\}$ is denoted by $v(\theta, q)$, and satisfies the following conditions. First, utilities are strictly increasing and strictly concave in q, and they satisfy

$$v_q(\theta_\ell, q) < v_q(\theta_h, q),$$

for all $q > 0$. That is, low-valuation firms have lower marginal utility than high-valuation firms and, correspondingly, demand less assets. Second, in order to apply differential arguments, we assume that, for both $\theta \in \{\theta_\ell, \theta_h\}$, $v(\theta, q)$ is three times continuously differentiable in $q > 0$ and satisfies the Inada conditions $v_q(\theta, 0) = +\infty$ and $v_q(\theta, \infty) = 0$. Finally, firms can produce (or consume) a non-storable numéraire good at constant marginal cost (utility) normalized to 1.

2.2 Liquidity shock

To model liquidity shocks we follow Weill (2004, 2007) and Duffie, Gârleanu, and Pedersen (2007). All financial firms are *ex-ante* identical: before the shock, each firm is in the high–valuation state, $\theta = \theta_h$, and holds s shares of the asset. At time zero, the liquidity shock hits a fraction $1 - \mu_{h,0}$ of financial firms, who make a switch to low–valuation, $\theta = \theta_\ell$. The switch from $\theta = \theta_h$ to $\theta = \theta_\ell$ induces a drop in utility flow, reflecting the increase in capital charges and additional regulatory costs of holding the asset induced by the liquidity shock. The shock, however, is transient. In practice firms can respond to liquidity shocks by hedging their positions, adjusting them, and raising capital. Once they have completed this process

[8] To simplify the exposition, for most stated equalities or inequalities between stochastic processes, we suppress the "almost surely" qualifier as well as the corresponding product measure over times and events.

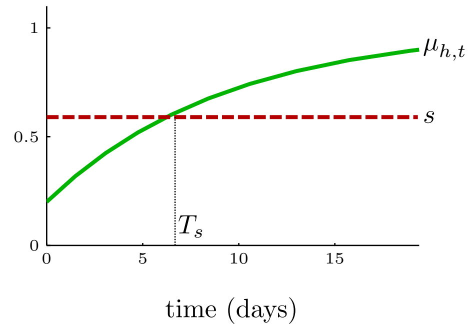

Figure 1: The measure of high-valuation institutions (plain green) and the asset supply (dashed red).

successfully, they recover from the shock and switch back to a high valuation for the asset. To model this, we assume that, for each firm, there is a random time at which it reverts to the high–valuation state, $\theta = \theta_h$, and then remains there forever. For simplicity, we assume that recovery times are exponentially distributed, with parameter γ, and independent across firms. Denote by $\mu_{h,t}$ the fraction of financial firms with high valuation at time t. By the law of large numbers, the flow of firms who become high valuation at time t is equal to[9] $\mu'_{h,t} = \gamma \left(1 - \mu_{h,t}\right)$, implying that:

$$1 - \mu_{h,t} = \left(1 - \mu_{h,0}\right) e^{-\gamma t}, \tag{1}$$

as illustrated in Figure 1. Because there is no aggregate uncertainty, in all what follow we will focus on equilibria in which aggregate outcomes (price, allocation, etc...) are deterministic functions of time.

2.3 Preference uncertainty

Each financial firm is represented in the market by one trader. As discussed above, the process by which the firm recovers from the liquidity shock is complex and costly. As a result, it takes time to collect all the data about this process, analyze the information and disseminate it to the traders. Correspondingly, we assume a trader does not observe the preference parameter θ of her firm in continuous time. We assume instead that, for each trader, there is a counting process, N_t, such that she updates her information about θ at each jump of N_t. For simplicity, we assume that traders' updating times are Poisson distributed with intensity ρ, independent

[9]For simplicity and brevity, we do not formally prove how the law of large numbers applies to our context. To establish the result precisely, one would have to follow Sun (2006), who relies on constructing an appropriate measure for the product of the agent space and the event space.

from each other and from everything else.[10]

2.4 Holding plans and intertemporal utilities

When a trader observes the current value of her θ at some time $t > 0$, she designs a new asset holding plan, $q_{t,u}$, for all subsequent times $u \geq t$ until her next updating time, at which point she designs a new holding plan, and so on. Each holding plan is implemented, in our centralized market, by the placement and updating of sequences of limit orders. At each point in time, the collection of the current limit orders of a trader determines her demand function.

Formally, letting $\mathcal{T} = \{(t,u) \in \mathbb{R}_+^2 \,:\, t \leq u\}$, the collection of *asset holding plans* is a stochastic process

$$
\begin{aligned}
q \,:\, & \mathcal{T} \times \Omega \to \mathbb{R}_+ \\
& (t,u,\omega) \mapsto q_{t,u}(\omega),
\end{aligned}
$$

which is adapted with respect to the filtration generated by θ_t and N_t. That is, a trader's asset holdings at time u can only depend on the information she received until time t, her last updating time: the history of her updating and valuation processes up to time t. Note that, given that there is no aggregate uncertainty and given our focus on equilibria with deterministic aggregate outcomes, we do not need to make the holding plan contingent on any aggregate information such as, e.g., the market price, since the later is a deterministic function of u. We impose, in addition, mild technical conditions ensuring that intertemporal values and costs of the holding plan are well defined: we assume that a trader must choose holding plans which are bounded, have bounded variation with respect to u for any t, and which generate absolutely integrable discounted utility flows. In all what follow, we will say that a holding plan is admissible if it satisfies these measurability and regularity conditions.

Now consider a trader's intertemporal utility. For any time $u \geq 0$, let τ_u denote the last updating time before u, with the convention that $\tau_u = 0$ if there has been no updating on θ since time zero. Note that τ_u has an atom of mass $e^{-\rho u}$ at $\tau_u = 0$, and a density $\rho e^{-\rho(u-t)}$ for $t \in (0,u]$.[11] At time u, the trader follows the plan she designed at time τ_u, so she holds a

[10] To simplify notations, we don't index the updating time processes of the different traders by subscripts specific to each trader. Rather we use the same generic notation, "N_t", for all traders.

[11] To see this, note that for any $t \in [0,u]$, the probability that $\tau_u \leq t$ is equal to the probability that there has been no updating time during $(t,u]$, i.e., to the probability of the event $N_u - N_t = 0$. Since the counting process for updating times follows a Poisson distribution, this probability is equal to $e^{-\rho(u-t)}$.

quantity $q_{\tau_u,u}$ of assets. Thus, the trader's *ex ante* intertemporal utility can be written:

$$V(q) = \mathbb{E}\left[\int_0^\infty e^{-ru} v(\theta_u, q_{\tau_u,u})\, du\right], \tag{2}$$

where the expectation is taken over θ_u and τ_u. Next, consider the intertemporal cost of buying and selling assets. During $[u, u + du]$, the trader follows the plan chosen at time τ_u, which prescribes that her holdings must change by $dq_{\tau_u,u}$. Denoting the price at time u by p_u, the cost of buying of selling asset during $[u, u + du]$ is, then, $p_u\, dq_{\tau_u,u}$. Therefore, the *ex-ante* intertemporal cost of buying and selling assets writes

$$C(q) = \mathbb{E}\left[\int_0^\infty e^{-ru} p_u dq_{\tau_u,u}\right], \tag{3}$$

which is well defined under natural regularity conditions about p_u.[12]

2.5 Market clearing

The market clearing condition requires that, at each date $u \geq 0$, aggregate asset holdings be equal to s, the asset supply. In our mass–one continuum setting, aggregate asset holdings are equal to the cross–sectional average asset holding. Moreover, by the law of large numbers, and given *ex–ante* identical traders, the cross-sectional average asset holding is equal to the expected asset holding of a representative trader. Hence, the market clearing condition at time u can be written:

$$\mathbb{E}\big[q_{\tau_u,u}\big] = s \tag{4}$$

for all $u \geq 0$, where the expectation is taken with respect to τ_u and to θ_{τ_u}, reflecting the aggregation of asset demands over a population of traders with heterogeneous updating times and uncertain preferences.

3 Equilibrium

An *equilibrium* is made up of an admissible holding plan q and of a price path p such that: i) given the price path p the holding plan q maximizes the intertemporal net utility $V(q) - C(q)$, where $V(q)$ and $C(q)$ are given by (2) and (3) and ii) the optimal holding plan is such that

[12]For example, it will be well defined in the equilibrium we study, where p_u is deterministic and continuous.

the market clearing condition (4) holds at all times. In this subsection we characterize the demands of traders for any given price path and then, substituting demands in the market–clearing condition, we show existence and uniqueness of equilibrium. We conclude the section by establishing that this equilibrium is socially optimal.

3.1 Asset demands

Focusing on equilibria in which the price path is deterministic, bounded, and continuously differentiable,[13] we define the holding cost of the asset at time u:

$$\xi_u = rp_u - \dot{p}_u, \tag{5}$$

which is equal to the cost of buying a share of the asset at time u and reselling it at $u + du$, i.e., the time value of money, rp_u, minus the capital gain, $\dot{p}_u$.

Lemma 1. *A trader's intertemporal net utility can be written:*

$$V(q) - C(q) = p_0 s + \mathbb{E}\left[\int_0^\infty e^{-ru}\Big\{\mathbb{E}\big[v(\theta_u, q_{\tau_u,u}) \,\big|\, \mathcal{F}_{\tau_u}\big] - \xi_u q_{\tau_u,u}\Big\}\, du\right]. \tag{6}$$

At time τ_u, her most recent updating time before time u, the trader received information θ_{τ_u} about her valuation, and she chose the holding plan $q_{\tau_u,u}$. Thus, she expects to derive utility $\mathbb{E}\left[v(\theta_u, q_{\tau_u,u})\big|\mathcal{F}_{\tau_u}\right]$ at time u, and to incur the opportunity cost $\xi_u q_{\tau_u,u}$.

Lemma 1 implies that an optimal holding plan can be found via optimization at each information set. Formally, a trader's optimal asset holding at time u solves:

$$q_{\tau_u,u} = \arg\max_q \mathbb{E}\left[v(\theta_u, q) \,\big|\, \mathcal{F}_{\tau_u}\right] - \xi_u q.$$

Let $\pi_{\tau_u,u}$ denote the probability that $\theta_u = \theta_h$ given the value of θ_{τ_u} observed at t. The trader's problem can be rewritten as

$$q_{\tau_u,u} = \arg\max_q \; \pi_{\tau_u,u} v(\theta_h, q) + (1 - \pi_{\tau_u,u})\, v(\theta_\ell, q) - \xi_u q,$$

[13] As argued above, deterministic price paths are natural given the absence of aggregate uncertainty. Further, in the environment that we consider, one can show that the equilibrium price must be continuous (see Biais, Hombert, and Weill, 2012b). The economic intuition is as follows. If the price jumps at time t, all traders who receive an updating opportunity shortly before t would want to "arbitrage" the jump: they would find it optimal to buy an infinite quantity of assets and re-sell these assets just after the jump. This would contradict market–clearing.

so the first-order necessary and sufficient condition is:

$$\pi_{\tau_u,u} v_q(\theta_h, q_{\tau_u,u}) + (1 - \pi_{\tau_u,u}) v_q(\theta_\ell, q_{\tau_u,u}) = \xi_u. \tag{7}$$

This equation means that each trader's expected marginal utility of holding the asset during $[u, u + du]$ is equal to the opportunity cost of holding the asset during that infinitesimal time interval. It implies the standard equilibrium condition that marginal utilities are equalized across traders. Analyzing the first order condition (7), we obtain the following lemma:

Lemma 2. *There exists a unique solution to (7), which is a function of ξ_u and $\pi_{\tau_u,u}$ only, and which we correspondingly denote by $D(\pi, \xi)$. The function $D(\pi, \xi)$ is strictly increasing in π and strictly decreasing in ξ, is twice continuously differentiable in (π, ξ), and it satisfies $\lim_{\xi \to 0} D(0, \xi) = \infty$ and $\lim_{\xi \to \infty} D(1, \xi) = 0$.*

Note that a trader's demand is increasing in the probability of being a high-valuation, π. This follows directly from the assumption that high-valuation traders have higher marginal utility than low-valuation.

3.2 Existence and uniqueness

Consider some time $u \geq 0$ and a trader whose most recent updating time is τ_u. If $\theta_{\tau_u} = \theta_h$, then the trader knows for sure that $\theta_u = \theta_h$ and so she demands $D(1, \xi_u)$ units of the asset at time u. If $\theta_{\tau_u} = \theta_\ell$, then

$$\pi_{\tau_u,u} \equiv \frac{\mu_{h,u} - \mu_{h,\tau_u}}{1 - \mu_{h,\tau_u}}, \tag{8}$$

and her demand is $D(\pi_{\tau_u,u}, \xi_u)$. Therefore, the market clearing condition (4) writes:

$$\mathbb{E}\left[\mu_{h,\tau_u} D(1, \xi_u) + \left(1 - \mu_{h,\tau_u}\right) D(\pi_{\tau_u,u}, \xi_u)\right] = s. \tag{9}$$

The first term in the expectation represents the aggregate demand of high-valuation traders, i.e., traders who discovered at their last updating time, τ_u, that their firm had a high valuation, $\theta_{\tau_u} = \theta_h$. Likewise, the second term represents the demand of low-valuation traders.

Note that aggregate demand, on the left-hand side of (9), inherits the properties of $D(\pi, \xi)$: it is continuous in (π, ξ), strictly decreasing in ξ, goes to infinity when $\xi \to 0$, and to zero when $\xi \to \infty$. Thus, this equation has a unique solution, ξ_u, which is easily shown to be a bounded and continuous function of time, u. The equilibrium price is obtained as the present discounted value of future holding costs:

$$p_t = \int_t^{\infty} e^{-r(u-t)} \xi_u \, du.$$

While the holding cost ξ_u measures the cost of buying the asset at u net of the benefit of reselling at $u + du$, the price p_t measures the cost of buying at t and holding until the end of time. Taking stock:

Proposition 1. *There exists a unique equilibrium. The holding cost at time u is the unique solution of (9), and is bounded and continuous. The asset holding of a time-τ_u high-valuation trader is $q_{h,u} \equiv D(1, \xi_u)$, and the asset holding of a time-τ_u low-valuation trader is $q_{\ell,\tau_u,u} \equiv D(\pi_{\tau_u,u}, \xi_u)$.*

3.3 Constrained efficiency

To study constrained efficiency we define a collection of holding plans to be *feasible* if it is admissible and if it satisfies the resource constraint, which is equivalent to the market-clearing condition (4). Furthermore, we say that a collection of holding plans, q, *Pareto dominates* some other collection of holding plans, q', if it is possible to generate a Pareto improvement by switching from q' to q while making time zero transfers among investors. Because utilities are quasi linear, q Pareto dominates q' if and only if $V(q) > V(q')$.

Proposition 2. *The holding plan arising in the equilibrium characterized in Proposition 1 is the unique maximizer of $V(q)$ among all feasible holding plans.*

The proposition reflects that, in our setup, there are no externalities. To assess the robustness of this welfare theorem, we consider (in supplementary appendix B.4) a simple variant of the model, with three stages: *ex-ante* banks choose how much effort to exert, to increase the precision of the information signal about their own type, *interim* banks receive their signal and trade in a centralized market, *ex-post* banks discover their types and payoffs realize. In this context, again, we find that the equilibrium is constrained Pareto efficient: both the choice of effort and the allocation coincide with the one that a social planner would choose.

4 Implications

4.1 Known preferences

To understand the implications of preference uncertainty, we first need to consider the benchmark in which traders continuously observe their valuation, θ_u. In that case, the last updating time is equal to the current time, $\tau_u = u$, and the market clearing condition (9) becomes:

$$\mu_{h,u} D(1, \xi_u^\star) + \left(1 - \mu_{h,u}\right) D(0, \xi_u^\star) = s. \tag{10}$$

Clearly, since $\mu_{h,u}$ is increasing and since $D(1,\xi) > D(0,\xi)$ for all ξ, it must be the case that both $\xi_u^\star$ and the price

$$p_t^\star = \int_t^\infty e^{-r(u-t)} \xi_u^\star \, du,$$

increase over time. Correspondingly, demand of high- and low-valuation traders, given by $D(1, \xi_u^\star)$ and $D(0, \xi_u^\star)$, must be decreasing over time.

The intuition is the following. High-valuation traders are more willing to hold the asset than low-valuation traders. In a sense, the high-valuation traders absorb the asset that low-valuation traders are not willing to hold, which can be interpreted as liquidity supply. As time goes by, the mass of traders with high valuation goes up. Thus, the amount of asset each individual trader of a given valuation needs to hold goes down. Correspondingly, traders' marginal valuation for the asset increases, and so does the price.

While this increase in price is perfectly predictable, it does not generate arbitrage opportunities. It just reflects the dynamics of the optimal allocation of the asset in a context where traders' willingness to hold the asset is finite. Such finite willingness to hold the asset can be interpreted as a form of "limit to arbitrage."While they know the price will for sure be higher in the future, investors are not willing to buy more of it now, because that would entail an opportunity cost of holding the asset (ξ_u) greater than their marginal valuation for the asset.

4.2 The impact of preference uncertainty on price

Let $\xi_u^\star$ be the opportunity cost of holding the asset when preferences are known. Starting from this benchmark, what would be the effect of preference uncertainty? Would it raise asset demand and asset prices?

A global condition for higher demand and higher prices. Given any holding cost ξ, preference uncertainty increases asset demand at time u, relative to the case of known preferences, if and only if:

$$\mathbb{E}\left[\mu_{h,\tau_u} D(1,\xi) + (1-\mu_{h,\tau_u}) D(\pi_{\tau_u,u},\xi)\right] > \mu_{h,u} D(1,\xi) + (1-\mu_{h,u}) D(0,\xi). \tag{11}$$

The left-hand side is the aggregate demand at time u under preference uncertainty, and the right-hand side is the aggregate demand under known preferences. Using the definition of $\pi_{\tau_u,u}$, this inequality can be rearranged into:

$$\mathbb{E}\left[\left(1-\mu_{h,\tau_u}\right) D(\pi_{\tau_u},\xi)\right] > \mathbb{E}\left[(1-\mu_{h,\tau_u})\left\{\pi_{\tau_u,u} D(1,\xi) + \left(1-\pi_{\tau_u,u}\right) D(0,\xi)\right\}\right]. \tag{12}$$

To interpret this inequality, note first that preference uncertainty has no impact on the demand of high-valuation time-τ_u traders, because they know for sure that they will keep a high valuation forever after τ_u. Thus, preference uncertainty only has an impact on the demand of the measure $1-\mu_{h,\tau_u}$ of time-τ_u low-valuation traders. The left-hand side of (12) is the time-u demand of these time-τ_u low-valuation traders, under preference uncertainty. The right-hand side is the demand of these same traders, but under known preferences.

Equation (12) reveals that preference uncertainty has two effects on asset demand, going in opposite directions. With known preferences, a fraction $\pi_{\tau_u,u}$ of time-τ_u low-valuation traders would have known for sure that they had a high valuation at time u: preference uncertainty *decreases* their demand, from $D(1,\xi)$ to $D(\pi_{\tau_u,u},\xi)$. But the complementary fraction, $1-\pi_{\tau_u,u}$, would have known for sure that they had a low valuation at time u: preference uncertainty *increases* their demand, from $D(0,\xi)$ to $D(\pi_{\tau_u,u},\xi)$.

Clearly one sees that inequality (12) holds for all u and τ_u if demand $D(\pi,\xi)$ is strictly concave in π.

Proposition 3. *If $D(\pi,\xi)$ is strictly concave in π for all ξ, then the holding cost and the price are strictly larger with preference uncertainty than with known preferences.*

Concavity implies that, under preference uncertainty, the increase in demand of low-valuation traders dominates the decrease for high-valuation traders.[14] To illustrate the proposition, con-

[14] This result is in line with those previously derived by Gârleanu (2009) and Lagos and Rocheteau (2009) in the context of OTC markets. In these papers, traders face a form of preference uncertainty, because they are uncertain about their stochastic utility flows in between two contact times with dealers. When OTC market frictions increase, inter-contact times are larger, and so is preference uncertainty. In line with Proposition 3, in

sider the iso-elastic utility:

$$v(\theta, q) = \theta \frac{q^{1-\sigma} - 1}{1 - \sigma}, \tag{13}$$

as in Lagos and Rocheteau (2009, Proposition 5 and 6). Then, after calculating demand, one sees that preference uncertainty increases prices if $\sigma > 1$, and decreases prices when $\sigma < 1$.

Note that the global concavity condition of Proposition 3 does not hold in other cases of interest. In particular, preferences in the spirit of Duffie, Gârleanu, and Pedersen (2005) tend to generate demands that are locally convex for low π and locally concave for high π.[15] To address such cases, we now develop a less demanding local condition which applies when $u \simeq 0$, i.e., just after the liquidity shock, or for all u when $\rho \to \infty$, i.e., when traders face small preference uncertainty.

A local condition for higher prices when $u \simeq 0$. Heuristically, just after the liquidity shock, $\tau_u = 0$ for most of the population, so we only need to study (12) for $\tau_u = 0$. Moreover, low-valuation traders only had a short time to switch to a high type, and so the probability $\pi_{0,u}$ is close to zero. Because $\pi_{0,u} \simeq \frac{\mu'_{h,0}}{1-\mu_{h,0}} \times u \simeq 0$ and demand is differentiable, we can make a first-order Taylor expansion of (12) in $\pi_{0,u}$, and the condition for higher holding cost becomes:

Proposition 4. *When $u > 0$ is small, time-u demand is larger under preference uncertainty, and so is the equilibrium holding cost ξ_u if:*

$$D_\pi(0, \xi_0^\star) > D(1, \xi_0^\star) - D(0, \xi_0^\star). \tag{14}$$

where $\xi_0^\star$ is the time-zero holding cost with known preferences.

The left-hand side of (14) represents the per-capita increase in demand for low types, and the right-hand side the per-capita decrease in demand for high types. The proposition shows that the holding cost is higher under preference uncertainty if: (i) $D_\pi(0, \xi)$ is large, i.e., low

Gârleanu (2009) the friction does not affect asset prices, because the flow of utility is linear in asset holdings and agent type. Also in line with Proposition 3, in Lagos and Rocheteau (2009) making the friction more severe increases the price if the utility function is sufficiently concave.

[15] In Duffie, Gârleanu, and Pedersen (2005), preferences are of the form $\theta \min\{1, q\}$. Then, when $\xi \in (\theta_\ell, \theta_h)$, demand is a step function of π: it is zero when $\pi\theta_h + (1-\pi)\theta_\ell < \xi$, equal to $[0, 1]$ when $\pi\theta_h + (1-\pi)\theta_\ell = \xi$, and equal to one when $\pi\theta_h + (1-\pi)\theta_\ell > \xi$. With a smooth approximation of $\min\{1, q\}$, demand becomes a smooth approximation of this step function. It tends to be convex for small π, rises rapidly when $\pi\theta_h + (1-\pi)\theta_\ell \simeq \xi$, and becomes concave for large π. In Section 4.4 below we offer a detailed study of equilibrium in a specification where asset demands are neither globally concave nor convex in π.

types react very strongly to changes in their expected valuation and (ii) $D(1,\xi) - D(0,\xi)$ is bounded, i.e., high types do not demand too much relative to low types.

Note that the demand shift of low- and high valuations are driven by different considerations. For the large population of low types, what matters is a change at the intensive margin: by how much each low type changes its demand in response to a small change in expected valuation, which is approximately equal to $\pi_{0,u} D_\pi(0, \xi_0^\star)$. For the small population of high types, on the other hand, what matters is a change at the extensive margin: by how much demand changes because a small fraction of high types turn into low types, which is approximately equal to $\pi_{0,u}\left[D(1,\xi) - D(0,\xi)\right]$. This distinction will be especially important in our discussion of trading volume, because we can have very large intensive margin changes even if holdings are bounded.

A local condition for higher prices when $\rho \to \infty$. When $\rho \to \infty$, most traders had their last updating time shortly before u, approximately at $\tau_u = u - \frac{1}{\rho}$. This is intuitively similar to the situation analyzed in the previous paragraph: when $u \simeq 0$, all traders had their last updating time shortly before u as well, at $\tau_u = 0$. Going through the same analysis, which can be thought of heuristically as using $\pi_{u-\frac{1}{\rho},u}$ instead of $\pi_{0,u}$, we arrive at:

Proposition 5. *For all $u > 0$ and for ρ large enough, time-u demand is larger under preference uncertainty, and so is the equilibrium holding cost if:*

$$D_\pi(0, \xi_u^\star) > D(1, \xi_u^\star) - D(0, \xi_u^\star), \tag{15}$$

where $\xi_u^\star$ is the time-u holding cost with known preferences. Moreover, the holding cost admits the first-order approximation:

$$\xi_u(\rho) = \xi_u^\star - \frac{\mu'_{h,u}}{\rho} \frac{D_\pi(0, \xi_u^\star) - [D(1, \xi_u^\star) - D(0, \xi_u^\star)]}{\mu_{h,u} D_\xi(1, \xi_u^\star) + \left(1 - \mu_{h,u}\right) D_\xi(0, \xi_u^\star)} + o_\alpha\left(\frac{1}{\rho}\right),$$

where $o_\alpha(1/\rho)$ is a function such that $\sup_{u \geq \alpha} |\rho o_\alpha(1/\rho)| \to 0$ as $\rho \to \infty$, for any $\alpha > 0$.

Condition (15) follows heuristically by replacing $\xi_0^\star$ by $\xi_u^\star$ in condition (14). To interpret the approximation formula, we first observe that, to a first-order approximation, the extra demand at time u induced by preference uncertainty can be written:

$$\frac{\mu'_{h,u}}{\rho} \left\{ D_\pi(0, \xi_u^\star) - \left[D(1, \xi_u^\star) - D(0, \xi_u^\star)\right] \right\} \tag{16}$$

by taking a first-order Taylor approximation of the difference between the left–hand–side and the right–hand–side of (12), for $\tau_u \simeq u - \frac{1}{\rho}$. Thus, the holding cost has to move by an amount equal to this extra demand, in equation (16), divided by the negative of the slope of the demand curve, $\mu_{h,u} D_\xi(1, \xi^\star_u) + (1 - \mu_{h,u}) D_\xi(0, \xi^\star_u)$.

4.3 The impact of preference uncertainty on volume

We now turn from the effects of preference uncertainty on prices to its consequences for trading volume. Does preference uncertainty increase or reduce volume? To study this question, we focus on the case where preference uncertainty is least likely to affect volume, as the friction is very small, i.e., it takes only a very short amount of time for traders to find out exactly what the objective of the financial firm is. To do so, we study the limit of the trading volume as ρ goes to infinity. One could expect that, as the friction vanishes, trading volume goes to its frictionless counterpart. We will show, however, that it is not the case, and we will offer an economic interpretation for that wedge

If the opportunity holding cost were equal to ξ^*_u (which is the price prevailing in the benchmark case in which preferences are known) then traders' holding plans would be:

$$q^\star_{\ell,t,u} \equiv D(\pi_{t,u}, \xi^\star_u) \quad \text{and} \quad q^\star_{h,u} \equiv D(1, \xi^\star_u).$$

Thus, when low–valuation traders know their preferences with certainty, they hold $q^\star_{\ell,u,u} = D(0, \xi^*_u)$ at all times. With these notations, (twice) the instantaneous trading volume is:[16]

$$2V^\star = \mu_{h,u} \left| \frac{dq^\star_{h,u}}{du} \right| + \left(1 - \mu_{h,u}\right) \left| \frac{dq^\star_{\ell,u,u}}{du} \right| + \mu'_{h,u} \left| q^\star_{h,u} - q^\star_{\ell,u,u} \right|. \tag{17}$$

The first and second terms account for the flow sale of high- and low-valuation traders. The last term accounts for the lumpy purchases of the flow $\mu'_{h,u}$ of traders who switch from low to high valuation.

With preference uncertainty, (twice) the instantaneous trading volume is

$$\begin{aligned} 2V = \mathbb{E}\Bigg[& \mu_{h,\tau_u} \left| \frac{dq_{h,u}}{du} \right| + \left(1 - \mu_{h,\tau_u}\right) \left| \frac{\partial q_{\ell,\tau_u,u}}{\partial u} \right| \\ & + \rho(1 - \mu_{h,\tau_u}) \Big\{ \pi_{\tau_u,u} \left| q_{h,u} - q_{\ell,\tau_u,u} \right| + (1 - \pi_{\tau_u,u}) \left| q_{\ell,u,u} - q_{\ell,\tau_u,u} \right| \Big\} \Bigg], \end{aligned} \tag{18}$$

[16] Equation (17) gives *twice* the volume because it double counts each trade as a sale and a purchase.

where the expectation, taken over τ_u, reflects the aggregation of trades over a population of agents with heterogeneous updating times.

The terms of the first line of equation (18) represent the flow trades of the traders who do not update their holding plans. The *partial* derivative with respect to u, $\frac{\partial q_{\ell,\tau_u,u}}{\partial u}$, reflects the fact that these traders follow a plan chosen at some earlier time, τ_u. In contrast, with known preferences, traders update their holding plans continuously so the corresponding term in equation (17) involves the *total* derivative, $\frac{dq^\star_{\ell,u,u}}{du} = \frac{\partial q^\star_{\ell,u,u}}{\partial t} + \frac{\partial q^\star_{\ell,u,u}}{\partial u}$.

