Trading and Liquidity with Limited Cognition*

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

November 4, 2010

Abstract

We study the reaction of traders and markets to an aggregate liquidity shock under cognition limits. While institutions recover from the shock at random times, traders observe the status of their institution only when their own information process jumps. This delay reflects the time it takes to collect and process information about positions, counterparties and risk exposure. Traders who find their institution has a low valuation place market sell orders, and then progressively buy back at relatively low prices, while simultaneously placing limit orders to sell later when the price will have recovered. We compare the case where algorithms enable traders to implement this strategy to that where traders can only place orders when their information process jumps. Our theoretical results are in line with empirical findings on order placements and algorithmic trading.

Keywords: Limit-orders, asset pricing and liquidity, bid-ask spread, algorithmic trading, limited cognition, sticky plans.

J.E.L. Codes: G12, D83.

---

*Many thanks, for helpful discussions and suggestions, to Andy Atkeson, Dirk Bergemann, Darrell Duffie, Emmanuel Farhi, Alfred Galichon, Thierry Foucault, Christian Hellwig, Hugo Hopenhayn, Vivien Lévy–Garboua, Johannes Horner, Boyan Jovanovic, Albert Menkveld, John Moore, Henri Pages, Thomas Philippon, Jean Charles Rochet, Ioanid Rosu, Larry Samuelson, Tom Sargent, Jean Tirole, Aleh Tsyvinski, Juusso Välimäki, Dimitri Vayanos, Adrien Verdelhan, and Glen Weyl; and 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, and New York University. Paulo Coutinho and Kei Kawakami provided excellent research assistance. We are grateful to the editor and four anonymous referees for comments. Bruno Biais benefitted from the support of the "Financial Markets and Investment Banking Value Chain Chair" sponsored by the Fédération Bancaire Française, and Pierre-Olivier Weill from the support of the National Science Foundation, grant SES-0922338.

†Toulouse School of Economics (CNRS, IDEI), `biais@cict.fr`.

‡HEC Paris, `johan.hombert@hec.fr`.

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

1 Introduction

> "The perception of the intellect extends only to the few things that are accessible to it and is always very limited" (Descartes, 1641, page 125)

To reach decisions, traders and investment managers must process large amounts of information. They must form expectations about their valuation for assets, find out about market conditions and assess the financial status of their own institution. To achieve the latter, they must evaluate the gross and net positions of the institution and the resulting risk exposure. They must also find out whether the institution complies with regulatory limits on positions. This requires collecting, processing and aggregating information across various counterparties, departments, instruments and markets. When investors have limited cognition, it takes time to complete this task.

This task is particularly challenging when the market is hit by an aggregate liquidity shock, whereby a significant fraction of the investors' population is affected by a change in its willingness and ability to hold assets. Such shocks can occur because of changes in the characteristics of assets, e.g., many institutions are required to sell bonds which lose their investment grade status, or sell stocks when they are de-listed from exchanges (Greenwood, 2005). Liquidity shocks can also reflect events affecting the overall financial situation of a population of investors, e.g., funds or banks experiencing large outflows or losses (Coval and Stafford, 2007, Berndt et al., 2005, Khandani and Lo, 2008). Around such shocks, the flow of information that traders must analyze is even greater than usual.

We thus address the following issues: How do traders and markets cope with liquidity shocks? What is the equilibrium price process after such shocks? How are trading and prices affected by cognition limits? Do the consequences of limited cognition vary with market mechanisms and technologies?

We consider an infinite horizon, continuous–time market with a continuum of rational, risk–neutral competitive investors. Investors derive a utility flow from holding the asset. To model the aggregate liquidity shock, we follow Duffie, Gârleanu, and Pedersen (2007) and Weill (2007) and assume that at time 0 the valuation for the asset drops for all investors. Then, as time goes by, some investors switch back to a high valuation. More precisely, each investor is associated with a Poisson process and switches back to high-valuation at the first jump in this process. Efficiency would require that low-valuation investors sell to high-valuation investors. Such efficient reallocation of the asset is delayed, however, because of cognition limits. To model the latter, in line with the rational inattention models of Mankiw and Reis (2002) and Gabaix and Laibson (2002), we assume investors engage in information collection and processing for some time and, only when this task is completed, observe the current valuation of their institution for the asset. Once they observe this refreshed information, they update their optimal holding plan, based on rational expectations about future variables and decisions.[1] In the same spirit as in Duffie, Gârleanu, and Pedersen (2005), we assume that investors observe such new information, and correspondingly revise their holding plans, at Poisson arrival times.[2] Note that, in our model, each investor is exposed to two Poisson processes: one concerns changes in his valuation for the asset, the other the timing of his information events. For simplicity, we assume these two processes are independent. Also for tractability we assume that these processes are independent across agents. Thus, by the law of large numbers, the aggregate state of the market changes deterministically with time. Correspondingly, the equilibrium price process is deterministic too. We show equilibrium existence and uniqueness. In equilibrium the price increases with time, reflecting that the market progressively recovers from the shock. Also, investors choose their holding plans to maximize the discounted sum of the difference between their expected utility flows and the (endogenous) opportunity cost of holding the asset. We show that the equilibrium is an information constrained Pareto optimum. This is because in our setup there are no pecuniary externalities, as the constraints on holdings imposed by cognition limits don't depend on prices.

While we first characterize the optimal policies of the agents in terms of abstract holding plans, we then show how these plans can be implemented in a realistic market setting, featuring an electronic order book, limit and market orders, and trading algorithms. The latter enable investors to conduct programmed trades while devoting their cognitive resources to investigating the liquidity status of their institution. In this context, traders who find out their institution is still subject to the shock, and correspondingly has a low valuation for the asset, sell a lump of their holdings, with a market sell order. They also program their trading algorithms to

[1]Thus, in the same spirit as in Mankiw and Reis (2002), traders have "sticky plans" and rationally take into account this stickyness.

[2]Note however that the interpretation is different. Duffie, Gârleanu, and Pedersen (2005) model the time it takes for traders to find a counterparty, while we model the time it takes them to collect and process information. This difference results in different outcomes. In Duffie, Gârleanu, and Pedersen investors don't trade between two jumps of their Poisson process. In our model they do, but based on imperfect information about their valuation for the asset. In a sense our model can be viewed as the dual of Duffie, Gârleanu, and Pedersen: in the latter traders continuously observe their valuation but are infrequently in contact with the market, in the former traders are continuously in contact with the market but infrequently observe refreshed information about their valuations.

then gradually buy back, as they expect their valuation to revert upward. Simultaneousy, they submit limit orders to sell the asset, to be executed later when the equilibrium price will have recovered. To the extent that they buy in the early phase of the liquidity cycle, and then sell towards the end of the cycle, the traders act as market makers.[3] The corresponding round–trip transactions reflect their optimal reaction to cognition limits. These transactions raise trading volume above the level it would reach under perfect cognition. While limits to cognition lengthen the time it takes market prices to fully recover from the shock, it does not necessarily amplify the initial price drop generated by that shock. Just after the shock, with perfect cognition the marginal investors knows she has low valuation, while with limited cognition she is imperfectly informed about her valuation, and realizes that, with some probability, it may have recovered.

We also study the case where trading algorithms are not available and, as traders devote their cognitive resources to investigating the liquidity status of their institution, they cannot reach trading decisions. In this case, traders can only place limit or market orders when their information process jumps. With increasing prices, this prevents them from buying in between jumps of their information process. When the liquidity shock is large, this constraint binds and reduces the efficiency of the equilibrium allocation. It does not necessarily amplify the price impact of the liquidity shock, however. Since traders anticipate they won't be able to buy back until their next information event, they sell less when they observe their valuation is low. Such a reduction in sales limits the selling pressure on prices. To an outside observer this might suggest that algorithms are destabilizing the market, by amplifying the price effects of liquidity shocks. Yet, in our model, the equilibrium with algorithmic trading is information constrained Pareto optimal.

The order placement policies generated by our model are in line with several stylized facts. Irrespective of whether agorithms are available or not, we find that successive traders place limit sell orders at lower and lower prices. Such undercutting is consistent with the empirical results of Biais, Hillion, and Spatt (1995), Griffiths, Smith, Turnbull, and White (2000) and Ellul, Holden, Jain, and Jennings (2007). Furthermore, our algorithmic traders both supply and consume liquidity, by placing market and limit orders, consistent with the empirical findings of Hendershott and Riordan (2010) and Brogaard (2010). Brogaard also finds that algorithms i)

[3]In doing so the act similarly to the market makers analyzed by Grossman and Miller (1988). Note however that, while in Grossman and Miller (1988) agents are exogenously assigned market making or market taking roles, in our model, agents endogenously choose to supply or demand liquidity, depending on the realisation of their own shocks.

don't tend to withdraw from the market after large liquidity shocks, ii) tend to provide liquidity by purchasing the asset after large price drops, and iii) in doing so profit from price reversals; all these features are in line with the implications of our model.

Our analysis of market dynamics when traders choose to place limit or market orders is related to the insightful papers of Parlour (1998), Foucault (1999), Foucault, Kadan, and Kandel (2005), Rosu (2009), Goettler, Parlour, and Rajan (2005, 2009). But we focus on different market frictions than they do. While they study strategic behaviour and/or asymmetric information under perfect cognition, we analyze competive traders with symmetric information under limited cognition. This enables us to study how the equilibrium interaction between the price process and order placement policies is affected by cognition limits and market instruments.

The next section presents the economic environment and the equilibrium prevailing under unlimited cognition. Section 3 presents the equilibrium prevailing with limited cognition. Sections 4 discusses the implementation of the abstract equilibrium holding plans with realistic market instruments such as limit and market orders and trading algorithms. Section 5 concludes. Proofs not given in the text are in the appendix, and a supplementary appendix offers additional material, helpful to further one's understanding of the model.

2 The economic environment

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).[4] There is an asset in positive supply $s \in (0,1)$ and the economy is populated by a $[0,1]$-continuum of infinitely-lived agents that we call "financial institutions" (funds, banks, insurers, etc...) discounting the future at the same rate $r > 0$.

Each institution can be in one of two states. Either it derives a high utility flow ("$\theta = h$") from holding any quantity $q \geq 0$ of the asset, or it derives a low utility flow ("$\theta = \ell$"), as illustrated in Figure 1. For high–valuation institutions, the utility flow per unit of time is $v(h,q) = q$, for all $q \leq 1$, and $v(h,q) = 1$, for all $q > 1$. For low–valuation institutions,

[4] 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.

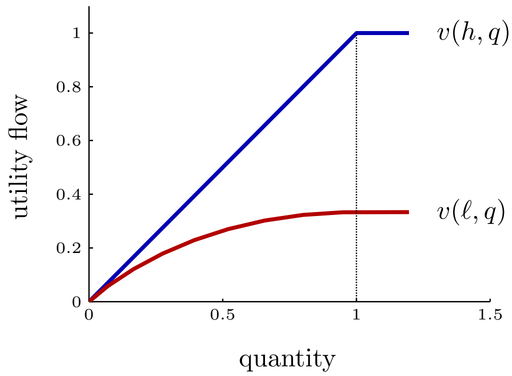

Figure 1: The utility flows of high– (in blue) and low–valuation (in red) investors, when $\sigma = 0.5$.

it is $v(\ell, q) = q - \delta \frac{q^{1+\sigma}}{1+\sigma}$, for all $q \leq 1$, and $v(\ell, q) = 1 - \delta/(1+\sigma)$, for all $q > 1$. The two parameters $\delta \in (0, 1]$ and $\sigma > 0$ capture the effect of a low liquidity status on utility flows. The parameter δ controls the level of utility: the greater is δ, the lower is the marginal utility flow of low–valuation institutions. The parameter σ, on the other hand, controls the curvature of low–valuation institutions utility flows. The greater is σ, the less willing they are to hold extreme asset positions.[5] Because of this concavity, it is efficient to spread holdings among low–valuation institutions. This is similar to risk–sharing between risk–averse agents, and as shown below will imply that equilibrium holdings take a rich set of values.[6] This is in line with Lagos and Rocheteau (2009) and Gârleanu (2009). Note that, even in the $\sigma \to 0$ limit, low–valuation investors' utility flow is reduced, by a factor $1 - \delta$, but in that case the utility is linear for $q \in (0, 1)$.[7]

In addition to deriving utility from the asset, institutions can produce (or consume) a non-storable numéraire good at constant marginal cost (utility) normalized to one.

2.2 Liquidity shock

To model liquidity shocks we follow Duffie, Gârleanu, and Pedersen (2007) and Weill (2007). Before the shock each institution is in the high–valuation state, $\theta = h$, and holds s shares of the asset. But, at time zero, a liquidity shock hits: all institutions make a transient switch to

[5]The curvature of low–valuations utilities contrasts with the constant positive marginal utility of high–valuation institutions have for $q < 1$. One could have introduced such curvatures for high–valuations too as in Lagos and Rocheteau (2009) or Gârleanu (2009) at the cost of reduced tractability, without qualitatively altering our results.

[6]Note also that the holding costs of low–valuation institutions are homothetic. This results in homogenous asset demand and, as will become clear later, facilitates aggregation.

[7]For the $\sigma \to 0$ limit, see our supplementary appendix, (Biais, Hombert, and Weill, 2010), Section III.

low–valuation, $\theta = \ell$. The economic motivation is the following.

First, consider what triggers the liquidity shock. It can be changes in the characteristics of assets, e.g., certain types of institutions, such as insurance companies or pension funds, are required to sell bonds which lose their investment grade status, or stocks which are delisted from exchanges or indices (see, e.g., Greenwood, 2005). Alternatively, liquidity shocks can reflect events affecting the overall financial situation of the institutions, e.g., funds experiencing large outflows or losses must sell some of their holdings (see Coval and Stafford, 2007) or banks incurring large losses are compelled by regulation to sell risky assets (see Berndt, Douglas, Duffie, Ferguson, and Schranz, 2005). The situation we have in mind is quite in line with the liquidity shock analyzed by Khandani and Lo (2008) who observe that, during the week of August 6th 2007, quantitative funds subject to margin calls and losses in credit portfolios had to rapidly unwind equity positions. This resulted in a sharp but transient drop in the S&P. But, by August 10th 2010 prices had in large part reverted.

Second, consider how institutions react to, and eventually recover from, the liquidity shock. In practice, they operate in several markets in addition to the one that is the focus of our model. For example they can have positions in credit default swaps (CDS), corporate bonds or mortgage based securities (MBS), all of which trade in rather illiquid over–the–counter (OTC) markets. When hit by the shock, each institution seeks to unwind some of these positions. It will also engage in financial structure adjustments, such as, e.g., new issuances. All this process is complex and takes time. But, once the institution has been able to arrange enough deals, it recovers from the liquidity shock.

To model this process we assume that, for each institution, there is a random time at which it reverts to the high–valuation state, $\theta = h$, and then remains there forever. For simplicity, we assume that recovery times are exponentially distributed, with parameter γ, and independent across investors. Hence, by the law of large numbers, the measure μ_{ht} of high–valuation investors at time t is equal to the probability of high–utility at that time conditional on low–utility at time zero.[8] Thus

$$\mu_{ht} = 1 - e^{-\gamma t}, \tag{1}$$

and we denote by T_s the time at which the mass of high–utility institutions equals the supply

[8] For simplicity and brevity, we don't formally prove how the law or 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.

of the asset, i.e., $\mu_{hT_s} = s$.