The terms on the second line of equation (18) represent the lumpy trades of the traders who update their holding plans. There is a flow $\rho\left(1-\mu_{h,\tau_u}\right)$ of time-τ_u low-valuation traders who update their holding plans. Out of this flow, a fraction $\pi_{\tau_u,u}$ find out that they have a high valuation, and make a lumpy adjustment to their holdings equal to $|q_{h,u} - q_{\ell,\tau_u,u}|$. The complementary fraction $1-\pi_{\tau_u,u}$ find out they have a low valuation and make the adjustment $|q_{\ell,u,u} - q_{\ell,\tau_u,u}|$.

To compare the volume with known versus uncertain preferences, we consider the $\rho\to\infty$ limit. As shown formally in the appendix, in order to evaluate this limit, we can replace $q_{h,u}$ and $q_{\ell,t,u}$ by their limits $q^\star_{h,u}$ and $q^\star_{\ell,t,u}$, and use the approximation $\tau_u \simeq u - \frac{1}{\rho}$. After a little algebra, we obtain that:

$$\lim_{\rho\to\infty} 2V = 2V^\infty = \mu_{h,u}\left|\frac{dq^\star_{h,u}}{du}\right| + \left(1-\mu_{h,u}\right)\left|\frac{\partial q^\star_{\ell,u,u}}{\partial u}\right| + \mu'_{h,u}\left|q^\star_{h,u} - q^\star_{\ell,u,u}\right| + (1-\mu_{h,u})\left|\frac{\partial q^\star_{\ell,u,u}}{\partial t}\right|.$$

Subtracting the volume with know preferences, $2V^\star$, we obtain:

$$2V^\infty - 2V^\star = \left(1-\mu_{h,u}\right)\left\{\left|\frac{\partial q^\star_{\ell,u,u}}{\partial t}\right| + \left|\frac{\partial q^\star_{\ell,u,u}}{\partial u}\right| - \left|\frac{dq^\star_{\ell,u,u}}{du}\right|\right\}, \tag{19}$$

which is positive by the triangle inequality, and strictly so if the partial derivatives are of opposite sign.

To interpret these derivatives, note that under preference uncertainty, at time u some low-valuation traders receive the bad news that they still have a low valuation while others receive no news.

The traders who receive bad news at time u switch from the time-τ_u to the time-u holding plan. When $\tau_u \simeq u - \frac{1}{\rho}$, they change their holdings by an amount proportional to the partial derivatives with respect to t, $\frac{\partial q^\star_{\ell,u,u}}{\partial t} < 0$. This derivative is negative, implying that these traders they sell upon receiving bad news.

The traders who receive no news trade the amount prescribed by their time-τ_u holding plan. When $\tau_u \simeq u - \frac{1}{\rho}$, this changes their holding by an amount proportional to the partial derivative with respect to u, $\frac{\partial q^\star_{\ell,u,u}}{\partial u}$. Thus, when $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$, traders with no news build up their inventories.

Under preference uncertainty, the changes in holdings due to $\frac{\partial q^\star_{\ell,u,u}}{\partial t}$ and $\frac{\partial q^\star_{\ell,u,u}}{\partial u}$ contribute separately to the trading volume, explaining the first two terms of (19). With known preferences, in contrast, all low-valuation traders are continuously aware that they have a low-valuation, and so their holdings change by an amount equal to the total derivative. This explains the last term of (19).

In the time series, the above analysis implies that, when $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$, low-valuation traders engage in round-trip trades. Consider a trader who finds out at two consecutive updating times, u and $u + \varepsilon$, that she has a low valuation. In between the two updating times, when ε is small, she builds up inventories since $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$. At the updating time $u + \varepsilon$, she receives bad news, switches holding plan, and thus sells, since $\frac{\partial q^\star_{\ell,u,u}}{\partial t} < 0$. Thus round trips arise only if $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$. Correspondingly, the next proposition states that preference uncertainty generates excess volume when $\rho \to \infty$ if and only if $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$.

Proposition 6. *As $\rho \to \infty$, the excess volume is equal to:*

$$V^\infty - V^\star = \left(1 - \mu_{h,u}\right) \max\left\{\frac{\partial q^\star_{\ell,u,u}}{\partial u}, 0\right\},$$

where $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$ if and only if:

$$D_\pi(0, \xi^\star_u) > \left[D(1, \xi^\star_u) - D(0, \xi^\star_u)\right] \frac{(1 - \mu_{h,u}) D_\xi(0, \xi^\star_u)}{\mu_{h,u} D_\xi(1, \xi^\star_u) + (1 - \mu_{h,u}) D_\xi(0, \xi^\star_u)}. \tag{20}$$

To understand why a trader may increase her holding shortly after an updating time, i.e., why $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$, note that:

$$(1 - \mu_{h,u}) \frac{\partial q^\star_{\ell,u,u}}{\partial u} = \left(1 - \mu_{h,u}\right) D_\pi(0, \xi^\star_u) \frac{\partial \pi_{t,u}}{\partial u} + \left(1 - \mu_{h,u}\right) D_\xi\left(0, \xi^\star_u\right) \frac{d\xi^\star_u}{du}. \tag{21}$$

The equation reveals two effects going in opposite directions. On the one hand, the first term is positive, reflecting the fact that a low-valuation trader expects that she may switch to a high-valuation, which increases her demand over time. On the other hand, the second term is negative because the price increases over time and, correspondingly, decreases demand. If low-valuation traders' demands are very sensitive to changes in expected valuation, then $D_\pi(0, \xi_u^\star)$ is large and preference uncertainty creates extra volume. A sufficient condition for this to be the case is that demand is weakly concave with respect to π. In the case of iso-elastic preferences, (13), this arises if $\sigma \geq 1$.

One sees that the condition (15) for preference uncertainty to increase demand is closely related to condition (20) for it to create excess volume. This is natural given that both phenomena can be traced back to low-valuation traders' willingness to increase their holdings as their probability of being high valuation increases. But the former condition turns out to be stronger than the later: excess volume is necessary but not sufficient for higher demand.

4.4 An analytical example

To illustrate our results and derive further implications, we now consider the following analytical example. We let $s \in (0,1)$, $\mu_{h,0} < s$, $\sigma > 0$ and we assume that preferences are given by:

$$v(\theta, q) = m(q) - \delta \mathbb{I}_{\{\theta=\theta_\ell\}} \frac{m(q)^{1+\sigma}}{1+\sigma}, \tag{22}$$

for some $\delta \in (0,1]$ and where

$$m(q) \equiv \begin{cases} 1 - \frac{\ln\left(1+e^{1/\varepsilon[1-q^{1-\varepsilon}/(1-\varepsilon)]}\right)}{\ln\left(1+e^{1/\varepsilon}\right)} & \text{if } \varepsilon > 0 \\ \min\{q, 1\} & \text{if } \varepsilon = 0. \end{cases}$$

When $\varepsilon > 0$ and is small, the function $m(q)$ is approximately equal to $\min\{q, 1\}$,[17] and it satisfies the smoothness and Inada conditions of Section 2.1, so all the results derived so far can be applied.

When $\varepsilon = 0$ the function $m(q)$ is exactly equal to $\min\{q, 1\}$ and so it no longer satisfies these regularity conditions. Nevertheless, existence and uniqueness can be established up to small adjustments in the proof. Moreover, equilibrium objects are continuous in ε, in the following sense:

Proposition 7. *As $\varepsilon \to 0$ the holding cost, holding plans and the asymptotic excess volume converge pointwise to their $\varepsilon = 0$ counterparts.*

[17] We follow Eeckhout and Kircher (2010) who use a closely related function to approximate a smooth but almost frictionless matching process.

This continuity result allows us to concentrate, for the remainder of this section, on the $\varepsilon = 0$ equilibrium, which can be solved in closed form.

When $\varepsilon = 0$, the marginal valuation of a high-valuation trader is equal to one as long as q is lower than 1, and equal to 0 for larger values of q. Hence, her demand is a step function of the holding cost, ξ:

$$D(1,\xi) = \begin{cases} 1 & \text{if } \xi < 1 \\ \in [0,1] & \text{if } \xi = 1 \\ 0 & \text{if } \xi > 1. \end{cases}$$

Also for $\varepsilon = 0$, the demand of a trader who expects to be of high valuation with probability $\pi < 1$ is:

$$D(\pi,\xi) = \min\left\{ \left(\frac{1-\xi}{\delta(1-\pi)} \right)^{\frac{1}{\sigma}}, 1 \right\}. \tag{23}$$

When, $\varepsilon = 0$ and $\sigma \to 0$, our specification nests the case analyzed in Duffie, Gârleanu, and Pedersen (2005) and the demand of low-valuation traders is a step function of both the holding cost and the probability π of having a high-valuation.[18] When $\sigma > 0$ our specification generates smoother demands for low-valuation traders, as with the iso-elastic specification (13) of Lagos and Rocheteau (2009). Note however that when preferences are as in (13), demand is either globally concave or convex in π, so that preference uncertainty either always increases or decreases prices. In contrast, for the specification we consider, in line with Duffie, Gârleanu, and Pedersen (2005), demands are neither globally concave nor convex in π. Correspondingly, we will show that preference uncertainty can increase prices in certain market conditions and decrease prices in others.

The equilibrium holding cost is easily characterized. First, $\xi_u \leq 1$ for otherwise aggregate demand would be zero. Second, $\xi_u = 1$ if and only if $u \geq T_f$, where T_f solves $\mathbb{E}\left[\mu_{h,\tau_{T_f}}\right] = s$. In other words, $\xi_u = 1$ if and only if the measure of traders who know that they have a high valuation, $\mathbb{E}\left[\mu_{h,\tau_u}\right]$, is greater than the asset supply, s. In that case, high-valuation traders absorb all the supply while holding $q_{h,u} \leq 1$, and therefore have a marginal utility $v_q(\theta_h, q_{h,u}) = 1$. Low-valuation traders, on the other hand, hold no asset. In this context

[18] See Addendum III in Biais, Hombert, and Weill (2012b) for a proof that the equilibrium is indeed continuous at $\sigma = 0$: precisely, we show that, as $\sigma \to 0$, equilibrium objects converges pointwise, almost everywhere, to their $\sigma = 0$ counterparts.

$p = 1/r$.

When $u < T_f$, then $\xi_u < 1$. All high-valuation traders hold one unit, and low-valuation traders hold positive amounts. The holding cost, ξ_u, is the unique solution of:

$$\mathbb{E}\left[\mu_{h,\tau_u} + \left(1 - \mu_{h,\tau_u}\right) D(\pi_{\tau_u,u}, \xi)\right] = s. \tag{24}$$

4.4.1 Known preferences

With known preferences, the above characterization can be applied by setting $\tau_u = u$ for all u. In this case T_f is the time T_s solving $\mu_{h,T_s} = s$. When $u \geq T_s$:

$$q^\star_{h,u} \in [0,1], \quad q^\star_{\ell,u,u} = 0, \quad \text{and } \xi^\star_u = 1,$$

that is, all assets are held by high-valuation traders, the holding cost is 1 and the price is $1/r$. When $u < T_s$:

$$q^\star_{h,u} = 1, \quad q^\star_{\ell,u,u} = \frac{s - \mu_{h,u}}{1 - \mu_{h,u}}, \quad \text{and } \xi^\star_u = 1 - \delta\left(q^\star_{\ell,u,u}\right)^\sigma < 1.$$

In this case, there are $\mu_{h,u}$ high-valuation traders who each hold one share, and $1 - \mu_{h,u}$ low-valuation traders who hold the residual supply $s - \mu_{h,u}$. The holding cost, $\xi^\star_u$, is equal to the marginal utility of a low-valuation trader and is less than one. Notice that the per-capita holding of low-valuation traders, $q^\star_{\ell,u,u}$, decreases over time. This reflects that, as time goes by, more and more firms recover from the shock, switch to $\theta = \theta_h$ and increase their holdings. As a result, the remaining low–valuation traders are left with less shares to hold.

4.4.2 Holding plans

With preference uncertainty, we already know that $q_{h,u} = 1$ for all $u < T_f$, $q_{h,u} \in [0,1]$ and $q_{\ell,\tau_u,u} = 0$ for all $u \geq T_f$. The only thing left to derive are the holdings of low-valuation traders when $u < T_f$.

Proposition 8. *Suppose preferences are given by (22) and that $\varepsilon = 0$. When $u < T_f$, low valuation traders hold $q_{\ell,\tau_u,u} = \min\left\{\left(1 - \mu_{h,\tau_u}\right)^{1/\sigma} Q_u, 1\right\}$, where Q_u is a continuous function such that $Q_0 = \frac{s-\mu_{h,0}}{(1-\mu_{h,0})^{1+1/\sigma}}$ and $Q_{T_f} = 0$. Moreover, Q_u is a hump-shaped function of u if condition (20) holds evaluated at $u = 0$, which, for the preferences given in (22) is equivalent* *to:*

$$q_{\ell,0,0} = q^{\star}_{\ell,0,0} = \frac{s-\mu_{h,0}}{1-\mu_{h,0}} > \frac{\sigma}{1+\sigma}, \tag{25}$$

Otherwise Q_u is strictly decreasing in u.

At time 0, all traders know their valuation for sure, so the allocation must be the same as with known preferences. In particular, $q_{\ell,0,0} = q^{\star}_{\ell,0,0} = \frac{s-\mu_{h,0}}{1-\mu_{h,0}}$, which determines Q_0. At time T_f high–valuation traders absorb the entire supply. Hence, $Q_{T_f} = 0$. For times $u \in (0, T_f)$, asset holdings are obtained by scaling down Q_u by $(1-\mu_{h,\tau_u})^{1/\sigma}$, where τ_u is the last updating time of the trader. This follows because low-valuation traders have iso-elastic holding costs, so their asset demands are homogenous. Correspondingly, if a low-valuation trader holds less than one unit, $q_{\tau_u,u} < 1$, then, substituting (8) in (23) we have:

$$q_{\ell,\tau_u,u} = D(\pi_{\tau_u,u}, \xi_u) = \left(1-\mu_{h,\tau_u}\right)^{1/\sigma} Q_u, \qquad \text{where } Q_u \equiv \left(\frac{1-\xi_u}{\delta(1-\mu_{h,u})}\right)^{1/\sigma}.$$

Otherwise $q_{\ell,\tau_u,u} = 1 < \left(1-\mu_{h,\tau_u}\right)^{1/\sigma} Q_u$. If Q_u is hump-shaped and achieves its maximum at some time T_ψ, then the holding plan of a trader with updating time $\tau_u \leq T_\psi$ will be hump-shaped, and the holding plan of a trader with updating time $\tau_u > T_\psi$ will be decreasing.

As shown in the previous section, (20) is the condition under which the holding plans of low valuation traders are increasing with time, near time zero. If this condition holds at time 0, it implies that holding plans defined at time 0 are hump–shaped. Because holdings plans at later times are obtained by scaling down time–0 holding plans, they also are hump–shaped.

Finally, note that, with the preference specification (22), the demand of high-valuation traders is inelastic when $\xi^{\star}_0 < 1$, i.e., $D_\xi(1, \xi^{\star}_0) = 0$. This implies that the condition under which preference uncertainty increases trading volume, (20), is equivalent to the simpler condition under which it increases demand, (15).

The closed form expression for equation (25) reveals some natural comparative static. When σ is small, (25) is more likely to hold. Indeed, utility is close to linear, $D_\pi(0, \xi^{\star}_0)$ is large, and traders' demands are very sensitive to changes in the probability of being high valuation. When s is large or when $\mu_{h,0}$ is small, (25) is also more likely to hold. In that case the liquidity shock is more severe. Hence, shortly after the initial aggregate shock, the inflow of traders who receive good news is not large enough to absorb the sales of the traders who currently receive bad news. In equilibrium, some of these sales are absorbed by traders who received the bad news at earlier

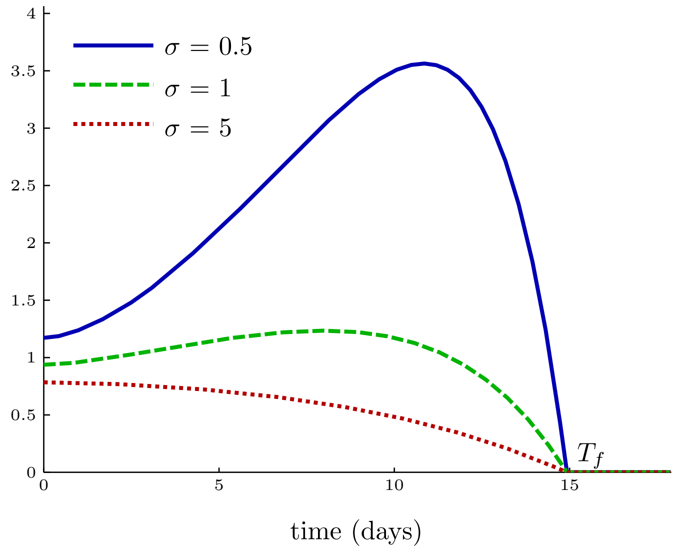

Figure 2: The function Q_u for various values of σ.

updating times, $\tau_u < u$. Indeed, these "early" low–valuation traders anticipate that, as time has gone by since their last updating time, $\tau_u < u$, their valuation is more and more likely to have reverted upwards $\pi_{\tau_u,u} > 0$. These traders find it optimal to buy if their utility is not too concave, i.e., if σ is not too high. Correspondingly, their holding plan can be increasing, and hence the function Q_u is hump-shaped, as depicted in Figure 2 for $\sigma = 0.5$ and 1. But for the larger value of $\sigma = 5$, the function Q_u is decreasing (the parameter values used for this figure are discussed in Section 4.4.5).

4.4.3 Trading volume

Proposition 8 provides a full characterization of the equilibrium holdings process, which can be compared to its counterpart without preference uncertainty. Holdings with known preferences are illustrated by the dash-dotted red curve in Figure 3: as long as a trader has not recovered from the shock, her holdings decline smoothly, and, as soon as she recovers, her holdings jump to 1. Holdings under preference uncertainty are quite different, as illustrated by the solid green curve in Figure 3. Consider a trader who is hit by a liquidity shock at time zero. After time zero, if (25) holds, the trader's holding plan, illustrated by the dotted blue curve, progressively buys back. If, at the next updating time, t_2, the trader learns that her valuation is still low, then she sells again. These round-trip trades continue until updating time t_6 when the trader finds out her valuation has recovered, at which point her holdings jump to 1. Thus, as we argued before, while the friction we consider implies less frequent observations of preferences,

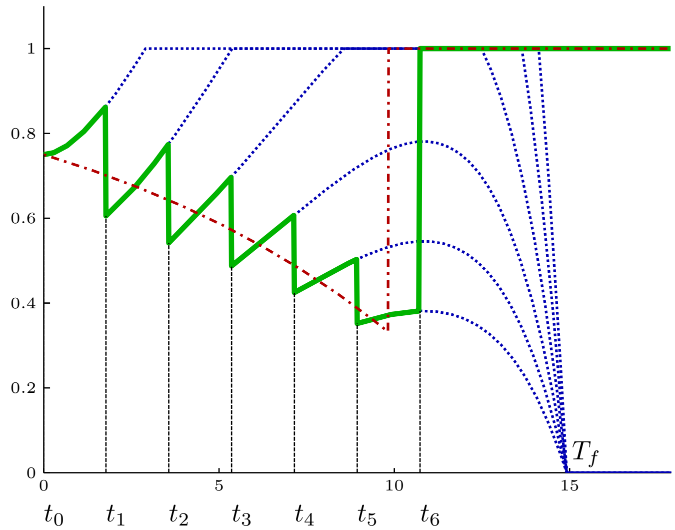

Figure 3: The realized holdings with continuous-updating are represented by the downward sloping curve (dash-dotted red). The realized holdings with infrequent updating are represented by the sawtooth curve (solid green). The consecutive holding plans with infrequent updating are represented by hump-shaped curves (dotted blue).

it does not induce less frequent *trading*, quite to the contrary. The hump–shaped asset holding plans shown in Figure 3 create round-trip trades and generate extra trading volume relative to the known preference case.[19]

As we know, this extra volume persists even in the $\rho \to \infty$ limit: although the round trip trades of a low-valuation trader become smaller and smaller, they occur more and more frequently. Note that, while the above analytical results on excess volume were obtained for the asymptotic case where ρ goes to infinity, Figure 3 illustrates that, even for finite ρ, preference uncertainty generates excess volume relative to the case where preferences are known.

Proposition 9. *Suppose preferences are given by (22) and that $\varepsilon = 0$. Then the asymptotic*

[19]As illustrated in Figure 3, both with known preferences and with preference uncertainty, the agent is continuously trading. Transactions costs, as analyzed by Constantinides (1986), Dumas and Luciano (1991) and Vayanos (1998), would reduce trading volume, as agents would wait until their positions get significantly unbalanced before engaging in trades. We conjecture that equilibrium dynamics would remain similar to that in Figure 3, except that holdings would be step functions. This would reduce the number of round–trips but not altogether eliminate them.

excess volume is equal to:

$$V^{\infty} - V^{\star} = \gamma(1-\mu_{h,u}) \max\left\{ D_{\pi}(0,\xi_u^{\star}) - \left[D(1,\xi_u^{\star}) - D(0,\xi_u^{\star})\right], 0 \right\}$$
$$= \gamma \max\left\{ \frac{s-\mu_{h,u}}{\sigma} - (1-s), 0\right\}.$$

The first equality involves, once again, the same terms as in condition (15): there is excess volume if demand is sufficiently sensitive to changes in the probability of being high valuation.

The second equality yields comparative statics of excess volume with respect to exogenous parameters. When σ decreases, demand becomes more sensitive to changes in the probability of being high valuation, and volume increases. When γ increases, low-valuation traders expect to change valuation faster, increase their demand by more, which increases excess volume.

4.4.4 Price

In the context of this analytical example we can go beyond the analysis of holding costs offered in the general case, and discuss detailed properties of the equilibrium price:

Proposition 10. *Suppose preferences are given by (22) and that $\varepsilon = 0$. Then, the price is continuously differentiable, strictly increasing for $u \in [0, T_f)$, and constant equal to $1/r$ for $u \geq T_f$. Moreover:*

- *For $u \in [T_s, T_f)$, the price is strictly lower than with known preferences.*
- *For $u \in [0, T_s]$, if (25) <u>does not hold</u>, then the price is strictly lower than with known preferences. But if s is close to 1 and σ is close to 0, then at time 0 the price is strictly higher than with known preferences.*

This proposition complements our earlier asymptotic results in various ways. First it characterizes the impact of preference uncertainty on price as opposed to holding cost; second, it offers results about the price path when ρ is finite; and third, it links price impact to fundamental parameters, such as s and σ.

The first bullet point follows because, from time T_s to time T_f, $\xi_u < \xi_u^{\star} = 1$. But it is not necessarily true for all $u \in [0, T_s)$. When (25) does not hold, then low-valuation traders do not create extra demand in between their updating times: to the contrary, they continue to sell their assets, and in equilibrium the price is smaller than its counterpart with known preferences. When (25) holds, then low-valuation traders increase their holdings in between updating times

Table 1: Parameter values

Parameter		Value
Discount rate	r	0.05
Updating intensity	ρ	250
Asset supply	s	0.8
Initial mass	$\mu_{h,0}$	0.2
Recovery intensity	γ	25
Utility cost	δ	1
Curvature of utility flow	σ	$\{0.5, 1, 5\}$

and the price at time zero can be larger than its counterpart with known preferences. This effect is stronger when low–valuation traders are marginal for a longer period, that is, when the shock is more severe (s close to one) and when their utility flow is not too concave (σ close to zero).

4.4.5 Empirical Implications

To illustrate numerically some key implications of the model, we select in Table 1 parameter values to generate effects comparable to empirical observations about liquidity shocks in large equity markets. Hendershott and Seasholes (2007) find liquidity price pressure effects of the order of 10 to 20 basis points, with duration ranging from 5 to 20 days. During the liquidity event described in Khandani and Lo (2008), the price pressure subsided in about 4 trading days. Adopting the convention that there are 250 trading days per year, setting γ to 25 means that an investor takes on average 10 days to switch back to high valuation. Setting the asset supply to $s = 0.8$ and the initial mass of high-valuation traders to $\mu_{h,0} = 0.2$ then implies that with continuous updating the time it takes the market to recover from the liquidity shock (as proxied by T_f) is approximately 15 days. For these parameter values, setting the discount rate to $r = 0.05$ and the holding cost parameter to $\delta = 1$ implies that the initial price pressure generated by the liquidity shock is between 10 and 20 basis points.[20] Finally, in line with the survey evidence cited in the introduction, we set the updating intensity to $\rho = 250$, i.e., we assume that each trader receives updated information about the utility flow she generates once

[20]Duffie, Gârleanu and Pedersen (2007) provide a numerical analysis of liquidity shocks in over–the–counter markets. They choose parameters to match stylized facts from illiquid corporate bond markets. Because we focus on more liquid electronic exchanges, we chose very different parameter values. For example in their analysis the price takes one year to recover while in ours it takes less than two weeks. While the price impact of the shock in our numerical example is relatively low, it would be larger for lower values of γ and r. For example if r were 10% and the recovery time 20 days, then the initial price impact of the shock would go up from 13 to 60 basis points. Note however that, with non-negative utility flow, the initial price impact of the shock is bounded above by $1 - e^{-rT}$.

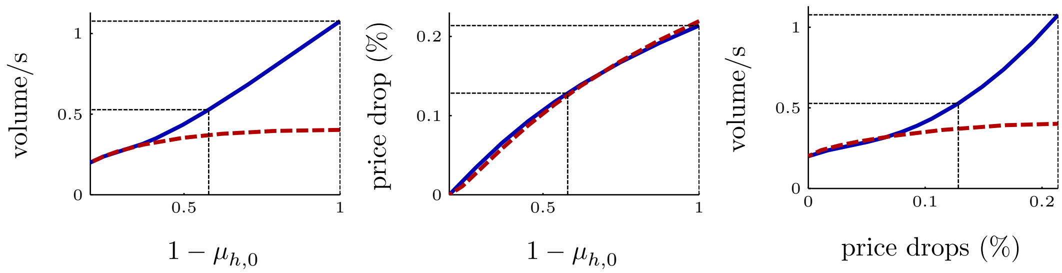

Figure 4: The relationship between size of the shock, $1-\mu_{h,0}$, and volume per unit of asset supply (left panel), between the size of the shock and price drop in percent (middle panel), and between price drop and volume (right panel), for known preferences (dashed red curves) vs. uncertain preferences (plain blue curves).

every day, on average.

Excess volume and liquidity shock. One of the main insights of our analysis is that preference uncertainty generates round trip trades, which in turn lead to excess trading volume after a liquidity shock. One natural measure of the size of the liquidity shock is the fraction of traders initially hit, $1-\mu_{h,0}$. The larger this fraction, the smaller $\mu_{h,0}$ and hence $\mu_{h,u}$ at any time u, and, by (26) in Proposition 9, the larger the excess volume. The left panel of Figure 4 illustrates this point by plotting total volume against initial price drop, for shocks of various sizes $(1-\mu_{h,0})$. As can be seen in the figure, the larger is $1-\mu_{h,0}$, the larger the total trading volume.

The middle panel of Figure 4 shows that the initial price drop, at time 0, is also increasing in the size of the liquidity shock. The right panel of Figure 4 illustrates these two points together by plotting total volume against initial price drop, a relationship that is perhaps easier to measure empirically. It shows that preference uncertainty generates a large elasticity of volume to price drop. Consider for instance an increase of $1-\mu_{h0}$ from 0.6 to one. The left panel of the figure indicates that the volume increases from 0.53 to 1.08, by about 105%. The price impact, on the other hand, increases from 0.13 to 0.21 basis points, by about 65%. Taken together, the elasticity of volume to price impact under preference uncertainty is around 1.6. As also shown in the figure, the elasticity with known preferences is an order of magnitude smaller, about 0.11.

Trading patterns in a cross-section of traders. While the above discussion bears on the empirical implications of our model for a cross–section of liquidity shocks, our analysis also delivers implications for the cross–section of traders within one liquidity shock.

Ex–ante, all traders are identical, but ex–post they differ, because they had different sample paths of valuations and information updates. Traders whose valuation recover early and who observe this rapidly, buy the asset early in the liquidity cycle, and then hold it. In contrast, traders whose valuations remain low throughout the major part of the liquidity cycle, and who have many information updates, engage in many round trips. These round trip trades are costly. They involve sequences of block sales at early stages in the liquidity cycle, when the price is still low, and an eventual block purchase towards the end of the cycle, when the price is high. Correspondingly, as shown in Appendix B.2, our model predicts that traders who engaged in many round–trips tend to earn lower trading profits than those with less frequent trades.[21]

The left panel of Figure 5 illustrates, in our numerical example, the model-generated relationship between trading profits and the number of round–trips (measured by the number of times two consecutive trades by the same trader were of opposite signs, e.g., a purchase followed by a sale.) Each blue dot represents the trading profits and number of round–trips of one trader in a cross-section of 500 traders, under preference uncertainty. The cross-section is representative in the sense that traders' characteristics (initial type, recovery time, and updating times) are drawn independently according to their "true," model-implied, probability distribution. The figure reveals that, with preference uncertainty, there is strong negative relationship between trading profits and the number of round–trips. In contrast, with known preferences (as illustrated by the red x-marks in the figure), there is no cross–sectional variation in the number of round–trips, and therefore no such relation. Hence, for the cross–section of traders, our model generates qualitatively different predictions for the known preferences and uncertain preferences cases. The middle panel of Figure 5 illustrates the negative relation between the number of round–trips and average inventories. Note that, once again, the pattern arising under preference uncertainty (blue dots) is significantly different from that arising with known preferences (red x-marks).

Since traders whose valuations recover early buy early and keep large holdings throughout the cycle, their behavior can be interpreted as liquidity supply. Put together, the negative relations i) between number of round–trips and trading profits, and ii) number of round–trips and average inventory, imply a positive relation between average inventory and trading profits. It is illustrated in the right panel of Figure 5. The figure illustrates that, in our model, liquidity

[21] Trading costs, here, are understood as reflecting only the proceeds from sales minus the cost of purchases, in the same spirit as in (3). That is, they don't factor in the utility flow $v(\theta, q)$ earned by the financial firm.

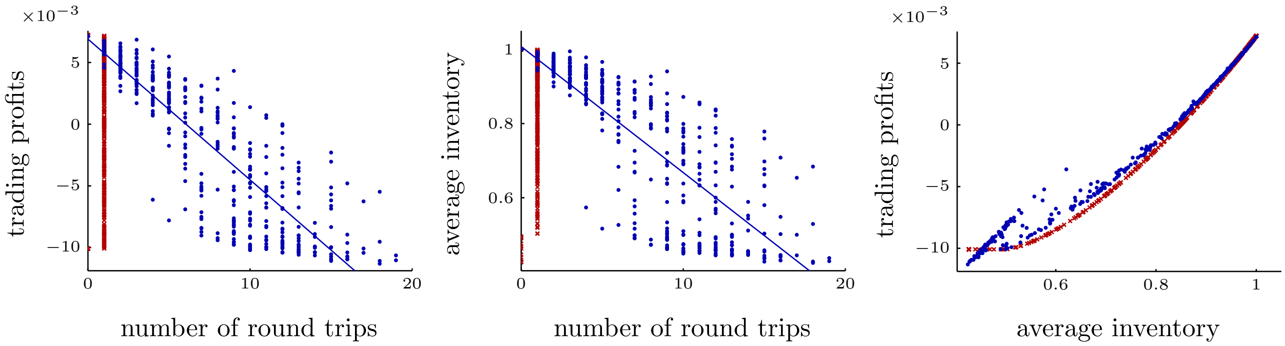

Figure 5: The relationship between individual trading volume and trading profits (left panel), between individual average inventory and trading profits (middle panel), and between individual average inventory and trading profits (right panel), for known preferences (red x-marks) vs. uncertain preferences (blue dots).

supply is profitable in equilibrium.

5 Conclusion

Information collection, processing and dissemination in financial institutions is challenging, as emphasized in practitioners' surveys and consultants' reports Ernst & Young (2012), Institute for International Finance (2011), Mehta et al. (2012). Completing these tasks is necessary for financial institutions to assess the bite of the regulatory and financial constraints they face, and their corresponding constrained optimal positions. As long as traders are not perfectly informed about the optimal position for their institution, they face preference uncertainty. We analyze optimal trading and equilibrium pricing in this context.

We focus on liquidity shocks, during which preference uncertainty is likely to matter most. Preference uncertainty generates allocative inefficiency, but need not reduce prices. As traders progressively learn about the preferences of their institution they conduct round–trip trades. This generates excess trading volume relative to the frictionless case. In a cross–section of liquidity shocks, the initial price drop is positively correlated with total trading volume. Across traders, the number of round–trips of a trader is negatively correlated with her trading profits.

While information collection, processing and dissemination frictions within financial institutions are very important in practice, to the best of our knowledge, Vayanos (2003) offers the only previous theoretical analysis of this issue. This seminal paper studies the optimal way to organize the firm to aggregate information. It therefore characterizes the endogenous structure of the information factored into the decisions of the financial institution, but it does not study the consequences of this informational friction for market equilibrium prices. Thus, the present paper complements Vayanos (2003), since we take as given the informational friction, but study its consequences for market pricing and trading. It would be interesting, in further research, to combine the two approaches: endogenize the organization of the firm and the aggregation of information, as in Vayanos (2003), *and* study the consequences of the resulting informational structure for market equilibrium, as in the present paper.

Another important, but challenging, avenue of further research would be to take into account interconnections and externalities among institutions. In the present model, individual valuations (θ_h or θ_ℓ) are exogenous. In practice, however, these valuations could be affected by others' actions. To study this, one would need a microfoundation for the endogenous determination of the valuations θ_h and θ_ℓ. For example, in an agency theoretic context, valuations could be affected by the pledgeable income of an institution (see Tirole, 2006, and Biais, Heider, and Hoerova, 2013). Thus, price changes, reducing the value of the asset held by an institution (or increasing its liabilities), would reduce its pledgeable income. In turn, this would reduce its ability to invest in the asset, which could push its valuation down to θ_ℓ. The analysis of the dynamics equilibrium prices and trades in this context is left for further research.

References

Gara Afonso and Ricardo Lagos. Trade dynamics in the federal funds market. Working paper, Federal Reverve Bank of New York and NYU, 2011. 6

George A. Akerlof. The market for lemons : Quality uncertainty and the market mechanism. *The Quaterly Journal of Economics*, 84(3):359–369, 1970. 7

Franklin Allen and Douglas Gale. Stock-price manipulation. *Review of Financial Studies*, 5: 504–529, 1992. 7

Fernando Alvarez, Luigi Guiso, and Francesco Lippi. Durable consumption and asset management with transaction and observation costs. *American Economic Review*, Forthcoming, 2010. 6

Fernando Alvarez, Francesco Lippi, and Luigi Paciello. Optimal price setting with observation and menu costs. *Quarterly Journal of Economics*, 126:1909–1960, 2011. 6

Tom M. Apostol. *Mathematical Analysis*. Addison-Wesley, 1974. 61

Andrew Atkeson, Andrea Eisfeldt, and Pierre-Olivier Weill. The market for otc credit derivatives. Working paper UCLA, 2012. 6

Ana Babus and Péter Kondor. Trading and information diffusion in over-the-counter markets. Working paper, Central European University and Federal Reserve Bank of Chicago, 2012. 6

Basel Committee on Banking Supervision. Findings on the interaction of market and credit risk. *Bank for International Settlement Working Paper*, (16), 2009. 2

Antje Berndt, Rohan Douglas, Darrell Duffie, Mark Ferguson, and David Schranz. Measuring default risk premia from default swap rates and edfs. Working paper, Stanford University, 2005. 3

Bruno Biais, Johan Hombert, and Pierre-Olivier Weill. Pricing and liquidity with sticky trading plans. Working paper, TSE, HEC, UCLA, 2012a. 4

Bruno Biais, Johan Hombert, and Pierre-Olivier Weill. Supplementary material for "pricing and liquidity with sticky plans". Working Paper, TSE, HEC, UCLA, 2012b. 12, 23

Bruno Biais, Florian Heider, and Marie Hoerova. An incentive theory of counterparty risk, margins, and clearing. Working paper, ECB and TSE, 2013. 33

Markus K. Brunnermeier. Information leakage and market efficiency. *Review of Financial Studies*, 18:417–457, 2005. 7

Markus K. Brunnermeier and Lasse H. Pedersen. Funding liquidity and market liquidity. *Review of Financial Studies*, 22:2201–2238, 2009. 7

Michael Carter and Bruce Van Brunt. *The Lebesgue-Stieltjes Integral: a Practical Introduction*. Springer-Verlag, New York, 2000. 38

George M. Constantinides. Capital market equilibrium with transaction costs. *Journal of Political Economy*, 94:842–862, 1986. 27

Joshua D. Coval and Erik Stafford. Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, 86:479–512, 2007. 3

Darrell Duffie, Nicolae Gârleanu, and Lasse H. Pedersen. Over-the-counter markets. *Econometrica*, 73:1815–1847, 2005. 6, 17, 23

Darrell Duffie, Nicolae Gârleanu, and Lasse H. Pedersen. Valuation in over-the-counter markets. *Review of Financial Studies*, 20:1865–1900, 2007. 3, 8

Bernard Dumas and Elisa Luciano. An exact solution to a portfolio choice problem under transactions costs. *Journal of Finance*, 46:577–595, 1991. 27

Jan Eeckhout and Philipp Kircher. Sorting and decentralized price competition. *Econometrica*, 78:539–574, 2010. 22

Ernst & Young. Progress in financial risk management: a survey of major financial institutions. Global Banking & Capital Market Center, 2012. 2, 32

Xavier Gabaix and David Laibson. The 6d bias and the equity premium puzzle. *NBER Macroeconomics Annual*, 16:257–312, 2002. 6

Nicolae Gârleanu. Portfolio choice and pricing in illiquid markets. *Journal of Economic Theory*, 144:532–564, 2009. 3, 6, 16, 17

Robin Greenwood. Short and long-term demand curves for stocks: Theory and evidence on the dynamics of arbitrage. *Journal of Financial Economics*, 75:607–650, 2005. 3

Denis Gromb and Dimitri Vayanos. Equilibrium and welfare in markets with financially constraint arbitrageurs. *Journal of Financial Economics*, 66:361–407, 2002. 7

Denis Gromb and Dimitri Vayanos. Limits of arbitrage: The state of the theory. *Annual Review of Financial Economics*, 2:251–275, 2010. 7

Sanford J. Grossman and Merton H. Miller. Liquidity and market structure. *Journal of Finance*, 43:617–637, 1988. 7

Sanford J. Grossman and Joseph E. Stiglitz. On the impossibility of informationally efficient markets. *American Economic Review*, 70:393–408, 1980. 76

Terry Hendershott and Mark S. Seasholes. Marketmaker inventories and stock prices, working paper. *American Economic Review (P&P)*, 97:210–214, 2007. 29

Institute for International Finance. Risk it and operations: Strengthening capabilities. Report, 2011. 2, 32

Amir E. Khandani and Andrew W. Lo. What happened to the quants in August 2007? *Forthcoming, Journal of Financial Markets*, 2008. 29

Ricardo Lagos and Guillaume Rocheteau. Liquidity in asset markets with search frictions. *Econometrica*, 77:403–426, 2009. 3, 6, 16, 17, 23

Ricardo Lagos, Guillaume Rocheteau, and Pierre-Olivier Weill. Crises and liquidity in otc markets. *Journal of Economic Theory*, 146:2169–2205, 2011. 6

Anthony W. Lynch. Decision frequency and synchronization across agents: Implications for aggregate consumption and equity return. *Journal of Finance*, 51:1479–1497, 1996. 6

Gregory Mankiw and Ricardo Reis. Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *Quarterly Journal of Economics*, 4(1295-1328), 2002. 6

Amit Mehta, Max Neukirchen, Sonja Pfetsch, and Thomas Poppensieker. Managing market risk:today and tomorrow. McKinsey Working Papers on Risk, Number 32, 2012. 2, 32

Emiliano Pagnotta and Thomas Philippon. Competiting on speed. Working paper, NYU Stern, 2011. 6

Philip Protter. *Stochastic Integration and Differential Equations*. Springer-Verlag, New York, 1990. 8

Ricardo Reis. Inattentive producers. *Review of Economic Studies*, 3:793–821, 2006a. 6

Ricardo Reis. Inattentive consumers. *Journal of Monetary Economics*, 8:1761–1800, 2006b. 6

Marzena Rostek and Marek Weretka. Dynamic thin markets. Technical report, University of Wisconsin, Madison, 2011. 84

Nancy L. Stokey and Robert E. Lucas. *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, 1989. 45

Jean Tirole. *The Theory of Corporate Finance*. Princeton University Press, 2006. 33

Dimitri Vayanos. Transaction costs and asset prices: A dynamic equilibrium model. *Review of Financial Studies*, 11:1–58, 1998. 27

Dimitri Vayanos. Strategic trading and welfare in a dynamic market. *Review of Economic Studies*, 66:219–254, 1999. 84

Dimitri Vayanos. The decentralization of information processing in the presence of interactions. *The Review of Economic Studies*, 70(3):667–695, 2003. 2, 32, 33

Pierre-Olivier Weill. Essays on liquidity in financial markets, ph.d. dissertation. Stanford University, 2004. 3, 8

Pierre-Olivier Weill. Leaning against the wind. *Review of Economic Studies*, 74:1329–1354, 2007. 3, 6, 8

A Proofs

A.1 Proof of Lemma 1

Let us begin by deriving a convenient expression for the intertemporal cost of buying and selling assets. For this we let $\tau_0 \equiv 0 < \tau_1 < \tau_2 < \ldots$ denote the sequence of updating times. For accounting purposes, we can always assume that, at her n-th updating time, the investor sells all of her assets, q_{τ_{n-1},τ_n}, and purchases a new initial holding q_{τ_n,τ_n}. Thus, the expected inter-temporal cost of following the successive holding plans can be written:

$$C(q) = \mathbb{E}\left[-p_0 s + \sum_{n=0}^{\infty}\left\{e^{-r\tau_n} p_{\tau_n} q_{\tau_n,\tau_n} + \int_{\tau_n}^{\tau_{n+1}} p_u \, dq_{\tau_n,u} e^{-ru} - e^{-r\tau_{n+1}} p_{\tau_{n+1}} q_{\tau_n,\tau_{n+1}}\right\}\right].$$

Given that p_u is continuous and piecewise continuously differentiable, and that $u \mapsto q_{\tau_n,u}$ has bounded variations, we can integrate by part (see Theorem 6.2.2 in Carter and Van Brunt, 2000), keeping in mind that $d/du(e^{-ru}p_u) = -e^{-ru}\xi_u$. This leads to:

$$C(q) = \mathbb{E}\left[-p_0 s + \sum_{n=0}^{\infty} \int_{\tau_n}^{\tau_{n+1}} e^{-ru} \xi_u q_{\tau_n,u} \, du\right] = -p_0 s + \mathbb{E}\left[\int_0^{\infty} e^{-ru} \xi_u q_{\tau_u,u} \, du\right].$$

In the above, the first equality follows by adding and subtracting $q_{0,u} = s$, and by noting that $q_{0,u}$ is constant; the second equality follows by using our "τ_u" notation for the last updating time before u. With the above result in mind, we find that we can rewrite the intertemporal payoff net of cost as:

$$\begin{aligned} V(q) - C(q) &= p_0 s + \mathbb{E}\left[\int_0^{\infty} e^{-ru}\Big(v(\theta_u, q_{\tau_u,u}) - \xi_u q_{\tau_u,u}\Big) \, du\right] \\ &= p_0 s + \mathbb{E}\left[\int_0^{\infty} e^{-ru} \mathbb{E}\big[v(\theta_u, q_{\tau_u,u}) - \xi_u q_{\tau_u,u} \,|\, \mathcal{F}_{\tau_u}\big] \, du\right], \end{aligned}$$

after switching the order of summation and applying the law of iterated expectations. □

A.2 Proof of Lemma 2

Because of the Inada conditions, the left–hand–side of (7) goes to infinity when $q_{\tau_u,u}$ goes to 0, and to 0 when $q_{\tau_u,u}$ goes to infinity. Because of the strict concavity of $q \mapsto v(\theta, q)$, the left–hand–side of (7) is strictly decreasing for all $q_{\tau_u,u} \in (0, \infty)$. Hence there exists a unique solution to (7). Since this solution only depends on ξ_u and $\pi_{\tau_u,u}$, we denote it by: $D(\pi, \xi)$. Since $v_q(\theta_h, q) > v_q(\theta_\ell, q)$, when $\pi_{\tau_u,u}$ is raised the left–hand–side of (7) is shifted upwards, while the left–hand–side is shifted upward when ξ_u is raised. Hence, $D(\pi, \xi)$ is strictly increasing in π and decreasing in ξ. Finally, since $v(\theta, q)$ is three times continuously differentiable, $D(\pi, \xi)$ is twice continuously differentiable. When $\pi = 0$ and ξ goes to 0, (7) implies that $v_q(\theta_\ell, D(0, \xi))$ goes to 0, therefore $D(0, \xi)$ goes to infinity by the Inada conditions. Similarly, When $\pi = 1$ and ξ goes to infinity, (7) and the Inada conditions imply that

$D(1, \xi)$ goes to zero. $\square$

A.3 Proof of Proposition 1

Aggregate demand is

$$e^{-\rho u} \left[\mu_{h,0} D(1, \xi) + (1 - \mu_{h,0}) D(\pi_{0,u}, \xi) \right] + \int_0^u \rho e^{-\rho(u-t)} \left[\mu_{h,t} D(1, \xi) + (1 - \mu_{h,t}) D(\pi_{t,u}, \xi) \right] \, dt.$$

Clearly, this is a strictly decreasing and continuous function of ξ, which is greater than s if $\xi = v_q(\theta_\ell, s)$, and smaller than s if $\xi = v_q(\theta_h, s)$. Thus, we can can apply the intermediate value theorem to establish that a unique equilibrium holding cost exists. Because aggregate demand is continuous in (ξ, u) and because ξ_u is bounded, ξ_u must be continuous in u. $\square$

A.4 Proof of Proposition 2

Let us begin with a preliminary remark. By definition, any feasible allocation q' satisfies the market–clearing condition $\mathbb{E}\left[q'_{\tau_u,u}\right] = s$. Taken together with the fact that $\xi_u = rp_u - \dot{p}_u$ is deterministic, this implies:

$$\begin{aligned} C(q') &= -p_0 s + \mathbb{E}\left[\int_0^\infty e^{-ru} \xi_u q'_{\tau_u,u} \, du\right] = -p_0 s + \int_0^\infty e^{-ru} \mathbb{E}\left[q'_{\tau_u,u}\right] \xi_u \, du \\ &= -p_0 s + \int_0^\infty e^{-ru} \xi_u s \, du = 0. \end{aligned} \tag{26}$$

With this in mind, consider the equilibrium asset holding plan of Proposition 1, q, and suppose it does not solve the planning problem. Then there is a feasible asset holding plan q' that achieves a strictly higher value, i.e.,

$$V(q') > V(q).$$

But, we just showed above that $C(q') = 0$. Subtracting the zero inter-temporal cost from both sides, we obtain that $V(q') - C(q') > V(q) - C(q)$, which contradicts individual optimality. Uniqueness of the planning solution follows because the planner's objective is strictly concave and the constraint set is convex. $\square$

A.5 Proof of Proposition 3

Note that:

$$\pi_{\tau_u,u} = \frac{\mu_{hu} - \mu_{h\tau_u}}{1 - \mu_{h\tau_u}} = \frac{1 - \mu_{hu}}{1 - \mu_{h\tau_u}} \times 0 + \frac{\mu_{hu} - \mu_{h\tau_u}}{1 - \mu_{h\tau_u}} \times 1.$$

Hence, if $D(\pi, \xi)$ is strictly concave in π, we have that:

$$\begin{aligned}
&\mu_{h\tau_u} D(1, \xi) + \left(1 - \mu_{h\tau_u}\right) D(\pi_{\tau_u,u}, \xi) \\
>&\mu_{h\tau_u} D(1, \xi) + \left(1 - \mu_{h\tau_u}\right) \left[\frac{1 - \mu_{hu}}{1 - \mu_{h\tau_u}} D(0, \xi) + \frac{\mu_{hu} - \mu_{h\tau_u}}{1 - \mu_{h\tau_u}} D(1, \xi)\right] \\
=&\mu_{hu} D(1, \xi) + (1 - \mu_{hu}) D(0, \xi).
\end{aligned}$$

Taking expectations with respect to τ_u on the left-hand side, we find that, for all ξ, demand is strictly higher with preference uncertainty. The result follows. □

A.6 Proof of Proposition 4

Consider a first-order Taylor expansion of aggregate asset demand under preference uncertainty when $u \simeq 0$, evaluated at $\xi_0^\star$:

$$\begin{aligned}
&e^{-\rho u} \left[\mu_{h,0} D(1, \xi_0^\star) + \left(1 - \mu_{h,0}\right) D(\pi_{0,u}, \xi_0^\star)\right] \qquad (27) \\
&\qquad\qquad + \int_0^u \rho e^{-\rho(u-t)} \left[\mu_{h,t} D(1, \xi_0^\star) + (1 - \mu_{h,t}) D(\pi_{0,t}, \xi_0^\star)\right] dt \\
=&[1 - \rho u] \left[\mu_{h,0} D(1, \xi_0^\star) + \left(1 - \mu_{h,0}\right) D(0, \xi_0^\star) + D_\pi(0, \xi_0^\star) \mu'_{h0} u\right] \\
&\qquad\qquad + \rho u \left[\mu_{h,0} D(1, \xi_0^\star) + (1 - \mu_{h,0} D(0, \xi_0^\star)\right] + o(u) \\
=&s + D_\pi(0, \xi_0^\star) \mu'_{h,0} u + o(u). \qquad (28)
\end{aligned}$$

The second equality follows from the fact that, by definition, $\mu_{h,0} D(1, \xi_0^\star) + \left(1 - \mu_{h,0}\right) D(0, \xi_0^\star) = s$. Now consider a first-order Taylor expansion of aggregate demand under known preferences:

$$\mu_{h,u} D(1, \xi_0^\star) + (1 - \mu_{h,u}) D(0, \xi_0^\star) = s + [D(1, \xi_0^\star) - D(0, \xi_0^\star] \, \mu'_{h,0} u + o(u). \qquad (29)$$

Clearly, for small u, the aggregate asset demand at $\xi_0^\star$ is larger under preference uncertainty if:

$$D_\pi(0, \xi_0^\star) > D(1, \xi_0^\star) - D(0, \xi_0^\star).$$

Since $\xi_u^\star$ is continuous at $u = 0$, this condition also ensures that, as long as u is small enough, aggregate demand at $\xi_u^\star$ is larger under preference uncertainty. The result follows. □

A.7 Asymptotics: Propositions 5 and 6

A.7.1 Preliminary results

Our maintained assumption on $v(\theta, q)$ implies that equilibrium holding costs will remain in the compact $[\underline{\xi}, \overline{\xi}]$, where $\underline{\xi}$ and $\overline{\xi}$ solve:

$$D(0, \underline{\xi}) = s \quad \text{and} \quad D(1, \overline{\xi}) = s.$$

The net demand at time u of traders who had their last information update at time $t < u$ is:

$$\mathcal{D}(t, u, \xi) \equiv \mu_{h,t} D(1, \xi) + \left(1 - \mu_{h,t}\right) D(\pi_{t,u}, \xi) - s.$$

It will be enough to study net demand over the domain $\Delta \times [\underline{\xi}, \overline{\xi}]$, where $\Delta \equiv \{(t, u) \in \mathbb{R}_+^2 \,:\, t \leq u\}$.

Lemma 3 (Properties of net demand)**.** *The net demand $\mathcal{D}(t, u, \xi)$ is twice continuously differentiable over $\Delta \times [\underline{\xi}, \overline{\xi}]$, with bounded first and second derivatives. Moreover $\mathcal{D}_\xi(t, u, \xi) < 0$ and is bounded away from zero.*

All results follow from direct calculations of first and second derivatives. The details can be found in Appendix B.1.1, page 56. Next, we introduce the following notation. For any $\alpha > 0$, we let $\Delta_\alpha \equiv \{(t, u) \in \mathbb{R}_+^2 \,:\, t \leq u \text{ and } u \geq \alpha\}$. Fix some function $g(\rho)$ such that $\lim_{\rho\to\infty} g(\rho) = 0$. We say that a function $h(t, u, \rho)$ is a $o_\alpha [g(\rho)]$ if it is bounded over $\Delta \times \mathbb{R}_+$, and if:

$$\lim_{\rho\to\infty} \frac{h(t, u, \rho)}{g(\rho)} = 0, \text{ uniformly over } (t, u) \in \Delta_\alpha.$$

To establish our asymptotic results we repeatedly apply the following Lemma.

Lemma 4. *Suppose $f(t, u, \rho)$ is twice continually differentiable with respect to t, and that $f(t, u, \rho)$, $f_t(t, u, \rho)$ and $f_{t,t}(t, u, \rho)$ are all bounded over $\Delta \times \mathbb{R}_+$. Then, for all $\alpha > 0$,*

$$\mathbb{E}\left[f(\tau_u, u, \rho)\right] = e^{-\rho u} f(0, u, \rho) + \int_0^u e^{-\rho(u-t)} \rho f(t, u, \rho)\, dt = f(u, u, \rho) - \frac{1}{\rho} f_t(u, u, \rho) + o_\alpha\left(\frac{1}{\rho}\right).$$

This follows directly after two integration by parts:

$$\int_0^u \rho e^{-\rho(u-t)} f(t, u, \rho)\, dt = f(u, u, \rho) - \frac{1}{\rho} f_t(u, u, \rho) + e^{-\rho u}\left[-f(0, u, \rho) + \frac{1}{\rho} f_t(0, u, \rho)\right] + \frac{1}{\rho^2} \int_0^u \rho e^{-\rho(u-t)} f_{t,t}(t, u, \rho)\, dt. \quad \square$$

We sometimes also use a related convergence result that apply under weaker conditions:

Lemma 5. *Suppose that $f(t, u)$ is bounded, and continuous $t = u$. Then* $\lim_{\rho\to\infty} \mathbb{E}\left[f(\tau_u, u)\right] = f(u, u)$.

Let M be an upper bound of $f(t,u)$ and, fixing $\varepsilon > 0$, let η be such that $|f(t,u) - f(u,u)| < \varepsilon$ for all $t \in [u-\eta, u]$. We have:

$$|\mathbb{E}[f(\tau_u,u)] - f(u,u)| \leq \mathbb{E}\left[|f(\tau_u,u) - f(u,u)|\,\mathbb{I}_{\{\tau_u \in [0,u-\eta)\}}\right] + \mathbb{E}\left[|f(\tau_u,u) - f(u,u)|\,\mathbb{I}_{\{\tau_u \in [u-\eta,u]\}}\right]$$
$$\leq 2Me^{-\rho\eta} + \varepsilon,$$

since the probability that $\tau_u \in [0, u-\eta]$ is equal to $e^{-\rho\eta}$. The result follows by letting $\rho \to \infty$. □

A.7.2 Proof of Proposition 5

Let $\xi_u(\rho)$ denote the market clearing holding cost at time u when the preference uncertainty parameter is ρ, i.e., the unique solution of:

$$e^{-\rho u}\mathcal{D}(0,u,\xi) + \int_0^u \rho e^{-\rho(u-t)}\mathcal{D}(t,u,\xi)\,dt = 0.$$

Note that $\xi_u(\rho) \in [\underline{\xi}, \overline{\xi}]$. Similarly, let $\xi_u^\star$ denote the frictionless market clearing holding cost, solving $\mathcal{D}(u,u,\xi) = 0$, which also belongs to $[\underline{\xi}, \overline{\xi}]$.

The first step is to show that $\xi_u(\rho)$ converges point wise towards $\xi_u^\star$. For this we note that $\xi_u(\rho)$ belongs to the compact $[\underline{\xi}, \overline{\xi}]$ so it admits at least one convergence subsequence, with a limit that we denote by $\hat{\xi}_u$. Now use Lemma 4 with $f(t,u,\rho) = \mathcal{D}(t,u,\xi_u(\rho))$, and recall from Lemma 3 that $\mathcal{D}(t,u,\xi)$ has bounded first and second derivatives. This implies that:

$$0 = e^{-\rho u}\mathcal{D}(0,u,\xi_u(\rho)) + \int_0^u \rho e^{-\rho(u-t)}\mathcal{D}(t,u,\xi_u(\rho))\,dt = \mathcal{D}(u,u,\xi_u(\rho)) + o_\alpha(1). \tag{30}$$

Letting ρ go to infinity we obtain, by continuity, that $\mathcal{D}(u,u,\hat{\xi}_u) = 0$, implying that $\hat{\xi}_u = \xi_u^\star$. Therefore, $\xi_u^\star$ is the unique accumulation point of $\xi_u(\rho)$, and so must be its limit.

The second step is to show that $\xi_u(\rho) = \xi_u^\star + o_\alpha(1)$. For this we use again (30), but with a first-order Taylor expansion of $\mathcal{D}(u,u,\xi_u(\rho))$. This gives:

$$0 = \mathcal{D}(u,u,\xi_u^\star) + \mathcal{D}_\xi(u,u,\hat{\xi}_u(\rho))\,[\xi_u(\rho) - \xi_u^\star] + o_\alpha(1),$$

where $\hat{\xi}_u$ lies in between $\xi_u^\star$ and $\xi_u(\rho)$. Given that $\mathcal{D}(u,u,\xi_u^\star) = 0$ by definition, and that $\mathcal{D}_\xi$ is bounded away from zero, the result follows.