2.3 Equilibrium without cognition limits

Consider the benchmark case where the institutions continuously observe θ. To find the competitive equilibrium, it is convenient to solve first for efficient asset allocations, and then find the price path which decentralizes these efficient allocations in a competitive equilibrium.[9]

In the efficient allocation, for $t > T_s$, all assets are held by high–valuation institutions, and all marginal utilities are equalized. Indeed, with an (average) asset holding equal to $s/\mu_{ht} < 1$, the marginal utility is 1 for high–valuation institutions, while with zero asset holdings marginal utility is $v_q(\ell, 0) = 1$ for low–valuation institutions. In contrast, for $t \leq T_s$, we have $\mu_{ht} \leq s$, and each high–valuation institution holds one unit of the asset while the residual supply, $s - \mu_{ht}$, is held by low–valuation institution. The asset holding per low–valuation institution is thus:

$$q_t = \frac{s - \mu_{ht}}{1 - \mu_{ht}}. \tag{2}$$

This is an optimal allocation because all high–valuation institutions are at the corner of their utility function: reducing their holdings would create a utility loss of 1, while increasing their holdings would create zero utility. Low valuation institutions, on the other hand, have holdings in $[0, 1)$, so their marginal utility is strictly positive and less than 1.

For $t \leq T_s$, as soon as an institution switches from $\theta = \ell$ to $\theta = h$, its holdings jump from q_t to 1, while as long as her valuation remains low, it holds q_t, given in (2), which smoothly declines with time. This decline reflects that, as time goes by, more and more institutions recover from the shock, switch to $\theta = h$ and increase their holdings. As a result, the remaining low valuation institutions are left with less shares to hold.

Equilibrium prices reflect the cross–section of valuations across institutions. In our setting, by the law of large numbers, there is no aggregate uncertainty and this cross–section is deterministic. Hence, the price also is deterministic. For $t \leq T_s$, it is equal to the present value of a low–valuation institution's marginal utility flow:

$$p_t = \int_t^{\infty} e^{-r(z-t)} v_q(\ell, q_z)\, dz,$$

[9]Note that, with quasi–linear utilities and unlimited cognition, in all Pareto efficient allocation of assets and numéraire goods, the asset allocation maximizes, at each time, the equally weighted sum of the institutions' utility flows for the asset, subject to feasibility.

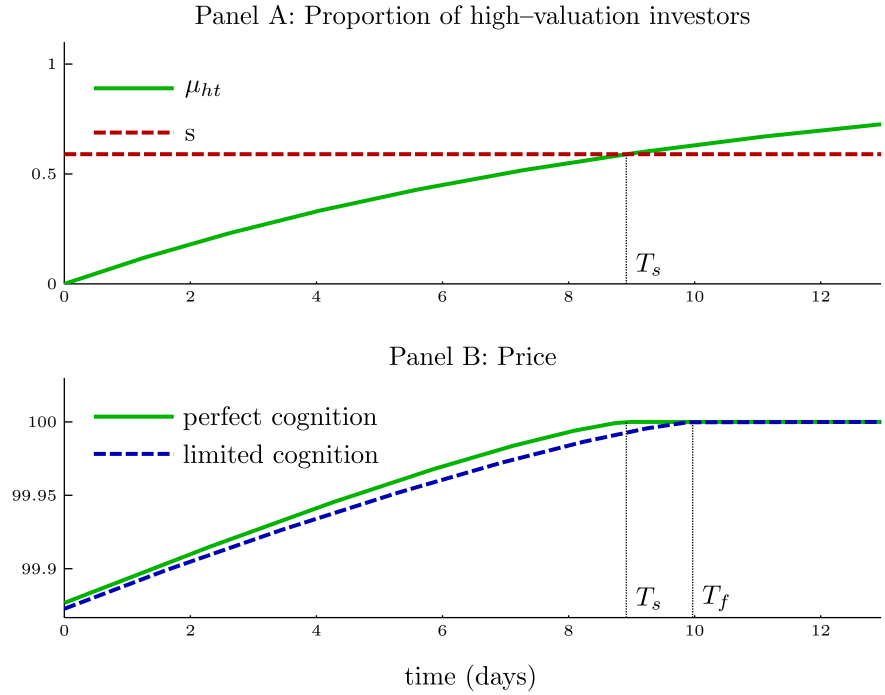

Figure 2: Population of high–valuation investors (Panel A) and price dynamics when $\sigma = 0.3$ (Panel B).

where q_z is given in (2). Taking derivatives with respect to t, we find that the price solves the Ordinary Differential Equation (ODE):

$$v_q(\ell, q_t) = rp_t - \dot{p}_t \equiv \xi_t. \tag{3}$$

The left-hand side of (3) is the institution's marginal utility flow over $[t, t + dt]$. The right-hand side is the opportunity cost of holding the asset: it is the cost of buying a share of the asset at time t and reselling it at $t + dt$, i.e., the time value of money, rp_t, minus the capital gain, $\dot{p}_t$. Finally, when $t \geq T_s$, $v_q(\ell, q_z) = v_q(\ell, 0) = 1$ and the price is $p_t = 1/r$.

Thus, the price increases deterministically towards $1/r$, as the holdings of low valuation institutions go to zero and their marginal utility increases. Institutions do not immediately bid up this predictable price increase because the demand for the asset builds up slowly: on the intensive margin, high–valuation institutions derive no utility flow if they hold more than one unit; and, on the extensive margin, the recovery from the aggregate liquidity shock occurs progressively as institutions switch back to high utility flows. Thus, there are "limits to arbitrage" in our model, in line with the empirical evidence on the predictable patterns of price drops and

Table 1: Parameter values

Parameter		Value
Intensity of information event	ρ	250
Asset supply	s	0.59
Recovery intensity	γ	25
Discount rate	r	0.05
Utility cost	δ	1
Curvature of utility flow	σ	$\{0.3, 0.5, 1.5\}$

reversals around liquidity shocks.[10]

Throughout the paper we will illustrate our results with numerical computations based on the parameter values shown in Table 1. We take the discount rate to be $r = 0.05$, in line with Duffie, Gârleanu, and Pedersen (2007). We pick the liquidity shock parameters to match empirical observations from large equity markets. Hendershott and Seasholes (2007) and Hendershott and Menkveld (2010) 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 of Khandani and Lo (2008), the price pressure subsides in about 4 trading days. Adopting the convention that there are 250 trading days per year, setting γ to 25 means that an institution takes on average 10 days to switch to high valuation. Setting the asset supply s to 0.59 then implies that with unlimited cognition the time it takes the market to recover from the liquidity shock (T_s) is around 9 days, as illustrated in Figure 2 Panel A.

Furthermore, for these parameter values, setting $\delta = 1$ implies the initial price pressure generated by the liquidity shock is between 10 and 20 basis points, as illustrated illustrated in Figure 2 Panel B.[11]

3 Equilibrium with limited cognition

3.1 Limited cognition

Each institution is represented in the market by one trader.[12] To determine optimal asset holdings, the trader must analyze the liquidity status of her institution. This task is cognitively

[10]See, e.g., for short–lived shocks the empirical findings of Hendershott and Seasholes (2007), Hendershott and Menkveld (2010) and Khandani and Lo (2008).

[11]Duffie, Gârleanu and Pedersen (2007) provide a numerical analysis of liquidity shocks in OTC markets. They chose parameters to match stylized facts from illiquid corporate bond markets. Because we focus on more liquid electronic exchanges, we chose parameter values different from theirs. For example in their analysis the price takes one year to recover while in ours it takes less than two weeks.

[12]For simplicity we abstract from agency issues and assume the trader maximizes the inter-temporal expected utility of the institution.

challenging. As mentioned in the previous section, to recover from the shock the institution engages in several financial transactions in a variety of markets, some of them complex, opaque and not computerized. Evaluating the liquidity status of the institution requires collecting, analyzing and aggregating information about the resulting positions. Our key assumption is that, because of limited cognition and information processing constraints, the trader cannot continuously and immediately observe the liquidity status of the institution.[13] Instead, we assume there is a counting process N_t such that the trader observes θ_t at each jump of N_t (and only then).[14] At the jumps of her information process N_t the trader submits a new optimal trading plan, based on rational expectations about $\{N_u, \theta_u : u \geq t\}$, and her future decisions. This is in line with the rational inattention model of Mankiw and Reis (2002). For simplicity, the traders' information event processes are assumed to be Poisson distributed, with intensity ρ, and independent from each others as well as from the times at which institutions emerge from the liquidity shock.[15]

3.2 Conditions on asset holding plans and prices

When an information event occurs at time $t > 0$, a trader designs an updated asset holding plan, $q_{t,u}$, for all subsequent times $u \geq t$ until the following information event.

Formally, denoting $D = \{(t, u) \in \mathbb{R}_+^2 : t \leq u\}$, we let an *asset holding plan* be a bounded and measurable stochastic process

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

satisfying the following two conditions:

Condition 1. *For each $u \geq t$, the stochastic process $(t, \omega) \mapsto q_{t,u}(\omega)$ is $\mathcal{F}_t$-predictable, where $\{\mathcal{F}_t\}_{t\geq 0}$ is the filtration generated by N_t and θ_t.*

[13]Regulators have recently emphasized the difficulty to come up with an integrated measurement of all relevant risk exposures within a financial institution (see Basel Commitee on Banking Supervision, 2009). Academic research has also underscored the difficulties associated with the aggregation of information dispersed in several departments of the financial institution (see Vayanos, 2003).

[14]The time between jumps create delays in obtaining fresh information about θ, which be interpreted as the time it takes to the risk management unit or head of strategy to aggregate all relevant information and disseminate it to the traders.

[15]For simplicity, we don't index the information processes of he differen traders by subscripts specific to each trader. Rather we use the same generic notation N for all traders.

Condition 2. *For each (t, ω), the function $u \mapsto q_{t,u}(\omega)$ has bounded variations.*

Condition 1 means that the plan designed at time t, $q_{t,u}$, can only depend on the trader's time-t information about her institution: the history of her information-event counting process, and of her institution utility status process up to, but not including, time t.[16] Condition 2 is a technical regularity condition ensuring that the present value of payments associated with $q_{t,u}$ is well defined.

To simplify notations in what follows we suppress the explicit dependence of asset holding plans on ω. For any time $u \geq 0$, let τ_u denote the time of the last information event before u, with the convention that $\tau_u = 0$ if no information event occurred. Correspondingly $q_{0,u}$ represents the holdings of a trader who had no information event by time u and thus no opportunity to update her holding plan. Given that all traders start with the same holdings at time zero, we have $q_{0,u} = q_{0,0} = s$ for all $u \geq 0$.

At this stage of the analysis, we assume that traders have access to a rich enough menu of market instruments to implement any holding plan satisfying Conditions 1 and 2. We address the implementation question in Section 4, where we analyze what types of market instruments are needed to implement equilibrium holding plans, and what equilibrium arises when the menu of market instrument is not rich enough.

The last technical condition concerns the price path:

Condition 3 (Well-behaved price path)**.** *The price path is bounded, deterministic and continuously differentiable (C^1).*

As in the unbounded cognition case, because there is no aggregate uncertainty, it is natural to focus on deterministic price paths. Furthermore, in the environment we consider the equilibrium price must be continuous, as formally shown in our supplementary appendix (see Biais, Hombert, and Weill, 2010, Section VI). The economic intuition is the following. If the price were to jump at time t, all traders who experience an information event shortly before t would want to "arbitrage" the jump: they would find it optimal to buy an infinite quantity of asset and re-sell these assets just after the jump. This would contradict market–clearing. Finally,

[16]We add the "not including" qualifier because the asset holding plans are assumed to be $\mathcal{F}_t$-predictable instead of $\mathcal{F}_t$-measurable. This predictability assumption is standard for dynamic optimization problems involving decisions at Poisson arrival times (see Chapter VII of Brémaud, 1981). For much of the paper, however, we won't need to distinguish between $\mathcal{F}_t$–predicability and $\mathcal{F}_t$–measurability. This is because the probability that the trader type switches exactly at the same time an information event occurs is of second order. Therefore, adding or removing the type information accruing exactly at information events leads to almost surely identical optimal trading decisions.

the condition that the price be bounded is imposed to rule out bubbles (see Lagos, Rocheteau, and Weill, 2007, for a proof that bubbles can't arise in a closely related environment).

3.3 Intertemporal payoffs

We now derive the intertemporal payoff generated by an asset holding plan $q_{t,u}$. The intertemporal expected utility from the holding plan $q_{t,u}$ can be written (gross of the corresponding payments, which we consider later):

$$\mathbb{E}\left[\int_0^\infty e^{-ru} v(\theta_u, q_{\tau_u,u}) du\right]. \tag{4}$$

For any $t \leq u$, the probability that $\tau_u \leq t$ is the probability that $N_u - N_t = 0$, which is equal to $e^{-\rho(u-t)}$. Thus, the distribution of τ_u has an atom of mass $e^{-\rho u}$ at $t = 0$, and then the density $\rho e^{-\rho(u-t)}$ for $t \in (0, u]$. Hence, after applying Baye's rule, (4) rewrites as:

$$\int_0^\infty e^{-ru}\left\{e^{-\rho u}\mathbb{E}\left[v(\theta_u, q_{0,u}) \,|\, \tau_u = 0\right] + \int_0^u \rho e^{-\rho(u-t)}\mathbb{E}\left[v(\theta_u, q_{t,u}) \,|\, \tau_u = t\right]\, dt\right\} du. \tag{5}$$

To simplify this expression, we rely on the following lemma.

Lemma 1. $\mathbb{E}\left[v(\theta_u, q_{t,u}) \,|\, \tau_u = t\right] = \mathbb{E}\left[v(\theta_u, q_{t,u})\right]$ *for all* $t \geq 0$

The lemma is clearly true for $t = 0$ since $q_{0,u} = s$ for all u, and since the information event process is independent from the type process. In Appendix A.1, we show that it also holds for $t > 0$. Intuitively, this is because of two facts. First, as noted above, the information event process is independent from the type process. Second, $\{\tau_u = t\} = \{N_t - N_{t^-} = 1 \text{ and } N_u - N_t = 0\}$ only depends on increments of the information process at and after t, which are independent from the trader's information one instant before t, and hence independent from the predictable process $q_{t,u}$.

Relying on Lemma 1, (5) becomes:

$$\begin{aligned}
&\int_0^\infty e^{-(r+\rho)u}\mathbb{E}\left[v(\theta_u, s)\right] du + \int_0^\infty e^{-ru} \int_0^u \rho e^{-\rho(u-t)}\mathbb{E}\left[v(\theta_u, q_{t,u})\right]\, dt\, du \\
=&\int_0^\infty e^{-(r+\rho)u}\mathbb{E}\left[v(\theta_u, s)\right] du + \mathbb{E}\left[\int_0^\infty e^{-rt} \int_t^\infty e^{-(r+\rho)(u-t)}\mathbb{E}_t\left[v\left(\theta_u, q_{t,u}\right)\right]\, du\, \rho dt\right],
\end{aligned}$$

where $\mathbb{E}_t[\cdot]$ refers to the expectation conditional upon $\mathcal{F}_t$ and the second line follows from changing the order of integration and applying the law of iterated expectations. Applying the same logic to the expected present value of payments associated with a given holding plan, one obtains the next Lemma:

Lemma 2. *The inter-temporal payoffs associated with the holding plan $q_{t,u}$ is, up to a constant:*

$$V(q) = \mathbb{E}\left[\int_0^\infty e^{-rt}\int_t^\infty e^{-(r+\rho)(u-t)}\left\{\mathbb{E}_t\left[v(\theta_u, q_{t,u})\right] - \xi_u q_{t,u}\right\}du\,\rho dt\right], \tag{6}$$

where ξ_u is defined in (3).