Now, for the last step, we use again to Lemma 4, with $f(t,u,\rho) = \mathcal{D}(t,u,\xi_u(\rho))$:

$$0 = \mathcal{D}(u,u,\xi_u(\rho)) - \frac{1}{\rho}\mathcal{D}_t(u,u,\xi_u(\rho))$$
$$+ e^{-\rho u}\frac{1}{\rho}\mathcal{D}_t(0,u,\xi_u(\rho)) + \frac{1}{\rho^2}\int_0^u \rho e^{-\rho t}\mathcal{D}_{t,t}(t,u,\xi(\rho))\,dt.$$

Clearly, by our maintained assumptions on $\mathcal{D}(t,u,\xi)$, the terms on the second line add up to a $o_\alpha(1/\rho)$.

Also, recall that $\xi_u(\rho) = \xi_u^\star + o_\alpha(1)$ and note that $\mathcal{D}_t(t, u, \xi)$ has bounded derivatives and thus is uniformly continuous over $\Delta \times [\underline{\xi}, \overline{\xi}]$, implying that $\mathcal{D}_t(u, u, \xi_u(\rho)) = \mathcal{D}_t(u, u, \xi_u^\star) + o_\alpha(1)$. Taken together, we obtain:

$$\begin{aligned} 0 =& \mathcal{D}(u, u, \xi_u(\rho)) - \frac{1}{\rho}\mathcal{D}_t(u, u, \xi_u^\star) + o_\alpha\left(\frac{1}{\rho}\right) \\ =& \mathcal{D}(u, u, \xi_u^\star) + \mathcal{D}_\xi(u, u, \xi_u^\star)\left[\xi_u(\rho) - \xi_u^\star\right] \\ & \qquad + \frac{1}{2}\mathcal{D}_{\xi,\xi}(u, u, \hat{\xi}_u(\rho)\left[\xi_u(\rho) - \xi_u^\star\right]^2 - \frac{1}{\rho}\mathcal{D}_t(u, u, \xi_u^\star) + o_\alpha\left(\frac{1}{\rho}\right), \end{aligned}$$

for some $\hat{\xi}_u(\rho)$ in between $\xi_u^\star$ and $\xi_u(\rho)$. Keeping in mind that $\mathcal{D}(u, u, \xi_u^\star) = 0$, we obtain:

$$\left[\xi_u(\rho) - \xi_u^\star(\rho)\right]\left\{1 + \frac{\mathcal{D}_{\xi,\xi}(u, u, \hat{\xi}_u(\rho))}{2\,\mathcal{D}_\xi(u, u, \xi_u^\star)}\left[\xi_u(\rho) - \xi_u^\star\right]\right\} = \frac{1}{\rho}\frac{\mathcal{D}_t(u, u, \xi_u^\star)}{\mathcal{D}_\xi(u, u, \xi_u^\star)} + o_\alpha\left(\frac{1}{\rho}\right).$$

The term in the curly bracket on the left-hand size is $1 + o_\alpha(1)$, and the result follows after substituting the explicit expression of $\mathcal{D}_t(u, u, \xi_u^\star)$ and $\mathcal{D}_\xi(u, u, \xi_u^\star)$. □

A.7.3 Proof of Proposition 6

We first need to establish further asymptotic convergence results. First:

Lemma 6. *For large ρ, the time derivative of the holding cost admits the approximation:*

$$\frac{d\xi_u(\rho)}{du} = \frac{d\xi_u^\star}{du} + o_\alpha(1).$$

The proof is in Appendix B.1.2, page 58. Now, using Proposition 5 and 6, it follows that:

Lemma 7. *The holding plans and their derivatives converge α-uniformly to their frictionless counterparts:*

$$\begin{aligned} q_{\ell,t,u} &= q_{\ell,t,u}^\star + o_\alpha(1), \quad \frac{\partial q_{\ell,t,u}}{\partial t} = \frac{\partial q_{\ell,t,u}^\star}{\partial t} + o_\alpha(1), \quad \frac{\partial q_{\ell,t,u}}{\partial u} = \frac{\partial q_{\ell,t,u}^\star}{\partial u} + o_\alpha(1) \\ q_{h,u} &= q_{h,u}^\star + o_\alpha(1), \quad \frac{dq_{h,u}}{du} = \frac{dq_{h,u}^\star}{du} + o_\alpha(1). \end{aligned}$$

The proof is in Appendix B.1.3, page 59. With these results in mind, let us turn to the various components of the volume, in equation (18). The first term of equation (18) is:

$$\mathbb{E}\left[\mu_{h,\tau_u}\left|\frac{dq_{h,u}}{du}\right|\right] = \mathbb{E}\left[\mu_{h,\tau_u}\left|\frac{dq_{h,u}^\star}{du}\right| + o_\alpha(1)\right] = \mu_{h,u}\left|\frac{dq_{h,u}^\star}{du}\right| + o(1),$$

where the first equality follows from Lemma 7. The second equality follows from Lemma 4 and from the observation that, by dominated convergence, $\mathbb{E}\left[o_\alpha(1)\right] \to 0$ as $\rho \to \infty$.[22] The second term of

[22] Indeed for $g(t, u, \rho) = o_\alpha(1)$, $|\mathbb{E}\left[g(\tau_u, u, \rho)\right]| \leq \sup_{t\in[0,\alpha]} |g(t, u, \rho)|e^{-\rho(u-\alpha)} + \int_\alpha^u \rho|g(t, u, \rho)|e^{-\rho(u-t)}\,dt$. The

equation (18) is:

$$\mathbb{E}\left[\left(1-\mu_{h,\tau_u}\right)\left|\frac{\partial q_{\ell,t,u}}{\partial u}\right|\right] = \mathbb{E}\left[\left(1-\mu_{h,\tau_u}\right)\left|\frac{\partial q^\star_{\ell,\tau_u,u}}{\partial u}\right| + o_\alpha(1)\right] \to \left(1-\mu_{h,u}\right)\left|\frac{\partial q^\star_{\ell,t,u}}{\partial u}\right|,$$

by an application of Lemma 5. For the first term on the second line of equation (18) is:

$$\begin{aligned}&\rho\mathbb{E}\left[(1-\mu_{h,\tau_u})\pi_{\tau_u,u}|q_{h,u}-q_{\ell,\tau_u,u}|\right] = \rho\mathbb{E}\left[\left(\mu_{h,u}-\mu_{h,\tau_u}\right)\left(q_{h,u}-q_{\ell,\tau_u,u}\right)\right]\\ =&\mu'_{h,u}\left(q_{h,u}-q_{\ell,u,u}\right)+o_\alpha(1) \to \mu'_{h,u}\left(q^\star_{h,u}-q^\star_{\ell,u,u}\right),\end{aligned}$$

where the first equality follows from the definition of $\pi_{t,u}$ and from the observation that $q_{h,u} \geq q_{\ell,\tau_u,u}$, and where the second equality follows from an application of Lemma 4. The limit follows from Lemma 7. Finally, using the same logic, the last term on the second line of equation (18) is:

$$\begin{aligned}&\rho\mathbb{E}\left[\mu_{h,\tau_u}\left(1-\pi_{\tau_u,u}\right)|q_{\ell,u,u}-q_{\ell,\tau_u,u}|\right] = \rho\mathbb{E}\left[\left(1-\mu_{h,u}\right)\left(q_{\ell,\tau_u,u}-q_{\ell,u,u}\right)\right]\\ =&\left(1-\mu_{h,u}\right)\frac{\partial q_{\ell,u,u}}{\partial t}+o_\alpha(1) \to \left(1-\mu_{h,u}\right)\frac{\partial q^\star_{\ell,u,u}}{\partial t}.\end{aligned}$$

Collecting terms, we obtain the desired formula for $V^\infty - V^\star$. Next, consider the necessary and sufficient condition for $\frac{\partial q^\star_{\ell,u,u}}{\partial u} > 0$. We first note that:

$$\frac{\partial q^\star_{\ell,u,u}}{\partial u} = D_\pi(0,\xi^\star_u)\frac{\partial \pi_{u,u}}{\partial u} + D_\xi(0,\xi^\star_u)\frac{d\xi^\star_u}{du}.$$

The holding cost $\xi^\star_u$ solves:

$$\mu_{h,u}D(1,\xi)+(1-\mu_{h,u})D(0,\xi) = s \Rightarrow \frac{d\xi^\star_u}{du} = \frac{\mu'_{h,u}\left[D(1,\xi^\star_u)-D(0,\xi^\star_u)\right]}{\mu_{h,u}D_\xi(1,\xi^\star_u)+(1-\mu_{h,u})D_\xi(0,\xi^\star_u)},$$

using the Implicit Function Theorem. The result follows by noting that $\mu'_{h,u} = \gamma(1-\mu_{h,u})$ and $\frac{\partial \pi_{u,u}}{\partial u} = \gamma$. □

A.8 Proof of Proposition 7

To clarify the exposition, our notations in this section are explicit about the fact that utility functions and equilibrium objects depend on ε: e.g., we write $m(q,\varepsilon)$ instead of $m(q)$, $\xi_u(\varepsilon)$ instead of ξ_u, etc... We begin with preliminary properties of the function $m(q,\varepsilon)$.

Lemma 8. *The function $m(q,\varepsilon)$ is continuous in $(q,\varepsilon) \in [0,\infty)^2$, and satisfies $m(0,\varepsilon) = 0$ and $\lim_{q\to\infty} m(q,\varepsilon) = 1$. For $\varepsilon > 0$, it is strictly increasing, strictly concave, three time continuously differentiable over $(q,\varepsilon) \in [0,\infty)\times(0,\infty)$, and satisfies Inada conditions $\lim_{q\to 0} m_q(q,\varepsilon) = \infty$, and*

first term goes to zero because $g(t,u,\rho)$ is bounded, and the second one goes to zero because $g(t,u,\rho)$ converges uniformly to 0 over $[\alpha,u]$.

$\lim_{q\to\infty} m_q(q,\varepsilon) \to 0$. *Lastly, its first and second derivatives satisfy, for all $q>0$:*

$$\lim_{\varepsilon\to 0} m_q(q,\varepsilon) = \begin{cases} 1 & \text{if } q<1 \\ 1/(1+e) & \text{if } q=1 \\ 0 & \text{if } q>1 \end{cases} \qquad \lim_{\varepsilon\to 0} m_{qq}(q,\varepsilon) = \begin{cases} 0 & \text{if } q<1 \\ -\infty & \text{if } q=1 \\ 0 & \text{if } q>1 \end{cases}$$

The proof is in Appendix B.1.4, page 59. Having shown that $m(q,\varepsilon)$ is continuous even at points such that $\varepsilon = 0$, we can on to apply the Maximum Theorem (see, e.g., Stokey and Lucas, 1989, Theorem 3.6) to show that demands are continuous in all parameters. We let:

$$D(\pi,\xi,\varepsilon) = \arg\max_{\xi q \leq 2} m(q,\varepsilon) - \delta\,(1-\pi)\,\frac{m(q,\varepsilon)^{1+\sigma}}{1+\sigma} - \xi q.$$

Note that the constraint $q \leq 2/\xi$ is not binding: when $q > 2/\xi$, the objective is strictly negative since $m(q,\varepsilon) \leq 1$, and so it can be improved by choosing $q=0$. Thus, the maximization problem satisfies the condition of the Theorem of the Maximum: the objective is continuous in all variables, and the constraint set is a compact valued continuous correspondence of ξ. This implies that $D(\pi,\xi,\varepsilon)$ is non-empty, compact valued, and upper hemi continuous. Moreover, the demand is single-valued in all cases except when $\pi=1$, $\xi=1$, and $\varepsilon=0$, in which case it is equal to $[0,1]$.

Existence and uniqueness of equilibrium. When $\varepsilon > 0$, this follows directly from Proposition 1. When $\varepsilon = 0$, an equilibrium holding cost solves:

$$\mathbb{E}\left[\mu_{h,\tau_u} q_{h,\tau_u,u} + \left(1-\mu_{h,\tau_u}\right) D(\pi_{\tau_u,u},\xi,0)\right] = s.$$

for some $q_{h,\tau_u,u} \in D(1,\xi,0)$. One verifies easily that, for $\pi=1$, $D(1,\xi,0)=1$ for all $\xi \in (0,1)$, $D(1,\xi,0)=[0,1]$ for $\xi=1$, and $D(1,\xi,0)=0$ for all $\xi>1$. For $\pi<1$, $D(\pi,\xi,0)=0$ for all $\xi \geq 1$. One sees that $\xi \leq 1$ or otherwise the market cannot clear. Then, there are two cases:

- If $\mathbb{E}\left[\mu_{h,\tau_u}\right] \geq s$, and $\xi < 1$, then $D(1,\xi,0)=1$ and the market cannot clear. Thus, the market clearing holding cost is $\xi_u = 1$, high-valuation holdings are indeterminate but must add up to s, and low-valuation holdings are equal to zero.
- If $\mathbb{E}\left[\mu_{h,\tau_u}\right] < s$, then $\xi<1$. Otherwise, if $\xi=1$, $D(\pi,\xi,0)=0$ for all $\pi<1$ and the market cannot clear. Because $D(1,\xi,0)=1$ and $D(\pi,\xi,0)$ is strictly decreasing, in this case as well there a unique market clearing holding cost, $\xi_u(0)$. High-valuation traders hold $q_{h,u}(0)=1$, and low-valuation traders hold $q_{\ell,\tau_u,u} = D(\pi_{\tau_u,u},\xi,0)$.

□

Convergence of holding costs. For all $\varepsilon \geq 0$, let $\xi_u(\varepsilon)$ be the market clearing holding cost at time u. For $\varepsilon > 0$, let $\underline{\xi}(\varepsilon) \equiv m_q(s,\varepsilon)\,(1-\delta m(s,\varepsilon)^\sigma)$ and $\overline{\xi}(\varepsilon) = m_q(s,\varepsilon)$, so that $D(1,\overline{\xi}(\varepsilon),\varepsilon) =$

$D(1, \underline{\xi}(\varepsilon), \varepsilon) = s$. Clearly, for all $\varepsilon > 0$, $\xi_u(\varepsilon) \in [\underline{\xi}(\varepsilon), \overline{\xi}(\varepsilon)]$. Moreover, from Lemma 8, it follows that $\lim_{\varepsilon \to 0} \underline{\xi}(\varepsilon) = 1 - \delta \lim_{\varepsilon \to 0} \overline{\xi}(\varepsilon) = 1$, so that $\underline{\xi}(\varepsilon)$ and $\overline{\xi}(\varepsilon)$ remain in some compact $[\underline{\xi}, \overline{\xi}]$, s.t. $\underline{\xi} > 0$.

Thus, $\xi_u(\varepsilon)$ has at least one convergence subsequence as $\varepsilon \to 0$, with some limit $\hat{\xi}_u$. When $\pi < 1$, we have by upper hemi continuity that a subsequence of $D(\pi, \xi_u(\varepsilon), \varepsilon)$ converges to some $\hat{q}_{h,u} \in D(1, \hat{\xi}_u, 0)$, and we have by continuity that $D(\pi, \xi_u(\varepsilon), \varepsilon) \to D(\pi, \hat{\xi}_u, 0)$, and when $\pi = 1$. Moreover, by dominated convergence, $\mathbb{E}\left[(1 - \mu_{h,\tau_u}) D(\pi_{\tau_u,u}, \xi_u(\varepsilon), \varepsilon)\right] \to \mathbb{E}\left[(1 - \mu_{h,\tau_u}) D(\pi_{\tau_u,u}, \hat{\xi}_u, 0)\right]$. Taken together, we obtain that:

$$\mathbb{E}\left[\mu_{h,\tau_u} \hat{q}_{h,u} + \left(1 - \mu_{h,\tau_u}\right) D(\pi_{\tau_u,u}, \hat{\xi}_u, 0)\right] = s,$$

where $\hat{q}_{h,u} \in D(1, \hat{\xi}_u, 0)$. Therefore, $\hat{\xi}_u$ is an equilibrium holding cost of the $\varepsilon = 0$ economy, which we know must be equal to $\xi_u(0)$. Thus, the unique accumulation point of $\xi_u(\varepsilon)$ is $\xi_u(0)$, and so it must be its limit. □

Convergence of holding plans. Let $q_{\ell,t,u}(\varepsilon) \equiv D(\pi_{t,u}, \xi_u(\varepsilon), \varepsilon)$. Let $q_{h,u}(\varepsilon) \equiv D(1, \xi_u, \varepsilon)$ whenever the correspondence is single-valued. When it is multi-valued, which only arises when $\varepsilon = 0$ and $\xi_u(0) = 1$, we let $q_{h,u} \equiv s / \mathbb{E}\left[\mu_{h,\tau_u}\right]$. We have two cases to consider:

- If $\pi \in [0, 1)$, or if $\pi = 1$ and $\xi_u(0) < 1$, then $D(\pi, \xi, \varepsilon)$ is singled valued and thus continuous in a neighborhood of $(\xi_u(0), 0)$. Therefore $D(\pi, \xi_u(\varepsilon), \varepsilon) \to D(\pi, \xi_u(0), 0)$, i.e., $\varepsilon > 0$ holding plan converge to their $\varepsilon = 0$ counterparts.

- If $\pi = 1$ and $\xi_u(0) = 1$, which occurs when $\mathbb{E}\left[\mu_{h,\tau_u}\right] \geq s$, we have:

  $$\mathbb{E}\left[\mu_{h,\tau_u}\right] q_{h,u}(\varepsilon) + \mathbb{E}\left[(1 - \mu_{h,\tau_u}) q_{\ell,\tau_u,u}(\varepsilon)\right] = s.$$

  But we have just shown that $q_{\ell,\tau_u,u}(\varepsilon) \to 0$. Moreover, $q_{\ell,\tau_u,u}(\varepsilon) \leq q_{h,u}(\varepsilon) \leq s / \mathbb{E}\left[\mu_{h,\tau_u}\right]$ otherwise the market would not clear. Thus, by dominated convergence, $\mathbb{E}\left[(1 - \mu_{h,\tau_u}) q_{\ell,\tau_u,u}(\varepsilon)\right] \to 0$. From the market clearing condition, this implies that $q_{h,u}(\varepsilon) \to q_{h,u}(0) = s / \mathbb{E}\left[\mu_{h,\tau_u}\right]$.

□

Convergence of the volume. Fixing some $\varepsilon > 0$ positive and small enough, the specification of preferences satisfy our basic regularity conditions. So, we have that the excess volume when $\rho \to \infty$ converges to:

$$\max\left\{\left(1 - \mu_{h,u}\right) \frac{\partial q^{\star}_{\ell,u,u}(\varepsilon)}{\partial u}, 0\right\}$$

where, as before, we use the star "$\star$" to index holding plans and holding costs in the equilibrium with known preferences. Next, we will show that:

$$\lim_{\varepsilon \to 0} \max \left\{ \left(1 - \mu_{h,u}\right) \frac{\partial q^\star_{\ell,u,u}(\varepsilon)}{\partial u}, 0 \right\} = \gamma \max \left\{ \frac{s - \mu_{h,u}}{\sigma} - (1-s), 0 \right\},$$

which is equal to the excess volume in the $\varepsilon = 0$ equilibrium, as is shown formally later in Appendix A.9.3 For this we recall that:

$$q^\star_{\ell,t,u}(\varepsilon) = D(\pi_{t,u}, \xi^\star_u(\varepsilon)) \Rightarrow \frac{\partial q^\star_{\ell,u,u}}{\partial u} = D_\pi(0, \xi^\star_u(\varepsilon)) \frac{\partial \pi_{u,u}}{\partial u} + D_\xi\left(0, \xi^\star_u(\varepsilon)\right) \frac{d\xi^\star_u(\varepsilon)}{du}. \tag{31}$$

The holding cost $\xi^\star_u(\varepsilon)$ solves:

$$\mu_{h,u} D(1, \xi, \varepsilon) + (1 - \mu_{h,u}) D(0, \xi, \varepsilon) = s.$$

Thus, by the Implicit Function Theorem, the time derivative of $\xi^\star_u(\varepsilon)$ is:

$$\frac{d\xi^\star_u(\varepsilon)}{du} = - \frac{\mu'_{h,u} \left[D(1, \xi^\star_u(\varepsilon), \varepsilon) - D(0, \xi^\star_u(\varepsilon), \varepsilon) \right]}{\mu_{h,u} D_\xi(1, \xi^\star_u(\varepsilon), \varepsilon) + \left(1 - \mu_{h,u}\right) D_\xi(0, \xi^\star_u(\varepsilon), \varepsilon)}.$$

For any (π, ξ, ε), with $\varepsilon > 0$, the demand $D(\pi, \xi, \varepsilon)$ is the unique solution of

$$0 = H(q, \pi, \xi, \varepsilon) - \xi, \quad \text{where} \quad H(q, \pi, \xi, \varepsilon) \equiv m_q(q, \varepsilon) \left[1 - \delta(1-\pi) m(q, \varepsilon)^\sigma\right].$$

Using the Implicit Function Theorem, $D_\pi(\pi, \xi, \varepsilon) = -H_\pi / H_q$ and $D_\xi(\pi, \xi, \varepsilon) = -H_\xi / H_q$, all evaluated at $q = D(\pi, \xi, \varepsilon)$ and (π, ξ, ε). To evaluate the limit of (31) as $\varepsilon \to 0$, we start with some preliminary asymptotic results.

Step 1: preliminary results. By continuity of $m(q, \varepsilon)$ we have:

$$\lim_{\varepsilon \to 0} m(q^\star_{\ell,u,u}(\varepsilon), \varepsilon) = m(q^\star_{\ell,u,u}(0), 0) = q^\star_{\ell,u,u}(0) \tag{32}$$

$$\lim_{\varepsilon \to 0} m(q^\star_{h,u}(\varepsilon), \varepsilon) = m(q^\star_{h,u}(0), 0) = q^\star_{h,u}(0), \tag{33}$$

where, on both lines, the second equality follows by noting that, since $\xi^\star_u(0) > 0$ (otherwise the market would not clear), we have $q^\star_{\ell,u,u}(\varepsilon) \leq 1$ and $q^\star_{h,u}(0) \leq 1$. The first-order condition of low-valuation traders is:

$$\xi^\star_u(\varepsilon) = m_q(q^\star_{\ell,u,u}(\varepsilon), \varepsilon) \left(1 - \delta m(q^\star_{\ell,u,u}(\varepsilon), \varepsilon)^\sigma\right).$$

Taking $\varepsilon \to 0$ limits on both sides we obtain that:

$$\xi^\star_u(0) = \lim_{\varepsilon \to 0} m_q(q^\star_{\ell,u,u}(\varepsilon), \varepsilon) \left(1 - \delta \left(q^\star_{\ell,u,u}(0)\right)^\sigma\right).$$

When $\varepsilon = 0$, since $q^\star_{\ell,u,u}(0) \leq s < 1$, the first-order condition of a low-valuation trader is $\xi^\star_u(0) = \left(1 - \delta\left(q^\star_{\ell,u,u}(0)\right)^\sigma\right)$. Hence:

$$\lim_{\varepsilon \to 0} m_q(q^\star_{\ell,u,u}(\varepsilon), \varepsilon) = 1. \tag{34}$$

Using (54) and the fact that $q^\star_{\ell,u,u}(0) \leq s < 1$, this implies in turns that:

$$\lim_{\varepsilon \to 0} \left(q^\star_{\ell,u,u}(\varepsilon)\right)^{-\varepsilon} = 1. \tag{35}$$

Turning to the first-order condition of a high-valuation trader, we obtain:

$$\lim_{\varepsilon \to 0} m_q(q^\star_{h,u}(\varepsilon)) = \xi^\star_u(0). \tag{36}$$

If $\mu_{h,u} < s$, then $\xi^\star_u(0) < 1$ and it follows from the analytical expression of $m_q(q, \varepsilon)$, in equation (54), that:

$$\lim_{\varepsilon \to 0} e^{\frac{1}{\varepsilon}\left(1 - \frac{q^\star_{h,u}(\varepsilon)}{1-\varepsilon}\right)} < \infty.$$

Therefore the analytical expression of $m_{qq}(q, \varepsilon)$, in equation (55), implies that:

$$\lim_{\varepsilon \to 0} m_{qq}(q^\star_{h,u}(\varepsilon), \varepsilon) = -\infty \text{ if } \mu_{h,u} < s. \tag{37}$$

Lastly, when $\mu_{h,u} < s$, $\xi^\star_u(0) < 1$ implies that $q^\star_{\ell,u,u}(0) > 1$ and, using (55) together with (35) , that

$$\lim_{\varepsilon \to 0} m_{qq}(q^\star_{\ell,u}(\varepsilon), \varepsilon) = 0 \text{ if } \mu_{h,u} < s. \tag{38}$$

Step 2: limit of volume when $\mu_{h,u} < s$. We have

$$D_\pi(0, \xi^\star_u(\varepsilon)) = -\frac{\delta m_q m^\sigma}{m_{qq}(1 - \delta m^\sigma) - \delta\sigma m_q^2 m^{\sigma-1}}$$

where m, m_q and m_{qq} are all evaluated at $q^\star_{\ell,u,u}(\varepsilon)$ and ε. Using (32), (34), and (38), we obtain that:

$$\lim_{\varepsilon \to 0} D_\pi(0, \xi^\star_u(\varepsilon)) = \frac{q^\star_{\ell,u,u}(0)}{\sigma}.$$

A similar argument shows that:

$$\lim_{\varepsilon \to 0} D_\xi(0, \xi^\star_u(\varepsilon)) = \lim_{\varepsilon \to 0} \frac{1}{m_{qq}(1 - \delta m^\sigma) - \delta\sigma m_q^2 m^{\sigma-1}} = -\frac{1}{\delta\sigma\left(q^\star_{\ell,u,u}\right)^{\sigma-1}}.$$

Lastly, using (33), (36), and (37), we have

$$\lim_{\varepsilon\to 0} D_\xi(1,\xi_u^\star(\varepsilon)) = \lim_{\varepsilon\to 0} \frac{1}{m_{qq}(1-\delta m^\sigma) - \delta\sigma m_q^2 m^{\sigma-1}} = 0,$$

where m, m_q, and m_{qq} are evaluated at $q^\star_{h,u}(\varepsilon)$ and ε. Now using these limits in (31) we obtain that

$$\lim_{\varepsilon\to 0} \frac{\partial q^\star_{\ell,u,u}(\varepsilon)}{\partial u} = \gamma \frac{q^\star_{\ell,u,u}(0)}{\sigma} + \gamma \left(q^\star_{h,u}(0) - q^\star_{\ell,u,u}(0)\right).$$

When $\varepsilon = 0$ and $\mu_{h,u} < s$, $\xi^\star_u(0) < 1$ implying that $q^\star_{h,u}(0) = 1$ and, from market clearing, that $q^\star_{\ell,u,u}(0) = (s-\mu_{h,u})/(1-\mu_{h,u})$. Plugging these expressions into the above, we obtain the desired result.

Step 3: limit when $\mu_{h,u} \geq s$. In this case we note that

$$\frac{\partial q^\star_{\ell,u,u}(\varepsilon)}{\partial u} \leq \gamma D_\pi(0,\xi^\star_u(\varepsilon)) \leq \gamma \frac{\delta m_q m^\sigma}{-m_{qq}(1-\delta m^\sigma) + \delta\sigma m_q^2 m^{\sigma-1}} \leq \gamma \frac{\delta m_q m^\sigma}{\delta\sigma m_q^2 m^{\sigma-1}} = \frac{\gamma m}{\delta m_q} \to 0,$$

using (32) and (34), where m, m_q and m_{qq} are evaluated at $q^\star_{\ell,u,u}(\varepsilon)$ and ε. Clearly, this implies that

$$\lim_{\varepsilon\to 0} \max\left\{(1-\mu_{h,u})\frac{\partial q^\star_{\ell,u,u}(\varepsilon)}{\partial u}, 0\right\} = 0 = \gamma \max\left\{\frac{s-\mu_{h,u}}{\sigma} - (1-s), 0\right\},$$

given that $\mu_{h,u} \geq s$. □

A.9 Proof of Propositions 8, 9, and 10

A.9.1 A characterization of equilibrium object

Let

$$S_u \equiv s - \mathbb{E}\left[\mu_{h,\tau_u}\right] = e^{-\rho u}(s-\mu_{h,0}) + \int_0^u \rho e^{-\rho(u-t)}(s-\mu_{h,t})dt, \tag{39}$$

the gross asset supply in the hand of investors, minus the maximum (unit) demand of high–valuation investors. Keeping in mind that T_s is the time such that $\mu_{h,T_s} = s$, one sees that S_u has a unique root $T_f > T_s$. Indeed, $S_u > 0$ for all $u \in [0, T_s)$ and S_u goes to minus infinity when u goes infinity, so S_u has a root $T_f > T_s$. It is unique because $\dot{S}_{T_f} = \rho(s-\mu_{h,T_f}) < 0$ since $T_f > T_s$.