The interpretation of equation (6) is the following. The outer expectation sign takes expectation over all time–t histories. The "$\rho\, dt$" term in the outer integral is the probability that an information event occurs during $[t, t+dt]$. Conditional on the time–t history and on an information event occurring during $[t, t+dt]$, the inner integral is the discounted expected utility of the holding plan until the next information event. At each point in time this involves the difference between a trader's expected valuation for the asset, $\mathbb{E}_t[v(\theta_u, q_{t,u})]$, and the opportunity cost of holding it, ξ_u. Finally, the discount factor applied to time u is adjusted by the probability $e^{-\rho(u-t)}$ that the next information event occurs after u.

3.4 Market clearing

In all what follows we focus on the case where all traders choose the same holding plan, which is natural given that traders are *ex–ante* identical.[17] Of course, while traders choose *ex–ante* the same holding plan, *ex–post* they realize different histories of N_t and θ_t, and hence different asset holdings.

The market clearing condition requests that, at each date $u \geq 0$, the cross-sectional average asset holding be equal to s, the per-capita asset supply. 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 can be written:

$$\mathbb{E}\left[q_{\tau_u,u}\right] = s. \tag{7}$$

[17]By *ex–ante* identical we mean that traders start with the same asset holdings and have identically distributed processes for information event and utility status.

for all $u \geq 0$. Integrating as in Section 3.3 against the distribution of τ_u, and keeping in mind that $q_{0,u} = s$, leads to our next lemma:

Lemma 3. *The time-u market clearing condition, (7), writes:*

$$\int_0^u \rho e^{-\rho(u-t)} \Big\{ (1-\mu_{ht})\mathbb{E}\left[q_{t,u} \,|\, \theta_t = \ell\right] + \mu_{ht}\mathbb{E}\left[q_{t,u} \,|\, \theta_t = h\right] - s \Big\} dt = 0. \tag{8}$$

This lemma states that the aggregate net demand of traders who experienced at least one information event is equal to zero. The first multiplicative term in the integrand of (8), $\rho e^{-\rho(u-t)}$, is the density of time–t traders, i.e., traders whose last information event occurred at time $t \in (0, u]$. The first and second terms in the curly bracket are the gross demands of time–t low– and high–valuation traders respectively. The last term in the curly bracket is their gross supply. It is equal to s because information events arrive at random, which implies that the average holding of time–t traders just before their information event equals the population average.

3.5 Equilibrium existence and uniqueness

We define an equilibrium to be a pair (q, p) subject to Conditions 1, 2 and 3 and such that: *i*) given the price path, the asset holding plan maximizes $V(q)$ given in (6), and *ii*) the market clearing condition (8) holds at all times.

Going back to the value $V(q)$, in equation (6), and bearing in mind that a trader can choose any function $u \mapsto q_{t,u}$ subject to Conditions 1 and 2, it is clear that the trader inter-temporal problem reduces to point-wise optimization. That is, a trader whose last information event occurred at time t chooses her asset holding at time u, $q_{t,u}$, in order to maximize the difference between her expected valuation for the asset and the corresponding holding cost:

$$\mathbb{E}_t\left[v(\theta_u, q_{t,u})\right] - \xi_u q_{t,u}. \tag{9}$$

Now, for all traders, utilities are strictly increasing for $q_{t,u} < 1$ and constant for $q_{t,u} \geq 1$. So, if one trader finds it optimal to hold strictly more than one unit at time u, then it must be that $\xi_u \leq 0$, implying that all other traders find it optimal to hold more than one unit. Inspecting equation (8), one sees that in that case the market cannot clear since $s < 1$. We conclude that:

Lemma 4. *In equilibrium,* $q_{t,u} \in [0, 1]$ *for all traders.*

To obtain further insights on holding plans, consider first a time–t high–valuation trader, i.e., a trader who finds out at time t that $\theta_t = h$. She knows her valuation for the asset will stay high forever. Hence

$$\mathbb{E}_t \left[v(\theta_u, q_{t,u}) \right] = v(h, q_{t,u}), \forall u \geq t. \tag{10}$$

Next, consider a time–t low–valuation trader, i.e. a trader who finds out at time t that $\theta_t = \ell$. This trader anticipates that her utility status will remain low by time u with probability $(1 - \mu_{hu})/(1 - \mu_{ht})$. Hence:

$$\mathbb{E}_t \left[v(\theta_u, q_{t,u}) \right] = q_{t,u} - \delta \frac{1 - \mu_{hu}}{1 - \mu_{ht}} \frac{q_{t,u}^{1+\sigma}}{1 + \sigma}, \forall q_{t,u} \in [0, 1] \tag{11}$$

Comparing (10) and (11), one sees that, for all asset holdings in $(0, 1)$, high–valuation traders have a uniformly higher marginal utility than low–valuation traders. Consequently, if some low–valuation trader finds it optimal to hold some asset, i.e. $\theta_t = \ell$ and $q_{t,u} > 0$, then all high–valuation traders find it optimal to hold one unit, i.e. $\theta_t = h$ implies that $q_{t,u} = 1$. Together with market clearing, this implies that

$$S_u \equiv \int_0^u \rho e^{-\rho(u-t)} \left(s - \mu_{ht} \right) dt > 0. \tag{12}$$

Indeed, S_u represent the residual supply held by low–valuation traders: the gross asset supply brought by all traders minus the unit demand of high–valuation traders, integrating across all traders with at least one information event. Note that the converse is also true: if $S_u > 0$, then since $q_{t,u} \leq 1$ for high–valuation traders, in order to clear the market we must have that $q_{t,u} > 0$ for some low–valuation trader. Given that S_u is first strictly positive for low values of u and then strictly negative for large values of u, we obtain:

Lemma 5. *Let T_f be the unique strictly positive solution of $S_u = 0$. Then:*

- *if $u \in (0, T_f)$ then, for all $t \in (0, u]$, $\theta_t = h$ implies $q_{t,u} = 1$;*
- *if $u \in [T_f, \infty)$ then, for all $t \in (0, u]$, $\theta_t = \ell$ implies $q_{t,u} = 0$.*

Next, consider the demand of high–valuation traders when $u > T_f$. We know from the previous lemma that low–valuation traders hold no asset. Thus, high–valuation traders must hold all the asset supply. Moreover, since $S_u < 0$, the market-clearing condition implies that some high–valuation traders must hold strictly less than one share. Keeping in mind that high–valuation traders have the same linear utility flow over $[0, 1]$, this implies they must be indifferent between any holding in $[0, 1]$. Thus we can state the following lemma.

Lemma 6. *For all $u > T_f$, the average asset holding of a high–valuation trader is*

$$\frac{\int_0^u \rho e^{-\rho(u-t)} s \, dt}{\int_0^u \rho e^{-\rho(u-t)} \mu_{ht} \, dt},$$

but the distribution of asset holdings across high–valuation traders is indeterminate.

All what is left to determine, then, is the demand of low–valuation traders when $u < T_f$. Taking first-order conditions when $\theta_t = \ell$ in (9), we obtain, given $q_{t,u} \in [0, 1]$:

$$q_{t,u} = 0 \quad \text{if } \xi_u \geq 1 \tag{13}$$

$$q_{t,u} = 1 \quad \text{if } \xi_u \leq 1 - \delta \frac{1 - \mu_{hu}}{1 - \mu_{ht}} \tag{14}$$

$$q_{t,u} = (1 - \mu_{ht})^{1/\sigma} Q_u \quad \text{if } \xi_u \in \left(1 - \delta \frac{1 - \mu_{hu}}{1 - \mu_{ht}}, 1\right), \quad \text{where } Q_u \equiv \left[\frac{1 - \xi_u}{\delta(1 - \mu_{hu})}\right]^{1/\sigma}. \tag{15}$$

Equation (13) states that low–valuation traders hold zero unit if the opportunity cost of holding the asset is greater than 1, their highest possible marginal utility, which arises when $q = 0$. Equation (14) states that low–valuation traders hold one unit if the opportunity cost of holding the asset is below the lowest possible marginal utility, which arises when $q = 1$. Lastly, equation (15) pins down a low–valuation trader's holdings in the intermediate interior case by equating to 0 the derivative of (9) with respect to $q_{t,u}$.

As argued earlier, before time T_f the holdings of some low–valuation trader must be strictly greater than 0: thus, by (13), we have $\xi_u < 1$. This implies that their holdings are determined by either (14) or 15. By the definition of Q_u, $\xi_u \leq 1 - \delta(1 - \mu_{hu})/(1 - \mu_{ht})$ if and only if $(1 - \mu_{ht})^{1/\sigma} Q_u \geq 1$. Hence, the asset demand defined by (14) and (15) can be written as

$$q_{t,u} = \min\{(1 - \mu_{ht})^{1/\sigma} Q_u, 1\}. \tag{16}$$

Substituting the demand from (16) into the market-clearing condition (8) and using the definition of S_u in (12), the following lemma obtains.

Lemma 7. *If $u \in (0, T_f)$, then for all $t \in (0, u]$, $\theta_t = \ell$ implies $q_{t,u} = \min\{(1 - \mu_{ht})^{1/\sigma} Q_u, 1\}$*

where:

$$\int_0^u (1-\mu_{ht}) \min\{(1-\mu_{ht})^{1/\sigma} Q_u, 1\} \rho e^{-\rho(u-t)}\, dt = S_u. \tag{17}$$

Equation (17) is a one-equation-in-one-unknown for Q_u that is shown in the proof appendix to have a unique solution. Taken together, Lemmas 5 to 7 imply:

Proposition 1. *There exists an equilibrium. The equilibrium asset allocation is unique up to the distribution of asset holdings across high–valuation traders after T_f, and is characterized by Lemma 5-7. The equilibrium price path is unique. It is increasing until T_f, constant thereafter, and solves the following ordinary differential equation:*

$$u \in (0, T_f): \quad rp_u - \dot{p}_u = 1 - \delta(1-\mu_{hu}) Q_u^{\sigma} \tag{18}$$

$$u \in [T_f, \infty): \quad p_u = \frac{1}{r}. \tag{19}$$

As in the perfect cognition case, the price deterministically increases until it reaches $1/r$. One difference is that, while under perfect cognition this recovery occurred at time T_s, with limited cognition it occurs at the later time $T_f > T_s$. For $u < T_f$, the time–u low–valuation traders are the marginal investors, and the equilibrium price is such that their marginal valuation is equal to the opportunity cost of holding the asset, as stated by (18). For $u > T_f$, the entire supply is held by high–valuation investors. Thus the equilibrium price is equal to the present value of their utility flow, as stated by (19).[18] The proposition is illustrated in Figure 2, Panel B, which plots the equilibrium price under limited cognition.

Note that for this numerical analysis we set the intensity of information events ρ to 250, which means that traders observe refreshed information on θ on average once a day.

3.6 Equilibrium properties

3.6.1 Welfare

To study welfare we define an asset holding plan to be *feasible* if it satisfies Conditions 1 and 2 as well as the resource constraint, which is is equivalent to the market-clearing condition (7). Furthermore, we say that an asset holding plan q *Pareto dominates* some other holding plan q'

[18] They must be indifferent between trading or not. This indifference condition implies that $1 - rp_u + \dot{p}_u = 0$. And, $p_u = 1/r$ is the only bounded and positive solution of this ODE.

if it is possible to generate a Pareto improvement by switching from $q^{'}$ to q while making time zero transfers among traders. Adapting the argument of footnote 9 to the limited cognition case, one easily sees that q Pareto dominates $q^{'}$ if and only if $W(q) > W(q^{'})$, where

$$W(q) = \mathbb{E}\left[\int_0^\infty e^{-ru} v(\theta_{\tau_u}, q_{\tau_u,u})\, du\right]. \tag{20}$$

The next proposition states that in our model the first welfare theorem holds:

Proposition 2. *The holding plan arising in the equilibrium characterized in Proposition 1 maximizes $W(q)$ among all feasible holding plans.*

The intuition for this result is that, in our setup, there are no "pecuniary externalities," in that the holdings constraints imposed by limited cognition (and expressed in conditions 1 and 2) do not depend on prices. These constraints translate into simple restrictions on the commodity space (conditions 1 and 2), allowing us to apply the standard proof of the first welfare theorem (see Mas-Colell, Whinston, and Green, 1995, Chapter 16)

3.6.2 Holdings

As stated in equation (16), for a trader observing at t that her valuation is low, the optimal holdings at time $u > t$ are (weakly) increasing in Q_u. Relying on the market clearing condition, the next proposition spells out the properties of Q_u.

Proposition 3. *The function Q_u is continuous, and such that $Q_{0^+} = s$ and $Q_{T_f} = 0$. Moreover, if*

$$s \leq \frac{\sigma}{1+\sigma} \tag{21}$$

Q_u is strictly decreasing with time. Otherwise, it is hump-shaped.

The economic intuition is the following. At time 0^+ the mass of traders with high–valuation is negligible. Therefore low–valuation investors have to absorb the entire supply. Hence, $Q_{0^+} = s$. At time T_f high valuation traders absorb all the supply. Hence, $Q_{T_f} = 0$.

When the per–capita supply of assets concerned by the shock s is low, so that (21) holds, the incoming flow of high–valuation traders reaching a decision at a given point in time is always large enough to accommodate the supply from low–valuation traders. Correspondingly,

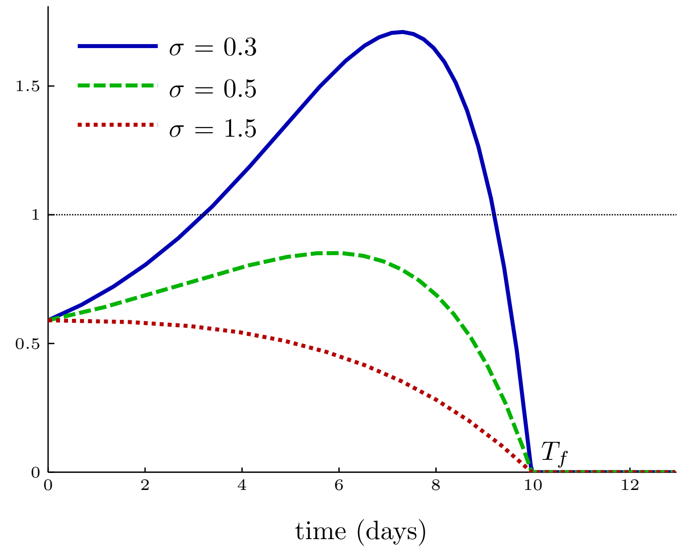

Figure 3: The function Q_u for various values of σ

in equilibrium low valuation traders sell a lump of their assets when they reach a decision and then smoothly unwind their inventory until the next information event.