We already know from the text that, when $u \geq T_f$, $\xi_u = 1$, $q_{h,u} \in [0,1]$ and $q_{\ell,\tau_u,u} = 0$.[23] Now consider $u \in [0, T_f)$. In that case, $S_u > 0$ and we already know that $\xi_u < 1$, which implies that

[23] The holding of high-valuation trader is indeterminate. However, it is natural to assume that they have identical holdings, $q_{h,u} = s/\mathbb{E}\left[\mu_{h,\tau_u}\right]$. Indeed, we have seen that this is the limit of high-valuation investors' holdings as $\varepsilon \to 0$.

high–valuation investors hold one unit, $q_{h,u} = 1$. Replacing expression (8) for $\pi_{\tau_u,u}$ into the asset demand (23) we obtain that:

$$q_{\ell,\tau_u,u} = \min\{(1-\mu_{h,\tau_u})^{1/\sigma} Q_u, 1\}, \qquad \text{where } Q_u \equiv \left(\frac{1-\xi_u}{\delta(1-\mu_{h,u})}\right)^{1/\sigma}.$$

This is the formula for holding plans in Proposition 8 Plugging this back into the market clearing condition (24), we obtain that Q_u solves:

$$e^{-\rho u}(1-\mu_{h,0})\min\{(1-\mu_{h,0})^{1/\sigma}Q_u, 1\} + \int_0^u \rho e^{-\rho(u-t)}(1-\mu_{h,t})\min\{(1-\mu_{h,t})^{1/\sigma}Q_u, 1\}\,dt = S_u. \tag{40}$$

The left-hand side of (40) is continuous, strictly increasing for $Q_u < (1-\mu_{h,u})^{-1/\sigma}$ and constant for $Q_u \geq (1-\mu_{h,u})^{-1/\sigma}$. It is zero when $Q_u = 0$, and strictly larger than S_u when $Q_u = (1-\mu_{h,u})^{-1/\sigma}$ since $s < 1$. Therefore, equation (40) has a unique solution and the solution satisfies $0 < Q_u < (1-\mu_{h,u})^{-1/\sigma}$.

Now, turning to the price, the definition of Q_u implies that the price solves $rp_u = 1 - \delta(1 - \mu_{h,u})Q_u^\sigma + \dot{p}_u$ for $u < T_f$. For $u \geq T_f$, the fact that high-valuation traders are indifferent between any asset holdings in $[0,1]$ implies that $rp_u = 1 + \dot{p}_u$. But the price is bounded and positive, so it follows that $p_u = 1/r$. Since the price is continuous at T_f, this provides a unique candidate equilibrium price path. Clearly this candidate is C^1 over $(0, T_f)$ and (T_f, ∞). To show that it is continuously differentiable at T_f note that, given $Q_{T_f} = 0$ and $p_{T_f} = 1/r$, the ODE $rp_u = 1 - \delta(1-\mu_{h,u})Q_u^\sigma + \dot{p}_u$ implies that $\dot{p}_{T_f^-} = 0$. Obviously, since the price is constant after T_f, $\dot{p}_{T_f^+} = 0$ as well. We conclude that $\dot{p}_u$ is continuous at $u = T_f$ as well.

Next, we show that the candidate equilibrium thus constructed is indeed an equilibrium. For this recall that $0 < Q_u < (1-\mu_{h,u})^{-1/\sigma}$, which immediately implies that $0 < 1 - rp_u + \dot{p}_u < 1$ for $u < T_f$. It follows that high-valuation traders find it optimal to hold one unit. Now one can directly verify that, for $u < T_f$, the problem of low-valuation traders is solved by $q_{t,u} = \min\{(1-\mu_{h,t})^{-1/\sigma}Q_u, 1\}$. For $u \geq T_f$, $1 - rp_u + \dot{p}_u = 0$ and so the problem of high-valuation traders is solved by any $q_{t,u} \in [0,1]$, while the problem of low-valuation traders is clearly solved by $q_{t,u} = 0$. The asset market clears at all dates by construction.

A.9.2 Concluding the proof of Proposition 8: the shape of Q_u

We begin with preliminary results that we use repeatedly in this appendix. For the first preliminary result, consider equation (40) after removing the min operator in the integral:

$$e^{-\rho u}(1-\mu_{h,0})^{1+1/\sigma}\overline{Q}_u + \int_0^u \rho e^{-\rho(u-t)}(1-\mu_{h,t})^{1+1/\sigma}\overline{Q}_u\,dt = S_u$$
$$\iff \quad \overline{Q}_u = \frac{(s-\mu_{h,0}) + \int_0^u \rho e^{\rho t}(s-\mu_{h,t})\,dt}{(1-\mu_{h,0})^{1+1/\sigma} + \int_0^u \rho e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma}\,dt}. \tag{41}$$

Now, whenever $(1-\mu_{h,0})^{1/\sigma}\overline{Q}_u \leq 1$, it is clear that $\overline{Q}_u$ also solves equation (40). Given that the solution of (40) is unique it follows that $Q_u = \overline{Q}_u$. Conversely if $Q_u = \overline{Q}_u$, subtracting (40) from (41) shows that:

$$\Big((1-\mu_{h,0})^{1/\sigma}\overline{Q}_u - \min\{(1-\mu_{h,0})^{1/\sigma}\overline{Q}_u, 1\}\Big)$$
$$+\int_0^u \rho e^{-\rho(u-t)}(1-\mu_{h,t})\Big((1-\mu_{h,t})^{1/\sigma}\overline{Q}_u - \min\{(1-\mu_{h,t})^{1/\sigma}\overline{Q}_u, 1\}\Big)\,dt = 0.$$

Since the first term and the integrand are positive, this can only be true if $(1-\mu_{h,0})^{1/\sigma}\overline{Q}_u \leq 1$. Taken together, we obtain:

Lemma 9 (A useful equivalence). *$\overline{Q}_u \leq (1-\mu_{h,0})^{-1/\sigma}$ if and only if $Q_u = \overline{Q}_u$.*

The next Lemma, proved in Section B.1.5, page 60, provides basic properties of $\overline{Q}_u$:

Lemma 10 (Preliminary results about $\overline{Q}_u$). *The function $\overline{Q}_u$ is continuous, satisfies $\overline{Q}_0 = \frac{s-\mu_{h,0}}{(1-\mu_{h,0})^{1+1/\sigma}}$ and $\overline{Q}_{T_f} = 0$. It is strictly decreasing over $(0, T_f]$ if $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} \leq \frac{\sigma}{1+\sigma}$ and hump-shaped otherwise.*

Taken together, Lemma 9 and Lemma 10 immediately imply that

Lemma 11. *The function Q_u satisfies $Q_0 = \frac{s-\mu_{h,0}}{(1-\mu_{h,0})^{1+1/\sigma}}$, $Q_{T_f} = 0$. If $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} \leq \frac{\sigma}{1+\sigma}$, then it is strictly decreasing over $(0, T_f]$. If $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} > \frac{\sigma}{1+\sigma}$ and $\overline{Q}_u \leq (1-\mu_{h,0})^{-1/\sigma}$ for all $u \in (0, T_f]$, Q_u is hump-shaped over $(0, T_f]$.*

The only case that is not covered by the Lemma is when $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} > \frac{\sigma}{1+\sigma}$ and $\overline{Q}_u > (1-\mu_{h,0})^{-1/\sigma}$ for some $u \in (0, T_f]$. In this case, note that for u small and u close to T_f, we have that $\overline{Q}_u < (1-\mu_{h,0})^{-1/\sigma}$. Given that $\overline{Q}_u$ is hump-shaped, it follows that the equation $\overline{Q}_u = (1-\mu_{h,0})^{-1/\sigma}$ has two solutions, $0 < T_1 < T_2 < T_f$. For $u \in (0, T_1]$ (resp. $u \in [T_2, T_f]$), $\overline{Q}_u \leq (1-\mu_{h,0})^{-1/\sigma}$ and is increasing (resp. decreasing), and thus Lemma 9 implies that $Q_u = \overline{Q}_u$ and increasing (resp. decreasing) as well. We first establish that Q_u is piecewise continuously differentiable. Let $\Psi(Q) \equiv \inf\{\psi \geq 0 \,:\, (1-\mu_{h,\psi})^{1/\sigma} Q \leq 1\}$, and $\psi_u \equiv \Psi(Q_u)$. Thus, $\psi_u > 0$ if and only if $u \in (T_1, T_2)$. We have:

Lemma 12. *Q_u is continuously differentiable except in T_1 and T_2, and*

$$Q'_u = \frac{\rho e^{\rho u}\left(s-\mu_{h,u}-(1-\mu_{h,u})^{1+1/\sigma}Q_u\right)}{\mathbb{I}_{\{\psi_u=0\}}(1-\mu_{h,0})^{1+1/\sigma} + \int_{\psi_u}^u \rho e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma}\,dt}. \tag{42}$$

The proof is in Appendix B.1.6, page 61. In particular, (42) implies that $Q_{T_1^+}$ has the same sign as $Q_{T_1^-}$, which is positive, and that $Q_{T_2^-}$ has the same sign as $Q_{T_2^+}$, which is negative. Thus, Q_u changes sign at least once in (T_1, T_2). To conclude, in Section B.1.7, page 62, we establish:

Lemma 13. *The derivative Q'_u changes sign only once in (T_1, T_2).*

A.9.3 Proof of Proposition 9: asymptotic volume

In Section B.1.8, page 63 in the supplementary appendix, we prove the following asymptotic results:

Lemma 14. *As ρ goes to infinity:*

$$T_f(\rho) \downarrow T_s \tag{43}$$

$$Q_u(\rho) = \frac{s-\mu_{h,u}}{(1-\mu_{h,u})^{1+1/\sigma}} - \frac{1}{\rho}\frac{\gamma}{(1-\mu_{h,u})^{1/\sigma}}\left[\left(1+\frac{1}{\sigma}\right)\frac{s-\mu_{h,u}}{1-\mu_{h,u}} - 1\right] + o\left(\frac{1}{\rho}\right), \quad \forall u \in [0, T_s] \tag{44}$$

$$T_\psi(\rho) = \arg\max_{u\in[0,T_f(\rho)]} Q_u(\rho) \longrightarrow \arg\max_{u\in[0,T_s]} \frac{s-\mu_{h,u}}{(1-\mu_{h,u})^{1+1/\sigma}}. \tag{45}$$

With this in mind we can study the asymptotic behavior of volume.

Basic formulas. As above, let T_ψ denote the arg max of the function Q_u. For any time $u < T_\psi$ and some time interval $[u, u+du]$, the only traders who *sell* are those who have an updating time during this time interval, and who find out that they have a *low* valuation. Thus, trading volume during $[u, u+du]$ can be computed as the volume of assets sold by these investors, as follows. Just before their updating time, low–valuation investors hold on average:

$$\mathbb{E}\left[q_{\ell,\tau_u,u}\right] = e^{-\rho u} q_{\ell,0,u} + \int_0^u \rho e^{-\rho(u-t)} q_{\ell,t,u}\, dt. \tag{46}$$

Instantaneous trading volume is then:

$$V_u = \rho(1-\mu_{h,u})\Big(\mathbb{E}\left[q_{\ell,\tau_u,u}\right] - q_{\ell,u,u}\Big), \tag{47}$$

where $\rho(1-\mu_{h,u})$ is the measure of low–valuations investors having an updating time, the term in large parentheses is the average size of low–valuation investors' sell orders, and $q_{\ell,u,u}$ is their asset holding right after the updating time.

For any time $u \in (T_\psi, T_f)$ and some time interval $[u, u+du]$, the only traders who *buy* are those who have an updating time during this time interval, and who find out that they have switched from a low to a high valuation. Trading volume during $[u, u+du]$ can be computed as the volume of assets purchased by these traders:

$$V_u = \rho\mathbb{E}\left[(1-\mu_{h,\tau_u})\pi_{\tau_u,u}\left(1-q_{\ell,\tau_u,u}\right)\right] = \rho\mathbb{E}\left[\left(\mu_{h,u}-\mu_{h,\tau_u}\right)\left(1-q_{\ell,\tau_u,u}\right)\right], \tag{48}$$

where the second equality follows by definition of $\pi_{\tau_u,u}$.

Finally, for $u > T_f$, the trading volume is not zero since high–valuation traders continue to buy from the low valuation investors having an updating time:

$$V_u = \rho(1-\mu_{hu})\mathbb{E}\left[q_{\ell,\tau_u,u}\right]. \tag{49}$$

Taking the $\rho \to \infty$ limit. We first note that $q_{\ell,u,u}(\rho) = \min\{(1-\mu_{h,u})^{1/\sigma} Q_u, 1\} = (1-\mu_{h,u})^{1/\sigma} Q_u(\rho)$. Next, we need to calculate an approximation for:

$$\mathbb{E}\left[q_{\ell,\tau_u,u}\right] = q_{\ell,0,u} e^{-\rho u} + \int_0^u \min\{(1-\mu_{h,t})^{1/\sigma} Q_u(\rho), 1\} \rho e^{-\rho(u-t)}\, dt.$$

For this we follow the same calculations leading to equation (63) in the proof of Lemma 14, but with $f(t,\rho) = \min\{(1-\mu_{h,t})^{1/\sigma} Q_u(\rho), 1\}$. This gives:

$$\begin{aligned}\mathbb{E}\left[q_{\ell,\tau_u,u}\right] &= f(0,u) e^{-\rho u} + \int_0^u \rho e^{-\rho(u-t)} f(t,\rho)\, dt = f(u,\rho) - \frac{1}{\rho} f_t(u,\rho) + o\left(\frac{1}{\rho}\right) \\ &= (1-\mu_{h,u})^{1/\sigma} Q_u(\rho) + \frac{1}{\rho}\frac{\gamma}{\sigma}\frac{s-\mu_{h,u}}{1-\mu_{h,u}} + o\left(\frac{1}{\rho}\right) = \frac{s-\mu_{h,u}}{1-\mu_{h,u}} + \frac{\gamma}{\rho}\frac{1-s}{1-\mu_{h,u}} + o\left(\frac{1}{\rho}\right),\end{aligned}$$

where the second line follows from plugging equation (44) into the first line. Substituting this expression into equations (47) and (48), we find after some straightforward manipulation that, when ρ goes to infinity, $V_u \to \gamma(s-\mu_{h,u})/\sigma$ for $u < T_\psi(\infty)$, $V_u \to \gamma(1-s)$ for $u \in (\lim T_\psi(\infty), T_s)$, and $V_u \to 0$ for $u > T_s$.

The trading volume in the Walrasian equilibrium is equal to the measure of low–valuation investors who become high-valuation investors: $\gamma(1-\mu_{h,u})$, times the amount of asset they buy at that time: $1-(s-\mu_{h,u})/(1-\mu_{h,u})$. Thus the trading volume is $\gamma(1-s)$. To conclude the proof, note that after taking derivatives of $Q_u(\infty)$ with respect to u, it follows that

$$Q'_{T_\psi(\infty)}(\infty) = 0 \Leftrightarrow \frac{s-\mu_{hT_\psi(\infty)}}{\sigma} = 1-s$$

which implies in turn that $\gamma(s-\mu_{h,u})/\sigma > \gamma(1-s)$ for $u < T_\psi(\infty)$.

A.9.4 Proof of Proposition 10

We have already argued that the price is continuously differentiable. To prove that it is strictly increasing for $u \in [0, T_f)$, we let $\Delta_u \equiv (1-\mu_{h,u})^{1/\sigma} Q_u$ for $u \le T_f$, and $\Delta_u = 0$ for $u \ge T_f$. In Section B.1.9, page 65 in the supplementary appendix, we show that:

Lemma 15. *The function Δ_u is strictly decreasing over* $(0, T_f]$.

Now, in terms of Δ_u, the price writes:

$$p_u = \int_u^\infty e^{-r(y-u)} \left(1 - \delta \Delta_y^\sigma\right) dy = \int_0^\infty e^{-rz} \left(1 - \delta \Delta_{z+u}^\sigma\right) dz,$$

after the change of variable $y - u = z$. Since Δ_u is strictly decreasing over $u \in (0, T_f)$, and constant over $[T_f, \infty)$, it clearly follows from the above formula that p_u is strictly increasing over $u \in (0, T_f)$. Next:

First bullet point: when (25) does not hold. Given the ODEs satisfied by the price path, it suffices to show that, for all $u \in (0, T_s)$,

$$(1-\mu_{h,u})Q_u^\sigma > \left(\frac{s-\mu_{h,u}}{1-\mu_{h,u}}\right)^\sigma.$$

Besides, when condition (25) holds, it follows from Lemma 9 and Lemma 10 that:

$$Q_u = \overline{Q}_u = \frac{s-\mu_{h,0}+\int_0^u e^{\rho t}(s-\mu_{h,t})\,dt}{(1-\mu_{h,0})^{1+1/\sigma}+\int_0^u e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma}\,dt}.$$

Plugging the above and rearranging, we are left with showing that:

$$\begin{aligned} F_u = (s-\mu_{h,u})&\left[(1-\mu_{h,0})^{1+1/\sigma}+\int_0^u e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma}\,dt\right] \\ &-(1-\mu_{h,u})^{1+1/\sigma}\left[s-\mu_{h,0}+\int_0^u e^{\rho t}(s-\mu_{h,t})\right] < 0. \end{aligned}$$

But we know from the proof of Lemma 10, equation (56), page 60, that F_u has the same sign as Q'_u, which we know is negative at all $u > 0$ since $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} \leq \frac{\sigma}{\sigma+1}$.

Second bullet point: when condition (25) holds and when s is close to 1 and σ is close to 0. The price at time 0 is equal to:

$$p_0 = \int_0^{+\infty} e^{-ru}\xi_u\,du.$$

With known preferences, $\xi^\star_u = 1-\delta\left(\frac{s-\mu_{h,u}}{1-\mu_{h,u}}\right)^\sigma = 1-\delta\left(1-\frac{1-s}{1-\mu_{h,0}}e^{\gamma u}\right)^\sigma$ for $u < T_s$, and $\xi^\star_u = 1$ for $u > T_s$. Therefore $p^\star_0 = 1/r - \delta J^\star(s)$, where:

$$J^\star(s) \equiv \int_0^{T_s} e^{-ru}\left(1-\frac{1-s}{1-\mu_{h,0}}e^{\gamma u}\right)^\sigma\,du,$$

where we make the dependence of $J^\star(s)$ on s explicit. Similarly, with preference uncertainty, the price at time 0 is equal to $p_0 = 1/r - \delta J(s)$, where:

$$J(s) \equiv \int_0^{T_f} e^{-ru}(1-\mu_{h,u})Q_u^\sigma\,du.$$

We begin with a Lemma proved in Section B.1.10, page 67:

Lemma 16. *When s goes to 1, both $J^\star(s)$ and $J(s)$ go to $1/r$.*

Therefore, p_0 goes to $(1-\delta)/r$ both with continuous and infrequent updating. Besides, with continuous updating:

$$p_0(s_0) = (1-\delta)/r + \delta \int_{s_0}^1 J^{\star\prime}(s)\, ds,$$

and with infrequent updating:

$$p_0(s_0) = (1-\delta)/r + \delta \int_{s_0}^1 J'(s)\, ds.$$

The next two lemmas compare $J^{\star\prime}(s)$ and $J'(s)$ for s in the neighborhood of 1 when σ is not too large. The first Lemma is proved in Section B.1.11, page 67.

Lemma 17. *When s goes to 1:*

$$\begin{aligned}
&J^{\star\prime}(s) \sim \sigma \times \textit{constant} && \textit{if } r > \gamma,\\
&J^{\star\prime}(s) \sim \Gamma_1(\sigma) \log((1-s)^{-1}) && \textit{if } r = \gamma,\\
&J^{\star\prime}(s) \sim \Gamma_2(\sigma)(1-s)^{-1+r/\gamma} && \textit{if } r < \gamma,
\end{aligned}$$

where the constant terms $\Gamma_1(\sigma)$ and $\Gamma_2(\sigma)$ go to 0 when $\sigma \to 0$.

In this Lemma and all what follows $f(s) \sim g(s)$ means that $f(s)/g(s) \to 1$ when $s \to 1$. The second Lemma is proved in Section B.1.12, page 68 in the supplementary appendix:

Lemma 18. *Assume $\gamma + \gamma/\sigma - \rho > 0$. There exists a function $\tilde{J}'(s) \leq J'(s)$ such that, when s goes to 1:*

$$\begin{aligned}
&\tilde{J}'(s) \to +\infty && \textit{if } r > \gamma,\\
&\tilde{J}'(s) \sim \Gamma_3(\sigma) \log((1-s)^{-1}) && \textit{if } r = \gamma,\\
&\tilde{J}'(s) \sim \Gamma_4(\sigma)(1-s)^{-1+r/\gamma} && \textit{if } r < \gamma,
\end{aligned}$$

where the constant terms $\Gamma_3(\sigma)$ and $\Gamma_4(\sigma)$ go to strictly positive limits when $\sigma \to 0$.

Lemmas 17 and 18 imply that, if σ is close to 0, then $J'(s) > J^{\star\prime}(s)$ for s in the left-neighborhood of 1. The second point of the proposition then follows.

B Supplementary Appendix (not for publication)

This supplementary appendix provides ommitted proofs and establishes results to complement the main analysis.

B.1 Omitted proofs

B.1.1 Proof of Lemma 3

Twice continuous differentiability follows directly from the fact that $D(\pi, \xi)$, $\mu_{h,t}$ and $\pi_{t,u}$ are all twice continuously differentiable. Boundedness is proved from the following direct calculations. First, note that:

$$\pi_{t,u} = \frac{\mu_{h,u} - \mu_{h,t}}{1 - \mu_{h,t}} = 1 - \frac{1 - \mu_{h,u}}{1 - \mu_{h,t}},$$

which implies that:

$$\frac{\partial \pi_{t,u}}{\partial u} = \frac{\mu'_{h,u}}{1 - \mu_{h,t}} \quad \text{and} \quad \frac{\partial \pi_{t,u}}{\partial t} = -\mu'_{h,t} \frac{1 - \mu_{h,u}}{[1 - \mu_{h,t}]^2}. \tag{50}$$

First derivative with respect to t: $\mathcal{D}_t(t, u, \xi)$. Then we can calculate the partial derivative of $\mathcal{D}(t, u, \xi)$ with respect to t

$$\begin{aligned} \mathcal{D}_t(t, u, \xi) &= \mu'_{h,t} D(1, \xi) - \mu'_{h,t} D(\pi_{t,u}, \xi) + \left[1 - \mu_{h,t}\right] D_\pi(\pi_{t,u}, \xi) \frac{\partial \pi_{t,u}}{\partial t} \\ &= \mu'_{h,t} \left\{ D(1, \xi) - D(\pi_{t,u}, \xi) - \frac{1 - \mu_{h,u}}{1 - \mu_{h,t}} D_\pi(\pi_{t,u}, \xi) \right\}, \end{aligned} \tag{51}$$

which is bounded over the relevant range, $\Delta \times [\underline{\xi}, \overline{\xi}]$, because: $\mu'_{h,t} = \gamma e^{-\gamma t}$ is bounded; $(1 - \mu_{h,u})/(1 - \mu_{h,t}) = e^{-\gamma(u-t)}$ and so is bounded; $\pi_{t,u} \in [0, 1]$ and so is bounded; $D(\pi, \xi)$ and $D_\pi(\pi, \xi)$ are continuous over the compact $[0, 1] \times [\underline{\xi}, \overline{\xi}]$ and so are bounded as well.

First derivative with respect to u: $\mathcal{D}_u(t, u, \xi)$. We have:

$$\mathcal{D}_u(t, u, \xi) = \left[1 - \mu_{h,t}\right] D_\pi(\pi_{t,u}, \xi) \frac{\partial \pi_{t,u}}{\partial u} = \mu'_{h,u} D_\pi(\pi_{t,u}, \xi),$$

which is bounded over the relevant range for the same reasons as above.

First derivative with respect to ξ: $\mathcal{D}_\xi(t, u, \xi)$. We have

$$\mathcal{D}(t, u, \xi) = \mu_{h,t} D_\xi(1, \xi) + \left[1 - \mu_{h,t}\right] D_\xi(\pi_{t,u}, \xi), \tag{52}$$

which is bounded over the relevant range for the same reasons as above.

Second derivative with respect to (t, t). It is equal to:

$$\mathcal{D}_{t,t}(t, u, \xi) = \mu''_{h,t} \left\{ D(\pi, \xi) - D(\pi_{t,u}, \xi) - \frac{1-\mu_{h,t}}{1-\mu_{h,u}} D_\pi(\pi_{t,u}, \xi) \right\} + \mu'_{h,t} \frac{\mu'_{h,t}}{1-\mu_{h,t}} \left[\frac{1-\mu_{h,u}}{1-\mu_{h,t}} \right]^2 D_{\pi\pi}(\pi_{t,u}, \xi),$$

which is bounded over the relevant range for the same reason as above and after noting that $\mu'_h(t)/ [1 - \mu_h(t)] = \gamma$.

Second derivative with respect to (t, u). It is equal to:

$$\mathcal{D}_{t,u}(t, u, \xi) = -\mu'_{h,t}\mu'_{h,u} \frac{1-\mu_{h,u}}{\left[1-\mu_{h,t}\right]^2} D_{\pi,\pi}(\pi_{t,u}, \xi),$$

which is bounded over the relevant range for the same reasons as above.

Second derivative with respect to (t, ξ). It is equal to:

$$\mathcal{D}_{t,\xi}(t, u, \xi) = \mu'_{h,t} \left\{ D_\xi(1, \xi) - D_\xi(\pi_{t,u}, \xi) - \frac{1-\mu_{h,u}}{1-\mu_{h,t}} D_{\pi,\xi}(\pi_{t,u}, \xi) \right\},$$

which is bounded over the relevant range for the same reasons as above.

Second derivative with respect to (u, u): $\mathcal{D}_{u,u}(t, u, \xi)$.

$$\mathcal{D}_{u,u}(t, u, \xi) = \mu''_{h,u} D_\pi(\pi_{t,u}, \xi) + \mu'_{h,u} \frac{\mu'_{h,u}}{1-\mu_{h,t}} D_{\pi,\pi}(\pi_{t,u}, \xi).$$

which is bounded over the relevant range since $\mu'_{h,u}/ \left[1 - \mu_{h,t}\right] = \gamma e^{-\gamma(u-t)}$.

Second derivative with respect to (u, ξ): $\mathcal{D}_{u,\xi}(t, u, \xi)$.

$$\mathcal{D}_{u,\xi}(t, u, \xi) = \mu'_{h,u} D_{\pi,\xi}(\pi_{t,u}, \xi),$$

which is bounded over the relevant range.

Second derivatives with respect to (ξ, ξ): $\mathcal{D}_{\xi,\xi}(t, u, \xi)$.

$$\mathcal{D}_{\xi,\xi} = \mu_{h,t} D_{\xi,\xi}(\theta, \xi) + \left[1 - \mu_{h,t}\right] D_{\xi,\xi}(\pi_{t,u}, \xi),$$

which is bounded over the relevant range.

$\mathcal{D}_\xi(t, u, \xi)$ is bounded away from zero. This follows directly from the formula for $\mathcal{D}_\xi(t, u, \xi)$ because, on the one hand, $\mu_{h,t} \in [0, 1]$ and, on the other hand, $D_\xi(\pi, \xi)$ is continuous and thus bounded away from zero over the compact $[0, 1] \times [\underline{\xi}, \overline{\xi}]$.

B.1.2 Proof of Lemma 6

First, note that, by an application of the Implicit Function Theorem, the holding cost $\xi_u(\rho)$ is continuously differentiable, with a derivative that can be written

$$\frac{d\xi_u(\rho)}{du} = -\frac{A_u(\rho) + B_u(\rho)}{C_u(\rho)},$$

$$\text{where } A_u(\rho) = \rho\mathcal{D}(u,u,\xi_u(\rho)); \quad B_u(\rho) = e^{-\rho u}\mathcal{D}_u(0,u,\xi_u(\rho)) + \int_0^u \rho e^{-\rho(u-t)}\mathcal{D}_u(t,u,\xi_u(\rho))\,dt;$$

$$C_u(\rho) = e^{-\rho u}\mathcal{D}_\xi(0,u,\xi_u(\rho)) + \int_0^u \rho e^{-\rho(u-t)}\mathcal{D}_\xi(t,u,\xi_u(\rho))\,dt.$$

To obtain the asymptotic behavior of $A_u(\rho)$, we apply a second-order Taylor formula:

$$\begin{aligned}
A_u(\rho) =& \rho\left\{\mathcal{D}(u,u,\xi_u^\star) + \mathcal{D}_\xi(u,u,\xi_u^\star)\left[\xi_u - \xi_u^\star\right] + \frac{\mathcal{D}_{\xi\xi}(u,u,\hat{\xi}_u(\rho))}{2}\left[\xi_u - \xi_u^\star\right]^2\right\} \\
=& \rho\Bigg\{\mathcal{D}_\xi(u,u,\xi_u^\star)\left[\frac{1}{\rho}\frac{\mathcal{D}_t(u,u,\xi_u^\star)}{\mathcal{D}_\xi(u,u,\xi_u^\star)} + o_\alpha\left(\frac{1}{\rho}\right)\right] \\
&\qquad + \frac{\mathcal{D}_{\xi\xi}(u,u,\hat{\xi}_u(\rho))}{2}\left[\frac{1}{\rho}\frac{\mathcal{D}_t(u,u,\xi_u^\star)}{\mathcal{D}_\xi(u,u,\xi_u^\star)} + o_\alpha\left(\frac{1}{\rho}\right)\right]^2\Bigg\} \\
=& \mathcal{D}_t(u,u,\xi_u^\star) + o_\alpha(1),
\end{aligned}$$

where the second line follows after noting that $\mathcal{D}(u,u,\xi_u^\star) = 0$ by definition of $\xi_u^\star$ and after plugging in the approximation of Proposition 5.