In contrast, when s is large so that (21) fails to hold, the liquidity shock is more severe. Hence, shortly after the initial aggregate shock, the inflow of high–valuation traders is not large enough to absorb the sales of low–valuation traders who currently reach a decision. In equilibrium, some of these sales are absorbed by "early" low–valuation traders who reached a decision at time $t < u$ and have not had another information event. Indeed, these "early" low–valuation traders anticipate that, as time goes by, their institution is more likely to have recovered. Thus, their expected valuation (in the absence of an information event) increases with time and they find it optimal to buy if their utility is not too concave, i.e., if σ is not too high. Correspondingly, near time zero, Q_u is increasing, as depicted in Figure 3 for $\sigma = 0.3$ and 0.5.

Combining Lemma 5, Lemma 6, Lemma 7 and Proposition 3, one obtains a full characterization of the equilibrium holdings process, which can be compared to its counterpart in the unlimited cognition case. In both cases, agents initially hold s. But, when cognition is not limited, as long as an institution has not recovered from the shock, its holdings decline smoothly, and, as soon as it recovers, they jump to 1. Trading histories are quite different with limited cognition. First institutions' holdings remain constant until their first information event. Then, if at her first information event the trader learns that her institution has a low valuation, she sells a lump. After that and before the next jump of their information process, if (21) does not hold the trader progressively buys back, and then eventually sells out. This process continues until she finds out her valuation has recovered. Such round–trip trades don't arise in the unbounded cognition case.

3.6.3 Trading volume

Because they result in round–trip trades, hump–shaped asset holding plans generate extra trading volume relative to the unlimited cognition case. Specifically, consider a trader who, at two consecutive information events t_1 and t_2, discovers that she has a low–valuation. During the time period $(t_1, t_2]$ she trades an amount of asset equal to

$$\int_{t_1}^{t_2} \left| \frac{\partial q_{t_1,u}}{\partial u} \right| du + \left| q_{t_2,t_2} - q_{t_1,t_2} \right| . \tag{22}$$

The first term in (22) is the flow of trading between time t_1 and time t_2 dictated by $q_{t_1,u}$, the time–t_1 holding plan. The second term is the lumpy adjustment at time t_2.

Note that (16) implies that, wherever $|\partial q_{t_1,u}/\partial u|$ is not 0, it has the same sign as Q'_u. Note also that, because at time t_2 the trader observes that the institution has still not recovered, $q_{t_2,t_2} < q_{t_1,t_2}$. Hence, if Q_u is decreasing, (22) is equal to $q_{t_1,t_1} - q_{t_2,t_2}$. In contrast, if Q_u is increasing, the trading volume between t_1 and t_2 is

$$q_{t_1,t_2} - q_{t_1,t_1} + q_{t_1,t_2} - q_{t_2,t_2} = \underbrace{2\left(q_{t_1,t_2} - q_{t_1,t_1}\right)}_{\text{round trip trade}} + q_{t_1,t_1} - q_{t_2,t_2} \tag{23}$$

Since, $q_{t_1,t_2} > q_{t_1,t_1}$, (23) is greater than $q_{t_1,t_1} - q_{t_2,t_2}$. The first term in the equation is the extra volume created by the round–trip trade: the purchase of $q_{t_1,t_2} - q_{t_1,t_1}$ during (t_1, t_2) followed by a sale of the same quantity at time t_2. The resulting extra trading volume is illustrated in Figure 4. One might wonder if this extra volume goes to 0 when cognition frictions vanish. The next proposition shows that it is not case.

Proposition 4. *When ρ goes to infinity, the equilibrium price and allocation converge to their unlimited cognition counterparts. Relative to unlimited cognition, the extra volume at time t with limited cognition converges to*

$$\gamma \max \left\{ \frac{s - \mu_{ht}}{\sigma} - (1-s), 0 \right\} .$$

In particular, this asymptotic extra volume is zero if condition (21) holds, or if t is large enough,

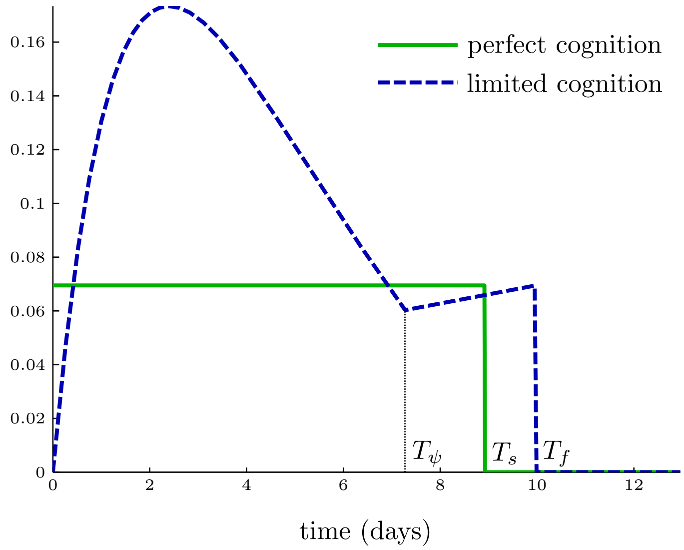

Figure 4: The trading volume unlimited (green) and limited cognition (blue), when $\sigma = 0.3$. In the figure, T_ψ denotes the argument maximum of Q_u.

and is strictly positive otherwise.

The proposition shows that if low–valuation traders' asset holdings are hump shaped, volume is larger in the $\rho \to \infty$ limit than with unlimited cognition. The proposition also reveals the crucial role of the curvature parameter, σ: As illustrated in Figure 3, when σ is small and utility is close to linear, the hump–shape pattern in holdings is quite pronounced. This leads to very large extra trading volume

3.6.4 Prices

Proposition 1 implies that, from time T_s to time T_f, the price under limited cognition is lower than its counterpart with unlimited cognition, $1/r$. By continuity, the price is lower under cognition limits just before T_s. The next proposition states that this ranking of prices holds at all times if $s < \sigma/(\sigma + 1)$, but not necessarily otherwise.

Proposition 5. *If (21) holds, then at all time the price is strictly lower with limited cognition. But if s is close to 1 and σ is close to 0, then at time 0 the price is strictly higher with limited cognition.*

That the price would be higher without cognition limits sounds intuitive. Unbounded cognition enables traders to continuously allocate the asset to those who value it the most. Such an efficient allocation could be expected to raise the price, and this is indeed what happens when (21) holds. But, as stated in Proposition 5, there are cases where the price can be higher when cognition is limited than when it is unbounded. The intuition is the following. Around time zero, marginal traders have low–valuation. With limited cognition, low–valuation traders have a higher marginal utility because they take into account the possibility that they may have switched to high–valuation. Consequently, they demand more assets, which tends to push up prices. This effect is stronger when low–valuation traders are "marginal" for a longer period, that is, when s and the shock is more severe, and when their utility flow is not too concave, that is, when σ is low.

4 Market technologies and order placement policies

So far, our characterization of equilibrium was cast in abstract terms, such as holding plans and market clearing. We now study how these holding plans can be implemented with realistic market instruments. In doing so, we focus on electronic order driven markets. Such venues are the major trading mechanism for stocks around the world (e.g., in the US Nasdaq and the NYSE, and in Europe Euronext, the London Stock Exchange and the Deutsche Börse.) In these markets, traders can place limit sell (resp. buy) orders requesting execution at prices at least as large (resp. low) as their limit price. These orders are stored in the book, until they are executed, canceled or modified. Traders can also place market orders, requesting immediate execution. A limit sell order standing in the book is executed, at its limit price, when hit by an incoming buy order (either a market order or a limit buy order with a higher price limit), if there are no unexecuted sell orders in the book at lower prices, or at the same price but at an earlier point in time. The case of a limit buy order is symmetric.

There are multiple possible implementations.[19] For instance, there is a trivial implementation, where all traders desiring to change their holdings continuously submit market orders or limit orders at the current equilibrium price. While it has the advantage of simplicity, this implementation is at odds with important stylized facts; for example it leads to an empty limit order book. To narrow down the set of possible implementations while giving rise to realistic dynamics, we restrict our attention to the case where market participants alter their trading strategies only when their information process jumps. This is quite natural in our framework, and it follows if, between two jumps of their information process, traders' cognitive resources are devoted (at least in large part) to the complex task of assessing the liquidity status of their

[19] This is similar to the multiplicity of implementations in optimal contracting models.

institution, which leaves little opportunity to alter trading plans.

4.1 Implementing the equilibrium with limit orders and algorithms

First we assume that, when their information process jumps, in addition to market and limit orders, traders can place trading algorithms. The latter are computer programs feeding orders in the market as time goes by, in response to pre-specified future changes in market variables.[20] In keeping with our limited–cognition assumption, we do not allow algorithms to condition their orders on changes in the liquidity status of the institution occurring between jumps of its information process. Thus, while both the limit orders and the orders triggered by algorithms satisfy condition 1, they enable trades to happen without direct human intervention while the trader is engaged in information collection and processing.[21]

It is straightforward to implement the holding plan of a high–valuation traders. Before T_f, they place a market (or marketable limit) order to buy as soon as their information process reveals their institution has recovered from the shock. At that point in time, since they now have their optimal holdings, they cancel any limit order they would have previously placed in the book.

The case of low–valuation traders is more intricate. Indeed, one must bear in mind that the equilibrium price given in Proposition 1 is strictly increasing over $(0, T_f)$. This implies that any limit order to buy submitted at time t is either immediately executed (if the limit price is greater than p_t) or never executed (if the limit price is lower than p_t). Consequently, if a trader only places (limit or market) orders when her information process jumps, she cannot implement increasing holding plans. In contrast, she can implement decreasing holding plans, by placing at time t limit orders to *sell* at price $p_u > p_t$. Now, Proposition 3 states that equilibrium holdings are decreasing if and only if condition (21) holds. This leads to the following Proposition:

Proposition 6. *The equilibrium characterized in Proposition 1 can be implemented by traders placing market and limit orders (only) when their information process jumps if and only if (21) holds.*

[20] Because we do not have any aggregate uncertainty, market-level variables (price, volume, quote...) are in our model deterministic functions of time. Thus, conditioning on time is enough to make the algorithm also depend on the state of the market. In a more general model with aggregate uncertainty, one would need to explicitly allow to depend on market-level variables at time u.

[21] This is in line with the view of Harris (2003), who argues that "Limit orders represent absent traders [enabling them] to participate in the markets while they attend to business elsewhere."

If (21) does not hold, asset holding plans are hump shaped. As argued above, in order to implement the increasing branch of the hump, a trader cannot use limit buy orders: instead, she must rely on algorithms.

For concreteness, consider the example illustrated in Figure 5. Interpret t_1 and t_2 as the time of two consecutive jumps of the information process of a trader, both revealing low valuation. At t_2 the trader's asset holdings undergo a discrete downward jump, from q_{t_1,t_2} to q_{t_2,t_2}. Then they continuously increase, reach a flat, and finally decrease again.

To implement this holding plan, the trader places a market sell order at time t_2 and also uses a schedule of limit sell orders. As shown on Figure 5, the limit sell orders placed at time t_2 start executing before those placed at time t_1. Since the equilibrium price process is increasing, this means that these orders are placed at lower prices than the previous ones. The figure also shows that the slope of the holdings of the agent declines less steeply for the holding plan set at time t_2 than for its t_1 counterpart. This reflects that the quantity offered in the book at these prices is lower for the plan set at t_2 than for the plan set at t_1. Thus, to implement the new holding plan, at t_2 the trader cancels some of the orders placed at t_1. Finally note that, at t_2, the trader also modifies the trading algorithm generating the purchases necessary to implement the increasing part of her holding plan. This can be interpreted in terms of human intervention resetting the parameters of the selling algorithm. This discussion is summarized in our next proposition.

Proposition 7. *Consider a trader observing low valuation at time $t < T_f$. If (21) does not hold she can implement the optimal holding plan arising in Proposition 1 by placing market– and limit-sell orders at t and programming her trading algorithm to trigger market buy orders at times $u > t$.*

Turning back to Figure 5, now consider t_1 and t_2 as the times at which the information processes of two traders jump, each time revealing low valuation. The figure illustrates that the late trader starts selling before the early trader implying that the limit sell orders of the late trader are placed at lower prices than the limit sell orders of the early trader, i.e., there is undercutting. To see why this is optimal, compare the time–u expected valuations of the time–t_1 trader and the time–t_2 trader, assuming that for both of them there has been no new information event by time u. For the trader who observed low valuation at time t_1 the probability that her valuation is high now is: $1-e^{-\gamma(u-t_1)}$. For the other trader it is $1-e^{-\gamma(u-t_2)}$, which is lower since $t_2 > t_1$. Hence the expected valuation of the time–t_1 trader is greater than

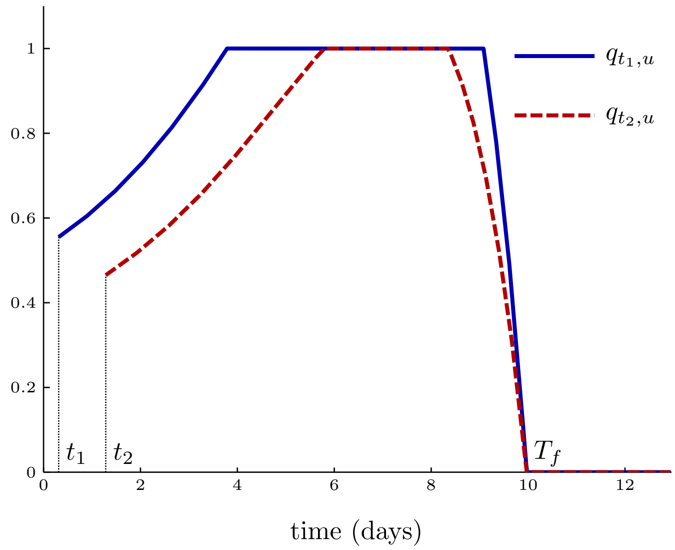

Figure 5: The holding plans of an early low–valuation trader and of a late low–valuation trader, when $\sigma = 0.3$

that of the time–t_2 trader. This is why the time–t_1 trader sells later.[22]

Taking stock of the above results, we now describe the overall market dynamics prevailing when (21) does not hold and $Q_u > 1$ for some u. There are four successive phases in the market, as illustrated in Figure 6.

- There exists a a time $T_1 < T_f$ such that, from time 0 to T_1, low valuation traders place limit sell orders at lower and lower prices, i.e., there is undercutting. These limit orders accumulate in the book, without immediate execution. Correspondingly, the best ask decreases and the depth on the ask side of the book increases. During this period, low valuation traders also place market sell orders, which are executed against buy orders stemming from high–valuation traders and algorithms.

- Denoting by $T_\psi \in (T_1, T_f)$ the time at which Q_u achieves its maximum, during $[T_1, T_{psi}]$ the best ask quote remains constant and the depth on the ask side of the book declines. Indeed, during this period, low valuation traders stop undercutting the best quote, while high valuation traders cancel their limit orders. During this second phase of the market, there are still no executions at the best ask, and trades are initiated by low valuation traders placing market sell orders.