Turning to $B_u(\rho)$, we first integrate by part to note that:

$$\begin{aligned}
B_u(\rho) &= \mathcal{D}_u(u,u,\xi_u(\rho)) - \frac{1}{\rho}\int_0^u \rho e^{-\rho(u-t)}\mathcal{D}_{u,t}(t,u,\xi_u(\rho))\,dt \\
&= \mathcal{D}_u(u,u,\xi_u(\rho)) + o_\alpha(1),
\end{aligned}$$

since, by Lemma 3, $\mathcal{D}(t,u,\xi)$ has bounded first and second derivatives. Given that $\mathcal{D}_u(t,u,\xi)$ has bounded first derivatives, it is uniformly continuous over $\Delta \times [\underline{\xi}, \overline{\xi}]$. Together with the fact that $\xi_u(\rho) = \xi_u^\star + o_\alpha(1)$, this implies that:

$$B_u(\rho) = \mathcal{D}_u(u,u,\xi_u^\star) + o_\alpha(1).$$

The same arguments applied to $C_u(\rho)$ show that:

$$C_u(\rho) = \mathcal{D}_\xi(u,u,\xi_u^\star) + o_\alpha(1).$$

Taken together, we obtain that

$$\frac{d\xi_u(\rho)}{du} = -\frac{\mathcal{D}_t(u,u,\xi_u^\star) + \mathcal{D}_u(u,u,\xi_u^\star)}{\mathcal{D}_\xi(u,u,\xi_u^\star)} + o_\alpha(1) = \frac{d\xi_u^\star}{du} + o_\alpha(1).$$

B.1.3 Proof of Lemma 7

Since $D(\pi,\xi)$ is uniformly continuous over $[0,1]\times[\underline{\xi},\overline{\xi}]$, and since $\xi_u(\rho) = \xi_u^\star + o_\alpha(1)$, it follows that $q_{\ell,t,u} = q_{\ell,t,u}^\star + o_\alpha(1)$, and $q_{h,u} = q_{h,u}^\star + o_\alpha(1)$. Using the same argument we obtain that:

$$\frac{\partial q_{\ell,t,u}}{\partial t} = D_\pi(\pi_{t,u},\xi_u(\rho))\frac{\partial \pi_{t,u}}{\partial t} = D_\pi(\pi_{t,u},\xi_u^\star)\frac{\partial \pi_{t,u}}{\partial t} + o_\alpha(1),$$

Next:

$$\begin{aligned}\frac{\partial q_{\ell,t,u}}{\partial u} &= D_\pi(\pi_{t,u},\xi_u(\rho))\frac{\partial \pi_{t,u}}{\partial u} + D_\pi(\pi_{t,u},\xi_u(\rho))\frac{d\xi_u(\rho)}{du}\\ &= D_\pi(\pi_{t,u},\xi_u^\star)\frac{\partial \pi_{t,u}}{\partial u} + D_\xi(\pi_{t,u},\xi_u^\star)\frac{d\xi_u^\star}{du} + o_\alpha(1),\end{aligned}$$

using the same argument as above as well as Lemma 6. Lastly,

$$\frac{dq_{h,u}}{du} = D_\pi(1,\xi_u(\rho))\frac{d\xi_u(\rho)}{du} = D_\xi(1,\xi_u^\star)\frac{d\xi_u^\star}{du} + o(1),$$

using the same argument as above.

B.1.4 Proof of Lemma 8

One easily verifies that $m(0,\varepsilon) = 0$ and $\lim_{q\to\infty} m(q,\varepsilon) = 1$. Clearly, $m(q,\varepsilon)$ is continuous over $(q,\varepsilon)\in[0,\infty)\times(0,\infty)$, so the only potential difficulty lies in proving continuity at all points $(q,0)$. For this consider $q\geq 0$ and a sequence $(q_n,\varepsilon_n)\to(q,0)$. We need to show that $m(q_n,\varepsilon_n)\to\min\{q,1\}$. If $q\geq 1$, the numerator of $1-m(q_n,\varepsilon_n)$ is positive and bounded above by $\ln(1+e^{\frac{-1}{1-\varepsilon}})$ which goes to $\ln(1+e^{-1})$, and the denominator goes to $+\infty$. Therefore, $1-m(q_n,\varepsilon_n)\to 0$ and so $m(q_n,\varepsilon_n)\to 1$. Consider now $q<1$. For $x>0$, let $\phi(x)\equiv x\ln\left(1+e^{\frac{1}{x}}\right)$ and let $\phi(0)=\lim_{x\to 0^+}\phi(x)=1$, so that $\phi(x)$ is extended by continuity at 0^+. We can then write:

$$1-m(q,\varepsilon) = \vartheta(q,\varepsilon)\frac{\phi\circ\psi(q,\varepsilon)}{\phi(\varepsilon)}, \text{ where } \vartheta(q,\varepsilon)\equiv\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right) \text{ and } \psi(q,\varepsilon)\equiv\varepsilon\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)^{-1}.$$

Clearly, for $q<1$ both $\vartheta(q,\varepsilon)$ and $\psi(q,\varepsilon)$ are continuous at $(q,0)$, with $\vartheta(q,0)=1-q$ and $\psi(q,0)=0$. The function $\phi(x)$ is continuous at 0 by construction. It then follows that

$$1-m(q_n,\varepsilon_n)\to 1-m(q,0)=1-q.$$

Next, consider the first and second derivatives of $m(q,\varepsilon)$:

$$m(q,\varepsilon) = 1 - \frac{\ln\left(1+e^{\frac{1}{\varepsilon}\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)}\right)}{\ln\left(1+e^{\frac{1}{\varepsilon}}\right)} \tag{53}$$

$$m_q(q,\varepsilon) = \frac{1}{\varepsilon \ln\left(1+e^{\frac{1}{\varepsilon}}\right)} q^{-\varepsilon} \frac{e^{\frac{1}{\varepsilon}\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)}}{1+e^{\frac{1}{\varepsilon}\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)}} > 0 \tag{54}$$

$$m_{qq}(q,\varepsilon) = \frac{-1}{\varepsilon \ln\left(1+e^{\frac{1}{\varepsilon}}\right)} \left[\varepsilon q^{-(1+\varepsilon)} \frac{e^{\frac{1}{\varepsilon}\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)}}{1+e^{\frac{1}{\varepsilon}\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)}} + \frac{q^{-2\varepsilon}}{\varepsilon} \frac{e^{\frac{1}{\varepsilon}\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)}}{\left(1+e^{\frac{1}{\varepsilon}\left(1-\frac{q^{1-\varepsilon}}{1-\varepsilon}\right)}\right)^2}\right] < 0 \tag{55}$$

Clearly, $m(q,\varepsilon)$ is increasing and concave, and three times continuously differentiable over $(q,\varepsilon) \in [0,\infty) \times (0,\infty)$. The limits of the first and second derivative follow from similar arguments as above.

B.1.5 Proof of Lemma 10

The continuity of $\overline{Q}_u$ is obvious. That $\overline{Q}_0 = (s-\mu_{h,0})/(1-\mu_{h,0})^{1+1/\sigma}$ follows from the definition of $\overline{Q}_u$, and $\overline{Q}_{T_f} = 0$ follows by definition of T_f. Next, after taking derivatives with respect to u we find that $\text{sign}\left[\overline{Q}'_u\right] = \text{sign}\left[F_u\right]$, where:

$$\begin{aligned} F_u \equiv &(s-\mu_{h,u})\left[(1-\mu_{h,0})^{1+1/\sigma} + \int_0^u \rho e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma} dt\right] \\ &-(1-\mu_{h,u})^{1+1/\sigma}\left[s-\mu_{h,0} + \int_0^u \rho e^{\rho t}(s-\mu_{h,t})dt\right], \end{aligned} \tag{56}$$

is continuously differentiable. Taking derivatives once more, we find that $\text{sign}\left[F'_u\right] = \text{sign}\left[G_u\right]$ where:

$$\begin{aligned} G_u \equiv &-\left[(1-\mu_{h,0})^{1+1/\sigma} + \int_0^u \rho e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma} dt\right] \\ &+\left(1+\frac{1}{\sigma}\right)(1-\mu_{h,u})^{1/\sigma}\left[s-\mu_{h,0} + \int_0^u \rho e^{\rho t}(s-\mu_{h,t})dt\right], \end{aligned} \tag{57}$$

is continuously differentiable. Now suppose that $\overline{Q}'_u = 0$. Then $F_u = 0$ and, after substituting (56) into (57):

$$G_u = \left[-\frac{(1-\mu_{h,u})^{1+1/\sigma}}{s-\mu_{h,u}} + \left(1+\frac{1}{\sigma}\right)(1-\mu_{h,u})^{1/\sigma}\right]\left[s-\mu_{h,0} + \int_0^u \rho e^{\rho t}(s-\mu_{h,t})dt\right]. \tag{58}$$

Thus,

R1. *Suppose that $F_u = 0$ for some $u \in [0, T_f)$. Then sign* $[F'_u] = $ *sign* $\left[\frac{s-\mu_{h,u}}{1-\mu_{h,u}} - \frac{\sigma}{1+\sigma}\right]$.

Now note that $G_0 = (1-\mu_{h,0})^{1+1/\sigma}\left(1+\frac{1}{\sigma}\right)\left(-\frac{\sigma}{1+\sigma}+\frac{s-\mu_{h,0}}{1-\mu_{h,0}}\right)$. Thus,

R2. *If* $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} \le \frac{\sigma}{1+\sigma}$, *then* $F_u < 0$ *for all* $u > 0$.

To see this, first note that, from application of the Mean Value Theorem (see, e.g., Theorem 5.11 in Apostol, 1974), it follows that $F_u < 0$ for small u. Indeed, since $F_0 = 0$, $F_u = uF'_v$, for some $v \in (0,u)$. But $\text{sign}\,[F'_v] = \text{sign}\,[G_v]$. Now, since $G_0 \le 0$, G_v is negative as long as u is small enough. But if F_u is negative for small u, it has to stay negative for all u. Otherwise, it would need to cross the x-axis from below at some $u > 0$, which is impossible given Result R1 and the assumption that $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} \le \frac{\sigma}{1+\sigma}$.

R3. *If* $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} > \frac{\sigma}{1+\sigma}$, *then* $F_u > 0$ *for small* u *and* F_u *changes sign only once in the interval* $(0, T_f)$.

$F_u > 0$ for small u follows from applying the same reasoning as in the above paragraph, since when $\frac{s-\mu_{h,0}}{1-\mu_{h,0}} > \frac{\sigma}{1+\sigma}$ we have $G_0 > 0$. Since $F_{T_s} < 0$, then F_u must cross zero at least once between 0 and T_s. The first time F_u crosses zero, it must be from above and hence with a negative slope. Thus, the expression in Result R1 for the sign of F'_u when $F_u = 0$ is negative and this expression is decreasing in u. Therefore, F_u cannot cross zero from below at a later time and hence it only crosses zero once. □

B.1.6 Proof of Lemma 12

To prove that Q_u is continuously differentiable except in T_1 and T_2, we apply the Implicit Function Theorem (see, e.g., Theorem 13.7 in Apostol, 1974). We note that (40) writes $K(u, Q_u) = 0$, where

$$K(u,Q) \equiv \Big((1-\mu_{h,0})\min\{(1-\mu_{h,0})^{1/\sigma}Q, 1\} + \mu_{h,0} - s\Big) + \int_0^u \rho e^{\rho t}\Big((1-\mu_{h,t})\min\{(1-\mu_{h,t})^{1/\sigma}Q, 1\} + \mu_{h,t} - s\Big)\,dt. \tag{59}$$

We consider first the case where $\overline{Q}_u > (1-\mu_{h,0})^{-1/\sigma}$ for some u. Recall that we defined $0 < T_1 < T_2 < T_f$ such that $\overline{Q}_{T_1} = \overline{Q}_{T_2} = (1-\mu_{h,0})^{-1/\sigma}$. Since $(1-\mu_{h,0})^{1/\sigma}Q_u < 1$ for $u < T_1$, we restrict attention to the domain $\{(u,Q) \in \mathbb{R}^2_+ : u < T_1 \text{ and } Q < (1-\mu_{h,0})^{-1/\sigma}\}$. In this domain, equation (59) can be written

$$K(u,Q) = \Big((1-\mu_{h,0})^{1+1/\sigma}Q + \mu_{h,0} - s\Big) + \int_0^u \rho e^{\rho t}\Big((1-\mu_{h,t})^{1+1/\sigma}Q + \mu_{h,t} - s\Big)\,dt.$$

To apply the Implicit Function Theorem, we need to show that $K(u,Q)$ is continuously differentiable. To see this, first note that the partial derivative of $K(u,Q)$ with respect to u is

$$\frac{\partial K}{\partial u} = \rho e^{\rho u}\Big((1-\mu_{h,u})^{1+1/\sigma}Q + \mu_{h,u} - s\Big)$$

and is continuous. The partial derivative with respect to Q is

$$\frac{\partial K}{\partial Q} = (1-\mu_{h,0})^{1+1/\sigma} + \int_0^u \rho e^{\rho t}\left(1-\mu_{h,t}\right)^{1+1/\sigma}\, dt$$

and is continuous and strictly positive. Therefore, we can apply the Implicit Function Theorem and state that

$$Q_u' = -\frac{\partial K/\partial u}{\partial K/\partial Q} = \frac{\rho e^{\rho u}\left(s-\mu_{h,u}-(1-\mu_{h,u})^{1+1/\sigma}Q\right)}{(1-\mu_{h,0})^{1+1/\sigma}+\int_0^u \rho e^{\rho t}\left(1-\mu_{h,t}\right)^{1+1/\sigma}\, dt}.$$

The same reasoning and expression for Q_u' obtain for $u > T_2$, as well as for all u in the case where $\overline{Q}_u \leq (1-\mu_{h,0})^{-1/\sigma}$ for all u.

The second domain to consider is $\{(u,Q) \in \mathbb{R}_+^2 \,:\, T_1 < u < T_2 \text{ and } Q > (1-\mu_{h,0})^{-1/\sigma}\}$. In this domain, equation (59) can be written, using the definition of $\Psi(Q)$,

$$K(u,Q) = (1-s) + \int_0^{\Psi(Q)} \rho e^{\rho t}(1-s)\, dt + \int_{\Psi(Q)}^u \rho e^{\rho t}\left((1-\mu_{h,t})^{1+1/\sigma}Q+\mu_{h,t}-s\right)\, dt.$$

The partial derivative of $K(u,Q)$ with respect to u is

$$\frac{\partial K}{\partial u} = \rho e^{\rho u}\left((1-\mu_{h,u})^{1+1/\sigma}Q+\mu_{h,u}-s\right)$$

and is continuous. Noting that $\Psi(Q)$ is differentiable, the partial derivative with respect to Q is

$$\frac{\partial K}{\partial Q} = \int_{\Psi(Q)}^u \rho e^{\rho t}\left(1-\mu_{h,t}\right)^{1+1/\sigma}\, dt$$

and is continuous since $\Psi(Q)$ is continuous. Moreover, since $\Psi(Q) < u$ in its domain, then $\partial K/\partial Q > 0$. Therefore, we can apply the Implicit Function Theorem and

$$Q_u' = -\frac{\partial K/\partial u}{\partial K/\partial Q} = \frac{\rho e^{\rho u}\left(s-\mu_{h,u}-(1-\mu_{h,u})^{1+1/\sigma}Q\right)}{\int_{\psi_u}^u \rho e^{\rho t}\left(1-\mu_{h,t}\right)^{1+1/\sigma}\, dt},$$

where we used that $\psi_u \equiv \Psi(Q_u)$. □

B.1.7 Proof of Lemma 13

For $u \in (T_1, T_2)$, we have $Q_u \neq \overline{Q}_u$ and therefore and therefore $\Psi(Q_u) = \psi_u > 0$. By definition of ψ_u, we also have

$$Q_u = (1-\mu_{h,\psi_u})^{-1/\sigma}. \tag{60}$$

Replacing into equation (42) for Q'_u of Lemma 12 , one obtains that:

$$\text{sign}\left[Q'_u\right] = \text{sign}\left[X_u\right] \text{ where } X_u \equiv s - \mu_{h,u} - (1-\mu_{h,u})\left(\frac{1-\mu_{h,u}}{1-\mu_{h,\psi_u}}\right)^{1/\sigma}.$$

As noted above, Q_u and thus X_u changes sign at least once over (T_1, T_2). Now, for any u_0 such that $X_{u_0} = 0$, we have $Q'_{u_0} = 0$ and, given (60), $\psi'_{u_0} = 0$. Taking the derivative of X_u at such u_0, and using $X_{u_0} = 0$, leads:

$$\text{sign}\left[X'_{u_0}\right] = \text{sign}\left[-1 + \left(1+\frac{1}{\sigma}\right)\left(\frac{1-\mu_{h,u_0}}{1-\mu_{h,\psi_{u_0}}}\right)^{1/\sigma}\right] = \text{sign}\left[Y_{u_0}\right],$$
$$\text{where } Y_u \equiv -1 + \left(1+\frac{1}{\sigma}\right)\frac{s-\mu_{h,u}}{1-\mu_{h,u}},$$

where the second equality follows by using $X_{u_0} = 0$. Now take u_0 to be the *first* time X_u changes sign during (T_1, T_2). Since $X_{u_0} = 0$, X_u strictly positive to the left of u_0, and X_u strictly negative to the right of u_0, we must have that $X'_{u_0} \leq 0$. Suppose, then, that X_u changes sign once more during (T_1, T_2) at some time u_1. The same reasoning as before implies that, at u_1, $X'_{u_1} \geq 0$. But this is impossible since Y_u is strictly decreasing. □

B.1.8 Proof of Lemma 14

Proof of the limit of $T_f(\rho)$, in equation (43). Recall that $T_f(\rho)$ solves $\mathbb{E}\left[\mu_{h,\tau_u}\right] = s$ and that $T_f(\rho) \geq T_s$. Note also that $\Pr(\tau_u \leq t) = \min\{e^{-\rho(u-t)}, 1\}$. Therefore, a increases in u and ρ induce first-order stochastic dominance shift. Since $\mu_{h,t}$ is increasing, it follows that $\mathbb{E}\left[\mu_{h,\tau_u}\right]$ is strictly increasing in u and ρ, and therefore that $T_f(\rho)$ is strictly decreasing in ρ. Thus, $T_f(\rho)$ admits a limit $T_f(\infty)$ as $\rho \to \infty$. Since $T_f(\rho)$ is greater than the limit, and since $\mathbb{E}\left[\mu_{h,\tau_u}\right]$ is increasing in u, we have: $\mathbb{E}\left[\mu_{h,\tau_u}\right] \leq s$ for $u = T_f(\infty)$. Taking the limit as $\rho \to \infty$ we find that $\mu_{h,T_f(\infty)} \leq s$ so that $T_f(\infty) \leq T_s$. Since $T_f(\rho) \geq T_s$, the result follows. □

Proof of the first–order expansion, in equation (44). Let

$$f(t,\rho) \equiv (1-\mu_{h,t})\min\left\{(1-\mu_{h,t})^{1/\sigma}Q_u(\rho), 1\right\} + \mu_{h,t} - s. \tag{61}$$

By its definition, $Q_u(\rho)$ solves: $\mathbb{E}\left[f(\tau_u,\rho)\right]$. Note that, for each ρ, $f(t,\rho)$ is continuously differentiable with respect to t except at $t = \psi_u(\rho)$ such that $(1-\mu_{h\psi(\rho)})^{1/\sigma}Q_u(\rho) = 1$. Thus, we can integrate the above by part and obtain:

$$0 = \int_0^u \rho e^{-\rho(u-t)} f(t,\rho)\,dt = f(u,\rho) - \int_0^u e^{-\rho(u-t)} f_t(t,\rho)\,dt, \tag{62}$$

where $f_t(t,\rho)$ denotes the partial derivative of $f(t,\rho)$ with respect to t. Now consider a sequence of ρ going to infinity and the associated sequence of $Q_u(\rho)$. Because $Q_u(\rho)$ is bounded above by $(1-\mu_{h,u})^{-1/\sigma}$, this sequence has at least one accumulation point $Q_u(\infty)$. Taking the limit in (62) along a subsequence converging to this accumulation point, we obtain that $Q_u(\infty)$ solves the equation

$$(1-\mu_{h,u})\min\{(1-\mu_{h,u})^{1/\sigma}Q_u(\infty),1\}+\mu_{h,u}-s=0.$$

whose unique solution is $Q_u(\infty)=(s-\mu_{h,u})/(1-\mu_{h,u})^{1+1/\sigma}$. Thus $Q_u(\rho)$ has a unique accumulation point, and therefore converges towards it. To obtain the asymptotic expansion, we proceed with an additional integration by part in equation (62):

$$\begin{aligned}0=&f(u,\rho)+\frac{1}{\rho}f_t(0,\rho)e^{-\rho u}+\frac{1}{\rho}\int_0^u f_{tt}(t,\rho)e^{-\rho(u-t)}\,dt\\&+\frac{1}{\rho}e^{-\rho(u-\psi_u(\rho))}\left[f_t(\psi_u(\rho)^+,\rho)-f_t(\psi_u(\rho)^-,\rho)\right].\end{aligned}$$

where the term on the second line arises because f_t is discontinuous at $\psi_u(\rho)$. Given that $Q_u(\rho)$ converges and is therefore bounded, the third, fourth and fifth terms on the first line are $o(1/\rho)$. For the second line we note that, since $Q_u(\rho)$ converges to $Q_u(\infty)$, $\psi_u(\rho)$ converges to $\psi_u(\infty)$ such that $(1-\mu_{h\psi_u(\infty)})^{1/\sigma}Q_u(\infty)=1$. In particular, one easily verifies that $\psi_u(\infty)<u$. Therefore $e^{-\rho(u-\psi_u(\rho))}$ goes to zero as $\rho\to\infty$, so the term on the second line is also $o(1/\rho)$. Taken together, this gives:

$$0=f(u,\rho)-\frac{1}{\rho}f_t(u,\rho)+o\left(\frac{1}{\rho}\right). \tag{63}$$

Equation (44) obtains after substituting in the expressions for $f(u,\rho)$ and $f_t(u,\rho)$, using that $\mu'_{h,t}=\gamma(1-\mu_{h,t})$. □

Proof of the convergence of the argmax, in equation (45). First one easily verify that $Q_u(\infty)$ is hump–shaped (strictly decreasing) if and only if $Q_u(\rho)$ is hump–shaped (strictly decreasing). So if $\frac{s-\mu_{h,0}}{1-\mu_{h,0}}\leq\frac{\sigma}{1+\sigma}$, then both $Q_u(\rho)$ and $Q_u(\infty)$ are strictly decreasing, achieve their maximum at $u=0$, and the result follows. Otherwise, if $\frac{s-\mu_{h,0}}{1-\mu_{h,0}}>\frac{\sigma}{1+\sigma}$, consider any sequence of ρ going to infinity and the associated sequence of $T_\psi(\rho)$. Since $T_\psi(\rho)<T_f(\rho)<T_f(0)$, the sequence of $T_{\psi(\rho)}$ is bounded and, therefore, it has at least one accumulation point, $T_{\psi(\infty)}$. At each point along the sequence, $T_\psi(\rho)$ maximizes $Q_u(\rho)$. Using equation (42) to write the corresponding first–order condition, $Q'_{T_\psi(\rho)}=0$, we obtain after rearranging that

$$Q_{T_\psi(\rho)}(\rho)=\frac{s-\mu_{h,T_\psi(\rho)}}{1-\mu_{h,T_\psi(\rho)}}=Q_{T_\psi(\rho)}(\infty)\geq Q_{T_\psi^*}(\rho).$$

where T_ψ^* denotes the unique maximizer of $Q_u(\infty)$. Letting ρ go to infinity on both sides of the equation, we find

$$Q_{T_\psi(\infty)}(\infty)\geq Q_{T_\psi^*}(\infty).$$

But since T_ψ^* is the unique maximizer of $Q_u(\infty)$, $T_\psi(\infty) = T_\psi^*$. Therefore, $T_\psi(\rho)$ has a unique accumulation point, and converges towards it. □

B.1.9 Proof of Lemma 15

Given that $\Delta_u = (1-\mu_{h,u})^{1/\sigma} Q_u$, we have

$$\Delta_u' = -\frac{\gamma}{\sigma}(1-\mu_{h,u})^{1/\sigma} Q_u + (1-\mu_{h,u})^{1/\sigma} Q_u'.$$

Using the formula (42) for Q_u', in Lemma 12, we obtain:

$$\begin{aligned} \text{sign}\left[\Delta_u'\right] &= \text{sign}\left[-\frac{\gamma}{\sigma}Q_u + Q_u'\right] \\ = \text{sign}\Big[&-\frac{\gamma}{\sigma}Q_u\left(\mathbb{I}_{\{\psi_u=0\}}(1-\mu_{h,0})^{1+1/\sigma} + \int_{\psi_u}^{u} \rho e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma}\,dt\right) \\ &+ \rho e^{\rho u}\left(s-\mu_{h,u}-(1-\mu_{h,u})^{1+1/\sigma}Q_u\right)\Big]. \end{aligned} \tag{64}$$

We first show:

R 4. $\Delta_u' < 0$ *for u close to zero.*

To show this result, first note that when u is close to zero, $\psi_u = 0$ and, by Lemma 9, $Q_0 = \overline{Q}_0 = \frac{s-\mu_{h,0}}{1-\mu_{h,0})^{1+1/\sigma}}$. Plugging in into (64), one obtains

$$\text{sign}\left[\Delta_0'\right] = -\frac{\gamma}{\sigma}(s-\mu_{h,0}) < 0. \tag{65}$$

Since $\psi_u = 0$ for u close to zero, the results follows by continuity. Next, we show:

R 5. *Suppose $\Delta_{u_0}' = 0$ for some $u_0 \in (0, T_f]$. Then, Δ_u is strictly decreasing at u_0.*

For this we first manipulate (64) as follows:

$$\begin{aligned}
&\text{sign}\left[\Delta_u'\right]\\
=&\text{sign}\Bigg[-\frac{\gamma}{\sigma}\frac{\Delta_u}{(1-\mu_{h,u})^{1/\sigma}}\left(\mathbb{I}_{\{\psi_u=0\}}(1-\mu_{h,0})^{1+1/\sigma}+\int_{\psi_u}^{u}\rho e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma}\,dt\right)+\\
&\qquad\qquad\qquad\qquad\qquad\qquad\rho e^{\rho u}\left(s-\mu_{h,u}-(1-\mu_{h,u})\Delta_u\right)\Bigg]\\
=&\text{sign}\Bigg[-\frac{\gamma}{\sigma}\Delta_u\left(\mathbb{I}_{\{\psi_u=0\}}e^{-\rho u}\left(\frac{1-\mu_{h,0}}{1-\mu_{h,u}}\right)^{1+1/\sigma}+\int_{\psi_u}^{u}\rho e^{-\rho(u-t)}\left(\frac{1-\mu_{h,t}}{1-\mu_{h,u}}\right)^{1+1/\sigma}dt\right)\\
&\qquad\qquad\qquad\qquad\qquad\qquad+\rho\left(\frac{s-\mu_{h,u}}{1-\mu_{h,u}}-\Delta_u\right)\Bigg]\\
=&\text{sign}\left[-\frac{\gamma}{\sigma}\Delta_u\left(\mathbb{I}_{\{\psi_u=0\}}e^{\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]u}+\int_{\psi_u}^{u}\rho e^{\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right](u-t)}\,dt\right)+\rho\left(1-(1-s)e^{\gamma u}-\Delta_u\right)\right]\\
=&\text{sign}\left[-\frac{\gamma}{\sigma}\Delta_u\left(\mathbb{I}_{\{\psi_u=0\}}e^{\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]u}+\int_{0}^{u-\psi_u}\rho e^{\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]t}\,dt\right)+\rho\left(1-(1-s)e^{\gamma u}-\Delta_u\right)\right]
\end{aligned}$$

and where we obtain the first equality after substituting in the expression for Q_u; the second equality after dividing by $(1-\mu_{h,u})e^{\rho u}$; the third equality by using the functional form $1-\mu_{h,t}=(1-\mu_{h,0})e^{-\gamma t}$; and the fourth equality by changing variable ($x=u-t$) in the integral. Now suppose $\Delta_u'=0$ at some u_0. From the above we have:

$$H_{u_0}\equiv-\frac{\gamma}{\sigma}\Delta_{u_0}\left(\mathbb{I}_{\{\psi_{u_0}=0\}}e^{\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]u_0}+\int_{0}^{u_0-\psi_{u_0}}\rho e^{\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]t}\,dt\right)+\rho\left(1-(1-s)e^{\gamma u_0}-\Delta_{u_0}\right)=0.$$

If $(1-\mu_{h,0})^{1/\sigma}Q_{u_0}<1$ then $\psi_{u_0}=0$ and $\psi_{u_0}'=0$. Together with the fact that $\Delta_{u_0}'=0$, this implies that

$$H_{u_0}'=-\frac{\gamma}{\sigma}\Delta_{u_0}\gamma\left(1+\frac{1}{\sigma}\right)e^{\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]u_0}-\rho(1-s)\gamma e^{\gamma u_0}<0.$$

If $(1-\mu_{h,0})^{1/\sigma}Q_{u_0}=1$, then $\psi_{u_0}=0$ and the left-derivative $\psi_{u_0^-}'=0$, so the same calculation implies that $H_{u_0^-}'<0$. If $(1-\mu_{h,0})^{1/\sigma}Q_{u_0}>1$ we first note that, around u_0,

$$Q_u=\left(1-\mu_{h,\psi_u}\right)^{-1/\sigma}\Rightarrow\Delta_u=\left(\frac{1-\mu_{h,\psi_u}}{1-\mu_{h,u}}\right)^{1/\sigma}=e^{-\gamma\frac{\psi_u-u}{\sigma}}.$$

So if $\Delta_{u_0}'=0$, we must have that $\psi_{u_0}'=1$. Plugging this back into H_{u_0}' we obtain that $H_{u_0}'=-\rho(1-s)\gamma e^{\gamma u_0}<0$. Lastly, if $(1-\mu_{h,0})^{1/\sigma}Q_{u_0}=1$, then the same calculation leads to $\psi_{u_0^+}=1$ and so $H_{u_0^+}<0$. In all cases, we find that H_{u_0} has strictly negative left- and right-derivatives when $H_{u_0}=0$. Thus, whenever it is equal to zero, Δ_u' is strictly decreasing. With Result R5 in mind, we then obtain:

R6. Δ_u' *cannot change sign over* $(0,T_f]$.

Suppose it did and let u_0 be the first time in $(0, T_f]$ where Δ'_u changes sign. Because Δ'_u is continuous, we have $\Delta'_{u_0} = 0$. But recall that $\Delta'_u < 0$ for $u \simeq 0$, implying that at $u = u_0$, Δ'_u crosses the x-axis from below and is therefore increasing, contradicting Result R5. $\square$

B.1.10 Proof of Lemma 16

With known preferences:

$$J^\star(s) = \int_0^{+\infty} \mathbb{I}_{\{u<T_s\}} e^{-ru} \left(1 - \frac{1-s}{1-\mu_{h,0}} e^{\gamma u}\right)^\sigma du.$$

Since, by definition $e^{\gamma T_s} \frac{1-s}{1-\mu_{h,0}} = 1$, we have that $T_s \to \infty$ when s goes to 1, and the integrand of $J^\star(s)$ converges pointwise towards e^{-ru}. Moreover, the integrand is bounded by e^{-ru}. Therefore, by an application of the Dominated Convergence Theorem, $J^\star(s)$ goes to $\int_0^{+\infty} e^{-ru}\, du = 1/r$ when $s \to 1$.

With preference uncertainty, for $u > 0$, we note that $Q_u(s)$ is an increasing function of s and is bounded above by $(1-\mu_{h,u})^{-1/\sigma}$. Letting $s \to 1$ in the market clearing condition (40) then shows that $Q_u \to (1-\mu_{h,u})^{-1/\sigma} > 1$. Using that $T_f > T_s$ goes to $+\infty$ when $s \to 1$, we obtain that the integrand of $J(s)$ goes to e^{-ru}. Moreover, the integrand is bounded by e^{-ru}. Therefore, by dominated convergence, $J(s)$ goes to $1/r$.

B.1.11 Proof of Lemma 17

In the market with continuous updating, we can compute:

$$J^{\star\prime}(s) = \int_0^{T_s} e^{-ru} \frac{\sigma e^{\gamma u}}{1-\mu_{h,0}} \left(1 - \frac{1-s}{1-\mu_{h,0}} e^{\gamma u}\right)^{\sigma-1} du + \frac{\partial T_s}{\partial s}\left(1 - \frac{1-s}{1-\mu_{h,0}} e^{\gamma T_s}\right)^\sigma. \tag{66}$$

The second term is equal to 0 since $e^{\gamma T_s} \frac{1-s}{1-\mu_{h,0}} = 1$. After making the change of variable $z = T_s - u$, keeping in mind that $e^{\gamma T_s} \frac{1-s}{1-\mu_{h,0}} = 1$, we obtain:

$$J^{\star\prime}(s) = \int_0^{T_s} e^{(r-\gamma)(z-T_s)} \frac{\sigma}{1-\mu_{h,0}} \left(1 - e^{-\gamma z}\right)^{\sigma-1} dz. \tag{67}$$

We then compute an approximation of $J^{\star\prime}(s)$ when $s \to 1$.

<u>When $r > \gamma$.</u> In this case we write:

$$J^{\star\prime}(s) = \int_0^{T_s/2} e^{(r-\gamma)(z-T_s)} \frac{\sigma}{1-\mu_{h,0}} \left(1 - e^{-\gamma z}\right)^{\sigma-1} dz + \int_{T_s/2}^{T_s} e^{(r-\gamma)(z-T_s)} \frac{\sigma}{1-\mu_{h,0}} \left(1 - e^{-\gamma z}\right)^{\sigma-1} dz.$$

The first term is less than $e^{-(r-\gamma)T_s/2} \frac{\sigma}{1-\mu_{h,0}} \int_0^{T_s/2} [1 - e^{-\gamma z}]\, dz$. The integrand goes to 1 as z goes to infinity and so, by Cesàro summation, the integral is equivalent to $T_s/2$, which is dominated by $e^{-(r-\gamma)T_s/2}$ as $s \to 1$ and $T_s \to \infty$. Thus, the first term converges to zero as $s \to 1$. The second term can be written:

$$\sigma \int_0^{\infty} \mathbb{I}_{\{u \le T_s/2\}} e^{-(r-\gamma)u} \left(1 - \frac{1-s}{1-\mu_{h,0}} e^{\gamma u}\right)^{\sigma-1} du.$$

Since T_s goes to infinity when s goes to 1, the integrand goes to , and is bounded by, $e^{-(r-\gamma)u}(1-\sqrt{\frac{1-s}{1-\mu_{h,0}}})^{\sigma-1}$. Therefore, by dominated convergence, $J^{\star\prime}(s)$ goes to $\frac{\sigma}{(1-\mu_{h,0})(r-\gamma)}$.

When $r = \gamma$. Then we have:

$$J^{\star\prime}(s) = \sigma \int_0^{T_s} \left(1 - e^{-\gamma z}\right)^{\sigma-1} dz.$$

The integrand goes to 1 when T_s goes to infinity. Thus, the Cesàro mean $I'(s)/T_s$ converges to σ, i.e.:

$$J^{\star\prime}(s) \sim \sigma T_s = -\frac{\sigma}{\gamma} \log\left(\frac{1-s}{1-\mu_{h,0}}\right).$$

When $r < \gamma$. In that case:

$$J^{\star\prime}(s) = \sigma e^{(\gamma-r)T_s} \int_0^{+\infty} \mathbb{I}_{\{z<T_s\}} e^{-(\gamma-r)z} \left(1 - e^{-\gamma z}\right)^{\sigma-1} dz,$$

The integrand in the second line goes to, and is bounded by, $e^{-(\gamma-r)z}(1-e^{-\gamma z})^{\sigma-1}$, which in integrable. Therefore, by dominated convergence, the integral goes to $\int_0^{+\infty} e^{-(\gamma-r)z} \left(1 - e^{-\gamma z}\right)^{\sigma-1} dz$ when s goes to 1. Finally, using that $e^{-\gamma T_s} = \frac{1-s}{1-\mu_{h,0}}$, we obtain:

$$J^{\star\prime}(s) \sim \sigma \left(\frac{1-\mu_{h,0}}{1-s}\right)^{1-r/\gamma} \int_0^{+\infty} e^{-(\gamma-r)z} \left(1 - e^{-\gamma z}\right)^{\sigma-1} dz.$$

B.1.12 Proof of Lemma 18

Throughout all the proof and the intermediate results therein, we work under the maintained assumption

$$\gamma + \gamma/\sigma - \rho > 0 \Longleftrightarrow \gamma + \sigma(\gamma - \rho) > 0, \tag{68}$$

which is without loss of generality since we want to compare prices when σ is close to zero. We start by differentiating $J(s)$:

$$J'(s) = \frac{\partial T_f}{\partial s} e^{-rT_f} e^{-\gamma T_f} Q^{\sigma}_{T_f^-} + \int_0^{T_f} e^{-ru} e^{-\gamma u} \frac{\partial Q_u^{\sigma}}{\partial s} du > \int_{T_1}^{T_2} e^{-ru} e^{-\gamma u} \frac{\partial Q_u^{\sigma}}{\partial s} du,$$

where the inequality follows from the following facts: the first term is zero since $Q_{T_f^-} = 0$; the integrand in the second term is positive since Q_u is increasing in s by equation (40); and $0 < T_1 < T_2 < T_f$ are defined as in the paragraph following Lemma 11, as follows. We consider that s is close to 1 so that $Q_u > 1$ for some u. Then, $T_1 < T_2$ are defined as the two solutions of $Q_{T_1} = Q_{T_2} = 1$. Note that T_1 and T_2 are also the two solutions of $\overline{Q}_{T_1} = \overline{Q}_{T_2}$. Because both Q_u and $\overline{Q}_u$ are hump shaped, we know that Q_u and $\overline{Q}_u$ are strictly greater than one for $u \in (T_1, T_2)$, and less than one otherwise. For $u \in (T_1, T_2)$, we can define $\psi_u > 0$ as in the paragraph following Lemma 11: $Q_u = (1 - \mu_{h\psi_u})^{-1/\sigma}$. By construction, $\psi_u \in (0, u)$, and, as shown in Section B.1.13:

$$\frac{\partial \psi_u}{\partial s} = \frac{\gamma + \sigma(\gamma - \rho)}{\gamma\rho} \frac{(1 - e^{-\rho u})\, e^{\gamma u}}{e^{-(\rho-\gamma)(u-\psi_u)} - e^{-(\gamma/\sigma)(u-\psi_u)}}. \tag{69}$$

Plugging $Q_u^\sigma = (1 - \mu_{h\psi_u})^{-1} = e^{\gamma\psi_u}$ in the expression of $J'(s)$, we obtain:

$$J'(s) > \frac{\gamma + \sigma(\gamma - \rho)}{\rho} \int_{T_1}^{T_2} e^{-ru} \frac{(1 - e^{-\rho u})\, e^{\gamma\psi_u}}{e^{-(\rho-\gamma)(u-\psi_u)} - e^{-(\gamma/\sigma)(u-\psi_u)}}\, du. \tag{70}$$

When $r > \gamma$. For this case fix some $\overline{u} > 0$ and pick s close enough to one so that that $Q_{\overline{u}} > 1$. Such s exists since, as argued earlier in Section B.1.10, for all $u > 0$, $Q_u \to (1 - \mu_{h,u})^{-1/\sigma}$ as $s \to 1$. Since the integrand in (70) is strictly positive, we have:

$$\begin{aligned} J'(s) &> \frac{\gamma + \sigma(\gamma - \rho)}{\rho} \int_0^{\overline{u}} \mathbb{I}_{\{u > T_1\}} e^{-ru} \frac{(1 - e^{-\rho u})\, e^{\gamma\psi_u}}{e^{-(\rho-\gamma)(u-\psi_u)} - e^{-(\gamma/\sigma)(u-\psi_u)}}\, du \\ &> \frac{\gamma + \sigma(\gamma - \rho)}{\rho} \frac{1}{e^{|\rho-\gamma|(\overline{u}-\psi_{\overline{u}})} - e^{-(\gamma/\sigma)(\overline{u}-\psi_{\overline{u}})}} \int_0^{\overline{u}} \mathbb{I}_{\{u > T_1\}} e^{-ru} \left(1 - e^{-\rho u}\right) e^{\gamma\psi_u}\, du. \end{aligned}$$

where the second line follows from the fact, proven is Section B.1.13, that $u - \psi_u$ is strictly increasing in u when $\psi_u > 0$. In Section B.1.13 we also prove that $T_1 \to 0$ and that, for all $u > 0$, $\psi_u \to u$ when s goes to 1. Therefore, in the above equation, the integral remains bounded away from zero, and the whole expression goes to infinity.

When $r \leq \gamma$. In this case we make the change of variable $z \equiv T_s - u$ in equation (70) and we use that $e^{-\gamma T_s} = \frac{1-s}{1-\mu_{h,0}}$:

$$\begin{aligned} J'(s) &> \frac{\gamma + \sigma(\gamma - \rho)}{\rho} \int_{T_s - T_2}^{T_s - T_1} \left(\frac{1 - s}{1 - \mu_{h,0}}\right)^{\frac{r}{\gamma}} e^{rz} \frac{\left(1 - e^{-\rho(T_s - z)}\right) e^{\gamma\psi_{T_s - z}}}{e^{-(\rho-\gamma)(T_s - z - \psi_{T_s - z})} - e^{-(\gamma/\sigma)(T_s - z - \psi_{T_s - z})}}\, dz \\ &> \frac{\gamma + \sigma(\gamma - \rho)}{\rho} \int_0^{+\infty} \mathbb{I}_{\{\max\{T_s - T_2, 0\} < z < T_s - T_1\}} \left(\frac{1 - s}{1 - \mu_{h,0}}\right)^{\frac{r}{\gamma}} \\ &\qquad e^{rz} \frac{\left(1 - e^{-\rho(T_s - z)}\right) e^{\gamma\psi_{T_s - z}}}{e^{-(\rho-\gamma)(T_s - z - \psi_{T_s - z})} - e^{-(\gamma/\sigma)(T_s - z - \psi_{T_s - z})}}\, dz. \end{aligned}$$

where the second line follows from the addition of the max operator in the indicator variable and the fact that the integrand is strictly positive. We show in Section B.1.13 that, if $\psi_{T_s - z} > 0$, then:

$$e^{\gamma \psi_{T_s-z}} > \begin{cases} \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}} \left(\frac{1-s}{1-\mu_{h,0}}\right)^{-1} e^{-\gamma z} & \text{if } \rho \neq \gamma, \\ e^{-(1+\sigma)} \left(\frac{1-s}{1-\mu_{h,0}}\right)^{-1} e^{-\gamma z} & \text{if } \rho = \gamma, \end{cases} \tag{71}$$

and:

$$\left(e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})} - e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}\right)^{-1} > \frac{\min\{\gamma, \rho\}}{\gamma + \sigma(\gamma-\rho)}. \tag{72}$$

When $\gamma \neq \rho$, we obtain:

$$J'(s) > \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}} \min\left\{\gamma/\rho, 1\right\} \left(\frac{1-s}{1-\mu_{h,0}}\right)^{-1+\frac{r}{\gamma}} \\ \times \int_0^{+\infty} \mathbb{I}_{\{\max\{T_s - T_2, 0\} < z < T_s - T_1\}} e^{-(\gamma-r)z} \left(1 - e^{-\rho(T_s - z)}\right) dz. \tag{73}$$

Consider first the case $\gamma < r$. In Section B.1.13 we show that $T_s - T_2 < 0$ when s is close to 1 and that T_1 goes to 0 when s goes to 1. Since T_s goes to infinity, these facts imply that the integrand goes to, and is bounded above by, $e^{-(\gamma-r)z}$ when $s \to 1$. Therefore, by dominated convergence, the integral goes to $1/(\gamma - r)$. A similar computation obtains when $\gamma = \rho$.

Consider now the case $\gamma = r$. When $\gamma \neq \rho$, equation (73) rewrites:

$$\begin{aligned} J'(s) &> \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}} \min\{\gamma/\rho, 1\} \int_{\max\{T_s-T_2,0\}}^{T_s-T_1} \left(1 - e^{-\rho(T_s-z)}\right) dz \\ &= \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}} \min\{\gamma/\rho, 1\} \left(T_s - T_1 - \max\{T_s - T_2, 0\} - \frac{e^{-\rho T_1} - e^{-\rho \min\{T_2, T_s\}}}{\rho}\right). \end{aligned}$$

Since $T_s - T_2 < 0$ and $T_1 \to 0$ when s goes to 1, the last term in large parenthesis is equivalent to $T_s = \log((1-s)^{-1})/\gamma$ when s goes to 1. A similar computation obtains when $\gamma = \rho$.

B.1.13 Intermediate results for the proofs of Lemma 16, 17 and 18

Derivative of the ψ_u function when $\psi_u > 0$. When $\psi_u > 0$, time–τ_u low–valuation investors hold $q_{\tau_u,u} = 1$ if $\tau_u < \psi_u$, and $q_{\tau_u,u} = (1-\mu_{h,\tau_u})^{1/\sigma}(1-\mu_{h\psi_u})^{-1/\sigma}$ if $\tau_u > \psi_u$. The market clearing condition (40) rewrites:

$$\begin{aligned} &1 - \mu_{h,0} + \int_0^{\psi_u} \rho e^{\rho t}(1-\mu_{h,t})\, dt + \int_{\psi_u}^{u} \rho e^{\rho t} (1 - (1-\mu_{h,0})\mu_{h,t})^{1+1/\sigma} (1-\mu_{h\psi_u})^{-1/\sigma}\, dt \\ =&s - \mu_{h,0} + \int_0^u \rho e^{\rho t}(s - \mu_{h,t})\, dt. \end{aligned} \tag{74}$$

We differentiate this equation with respect to s:

$$\frac{\partial \psi_u}{\partial s}\frac{\gamma}{\sigma}\int_{\psi_u}^{u} e^{\rho t}(1-\mu_{h,t})^{1+1/\sigma}(1-\mu_{h\psi_u})^{-1/\sigma}\,dt = \int_0^u e^{\rho t}\,dt.$$

After computing the integrals and rearranging the terms we obtain equation (69).

Limits of T_1 and T_2 when $s \to 1$. For any $u > 0$, when s is close enough to 1 we have $Q_u > 1$ and thus $T_1 < u < T_2$. Therefore $T_1 \to 0$ and $T_2 \to \infty$, when $s \to 1$. To obtain that $T_2 > T_s$ when s is close to 1, it suffices to show that $\overline{Q}_{T_s} > 1$ for s close to 1. After computing the integrals in equation (41) and using that $e^{-\gamma T_s} = \frac{1-s}{1-\mu_{h,0}}$, we obtain:

$$\overline{Q}_{T_s} = \frac{N_s}{D_s},$$

where

$$N_s = \begin{cases} (1-s)\frac{\gamma}{\rho-\gamma} + \frac{\gamma}{\gamma-\rho}(1-\mu_{h,0})\left(\frac{1-s}{1-\mu_{h,0}}\right)^{\rho/\gamma} & \text{if } \rho \neq \gamma \\ (1-s)\log\left(\frac{1-\mu_{h,0}}{1-s}\right) & \text{if } \rho = \gamma \end{cases}$$

$$D_s = \frac{\sigma(1-\mu_{h,0})^{1+1/\sigma}}{\gamma+\sigma(\gamma-\rho)}\left\{\gamma\left(1+\frac{1}{\sigma}\right)\left(\frac{1-s}{1-\mu_{h,0}}\right)^{\rho/\gamma} - \rho\left(\frac{1-s}{1-\mu_{h,0}}\right)^{1+1/\sigma}\right\}.$$

When $\gamma \leq \rho$, $\overline{Q}_{T_s}$ goes to infinity when s goes to 1. When $\gamma > \rho$, $\overline{Q}_{T_s}$ goes to $\frac{\gamma+\sigma(\gamma-\rho)}{\sigma(\gamma-\rho)(1-\mu_{h,0})^{1+1/\sigma}} > 1$.

Proof that $u - \psi_u$ is strictly increasing in u when $\psi_u > 0$. Rearranging (74), we obtain:

$$\frac{1-s}{1-\mu_{h,0}}e^{\rho u} = \int_{\psi_u}^{u} \rho e^{(\rho-\gamma)t}\,dt - e^{\frac{\gamma}{\sigma}\psi_u}\int_{\psi_u}^{u} \rho e^{\left[\rho-\gamma\left(1+\frac{1}{\sigma}\right)\right]t}\,dt.$$

When $\rho \neq \gamma$, calculating the integrals and reorganizing terms leads to

$$\frac{1-s}{1-\mu_{h,0}}\frac{e^{\gamma u}}{\rho} = \left(\frac{1}{\rho-\gamma} + \frac{1}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho}\right)\left(1-e^{-(\rho-\gamma)(u-\psi_u)}\right) - \frac{1}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho}\left(1-e^{-\frac{\gamma}{\sigma}(u-\psi_u)}\right) \quad (75)$$

Taking the derivative of the right-hand side with respect to $u - \psi_u$ we easily obtain that it is strictly increasing in $u - \psi_u$, given our parameter restriction that $\gamma > \sigma(\gamma - \rho)$. Since the right-hand side is strictly increasing in u, then $u - \psi_u$ is a strictly increasing function of u. When $\rho = \gamma$, the left-hand side stays the same and the right-hand side becomes

$$u - \psi_u + \frac{\sigma}{\gamma}\left(e^{-\frac{\gamma}{\sigma}(u-\psi_u)} - 1\right)$$

which is strictly increasing in $u - \psi_u$ as well, implying that $u - \psi_u$ is a strictly increasing function of u.

Proof that $\psi_u \to u$ when $s \to 1$. As noted earlier in Section B.1.10, for any u, $Q_u \to (1-\mu_{h,u})^{-1/\sigma}$ as $s \to 1$. Together with the defining equation of ψ_u, $Q_u = (1-\mu_{h\psi_u})^{1/\sigma}$, this implies that $\psi_u \to u$ as $s \to 1$.

Proof of equation (71). When $\gamma \neq \rho$, we make the change of variable $z \equiv T_s - u$ in the market clearing condition (75):

$$\frac{e^{-\gamma z}}{\rho} = \frac{1}{\rho-\gamma} - \left(\frac{1}{\rho-\gamma} + \frac{1}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho}\right) e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})} + \frac{1}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho} e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}, \quad (76)$$

where we have used that $e^{-\gamma T_s} = \frac{1-s}{1-\mu_{h,0}}$. This implies that:

$$\begin{aligned}\frac{e^{-\gamma z}}{\rho} &> \frac{1}{\rho-\gamma} - \left(\frac{1}{\rho-\gamma} + \frac{1}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho}\right) e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})} \\ &= \frac{1}{\rho-\gamma} - \frac{\frac{\gamma}{\sigma}}{(\rho-\gamma)\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]} e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})}\end{aligned}$$

Using $e^{-\gamma T_s} = \frac{1-s}{1-\mu_{h,0}}$ and doing some algebra, we arrive at:

$$\frac{\rho}{(\rho-\gamma)\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]} e^{(\rho-\gamma)\psi_{T_s-z}} > \left(\frac{1-s}{1-\mu_{h,0}}\right)^{-\frac{\rho-\gamma}{\gamma}} \frac{e^{-(\rho-\gamma)z}}{\rho-\gamma}.$$

Equation (71) for $\gamma \neq \rho$ follows. Finally, when $\gamma = \rho$, the same manipulations lead to:

$$e^{-\gamma z} = \gamma\left(T_s - z - \psi_{T_s-z}\right) - \sigma + \sigma e^{-\frac{\gamma}{\sigma}(T_s - z - \psi_{T_s-z})} \Rightarrow 1 > e^{-\gamma z} > \gamma(T_s - z - \psi_{T_s-z}) - \sigma.$$

Taking exponentials on both sides, and using $e^{-\gamma T_s} = \frac{1-s}{1-\mu_{h,0}}$, lead to equation (71) for $\gamma = \rho$.

Proof of equation (72). When $\gamma \neq \rho$, we write equation (76) as follows:

$$\frac{1}{\rho-\gamma} - \frac{e^{-\gamma z}}{\rho} = \left(\frac{1}{\rho-\gamma} + \frac{1}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho}\right) e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})} - \frac{1}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho} e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}$$

When $\rho > \gamma$, we add $-\frac{1}{\rho-\gamma} \times e^{-(\gamma/\sigma)(T_s - z - \psi_{T_s-z})}$, which is negative, to the right–hand side:

$$\begin{aligned}&\frac{1}{\rho-\gamma} - \frac{e^{-\gamma z}}{\rho} > \frac{\frac{\gamma}{\sigma}}{(\rho-\gamma)\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]}\left(e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})} - e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}\right) \\ \Longrightarrow \quad &\frac{1}{\rho-\gamma} > \frac{\frac{\gamma}{\sigma}}{(\rho-\gamma)\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]}\left(e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})} - e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}\right) \\ \Longrightarrow \quad &\left(e^{-(\rho-\gamma)(T_s - z - \psi_{T_s-z})} - e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}\right)^{-1} > \frac{\gamma}{\gamma+\sigma(\gamma-\rho)},\end{aligned}$$

where we can keep the inequality the same because $\rho > \gamma$. Equation (72) when $\rho > \gamma$ follows.

When $\rho < \gamma$, we can also add $-\frac{1}{\rho-\gamma} \times e^{-(\gamma/\sigma)(T_s - z - \psi_{T_s - z})}$ to the right hand side. But since this term is now negative, we obtain:

$$\begin{aligned}
& \frac{1}{\rho-\gamma} - \frac{e^{-\gamma z}}{\rho} < \frac{\frac{\gamma}{\sigma}}{(\rho-\gamma)\left[\gamma\left(1+\frac{1}{\sigma}\right)-\rho\right]}\left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})} - e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}\right) \\
\Longrightarrow \quad & 1 - e^{-\gamma z}\frac{\rho-\gamma}{\rho} > \frac{\frac{\gamma}{\sigma}}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho}\left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})} - e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}\right) \\
\Longrightarrow \quad & \frac{\gamma}{\rho} > \frac{\frac{\gamma}{\sigma}}{\gamma\left(1+\frac{1}{\sigma}\right)-\rho}\left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})} - e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}\right) \\
\Longrightarrow \quad & \left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})} - e^{-\frac{\gamma}{\sigma}\left(T_s-z-\psi_{T_s-z}\right)}\right)^{-1} > \frac{\rho}{\gamma+\sigma(\gamma-\rho)}.
\end{aligned}$$

where we use $e^{-\gamma z} < 1$ to move from the second to the third line. Equation (72), when $\rho > \gamma$, follows. Finally, when $\gamma = \rho$, equation (72) follows since $1 - e^{-(\gamma/\sigma)(T_s - z - \psi_{T_s - z})} < 1$.

B.2 Trading profits

Consider, in the analytical example, a trader who learns at some time T that she has a high valuation. Assume for simplicity that $T < T_f$ so that the investor find it optimal to hold 1 unit at this time. The trading profits can be defined as:

$$\Pi = -\int_0^T p_t dq_t.$$

After integrating by part we obtain:

$$\begin{aligned}\Pi &= -p_T q_T + p_0 s + \int_0^T \dot{p}_t q_t \, dt = -p_t + p_0 s + \int_0^T \dot{p}_t q_t \, dt \\ &= -p_0 + \int_0^T \dot{p}_t \, dt + p_0 s + \int_0^T \dot{p}_t q_t \, dt = -p_0(1-s) + \int_0^T \dot{p}_t \left[q_t - 1\right] \, dt.\end{aligned}$$

Now in term of holding plans this can be written:

$$\Pi = -p_0(1-s) + \int_0^T \dot{p}_u \left[q_{\ell,\tau_u,u} - 1\right] \, du < 0.$$

Note that trading profits are negative. This makes sense because, in this model, every trader who ends up purchasing before T_f is a net buyer: she starts with s and ends with 1. This is in contrast with models of liquidity provision, in which trading profits are positive.

Note also that, since traders are net buyers, the best way to minimize cost would be to buy immediately $1-s$ at time zero. Of course, although this maximizes trading profits, this strategy does not maximize inter temporal utility, because it requires the trader to incur large holding costs during the liquidity shock.

Next, let us calculate the expectations of Π conditional on the event that there are exactly n updates over $[0,T)$. For this we need to figure out the distribution of τ_u conditional on n updates over $[0,T)$. Note first that:

$$\begin{aligned}\text{Proba}(\tau_u \leq t \wedge N_T = n) &= \sum_{k=0}^{n} \text{Proba}(N_t = k \wedge N_u - N_t = 0 \wedge N_T - N_u = n-k) \\ &= \sum_{k=0}^{n} \frac{e^{-\rho t}(\rho t)^k}{k!} e^{-\rho(u-t)} \frac{e^{-\rho(T-u)}(\rho(T-u))^{n-k}}{(n-k)!} \\ &= \frac{e^{-\rho T}(\rho T)^n}{n!} \sum_{k=0}^{n} C_n^k \left(\frac{t}{T}\right)^k \left(\frac{T-u}{T}\right)^{n-k} \\ &= \text{Proba}(N_t = n) \left[1 - \frac{u-t}{T}\right]^n.\end{aligned}$$

Therefore the distribution of τ_u conditional on n updates over $[0, T)$ is

$$Pr(\tau_u \leq t \,|\, N_T = n) = \left[1 - \frac{u-t}{T}\right]^n.$$

One sees that an increase in n creates a first-order stochastic dominance shift in the distribution. This is intuitive: if there has been lots of updates, then it is more likely that the last update before u is close to u. Combined with the observation that $q_{\ell,t,u}$ is decreasing in t, this implies that the expectations of Π conditional on n updates before T is decreasing. This implies that $\mathbb{E}\left[\Pi \,|\, n, T \leq T_f\right]$ is decreasing in n.