- The first phase is between T_ψ and T_f. During this period, high valuation traders hit the

[22]This goes along the same lines as the intuition why low valuation traders placing trading plans at t will initially sell and then buy from agents placing their trading plans later.

limit sell orders outstanding in the book. Correspondingly, the depth on the ask side continues to decline, and the best ask price goes up.

- Finally, after time T_f, the market has recovered from the shock, and the price remains constant at $1/r$.

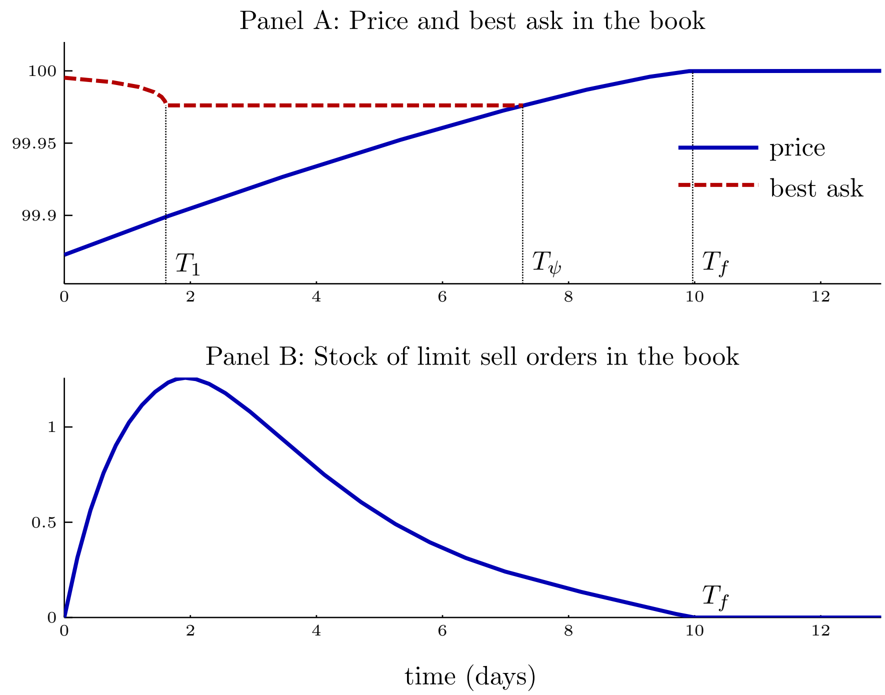

Figure 6: Price dynamics (Panel A) and limit order book activity (Panel B), when $\sigma = 0.3$

The algorithmic trading strategies generated by our model are in line with stylized facts and empirical findings. That algorithms progressively build up an increasing position via successive buy orders can be interpreted in terms of order splitting. That they buy progressively as the price deterministically trends upwards can be interpreted in terms of short–term momentum trading. That after the liquidity shock they build up long positions which they will eventually unwind is in line with the findings by Brogaard (2010) that algorithms don't withdraw after large price drops and benefit from price reversals. Our theoretical resuts are also is in line with the empirical findings by Hendershott and Riordan (2010) that trading algorithms provide liquidity when it is scarce and rewarded. Indeed, the strategies followed by our algo–traders (who buy initially while simultaneously placing limit orders to sell to be executed later) is a form of market–making, similar to that arising in Grossman and Miller (1988). Note however that, while in Grossman and Miller (1988) some market participants are exogenously assumed to be liquidity providers and other liquidity consumers, in our model all participants are identical ex–ante, yet they play different roles in the market because of differences in the realizations of their information and valuation processes.

4.2 Equilibrium when the menu of orders is not rich enough

What happens if trading algorithms are not available and traders can only place limit and market orders at the time of information events? Suppose for now that the equilibrium price is increasing (which will turn out to be the case in equilibrium). As mentioned in Subsection 4.1 if traders place orders only when their information process jumps, this precludes them from implementing increasing holding plans. Instead, they can only implement decreasing holding plans. Therefore, if condition (21) does not hold, the equilibrium will differ from that arising in Proposition 1. Instead, the equilibrium is as in the next proposition.

Proposition 8. *If $s > \sigma/(1+\sigma)$ and traders can only place limit and market orders when their information process jumps, there exists an equilibrium in which the price path is strictly increasing over $(0, T_f)$, and equal to $1/r$ over $[T_f, \infty)$. High-valuation traders, and low–valuation traders after T_f, follow the same asset holding plan as in Proposition 1. For low–valuation traders before T_f, the optimal holding plan $q_{t,u}$ is continuous in (t, u) and strictly less than 1. Moreover, there exists $T_\phi \in (0, T_f)$ and a strictly decreasing function $\phi : (0, T_\phi] \mapsto \mathbb{R}_+$, such that:*

- *If $t \in (0, T_\phi]$, then $q_{t,u}$ is constant for $u \in [t, \phi_t]$, and strictly decreasing for $u \in (\phi_t, T_f)$.*
- *If $t \in (T_\phi, T_f]$, then $q_{t,u}$ is strictly decreasing for $u \in (t, T_f)$.*

While the Proposition describes the equilibrium in abstract terms, in the the appendix we provide closed–form analytical solutions for all relevant equilibrium objects. One may wonder whether the equilibrium in Proposition 8 is unique. We provide a partial answer to this question in the supplementary appendix to this paper. We show that the equilibrium of Proposition 8 is unique in the class of Markov equilibria, i.e., equilibria where traders find it optimal to choose holding plans which only depend on the information-event time (the current aggregate state) and their current utility status (their current idiosyncratic state).[23]

[23] In contrast with the earlier work of Biais and Weill (2009), our proof does not make any *a priori* monotonicity restriction on the price path. Instead, we consider general and possibly non-monotonic price paths. We then

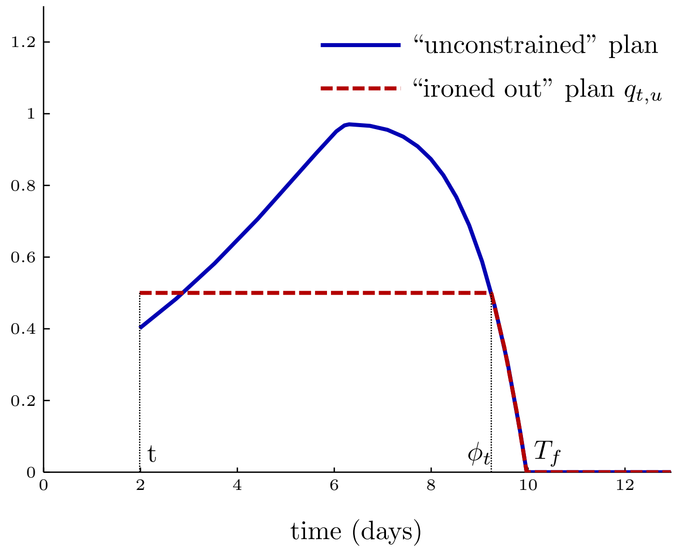

Figure 7: Holding plans of algorithmic versus limit order traders, when $\sigma = 0.3$, given the equilibrium price of Proposition 8.

The intuition for Proposition 8 is the following. When $s > \sigma/(1+\sigma)$, time–$t < T_\phi$ low–valuation traders would choose hump-shaped holding plans if their holdings were not constrained to be decreasing (the solid curve in Figure 7). Faced with the constraint of choosing a decreasing holding plan, they "iron" the increasing part of the hump-shaped plans (the dashed curve in Figure 7).

To implement the holding plans of the proposition, time–t low–valuation traders place market sell orders, as well as schedules of limit sell orders, which start executing at time ϕ_t. Proposition 8 also implies that, similarly to the case where traders could use algorithms, there is undercutting in equilibrium: since ϕ_t is strictly decreasing, successive traders place limit sell orders at lower and lower prices.[24]

Note that, both in Proposition 1 and Proposition 8, the time at which the price fully recovers from the liquidity shock is T_f, the time at which the residual supply S_u (defined in (12)) reaches 0. This is because S_u is a function of the total quantity of the asset brought to the market and of the unit demand of high–valuation traders, both of which are unaffected by whether low–valuation traders can use algorithms or not. The next proposition offers further insights

show, via elementary optimality and market clearing considerations, that the preference dynamics and the focus on Markov equilibria imply that the price path is continuous, strictly increasing until time T_f, and flat after time T_f.

[24]Note that in Proposition 8 we obtain undercutting for all (s, σ) such that $\sigma/(1+\sigma) < s$. In Proposition 1, by contrast, we obtain undercutting for the smaller set of parameters such that the maximum of Q_u is greater than one.

into the comparison of the price paths are arising in Proposition 8 and Proposition 1.

Proposition 9. *When $s > \sigma/(1+\sigma)$:*

- *The price arising in Proposition 1 is strictly lower than its counterpart in Proposition 8 around time zero for ρ close to* 0.
- *The price arising in Proposition 1 is strictly higher than its counterpart in Proposition 8 between T_ϕ and ϕ_0;*
- *The price is identical in Proposition 1 and in Proposition 8 after time ϕ_0.*

The Proposition shows that, shortly after the liquidity shock, the price can be lower when institutions use trading algorithms.[25] The intuition is the following. When they can't use algorithms, traders know they won't be able to buy back before their next information event. Hence they initially sell less, which reduces the selling pressure on the price. Consequently the price can be higher than when traders can use algorithms. For an outside observer this could suggest that algorithmic trading destabilizes markets, by amplifying the price drops due to liquidity shocks. However, inferring aggregate welfare effects based on such price movement is misleading. Indeed, Proposition 2 implies that the equilibrium arising with algorithms Pareto dominates that arising when traders can only place limit and market orders when their information process jumps.

Figure 8 plots the trading volume (upper panel) and the volume of limit orders outstanding in the book (lower panel) in Propositions 1 and 8.[26] Trading volume is higher when institutions can use algorithms. This reflects the additional trading volume generated by round trip trades. The stock of limit orders in the book has the same shape in Propositions 1 and 8. The book is filled progressively as limit orders to sell are placed by low–valuation traders. Then, cancellations and executions lead to a decrease in the stock of orders in the book. During the early phase, the amount of limit orders outstanding is higher in Proposition 1 than in Proposition 8. Indeed, with algorithms low valuation traders buy after their information even, this induces them to place more limit order to sell, to unwind this position towards the end of the price recovery process.

[25] While we are only able to establish this result analytically for small ρ, our numerical calculations suggests that it can hold for larger value of ρ. In particular, it does hold for the value $\rho = 250$ that we have chosen for the numerical calculations presented in this paper as well as in our supplementary appendix.

[26] Analytical formulas used in the calculations are gathered in Section X of our supplementary appendix.

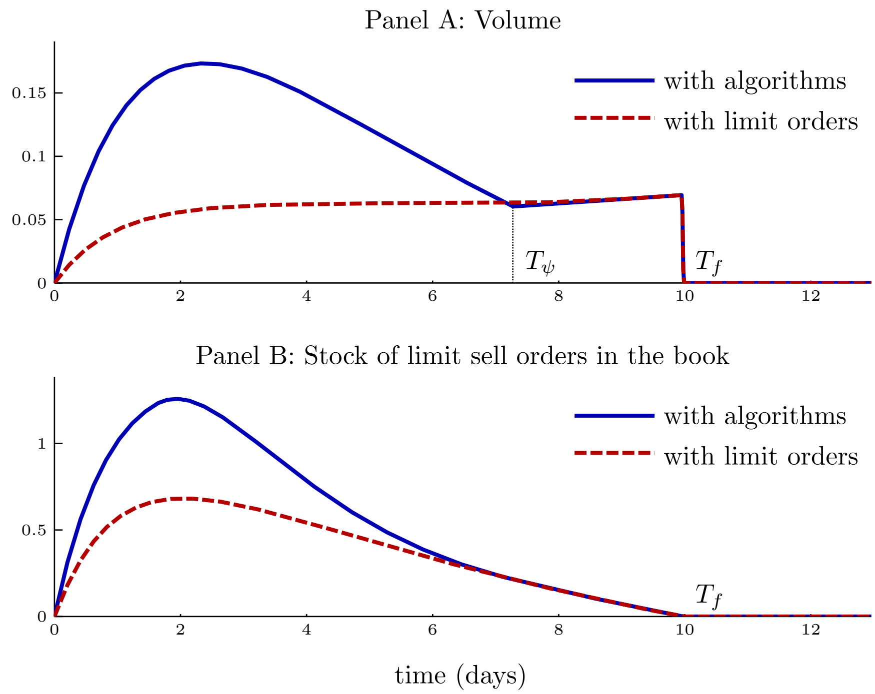

Figure 8: Trading volume (upper panel) and volume of limit orders outstanding in the book (lower panel), when $\sigma = 0.3$

5 Conclusion

This paper studies the reaction of traders and markets to liquidity shocks under cognition limits. As in earlier work, we model the aggregate liquidity shock as a transient decline in the valuation of the asset by all participants. While institutions recover from the shock at random times, traders with limited cognition observe the status of their institution only when their own information process jumps. We interpret this delay as the time it takes traders to collect and process information about positions, counterparties, risk exposure and compliance. After the aggregate liquidity shock, the equilibrium price immediately drops. It then gradually recovers. Traders who find their institution has a low valuation sell via market orders, and then progressively buy back, at relatively low prices, while simultaneously placing limit orders to sell later when the price will have recovered. We compare the case where the traders can use algorithms to trade while investigating the liquidity status of their institution, to the case where algorithms are not available and traders can only place market and limit orders when their information process jumps.

Our analysis suggests that trading algorithms can play a useful role by facilitating market–making. In the equilibrium we characterize, they can seem to destabilize markets, to the extent that they lead to lower prices than those prevailing without algorithms. Yet the equilibrium prevailing when traders can use algorithms Pareto dominates that prevailing if traders can only place orders when their information process jumps. In our setup the optimality of equilibrium reflects the absence of pecuniary externalities, due to the fact that constraints on holdings imposed by cognition limits don't depend on prices.

Note however that in our analysis there is no adverse selection. In a more general model where some "slow" traders could not use algorithms, while "fast traders" could, information asymmetries could arise.[27] In this context, algorithmic trading could inflict negative externalities on slow traders, and equilibrium might no longer be optimal. Biais, Foucault and Moinas (2010) analyze these issues in a one period model. Because of their static setup, however, they cannot consider rich dynamic order placement policies such as those arising in the present model. An interesting, but challenging avenue of further research would be to extend the present model to the asymmetric information case. In doing so one could take stock of the economic and methodological insights of Goettler, Parlour, and Rajan (2009) and Pagnotta (2010) who study dynamic order placement under asymmetric information.

[27]Consistent with the view that algorithmic traders have superior information Hendershott and Riordan (2010) and Brogaard (2010) find empirically that algorithmic trades have greater permanent price impact than slow trades and lead price discovery.