B.3 Private information about common values reduces trading volume

In general, private information about common values reduces trading volume, because it generates adverse selection. Below, we illustrate this point in a noisy rational expectations model, adapted from Grossman and Stiglitz (1980). The main difference is that, while in Grossman and Stiglitz (1980) there are noise traders, in the present case all investors are rational. Trading occurs, in equilibrium, because of endowment shocks generating potential gains from trade. This is an important difference for the analysis of trading volume with private information. Since noise traders do not optimize, they don't respond to increased adverse selection.

The model. Let us consider a simple version of Grossman and Stiglitz (1980). There is one asset with random payoff $v \sim \mathcal{N}(0, 1/\Psi_v)$. There are λ informed investors and $1-\lambda$ uninformed ones, all with Constant Absolute Risk Aversion (CARA) utility, α. Uninformed investors receive no signal and no endowment. Informed investors observe signal

$$v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}}, \tag{77}$$

and have random endowment s/λ, where $s \sim \mathcal{N}(0, 1/\Psi_s)$. As is standard, the common but random component of the endowment shock prevents uninformed investor from perfectly inferring informed investors' information from the asset price. The factor $1/\lambda$ keeps the aggregate supply equal to s as we vary the fraction of informed investors.

Equilibrium. To solve the model, we guess and verify that, to an uninformed investor, the price is observationally equivalent to a signal of the form:

$$v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}} - \frac{s}{\theta}, \tag{78}$$

for some $\theta > 0$ to be determined in equilibrium. Note in particular that the coefficient on s is negative: when they receive a larger endowment, the informed investors want to sell more. This puts downward pressure on the price. But uninformed investors do not know whether the downward pressure originates from an endowment shock or from adverse information about v. Thus, they will rationally interpret this negative price pressure as a noisy signal that the fundamental value has gone down.

Straightforward calculations show that the precision of the price signal, (78) is

$$\Psi_p = \Psi_\varepsilon \frac{\Psi_s \theta^2}{\Psi_s \theta^2 + \Psi_\varepsilon} < \Psi_\varepsilon.$$

Clearly, because of the noisy supply, the precision of the price signal, (78), is lower than that of informed investors' signal, (77). The demand of informed and uninformed investors can be written:

$$D_I = \frac{\mathbb{E}_I[v] - p}{\alpha \mathbb{V}_I[v]} - \frac{s}{\lambda}, \quad \text{and} \quad D_U = \frac{\mathbb{E}_U[v] - p}{\alpha \mathbb{V}_U[v]}.$$

Using Bayes' rule, and keeping in mind that the prior has mean zero, we obtain that the posterior mean of informed and uninformed investors are

$$\mathbb{E}_I[v] = \frac{\Psi_\varepsilon}{\Psi_\varepsilon + \Psi_v}\left[v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}}\right], \quad \text{and} \quad \mathbb{E}_U[v] = \frac{\Psi_p}{\Psi_p + \Psi_v}\left[v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}} - \frac{s}{\theta}\right].$$

The posterior variances of informed and uninformed investors are

$$\mathbb{V}_I[v] = (\Psi_v + \Psi_\varepsilon)^{-1}, \quad \text{and} \quad \mathbb{V}_U[v] = (\Psi_v + \Psi_p)^{-1}.$$

Therefore, the demand of informed and uninformed investors can be written:

$$\begin{aligned} D_I &= \frac{1}{\alpha}\left[\Psi_\varepsilon\left(v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}}\right) - (\Psi_\varepsilon + \Psi_v)\, p\right] - \frac{s}{\lambda} \\ D_U &= \frac{1}{\alpha}\left[\Psi_p\left(v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}} - \frac{s}{\theta}\right) - (\Psi_p + \Psi_v)\, p\right]. \end{aligned}$$

Solving for the price in $\lambda D_I + (1-\lambda) D_U = 0$, we obtain:

$$\begin{aligned} p = &- \frac{\alpha s}{\lambda \Psi_\varepsilon + (1-\lambda)\Psi_p + \Psi_v} \\ &+ \frac{\lambda \Psi_\varepsilon}{\lambda \Psi_\varepsilon + (1-\lambda)\Psi_p + \Psi_v}\left(v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}}\right) + \frac{(1-\lambda)\Psi_p}{\lambda \Psi_\varepsilon + (1-\lambda)\Psi_p + \Psi_v}\left(v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}} - \frac{s}{\theta}\right). \end{aligned}$$

After a couple of lines of algebra we see that our guess is verified iff:

$$\theta = \frac{\lambda \Psi_\varepsilon}{\alpha}.$$

The Volume. The aggregate demand from uninformed investors is

$$(1-\lambda) D_U = -\frac{\lambda(1-\lambda)\Psi_v (\Psi_\varepsilon - \Psi_p)}{\lambda \Psi_\varepsilon + (1-\lambda)\Psi_p + \Psi_s}\frac{1}{\alpha}\left\{v + \frac{\varepsilon}{\sqrt{\Psi_\varepsilon}} - \frac{s}{\theta}\right\}.$$

Without asymmetric information, it would be equal to $(1-\lambda)s$: indeed, the equilibrium allocation in this case dictates that there is full risk sharing, and hence that all investors, informed and uninformed, hold s shares of the assets.[24]

We would like to know whether this trading volume increases or decreases with asymmetric infor-

[24] Note that with symmetric information, the equilibrium volume is the same regardless of the level of risk (as long as it is positive). Indeed, with CARA agents, in the setup considered, the equilibrium allocation prescribes that agents share risk equally, regardless of their (positive) risk aversion and regardless of the level of risk.

mation. One sees that there are competing effects. On the one hand, the loading of the order flow, D_U, on s, is equal to

$$(1-\lambda)\left[1-\frac{\Psi_p}{\Psi_\varepsilon}\right]\frac{\Psi_v}{\lambda\Psi_\varepsilon+(1-\lambda)\Psi_p+\Psi_v}<1-\lambda.$$

That is, asymmetric information reduces the "fundamental" trading volume associated with hedging needs. For example, suppose that $v=\varepsilon=0$. Then, when s is positive, the informed investors want to sell assets, which puts downward pressure on the price. Uninformed investors rationnaly interpret the low price as a bad signal about the fundamental value of the asset, and reduce their demand relative to the full information case. In equilibrium, uninformed investors end up purchasing less asset from informed investors than they would have under symmetric infomaiton.

While there is less trading for fundamental "hedging" motives, there is now some speculative trading. For example, suppose that v is positive, but $\varepsilon=s=0$. Then both the informed and the uninformed investors receive a positive signal about the fundamental value of the asset. But the informed investor views his signal as more precise: hence, if the uninformed investor demand is positive, the informed demand will be positive as well. Thus, market clearing implies that the price must adjust so that uninformed demand must be negative, and informed demand must be positive.

Our main result is that:

Proposition 11. *The volume is smaller under asymmetric than under symmetric information:*

$$(1-\lambda)\mathbb{V}\left[D_U\right]<\frac{1-\lambda}{\Psi_s}.$$

To show this, we start from:

$$\mathbb{V}\left[D_U\right]=\left(\frac{\lambda(1-\lambda)\Psi_v(\Psi_p-\Psi_\varepsilon)}{\lambda\Psi_\varepsilon+(1-\lambda)\Psi_p+\Psi_v}\right)^2\frac{1}{\alpha^2}\left\{\frac{1}{\Psi_v}+\frac{1}{\Psi_\varepsilon}+\frac{1}{\theta^2\Psi_s}\right\}.$$

Substituting in $\alpha^2=\lambda^2\Psi_\varepsilon^2/\theta^2$:

$$\mathbb{V}\left[D_U\right]=\frac{(1-\lambda)^2}{\Psi_s}\left(\frac{\Psi_v(1-\Psi_p/\Psi_\varepsilon)}{\lambda\Psi_\varepsilon+(1-\lambda)\Psi_p+\Psi_v}\right)^2\left\{\frac{\theta^2\Psi_s}{\Psi_v}+\frac{\theta^2\Psi_s}{\Psi_\varepsilon}+1\right\}.$$

Now using the formula for Ψ_p we have that $\theta^2\Psi_s=\Psi_p/(1-\Psi_p/\Psi_\varepsilon)$. Plugging this in we have:

$$
\begin{aligned}
\mathbb{V}\left[D_U\right] &= \frac{(1-\lambda)^2}{\Psi_s}\left(\frac{\Psi_v(1-\Psi_p/\Psi_\varepsilon)}{\lambda\Psi_\varepsilon+(1-\lambda)\Psi_p+\Psi_v}\right)^2 \frac{\Psi_p(\Psi_v+\Psi_\varepsilon)+\Psi_v\Psi_\varepsilon(1-\Psi_p/\Psi_\varepsilon)}{\Psi_v\Psi_\varepsilon(1-\Psi_p/\Psi_\varepsilon)} \\
&= \frac{(1-\lambda)^2}{\Psi_s}\frac{\Psi_v(1-\Psi_p/\Psi_\varepsilon)}{(\lambda\Psi_\varepsilon+(1-\lambda)\Psi_p+\Psi_v)^2}\left(\Psi_p+\Psi_v\right) \\
&= \frac{(1-\lambda)^2}{\Psi_s}\times(1-\Psi_p/\Psi_\varepsilon)\times\frac{\Psi_v}{\lambda\Psi_\varepsilon+(1-\lambda)\Psi_p+\Psi_\varepsilon}\times\frac{\Psi_p+\Psi_v}{\Psi_p+\Psi_v+\lambda(\Psi_\varepsilon-\Psi_p)}
\end{aligned}
$$

Clearly, all terms multiplying $(1-\lambda)^2/\Psi_s$ are less than one, establishing the claim. □

B.4 Information collection effort

In this appendix we study a simple static variant of our model, with three stages: *ex-ante* banks choose how much information collection effort to exert, *interim* banks receive a signal about their preferences and trade in a centralized market, *ex-post* banks discover their types and payoffs realize. In this context, again, we find that the equilibrium is constrained Pareto efficient, i.e., both the choice of effort, and the allocation coincide with the one that a social planner would choose.

B.4.1 Setup

Consider a continuum of banks with utility $v(\theta, q)$ for holding an asset in supply s. Assume bank type can be either high or low, $\theta \in \{\theta_\ell, \theta_h\}$ and that the utility function satisfies the same regularity conditions as in the paper. There are three stages: *ex-ante* and *interim* and *ex-post*. In the first stage all banks start with endowment equal to s, and they invest in information collection efforts. In the second stage, banks receive a signal about their type and trade assets in a centralized market. In the third stage, banks discover their types and payoffs realize.

To model information collection effort, we assume that a bank can choose the probability ρ of knowing its type for sure. Namely, we assume that a bank observes its type exactly with probability ρ, i.e., it receives the signal $s = h$ if it has a high type, or $s = \ell$ if it has a low type. With the complementary probability, $1 - \rho$, the bank observes no signal, which we indicate using the shorthand $s = m$. Just as in our main dynamic model, banks who observe $s = m$ face preference uncertainty: they believe that they have a high type with probability μ, and a low type with probability $1 - \mu$.

Assume for now that all banks choose the same level of effort (we will argue later that this is without loss of generality). An allocation of asset is a vector $\{q_s\}_{s\in\{\ell,m,h\}}$, prescribing that a bank who observes signal $s \in \{\ell, m, h)$ holds a quantity q_s of assets. An allocation is feasible if

$$\rho\left[\mu q_h + (1-\mu)q_\ell\right] + (1-\rho)q_m = s, \tag{79}$$

where ρ is the level of effort chosen by banks.

B.4.2 Social planning problem

We define the social planning problem in two steps. First, given any level of effort, ρ, the planner solves, at the *interim* stage:

$$W(\rho) = \max_{\{q_s\}} \rho\left[\mu v(\theta_h, q_h) + (1-\mu)v(\theta_\ell, q_\ell)\right] + (1-\rho)\left[\mu v(\theta_h, q_m) + (1-\mu)v(\theta_\ell, q_m)\right],$$

subject to (79). At the *ex-ante* stage, the planner solves:

$$\max_{\rho\in[0,1]} W(\rho) - C(\rho),$$

where $C(\rho)$ is a continuously differentiable and strictly convex function of ρ. Clearly, since $W(\rho)$ is continuous by the theorem of the maximum, the *ex-ante* planner's problem has a solution.

Next, we show that this solution can be characterized by simple first-order conditions. First, standard arguments show that the *interim* problem is solved by:

$$q_h = D(1,\xi), \quad q_\ell = D(0,\xi), \text{ and } q_m = D(\mu,\xi), \tag{80}$$

where $D(\mu,\xi)$ is a demand function defined exactly as in the main body of the paper, and ξ solves:

$$\rho\left[\mu D(1,\xi) + (1-\mu)D(0,\xi)\right] + (1-\rho)D(\mu,\xi) = s. \tag{81}$$

Now consider $W(\rho)$, the social value of choosing effort, at the *ex-ante* stage. Our main result is:

Proposition 12. *The planner's problem is solved by the unique $\rho^\star$ such that*

$$W'(\rho^\star) \leq 0 \text{ if } \rho^\star = 0, W'(\rho^\star) = 0 \text{ if } \rho^\star \in (0,1), \text{ and } W'(\rho^\star) \geq 0 \text{ if } \rho^\star = 1, \tag{82}$$

where

$$\begin{aligned} W'(\rho) = {} & \left[\mu v(\theta_h, q_h) + (1-\mu)v(\theta_\ell, q_\ell)\right] - \left[\mu v(\theta_h, q_m) + (1-\mu)v(\theta_\ell, q_m)\right] \\ & - \xi\left[\mu q_h + (1-\mu)q_\ell - q_m\right], \end{aligned} \tag{83}$$

and $\{q_s\}$ and ξ jointly solve (80) and (81) given ρ.

The expression for $W'(\rho)$ is obtained by an application of the envelope theorem. Clearly, condition (83) is necessary for optimality. To show uniqueness and sufficiency, we take another round of derivative to obtain that:

$$W''(\rho) = -\frac{d\xi}{d\rho}\left[\mu q_h + (1-\mu)q_\ell - q_m\right] = \frac{\left[\mu q_h + (1-\mu)q_\ell - q_m\right]^2}{\rho\left[\mu D_\xi(1,\xi) + (1-\mu)D_\xi(0,\xi)\right] + (1-\rho)D\xi(\mu,\xi)} < 0.$$

In the above, the first equality follows because, when $\{q_s\}$ are given by (79), then marginal utilities are equal to ξ. The second equality follows by calculating $d\xi/d\rho$ explicitly using the implicit function theorem. □

Finally, we argue that our restriction that banks choose the same level of effort is without loss of generality. Notice indeed that, with heterogeneous ρ, the social welfare in the interim stage, W, only depends on the average ρ. Given convexity of the cost function, the planner strictly prefers to have all banks choose a common level of effort.

B.4.3 Equilibrium

We now study the equilibrium choice of information collection effort and show that it coincides with the social optimum. Suppose that other banks exert a level of information collection effort equal to

$\bar{\rho}$. As in the paper, the *interim* equilibrium is socially optimal given $\bar{\rho}$. This implies that the *interim* equilibrium price is the unique solution $\bar{\xi}$ of (81), and the asset holdings are given by (80). *Ex-ante*, each individual bank chooses its level of information collection effort, ρ, taking as given the information collection of others, $\bar{\rho}$, which determines the *interim* equilibrium price, $\bar{\xi}$. To an individual bank, the value of choosing ρ is:

$$V(\rho \,|\, \bar{\xi}) = \max_{\{q_s\}} \rho \left[\mu v(\theta_h, q_h) + (1-\mu) v(\theta_\ell, q_\ell)\right] + (1-\rho)\left[\mu v(\theta_h, q_m) + (1-\mu) v(\theta_\ell, q_m)\right]$$
$$- \bar{\xi} \left\{\rho \left[\mu q_h + (1-\mu) q_\ell\right] + (1-\rho) q_m\right\}.$$

A bank's *ex-ante* effort choice problem is:

$$\max_{\rho \in [0,1]} V(\rho \,|\, \bar{\xi}) - C(\rho).$$

An *ex-ante* equilibrium is defined as a pair $(\bar{\rho}, \bar{\xi})$ such that: (i) $\bar{\xi}$ is an *interim* equilibrium price given $\bar{\rho}$, and (ii) $\bar{\rho}$ solves the bank's *ex-ante* effort choice problem given $\bar{\xi}$. Our main result is:

Proposition 13. *There exists a unique ex-ante equilibrium. In this equilibrium, bank's effort collection choice is socially optimal, i.e.,* $\bar{\rho} = \rho^\star$.

To show this proposition, we first use the envelope theorem to assert that:

$$V'(\rho \,|\, \bar{\xi}) = \left[\mu v(\theta_h, q_h) + (1-\mu) v(\theta_\ell, q_\ell)\right] - \left[\mu v(\theta_h, q_m) + (1-\mu) v(\theta_\ell, q_m)\right]$$
$$- \bar{\xi} \left\{\mu q_h + (1-\mu) q_\ell - q_m\right\},$$

where $\{q_s\}$ solves (80) given $\bar{\xi}$. Since $\{q_s\}$ only depend on $\bar{\xi}$, which a bank takes as given, we have that $V''(\rho \,|\, \bar{\xi}) = 0$. Since the cost function $C(\rho)$ is strictly convex, it thus follows that the *ex-ante* effort choice problem is strictly concave, and its solution is uniquely characterized by the first-order condition. Clearly, one sees that the equilibrium condition coincides with the optimality condition of the planning problem. □

Notice again that we need not worry about asymmetric equilibria in which banks choose heterogenous levels of efforts: given the price that will prevail at the interim stage, a bank's effort choice problem is strictly concave, so it has a unique maximizer.

To illustrate the proposition we consider the following numerical example. We use iso-elastic preferences $v(\theta, q) = \theta q^{1-\sigma}/(1-\sigma)$, and we set $\sigma = 0.5$, $s = 0.5$, $\theta_h = 1$, and $\theta_\ell = 0.1$. We assume that $\mu = 0.5$ and that the cost of effort is:

$$C(\rho) = c \frac{\rho^{1+\gamma}}{1+\gamma},$$

where $\gamma = 0.1$ and the constant c is chosen so that the planner's problem is maximized at $\rho^\star = 0.5$. In Figure 6, the social value of information collection effort, $W(\rho) - C(\rho)$, is shown as the plain red

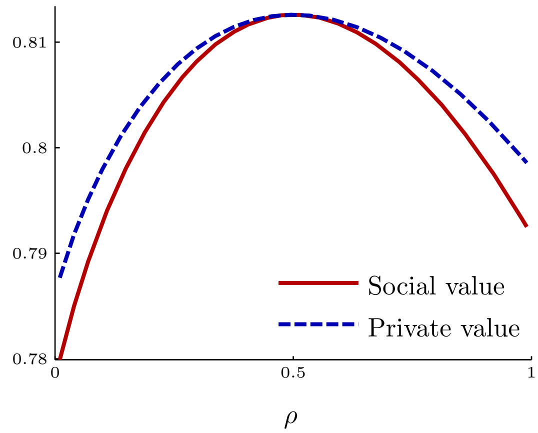

Figure 6: The social value (plain red) and private value (dashed blue) of information collection effort.

curve. The individual bank's private value of recovery effort given the equilibrium price $\xi^\star$ generated by $\rho^\star$, $V(\rho\,|\xi^\star) - C(\rho)$, is the dashed blue curve. One sees that the social value of effort differs from the social value. In particular, the social value is more concave than the private value: this is because the planner's value takes into account the impact of changing ρ on the (shadow) price of the asset, ξ, while an individual bank does not. However, one sees that the envelope theorem ensures that the private and social value coincide and are tangent to each other at $\rho = \rho^\star$.

B.5 Finite number of traders

In this appendix we offer some numerical calculations of an equilibrium when there is a finite number of traders, with and without preference uncertainty. We describe the evolution of traders' asset holdings and of the holding cost. Our calculations reveal that our main excess volume result continues to hold when there is a finite number of traders. In addition, since idiosyncratic preference shocks and updating times no longer average out, the model features a new source of holding cost volatility. Our calculations suggests that, relative to the known preference case with the same finite number of traders, preference uncertainty tends to mitigate this new source of volatility.

We consider a finite number N of traders but otherwise keep the model exactly as in the text. In particular, we continue to assume that traders behave competitively, as price takers. Studying price impact, along the line of Vayanos (1999) or Rostek and Weretka (2011) would introduce additional technical difficulties that go beyond the main objective of this appendix. Under price taking, the demand of trader $i \in \{1, \ldots, N\}$ at time u remains equal to $D(\pi_{\tau_u^i,u}, \xi_u)$, where τ_u^i denotes the last updating time of trader $i \in \{1, \ldots, N\}$ before the current time, u. What is different is the market clearing condition, which becomes:

$$\frac{1}{N}\sum_{i=1}^{N} D(\pi_{\tau_u^i,u}, \xi_u) = s.$$

One sees that, each trader's updating time before recovery becomes an aggregate shock: it changes that trader's demand and thus moves the price discretely.

Figure 7 shows the equilibrium holdings along a particular sample path of preference shocks and updating times. The number of traders is set to $N = 5$ and otherwise the parameters are the same as in our main parametric calculations. Equilibrium objects under preference uncertainty and known preferences are depicted by plain blue lines and dashed red lines, respectively. One sees clearly from the figure that the updating times of others become aggregate shocks and cause every trader to change its holdings. This is an additional source of trading volume, above and beyond the one identified in the continuum-of-traders case.

Figure 8 shows the cumulative volume along this particular sample path of shocks (left panel), as well as the average volume across $10,000$ sample paths (right panel). Both figures indicate that, just as in our main model, cumulative volume is larger with preference uncertainty than with known preferences.

Figure 9 shows the holding cost for the same particular sample path of shocks (left panel) as well as the average holding cost across 10,000 sample paths of shocks. One sees clearly from both panels that preference uncertainty tends to raise the holding cost at the inception of the liquidity shock, because traders who still have a low valuation believe they may have switched to a high valuation. One also sees that the full recovery is delayed, as traders need to wait for an updating time before being certain that they have a high valuation.

Finally, one may wonder what is the impact of having a finite number of traders on holding cost

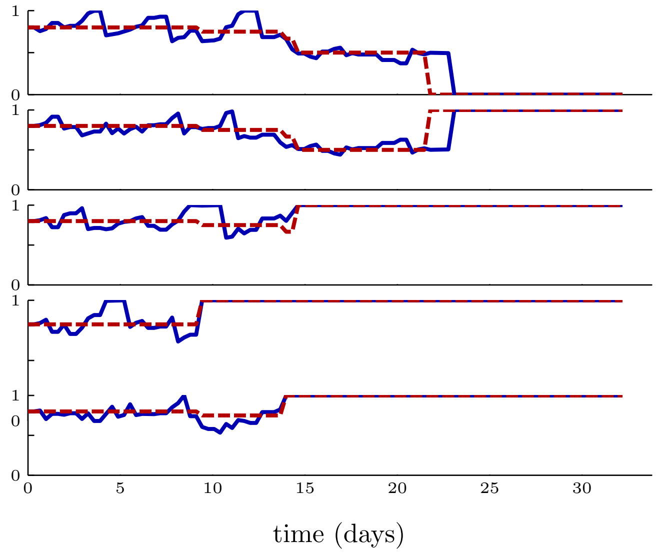

Figure 7: Holdings of 5 traders along a sample path of shocks, for known preferences (dashed red) vs. uncertain preferences (plain blue).

volatility, with and without preference uncertainty. One sees intuitively that, with known preferences, there are larger upward changes in holding costs. With preference uncertainty, there are many small changes in holding costs, upward and downward. Figure 10 confirms this observation by calculating the volatility of the percentage difference between the holding cost and the average holding cost across 10,000 simulations. The volatility with known preferences is higher, and peaks sooner, reflecting the large change in holding cost arising when sufficiently many traders have switched to high.

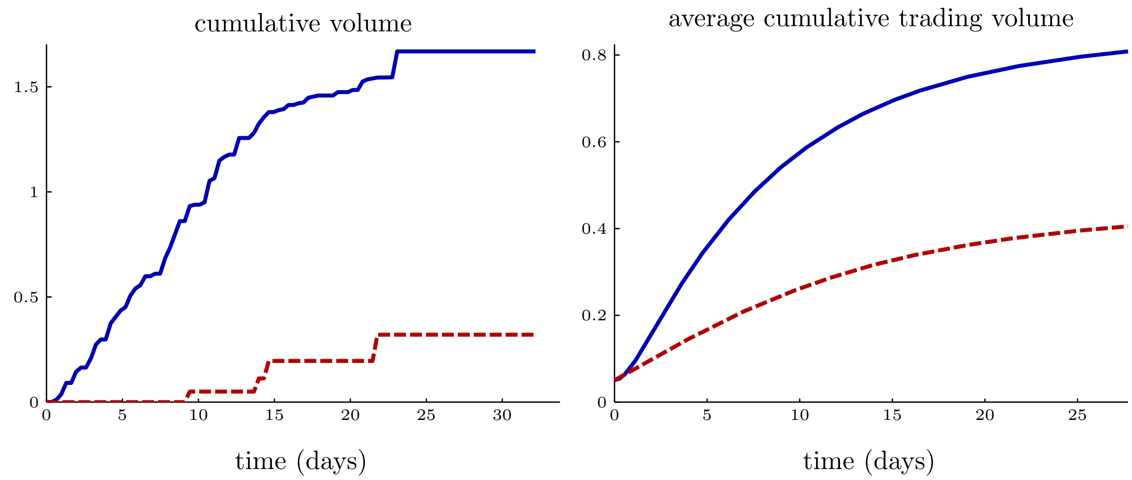

Figure 8: Cumulative trading volume along a sample path of shocks (left panel) and average cumulative trading volume across 10,000 sample paths of shocks, for known preferences (dashed red) vs. uncertain preferences (plain blue).

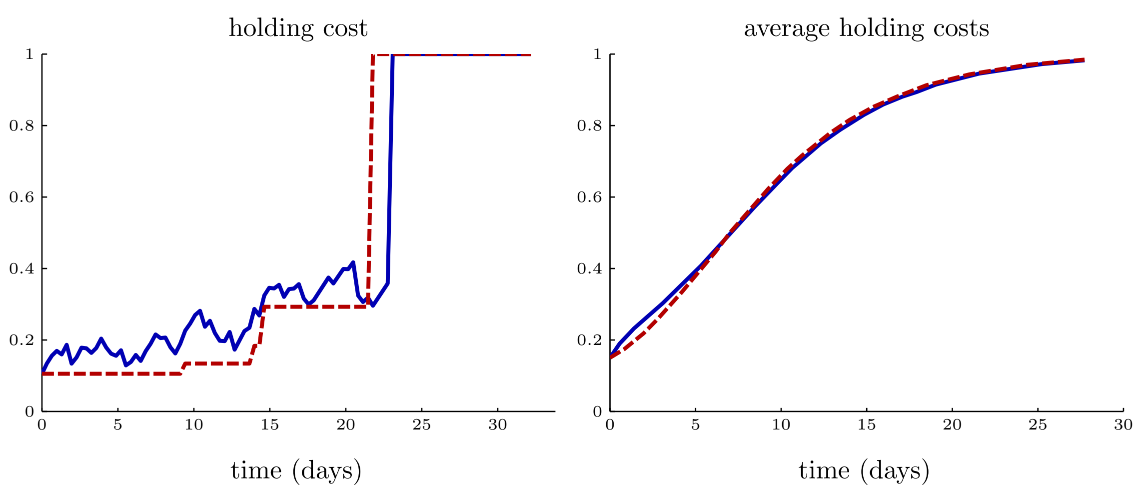

Figure 9: Holding cost along a sample path of shocks (left panel) and average holding cost across 10,000 sample paths of shocks, for known preferences (dashed red) vs. uncertain preferences (plain blue).

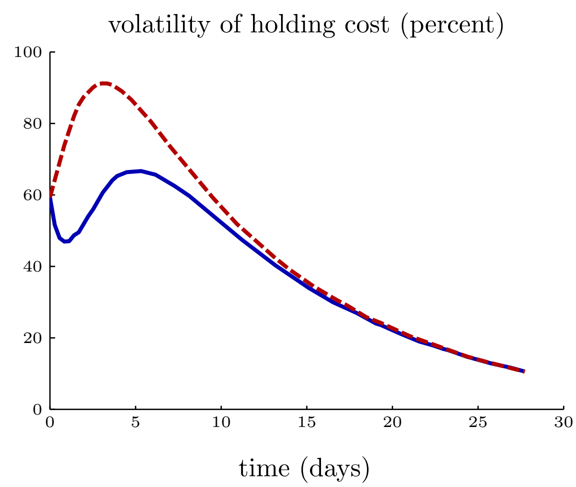

Figure 10: Volatility of holding costs across 10,000 sample paths of shocks, for known preferences (dashed red) vs. uncertain preferences (plain blue).