A Proofs

A.1 Proof of Lemma 1

We first note that, by the law of iterated expectations:

$$\mathbb{E}\left[v(\theta_u, q_{t,u}) \,|\, \tau_u = t\right] = \mathbb{E}\left[\mathbb{E}\left[v(\theta_u, q_{t,u}) \,|\, \mathcal{F}_{t^-}, \tau_u = t\right] \,\middle|\, \tau_u = t\right] \tag{A.1}$$

where, as usual, $\mathcal{F}_{t^-}$ is the sigma algebra generated by all the $\mathcal{F}_z$, $z < t$, and represents the trader information "one instant prior to t." Now recall that

$$v(\theta_u, q_{t,u}) = q_{t,u} - \mathbb{I}_{\{\theta_u = \ell\}} \delta \frac{q_{t,u}^{1+\sigma}}{1+\sigma}.$$

Therefore, the inner expectation on the right-hand side of (A.1) writes as:

$$\begin{aligned}\mathbb{E}\left[q_{t,u} - \mathbb{I}_{\{\theta_u = \ell\}} \delta \frac{q_{t,u}^{1+\sigma}}{1+\sigma} \,\middle|\, \mathcal{F}_{t^-}, \tau_u\right] &= q_{t,u} - \mathbb{E}\left[\mathbb{I}_{\{\theta_u = \ell\}} \,|\, \mathcal{F}_{t^-}, \tau_u\right] \frac{q_{t,u}^{1+\sigma}}{1+\sigma} \\ &= q_{t,u} - \mathbb{E}\left[\mathbb{I}_{\{\theta_u = \ell\}} \,|\, \mathcal{F}_{t^-}\right] \frac{q_{t,u}^{1+\sigma}}{1+\sigma} \end{aligned} \tag{A.2}$$

where the first equality follows because $q_{t,u}$ is $\mathcal{F}_t$-predictable, and thus measurable with respect to $\mathcal{F}_{t^-}$ (see Exercise E10, Chapter I, in Brémaud, 1981). The second equality, on the other hand, follows because the type process is independent from the information event process: this allows to freely add or remove any information generated by the information event process from the conditioning information.

Now the random variable of equation (A.2) is $\mathcal{F}_{t^-}$-measurable. Since $\{\tau_u = t\} = \{N_t - N_{t^-} = 1 \text{ and } N_u - N_t = 0\}$ and because the information event process has independent increment and is independent from the type process, it follows that $\{\tau_u = t\}$ is independent $\mathcal{F}_{t^-}$. Thus, the expectation of (A.2) conditional on $\{\tau_u = t\}$, is equal to its unconditional expectation, which proves the claim.

A.2 Proof of Lemma 2

All what is left to derive is the expression for the inter-temporal payments associated with a given holding plan $q_{t,u}$. For this we let $\tau_0 \equiv 0 < \tau_1 < \tau_2 < \ldots$ denote the sequence of information events. For accounting purposes, we can always assume that, at this n-th information event, the trader sells 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 the holding plan can be written:

$$\mathbb{E}_0\left[-p_{\tau_1} e^{-r\tau_1} q_{0,\tau_1} + \sum_{n=1}^{\infty} \left\{ e^{-r\tau_n} p_{\tau_n} q_{\tau_n,\tau_n} + \int_{\tau_n}^{\tau_{n+1}} p_u \, dq_{\tau_n,u} e^{-ru} - e^{-r\tau_{n+1}} p_{\tau_{n+1}} q_{\tau_n,\tau_{n+1}} \right\}\right].$$

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

$d/du(e^{-ru}p_u) = -\xi_u$. This leads to:

$$\mathbb{E}_0\left[-p_{\tau_1}e^{-r\tau_1}q_{0,\tau_1} + \sum_{n=1}^{\infty}\int_{\tau_n}^{\tau_{n+1}} e^{-ru}\xi_u q_{\tau_n,u}\,du\right] = -p_0 s + \mathbb{E}_0\left[\sum_{n=0}^{\infty}\int_{\tau_n}^{\tau_{n+1}} e^{-ru}\xi_u q_{\tau_n,u}\,du\right]$$
$$= \text{constant} + \mathbb{E}_0\left[\int_0^{\infty} e^{-ru}\xi_u q_{\tau_u,u}\,du\right] = \text{constant} + \mathbb{E}_0\left[\int_0^{\infty} e^{-rt}\int_t^{\infty}\rho e^{-(r+\rho)(u-t)}\xi_u q_{t,u}\,du\,dt\right].$$

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 contact time before u; the last equality follows from the The move from the third to the fourth line follows the same steps as for the present value of utility flows.

A.3 Preliminary results about Q_u

We start with preliminary results that we use repeatedly in this Appendix. For the first preliminary result, let $\Psi(Q) \equiv \inf\{\psi \geq 0 : (1-\mu_{h\psi})^{1/\sigma}Q \leq 1\}$, and $\psi_u \equiv \Psi(Q_u)$. We have:

Lemma A.1 (Preliminary results about Q_u)**.** *Equation (17) has a unique solution, Q_u. Moreover, $0 \leq Q_u < (1-\mu_{hu})^{-1/\sigma}$ and Q_u is continuously differentiable with*

$$Q'_u = \frac{e^{\rho u}(s-\mu_{hu}) - e^{\rho u}\left(1-\mu_{hu}\right)^{1+1/\sigma}Q_u}{\int_{\psi_u}^{u} e^{\rho t}\left(1-\mu_{ht}\right)^{1+1/\sigma}\,dt}. \tag{A.3}$$

In the proof of the Lemma, in Section A.11.1, we first show existence, uniqueness, and the inequality $Q_u < (1-\mu_{hu})^{1/\sigma}$ using the monotonicity and continuity of equation (17) with respect to Q. We then show that the solution is continuously differentiable using the Implicit Function Theorem.

For the second preliminary result, consider equation (17) after removing the min operator in the integral:

$$\int_0^u \rho e^{-\rho(u-t)}\left(1-\mu_{ht}\right)^{1+1/\sigma}\overline{Q}_u\,dt = S_u \Longleftrightarrow \quad \overline{Q}_u \equiv \frac{\int_0^u e^{\rho t}\left(s-\mu_{ht}\right)\,dt}{\int_0^u e^{\rho t}\left(1-\mu_{ht}\right)^{1+1/\sigma}\,dt}. \tag{A.4}$$

Now, whenever $\overline{Q}_u \leq 1$, it is clear that $\overline{Q}_u$ also solves equation (17). Given that the solution of (17) is unique it follows that $Q_u = \overline{Q}_u$. Conversely if $Q_u = \overline{Q}_u$, subtracting (17) from (A.4) shows that:

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

Since the integrand is positive, this can only be true if $(1-\mu_{ht})^{1/\sigma}\overline{Q}_u = \min\{(1-\mu_{ht})^{1/\sigma}\overline{Q}_u, 1\}$ for almost all t. Letting $t \to 0$ delivers $\overline{Q}_u \leq 1$. Taken together, we obtain:

Lemma A.2 (A useful equivalence)**.** $\overline{Q}_u \leq 1$ *if and only if* $Q_u = \overline{Q}_u$.

The next Lemma, proved in Section A.11.2, provides basic properties of $\overline{Q}_u$:

Lemma A.3 (Preliminary results about $\overline{Q}_u$)**.** *The function $\overline{Q}_u$ is continuous, satisfies $\overline{Q}_{0^+} = s$, $\overline{Q}_{T_f} = 0$. It is strictly decreasing over $(0, T_f]$ if $s \leq \sigma/(1+\sigma)$ and hump-shaped otherwise.*

A.4 Proof of Proposition 1

The asset holding plans are determined according to Lemmas 5-7. The asset holding plan of high-valuation traders are uniquely determined for $u \in (0, T_f]$, and are indeterminate for $u > T_f$. The holding plan of low-valuation investor is uniquely determined, as Lemma A.1 shows that equation (17) has a unique solution. Now, turning to the price, the definition of Q_u implies that the price solves $rp_u = 1 - \delta(1-\mu_{hu})Q_u^\sigma + \dot{p}_u$ 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 $p_{T_f} = 1/r$, ODE (18) 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 from Lemma A.1 that $0 < Q_u < (1-\mu_{hu})^{-1/\sigma}$, 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 $q_{t,u} = \min\{(1-\mu_{ht})^{-1/\sigma} Q_u, 1\}$. 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 the problem of low-valuation traders is clearly solved by $q_{t,u} = 0$. The asset market clears at all dates by construction.

We already know that the price is equal to $1/r$ for $t > T_f$ so the last thing to show is that it is strictly increasing for $t < T_f$. Letting $\Delta_u \equiv (1-\mu_{hu})^{1/\sigma} Q_u$ for $u \leq T_f$, and $\Delta_u = 0$ for $u \geq T_f$. In Section A.11.3 we show that:

Lemma A.4. *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 $u \in (0, T_f)$.

A.5 Proof of Proposition 2

Let us start with a preliminary remark. By definition, any feasible allocation $\tilde{q}$ satisfies the market–clearing condition $\mathbb{E}_0 \left[\tilde{q}_{\tau_u, u}\right] = s$. Taken together with the fact that $\xi_u = rp_u - \dot{p}_u$ is deterministic, this implies:

$$\mathbb{E}\left[\int_0^\infty e^{-ru} q_{\tau_u, u} \xi_u \, du\right] \tag{A.5}$$

$$= \int_0^\infty e^{-ru} \mathbb{E}\left[q_{\tau_u, u}\right] \xi_u \, du = \int_0^\infty e^{-ru} s \xi_u \, du = p_0 s. \tag{A.6}$$

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 of the objective (20). But, by (A.6), this asset holding plan has the same cost as the equilibrium asset holding plan. Subtracting the inter-temporal cost (A.6) from the inter-temporal utility (4), we obtain that $V(q') > V(q)$, which contradicts individual optimality.

A.6 Proof of Proposition 3

Taken together, Lemma A.2 and A.3 immediately imply that

Lemma A.5. *The function Q_u satisfies $Q_{0^+} = s$, $Q_{T_f} = 0$. If $s \leq \sigma/(1+\sigma)$, then it is strictly decreasing over $(0, T_f]$. If $s > \sigma/(1+\sigma)$ and $\overline{Q}_u \leq 1$ 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 $s > \sigma/(1+\sigma)$ and $\overline{Q}_u > 1$ for some $u \in (0, T_f]$. In that case, note that for u small and u close to T_f, we have that $\overline{Q}_u < 1$. Given that $\overline{Q}_u$ is hump-shaped, it follows that the equation $\overline{Q}_u = 1$ has two solutions, $0 < T_1 < T_2 < T_f$. For $u \in (0, T_1]$ (resp. $u \in [T_2, T_f]$), $\overline{Q}_u \leq 1$ and is increasing (resp. decreasing), and thus Lemma A.2 implies that $Q_u = \overline{Q}_u$ and increasing (resp. decreasing) as well.

In particular, since we know from Lemma A.1 that Q_u is continuously differentiable, we have that $Q'_{T_1} > 0$ and $Q'_{T_2} < 0$. Thus, Q_u changes sign at least once in (T_1, T_2). To conclude, in Section A.11.4 we establish:

Lemma A.6. *The derivative Q'_u changes sign only once in (T_1, T_2).*

A.7 Proof of Proposition 4

In Section A.11.5 we prove the following asymptotic results:

Lemma A.7. *As ρ goes to infinity:*

$$T_f(\rho) \downarrow T_s \tag{A.7}$$

$$Q_u(\rho) = \frac{s - \mu_{hu}}{(1-\mu_{hu})^{1+1/\sigma}} - \frac{1}{\rho} \frac{\gamma}{(1-\mu_{hu})^{1/\sigma}} \left[\left(1 + \frac{1}{\sigma}\right) \frac{s - \mu_{hu}}{1 - \mu_{hu}} - 1 \right] + o\left(\frac{1}{\rho}\right), \quad \forall u \in [0, T_s] \tag{A.8}$$

$$T_\psi(\rho) = \arg \max_{u \in [0, T_f(\rho)]} Q_u(\rho) \longrightarrow \arg \max_{u \in [0, T_s]} \frac{s - \mu_{hu}}{(1-\mu_{hu})^{1+1/\sigma}}. \tag{A.9}$$

With this in mind we can study the asymptotic behavior of price, allocation, and volume.

A.7.1 Asymptotic price

Proposition 1 shows that the price at all times $t \in [T_f, \infty)$ is equal to $1/r$. But we know from Lemma A.7 that $T_f(\rho) \to T_s$ as $\rho \to \infty$. Clearly, this implies that, for all $t \in (T_s, \infty)$, as $\rho \to \infty$ the price converges towards $1/r$, its unlimited cognition counterpart.

Integrating the ODEs of Proposition 1 shows that, at all times $t \in [0, T_s]$, the price is equal to:

$$p_t = \int_t^{T_f(\rho)} e^{-r(u-t)} \left(1 - \delta(1-\mu_{hu}) Q_u^\sigma(\rho)\right) du + \frac{e^{-r(T_f(\rho)-t)}}{r}.$$

But Lemma A.7 shows that $(1-\mu_{hu})Q_u^\sigma(\rho)$ converges point–wise towards $[(s-\mu_{hu})/(1-\mu_{hu})]^\sigma$. Moreover, we know that $(1-\mu_{hu})Q_u^\sigma \in [0, 1]$. It thus follows from dominated convergence that p_t converges to its unlimited cognition counterpart.

A.7.2 Asymptotic distribution of asset holdings

Let us start with some $u \leq T_s$. With unlimited cognition, traders whose valuation is low at time u hold $(s - \mu_{hu})/(1 - \mu_{hu})$, and traders whose valuation is high hold one share. With limited cognition, the time–u cross–sectional distribution of asset holdings across low–valuation traders is, by our usual law of large numbers argument, the time–u distribution of asset holding generated by the holding plan of a representative trader, conditional on $\theta_u = \ell$. Thus, to establish convergence of the distribution of asset holdings, we need to show that for all $\varepsilon > 0$, as $\rho \to \infty$:

$$\text{Proba}\left(\Big| \, q_{\tau_u,u} - \frac{s - \mu_{hu}}{1 - \mu_{hu}} \, \Big| > \varepsilon \, \Big| \, \theta_u = \ell \right) \to 0 \tag{A.10}$$

$$\text{Proba}\left(\big| \, q_{\tau_u,u} - 1 \, \big| > \varepsilon \, \Big| \, \theta_u = h \right) \to 0. \tag{A.11}$$

Proof of convergence for low valuation, $u \leq T_s$, in equation (A.10). We first introduce the following notation: for $t < T_f$ and $u \in [t, T_f)$, $q_{\ell,t,u}(\rho) = \min\{(1 - \mu_{ht})^{1/\sigma} Q_u(\rho), 1\}$ is the time-u asset holding of a time–t low–valuation trader. Now pick $u_\varepsilon < u$ such that, for all $t \in [u_\varepsilon, u]$,

$$\frac{s - \mu_{hu}}{1 - \mu_{hu}} \leq \left(\frac{1 - \mu_{ht}}{1 - \mu_{hu}} \right)^{1/\sigma} \frac{s - \mu_{hu}}{1 - \mu_{hu}} < \min\left\{ 1, \frac{s - \mu_{hu}}{1 - \mu_{hu}} + \frac{\varepsilon}{2} \right\}. \tag{A.12}$$

Then, note that, for $t \in [u_\varepsilon, u]$, by (A.8), as $\rho \to \infty$:

$$q_{\ell,t,u}(\rho) \to \min\left\{ \left(\frac{1 - \mu_{ht}}{1 - \mu_{hu}} \right)^{1/\sigma} \frac{s - \mu_{hu}}{1 - \mu_{hu}}, 1 \right\} = \left(\frac{1 - \mu_{ht}}{1 - \mu_{hu}} \right)^{1/\sigma} \frac{s - \mu_{hu}}{1 - \mu_{hu}} \tag{A.13}$$

since, by (A.12), the left–hand side of the "min" is less than one. Moreover, since $q_{\ell,t,u}(\rho)$ is decreasing in t, (A.13) implies that $q_{\ell,t,u}(\rho) \leq q_{\ell,u_\varepsilon,u}(\rho) < 1$ for ρ large enough. Put differently, for ρ large enough, $q_{\ell,t,u}(\rho) = (1 - \mu_{ht})^{1/\sigma} Q_u(\rho)$ for all $t \in [u_\varepsilon, u]$. Clearly, this implies that the convergence of $q_{\ell,t,u}(\rho)$ is uniform in $t \in [u_\varepsilon, u]$. Together with (A.12), this implies that:

$$\left| q_{\ell,t,u}(\rho) - \frac{s - \mu_{hu}}{1 - \mu_{hu}} \right| \leq \left| q_{\ell,t,u}(\rho) - \left(\frac{1 - \mu_{ht}}{1 - \mu_{hu}} \right)^{1/\sigma} \frac{s - \mu_{hu}}{1 - \mu_{hu}} \right| + \left| \left(\frac{1 - \mu_{ht}}{1 - \mu_{hu}} \right)^{1/\sigma} \frac{s - \mu_{hu}}{1 - \mu_{hu}} - \frac{s - \mu_{hu}}{1 - \mu_{hu}} \right| \leq \varepsilon,$$

for ρ large enough, for all $t \in [u_\varepsilon, u]$. Keeping in mind that all traders who have a low–valuation at time u must have had a low–valuation at their last information event, and going back to equation (A.10), this implies that, for ρ large enough:

$$\text{Proba}\left(\Big| \, q_{\tau_u,u} - \frac{s - \mu_{hu}}{1 - \mu_{hu}} \, \Big| > \varepsilon \, \Big| \, \theta_u = \ell \right) \leq \text{Proba}\left(\tau_u < u_\varepsilon \, \Big| \, \theta_u = \ell \right) = e^{-\rho(u - u_\varepsilon)} \to 0$$

as $\rho \to \infty$.

Proof of convergence for high valuation, $u \leq T_s$, in equation (A.11). With limited cognition, if $\theta_u = h$ and $\theta_{\tau_u} = h$, then $q_{\tau_u,u} = 1$. Thus, a necessary condition for $\theta_u = h$ and $|q_{\tau_u,u} - 1| > \varepsilon$ is that $\theta_{\tau_u} = \ell$. This implies that:

$$\text{Proba}\Big(\big|q_{\tau_u,u}-1\big|>\varepsilon\,\Big|\,\theta_u=h\Big)\leq\text{Proba}\Big(\theta_{\tau_u}=\ell\,\Big|\,\theta_u=h\Big)=e^{-\rho u}+\int_0^u\rho e^{-\rho(u-t)}\left(1-\frac{\mu_{ht}}{\mu_{hu}}\right)dt$$
$$=\int_0^u\frac{\mu'_{ht}}{\mu_{hu}}e^{-\rho(u-t)}\,dt\to 0.$$

where, in the first equality, $1-\mu_{ht}/\mu_{hu}$ is the probability of a low–valuation at time t conditional on a low–valuation at time u, and where the second equality follows after integrating by parts. Convergence of the integral to zero follows by dominated convergence, since the integrand is bounded and converges to zero for all $t<u$.

Convergence of the distribution of asset holdings for $u>T_s$. With unlimited cognition, all traders with a low–valuation at time $u>T_s$ hold zero asset. With limited cognition, low–valuation traders hold zero asset if $\tau_u\geq T_f(\rho)$ and $\theta_{\tau_u}=\ell$. Moreover, for ρ large enough and η small enough, $T_s<T_f(\rho)<T_s+\eta<u$. Thus:

$$\text{Proba}\big(q_{\tau_u,u}>\varepsilon\,\big|\,\theta_u=\ell\big)\leq\text{Proba}\big(\tau_u<T_f(\rho)\,\big|\,\theta_u=\ell\big)=e^{-\rho(u-T_f(\rho))}\leq e^{-\rho(u-T_s-\eta)}\to 0$$

as $\rho\to\infty$. Lastly, let us turn to traders with a high–valuation at time $u>T_s$. With unlimited cognition, the distribution of asset holdings is indeterminate with a mean of s/μ_{hu}. With limited cognition, take ρ large enough so that $T_f(\rho)<u$. The distribution of asset holdings is also indeterminate with mean

$$\frac{\int_0^u\rho e^{-\rho(u-t)}s\,dt}{\int_0^u\rho e^{-\rho(u-t)}\mu_{ht}\,dt}.$$

Integrating the numerator and denominator by part shows that, as $\rho\to\infty$, this mean asset holding converges to s/μ_{ht}, its unlimited cognition counterpart.

A.7.3 Asymptotic volume

Basic formulas. As before, 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 the one who have an information event 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 traders, as follows. Just before their information event, low–valuation traders hold on average:

$$\mathbb{E}\left[q_{\tau_u,u}\,|\,\theta_u=\ell\right]=e^{-\rho u}s+\int_0^u\rho e^{-\rho(u-t)}q_{\ell,t,u}\,dt,\qquad\text{(A.14)}$$

since a fraction $e^{-\rho u}$ of them have had no information event and still hold s unit of the asset and a density $\rho e^{-\rho(u-t)}$ of them had their last information event at time t and hold $q_{\ell,t,u}$. Instantaneous trading volume is then:

$$V_u=\rho(1-\mu_{hu})\Big(\mathbb{E}\left[q_{\tau_u,u}\,|\,\theta_u=\ell\right]-q_{\ell,u,u}\Big),\qquad\text{(A.15)}$$

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

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

$$V_u = \rho \mu_{hu} \left(1 - \mathbb{E}\left[q_{\tau_u, u} \,|\, \theta_u = h \right] \right), \tag{A.16}$$

where the time–u average asset holdings of high-valuation investors is found by plugging (A.14) into the market clearing condition:

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

For $u > T_f$, the trading volume is not zero since high–valuation traders continue to buy from the low valuation traders having an information event:

$$V_u = \rho (1 - \mu_{hu}) \mathbb{E}\left[q_{\tau_u, u} \,|\, \theta_u = \ell \right]. \tag{A.17}$$

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

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

For this we can follow the same calculations leading to equation (A.42) in the proof of Lemma A.7, but with $f(t, \rho) = \min\{(1-\mu_{ht})^{1/\sigma} Q_u(\rho), 1\}$. This gives:

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

where the second line follows from plugging equation (A.8) in the first line. Substituting this expression into equations (A.15) and (A.16), we find after some straightforward manipulation that, when ρ goes to infinity, $V_u \to \gamma(s - \mu_{hu})/\sigma$ 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: $\gamma(1 - \mu_{hu})$, times the amount of asset they buy at that time: $1 - (s - \mu_{hu})/(1 - \mu_{hu})$. 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 $\gamma(s - \mu_{hu})/\sigma > \gamma(1-s)$ for $u < T_\psi(\infty)$.

A.8 Proof of Proposition 5

A.8.1 First point: when (21) holds

Given the ODEs satisfied by the price path in both markets, it suffices to show that, for all $u < T_s$,

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

Besides, when $s \leq \sigma/(\sigma + 1)$, it follows from Lemma A.2 and Lemma A.3 that:

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

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

$$F_u = (s - \mu_{hu}) \int_0^u e^{\rho t}(1 - \mu_{ht})^{1+1/\sigma}\, dt - (1 - \mu_{hu})^{1+1/\sigma} \int_0^u e^{\rho t}(s - \mu_{ht}) < 0.$$

But we know from the proof of Lemma A.3, equation (A.35), page 34, that F_u has the same sign than Q'_u, which we know is negative at all $u > 0$ since $s \leq \sigma/(\sigma + 1)$.

A.8.2 Second point: when s is close to 1 and σ is close to 0

The price at time 0 is equal to:

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

where we make the dependence of s and σ explicit. With perfect cognition, $\xi_u = 1 - \delta((s - \mu_{hu})/(s - \mu_{hu}))^\sigma = 1 - \delta\left(1 - (1 - s)e^{\gamma u}\right)^\sigma$ for $u < T_s$, and $\xi_u = 1$ for $u > T_s$. Therefore $p_0 = 1/s - \delta I(s)$, where:

$$I(s) \equiv \int_0^{T_s} e^{-ru} \left(1 - (1 - s)e^{\gamma u}\right)^\sigma\, du,$$

where we make explicit the dependence on s. Similarly, with limited cognition, the price at time 0 is equal to $p_0 = 1/s - \delta J(s)$, where:

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

Lemma A.8. *When s goes to 1, both $I(s)$ and $J(s)$ go to $1/r$.*

Therefore, p_0 goes to $(1 - \delta)/r$ both with perfect cognition and with limited cognition. Besides, with perfect cognition:

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

and with limited cognition:

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

The next two lemmas compare $I'(s)$ and $J'(s)$ for s in the neighborhood of 1 when σ is not too large.

Lemma A.9. *When s goes to 1:*

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

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

In the Lemma and in all what follows $f(s) \sim g(s)$ means that $f(s)/g(s) \to 1$ when $s \to 1$.

Lemma A.10. *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 A.9 and A.10 imply that, if σ is close to 0, then $J'(s) > I'(s)$ for s in the left-neighborhood of 1. The second point of the proposition then follows.

A.9 Proofs of Proposition 8

A.9.1 The candidate equilibrium

Before formulating our guess, we need the following Lemma, proved in in Section A.11.10:

Lemma A.11. *Suppose $s > \sigma/(1+\sigma)$ and let $T_\phi > 0$ be the unique solution of:*

$$\frac{\mu_{hT_\phi}}{\sigma} = \left(1 + \frac{1}{\sigma}\right) s - 1. \tag{A.18}$$

Then $T_\phi < T_s$ and, for all $t \in [0, T_\phi)$ there exists a unique $\phi_t \in (T_\phi, T_f)$ such that

$$\int_t^{\phi_t} e^{\rho u} \left[(1-\mu_{hu})^{1+1/\sigma}(s-\mu_{ht}) - (1-\mu_{ht})^{1+1/\sigma}(s-\mu_{hu}) \right] du = 0. \tag{A.19}$$

In addition $t \mapsto \phi_t$ is continuously differentiable, strictly decreasing over $[0, T_\phi]$, and $\lim_{t \to T_\phi} \phi_t = T_\phi$.

For concision, we directly state below the analytical formula defining the candidate LOE. The reader interested in building up intuition for these formula may refer to Section VIII in Biais, Hombert, and Weill (2010), where we propose a heuristic construction of this candidate equilibrium.

The price path. The price path is continuous and solves the following ODEs. When $t \in (0, T_\phi)$:

$$rp_t - \dot{p}_t = \xi_t = 1 - \delta \left(\frac{s - \mu_{ht}}{1 - \mu_{ht}} \right)^{\sigma} + \delta \frac{d}{dt} \left[\left(\frac{s - \mu_{ht}}{(1 - \mu_{ht})^{1+1/\sigma}} \right)^{\sigma} \right] \int_t^{\phi_t} e^{-(r+\rho)(u-t)} (1 - \mu_{hu}) \, du, \tag{A.20}$$

When $t \in (T_\phi, \phi_0)$:

$$rp_t - \dot{p}_t = \xi_t = 1 - \delta \frac{1 - \mu_{ht}}{1 - \mu_{h\phi_t^{-1}}} \left(\frac{s - \mu_{h\phi_t^{-1}}}{1 - \mu_{h\phi_t^{-1}}} \right)^{\sigma}. \tag{A.21}$$

When $t \in (\phi_0, T_f)$:

$$rp_t - \dot{p}_t = \xi_t = 1 - \delta (1 - \mu_{ht}) \overline{Q}_t^{\sigma}. \tag{A.22}$$

where $\overline{Q}_t$ is defined in equation (A.4). Lastly, when $t \geq T_f$, $p_t = 1/r$. Clearly, together with the continuity conditions at the boundaries of each intervals (T_ϕ, ϕ_0, and T_f), the above ODEs uniquely define the price path.

High–valuation traders' holding plans. The asset holding plans of high–valuation traders are as follows. For $t \in [0, T_f)$, the time–t high–valuation trader holds $q_{t,u} = 1$ for all $u \geq t$. For $t \geq T_f$, the time–t high–valuation trader holds $q_{t,u} = s/\mu_{ht}$ for all $u \geq t$. Since these asset holding plan are constant, we guess that they are implemented by submitting market orders at time t.

Low–valuation traders' holding plans. The asset holding plans of low–valuation traders are as follows. For $t \in (0, T_\phi]$, the time–t low–valuation trader holds:

$$q_{t,u} = \frac{s - \mu_{ht}}{1 - \mu_{ht}} \quad \text{for all} \quad u \in [t, \phi_t] \tag{A.23}$$

$$= (1 - \mu_{ht})^{1/\sigma} Q_u \quad \text{for all} \quad u \in [\phi_t, T_f), \quad \text{where} \quad Q_u \equiv \left(\frac{1 - \xi_u}{\delta (1 - \mu_{hu})} \right)^{1/\sigma} \tag{A.24}$$

$$= 0 \quad \text{for all} \quad u \geq T_f, \tag{A.25}$$

and where ξ_u is defined in equations (A.20) and (A.21). For $t \in [T_\phi, T_f)$

$$q_{t,u} = (1 - \mu_{ht})^{1/\sigma} Q_u \quad \text{for all} \quad u \in [t, T_f) \tag{A.26}$$

$$= 0 \quad \text{for all} \quad u \geq T_f. \tag{A.27}$$

Lastly, for $t \geq T_f$ and $u \geq t$, $q_{t,u} = 0$. We guess that, in order to implement her holding plan at the information event time t, the trader submits market orders and a schedule of limit sell orders.

Note that, after plugging in the definition of ξ_{ϕ_t} given in equation (A.21), one sees that the above–defined asset holding plan is continuous at $u = \phi_t$. Since $Q_{T_f} = \overline{Q}_{T_f} = 0$, we also have continuity at $u = T_f$. Lastly in Lemma A.14 below, we will show that when $t \in [0, T_\phi)$ ($t \in [T_\phi, T_f)$) and $u \in [\phi_t, T_f]$ ($u \in [t, T_f]$), the holding plan $u \mapsto q_{t,u}$ is strictly decreasing.

The verification proof is organized as follows. Section A.9.2 provides preliminary properties of equilibrium objects. Section A.9.3 proves the optimality of the candidate asset holding plans. Section A.9.4 concludes by showing that the market clears at all dates.

A.9.2 Preliminary results

We start by noting that, by direct inspection of our guess, the candidate LOE coincides with the ATE at all times $t > \phi_0$. Next, a result we use repeatedly is:

Lemma A.12. *For* $t \in [0, T_\phi]$*,*

$$\frac{d}{dt}\left[\frac{s-\mu_{ht}}{(1-\mu_{ht})^{1+1/\sigma}}\right] \geq 0, \quad \textit{with an equality only if} \quad t = T_\phi.$$

The proof is in Section A.11.11. Next, we prove in Section A.11.12:

Lemma A.13. *For* $t \in (0, T_f)$*,* $rp_t - \dot{p}_t = \xi_t \in (0,1)$*.*

The Lemma ensures that Q_u, in equation (A.24), is well defined, and is also helpful to establish the optimality of trading strategies. To show that asset holding plans are decreasing when $u \geq \phi_t$, we will need the following Lemma, proved in Section A.11.13:

Lemma A.14. $u \mapsto Q_u$ *is continuous over* $(0, T_f)$*, strictly increasing over* $(0, T_\phi)$*, and strictly decreasing over* (T_ϕ, T_f)*.*

Equipped with the above Lemma, we show in Section A.11.14 another crucial property of the price path:

Lemma A.15. *The candidate equilibrium price,* p_t*, is continuously differentiable and strictly increasing over* $[0, T_f]$*.*

A.9.3 Asset holding plans are optimal

Execution times of limit orders. The first step is to verify that the candidate asset holding plans can be implemented using market and limit orders only, as explained in Section A.9.1. To that end, we start by deriving the execution times associated with alternative limit orders to buy or sell.

Limit orders to buy or sell submitted at $t \geq T_f$. Consider the case of limit sell orders (the case of limit buy order is symmetric). In the candidate equilibrium, the price is constant for all $t \geq T_f$. Clearly, this means that a limit order to sell at the ask price $a > p_{T_f}$ is never executed. By the price priority rule, a limit order to sell at price $a < p_{T_f}$ is executed immediately. We can always ignore such limit orders because they are clearly dominated by a market order to sell, which is also executed immediately, but has a strictly higher execution price, p_{T_f}. Lastly, we argue that a limit order to sell at price $a = p_{T_f}$ submitted at time $t \geq T_f$ is executed immediately, at time t. First we know from the price priority rule that this order is executed at or after time t. But note that, at all times $u > t \geq T_f$, the limit order book must be empty at price p_{T_f}: otherwise, by the time priority rule, the earliest submitted limit order at price p_{T_f} would be executed at time u, which would contradict the fact that asset holding plans are constant after T_f. Thus, by any time $u > t$, all limit orders to sell at price p_{T_f}, and in particular the ones submitted at time t, have been executed. Since this is true for all $u > t$, this implies that a limit order to sell at price p_{T_f} is executed immediately at time t, and is thus equivalent to a market order.

Limit orders to sell submitted at $t < T_f$. Lemma A.15 showed that the price path is strictly increasing over $[0, T_f]$. Therefore, the price–priority rule implies that: a limit order to sell at price $a > p_{T_f}$ is never executed; a limit order to sell at price $a \in [p_t, p_{T_f})$ is executed when the price reaches a, i.e. at the time u such that $p_u = a$; a limit order to sell at price $a = p_{T_f}$ is executed at or after time T_f. But, as noted in the previous paragraph, the limit order book is empty at all times $u > T_f$, and the same reasoning as before shows that this limit order has to be executed at or before T_f. Taken together, these two remarks imply that a limit order to sell at price $a = p_{T_f}$, submitted at time $t < T_f$, is executed exactly at time T_f.

As before, we can ignore limit orders to sell at price $a < p_t$. By the price priority rule, they are executed immediately at price $a < p_t$, and therefore are clearly dominated by a market order to sell at price p_t.

Limit orders to buy submitted at $t < T_f$. By the price–priority rule: a limit order to buy at price $b \leq p_t$ is never executed; limit order to buy at price $b > p_t$ is executed immediately at price b, and is thus clearly dominated by market order to buy at price p_t.

A re-statement of the trader's problem.

Based on the above, we show that the trader problem reduces to maximizing her utility from one information event to the next:

Lemma A.16. *In the candidate LOE, an asset holding plan $q_{t,u}$ solves the trader's problem if and only if it maximizes*

$$\int_t^{\infty} e^{-(r+\rho)(u-t)} \left\{ \mathbb{E}_t \left[v(\theta_u, q_{t,u}) \right] - \xi_u q_{t,u} \right\} du \tag{A.28}$$

for almost all $(t, \omega) \in \mathbb{R} \times \Omega$, and subject to the constraint that $u \mapsto q_{t,u}$ is decreasing for $u \in [t, t \vee T_f]$, and constant for $u \geq t \vee T_f$.

For the "if" part, note that the execution times of Section A.9.3 imply that, upon an information event at time t:

- a trader cannot submit a limit order to *buy* executed at time $u > t$;
- if $t < T_f$, a trader can submit a limit order to *sell* at any time $u \in (t, T_f]$;
- a trader cannot submit a limit order to *sell* at any time $u > T_f$.

Therefore, in the candidate LOE, a trader's asset holding plan $u \mapsto q_{t,u}$ can never increase, can decrease in an arbitrary fashion over $[t, t \vee T_f]$, and has to stay constant after $t \vee T_f$. This means that the value of the time–t asset holding plan, $u \mapsto q_{t,u}$, is less than the maximum value of the maximization problem in Lemma A.16. Therefore, an asset holding plan solves trader's problem if it solves the program of Lemma A.16 for almost all $(t, \omega) \in \mathbb{R} \times \Omega$.

To prove the "only if" part, we proceed by contrapositive. Suppose that $q_{t,u}$ does not maximize (A.28) for some positive measure set of $\mathbb{R} \times \Omega$. Then, consider the following change of holding plan: for all times and events in that set, switch to a plan that achieves a higher value in the objective (A.28), and keep your original holding plan the same otherwise. To see that the change of holding plan is feasible, note that the execution times of Section A.9.3 imply that, if the change of holding plan requires to modify orders submitted at some information event t_1, all of these modifications can be undone at any subsequent information event $t_2 > t_1$. Indeed a limit sell order at price $p_u \in (p_0, p_{T_f}]$ is executed at time u regardless of its submission time $t_1 < u$. Thus, if such an order is modified at t_1 and is still unfilled at time t_2, then $t_2 < u$ and the modification can be undone. Other kind of orders (limit sell orders at different prices, limit buy orders, or market orders) are either immediately executed at information event t_1, or never executed, so any modification at t_1 can be undone with market orders at t_2. Taken together, this shows that the change of holding plan is feasible and, clearly, it increases the value of the trader's objective.

Optimality of high–valuation traders' asset holding plans. By construction, the candidate asset holding plan is constant for all u given t, and so it can be implemented by only submitting market orders at each information event. To show that the candidate asset holding plan is optimal, it suffices to show that it maximizes a high–valuation trader's utility flow at each time u. To see that it is indeed the case, recall from Lemma A.13 that $\xi_u \in (0,1)$ for $u \in (0, T_f)$, and that by construction we have that $\xi_u = 1$ for $u > T_f$. Therefore, the flow utility

$$v(h,q) - \xi_u q = \min\{q, 1\} - \xi_u q,$$

is maximized by $q = 1$ if $u < T_f$, and by any $q \in [0,1]$ if $u \geq T_f$. In particular, $q_{t,u} = 1$ and, as long as $t \geq T_f > T_s$, $q_{t,u} = s/\mu_{ht} \leq 1$ maximizes the flow utility for $u \geq T_f$. Thus, the candidate asset holding plan is optimal for high–valuation traders.

Optimality of low–valuation traders' asset holding plans After time T_f, the price is constant and equal to $1/r$, and so the opportunity cost is $\xi_u = 1$. Given that $v_q(\ell, 0) = v(h, 0) = 1$ and $v_q(\ell, q) < 1$ for $q > 0$, this immediately implies that $q = 0$ is the asset holding maximizing the flow payoff, $\mathbb{E}\left[v_q(\theta_u, q) \,|\, \theta_t = \ell\right] - \xi_u q$, for any $t < u$. Clearly, given the execution time of Section A.9.3, a low–valuation trader can always implement this zero asset holding for all $u \geq T_f$. Indeed, if $t < T_f$, she just needs to submit a limit order to sell all her remaining assets at price p_{T_f}. If $t > T_f$, she just needs to submit a market order to sell all her assets at time t. We conclude that we can restrict attention to holding plans such that $q_{t,u} = 0$ for all $u \geq T_f$. Plugging this restriction into the objective (A.28), we find that the low–valuation trader's problem at time $t < T_f$ reduces to choosing $\tilde{q}_{t,u}$ in order to maximize

$$\int_t^{T_f} e^{-(r+\rho)(u-t)} M(u, \tilde{q}_{t,u})\, du \tag{A.29}$$

$$\text{where } M(u,q) \equiv \mathbb{E}\left[v(\theta_u, q) \,|\, \theta_t = \ell\right] - \xi_u q = (1 - \xi_u)\, q - \delta \frac{1 - \mu_{hu}}{1 - \mu_{ht}} \frac{q^{1+\sigma}}{1+\sigma}, \tag{A.30}$$

subject to the constraint that $u \mapsto \tilde{q}_{t,u}$ is decreasing over $[t, T_f]$.

We now verify that our candidate asset holding plan solves this optimization problem. First, we note that, by Lemma A.14, $q_{t,u}$ is, as required, decreasing over $[t, T_f]$. Second, we show below that two sufficient conditions for optimality are

$$(i) \qquad \int_t^{T_f} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u}) q_{t,u}\, du = 0,$$

$$(ii) \qquad \int_t^{T_f} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u}) \tilde{q}_{t,u}\, du \leq 0 \quad \text{for any decreasing function } u \mapsto \tilde{q}_{t,u},$$

where the partial derivative is well defined because, since $q_{t,u}$ is decreasing, $q_{t,u} \leq q_{t,t} = (s - \mu_{ht})/(1 - \mu_{ht}) < 1$.

To see why condition (i) and (ii) are sufficient, consider another decreasing holding plan $\tilde{q}_{t,u}$. We have

$$\int_t^{T_f} e^{-(r+\rho)(u-t)} \left[M(u, q_{t,u}) - M(u, \tilde{q}_{t,u})\right] du \geq \int_t^{T_f} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u})(q_{t,u} - \tilde{q}_{t,u})\, du \geq 0,$$

where the first inequality follows from the concavity of $q \mapsto M(u,q)$, and the second inequality follows from (i) and (ii). The proof that condition (i) and (ii) hold follows from algebraic manipulations that we gather in Section A.11.15.

A.9.4 The market clears at all times

For all $u \leq T_f$, high–valuation traders who had at least one information event hold $q_{t,u} = 1$. Plugging this in equation (8) and rearranging, this leads to the market–clearing condition:

$$\int_0^u \rho e^{-\rho(u-t)}(1-\mu_{ht})\mathbb{E}\left[q_{t,u} \,|\, \theta_t = \ell\right] dt = \int_0^u \rho e^{-\rho(u-t)}(s-\mu_{ht})\, dt. \tag{A.31}$$

Market clearing at $u \in (0, T_\phi)$**.** Then, for all $t \leq u$, it follows from equation (A.23) that low–valuation traders hold $q_{t,u} = (s-\mu_{ht})/(1-\mu_{ht})$ and so clearly the market–clearing condition (A.31) holds.

Market clearing at $u \in (T_\phi, \phi_0)$**.** Then, it follows from equation (A.23) and (A.24) that low–valuation traders hold $q_{t,u} = (s-\mu_{ht})/(1-\mu_{ht})$ if $t \leq \phi_u^{-1}$, and $(1-\mu_{ht})^{1/\sigma} Q_u$ if $t \in [\phi_u^{-1}, u]$. Thus, the left–hand side of (A.31) writes:

$$\int_0^{\phi_u^{-1}} \rho e^{-\rho(u-t)}(1-\mu_{ht})\frac{s-\mu_{ht}}{1-\mu_{ht}}\, dt + \int_{\phi_u^{-1}}^u \rho e^{-\rho(u-t)}(1-\mu_{ht})^{1+1/\sigma} Q_u\, dt. \tag{A.32}$$

Plugging ξ_u, as defined in equation (A.21), into the definition of Q_u, as defined in equation (A.24), we obtain:

$$Q_u = \frac{s - \mu_{h\phi_u^{-1}}}{(1-\mu_{h\phi_u^{-1}})^{1+1/\sigma}}$$

Next, using the implicit equation (A.19) defining ϕ_u^{-1}, we obtain that:

$$Q_u = \frac{\int_{\phi_u^{-1}}^u \rho e^{-\rho(u-t)}(s-\mu_{ht})\, dt}{\int_{\phi_u^{-1}}^u \rho e^{-\rho(u-t)}(1-\mu_{ht})^{1+1/\sigma}\, dt}.$$

Plugging in (A.32) shows that the market–clearing condition (A.31) holds.

Market clearing at $u \in (\phi_0, T_f)$**.** First note that by equation (A.22), we have $Q_u = \overline{Q}_u$, where $\overline{Q}_u$ is defined in equation (A.4). Then, the demand from low-valuation investors is

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

by definition of $\overline{Q}_u$,

Market clearing at $u > T_f$**.** The market clears by construction of the holding plans of high– and low–valuation investors after time T_f.

A.10 Proof of Proposition 9

For this proof the superscript ATE (algorithmic trading equilibrium) refers to the equilibrium objects of Proposition 1, while the superscript LOE (limit order equilibrium) refers to the equilibrium object of Proposition 8.

A.10.1 First point: $p_u^{ATE} = p_u^{LOE}$ for $u \geq \phi_0$

By construction $p_u^{ATE} = p_u^{LOE} = 1/r$ for $u \geq T_f$, and $\xi_u^{ATE} = \xi_u^{LOE}$ for $u \in (\phi_0, T_f)$. This immediately imply that $p_u^{ATE} = p_u^{LOE}$ for all $u \geq \phi_0$.

A.10.2 Second point: $p_u^{ATE} > p_u^{LOE}$ for $u \in (T_\phi, \phi_0)$

To obtain the result, we prove the following Lemma.

Lemma A.17. *Let $u \in (T_\phi, \phi_0)$. If $Q_u^{ATE} \geq Q_u^{LOE}$, then time–t low–valuation traders' asset holdings satisfy $q_{t,u}^{ATE} \geq q_{t,u}^{LOE}$ for all $t \in (0, u)$ with a strict equality $t \in (0, \phi_u^{-1})$.*

The Lemma is proved in Section A.11.16. Since high–valuation traders hold 1 unit of the asset in both the ATE and the LOE, it cannot be that both markets clear at a date $u \in (T_\phi, \phi_0)$ such that $Q_u^{ATE} \geq Q_u^{LOE}$. Thus, $Q_u^{ATE} < Q_u^{LOE}$ for all $u \in (T_\phi, \phi_0)$. From the definition of Q_u^{ATE}, we have $\xi_u^{ATE} = 1 - (1 - \mu_{hu}) {Q_u^{ATE}}^\sigma$. From (A.25), in the LOE we have $\xi_u^{LOE} = 1 - (1 - \mu_{hu}) {Q_u^{LOE}}^\sigma$. Thus $\xi_u^{ATE} > \xi_u^{LOE}$ for $u \in (T_\phi, \phi_0)$, and $\xi_u^{ATE} = \xi_u^{LOE}$ for $u \geq \phi_0$. This clearly implies that $p_u^{ATE} > p_u^{LOE}$ for $u \in (T_\phi, \phi_0)$.

A.10.3 Third point: $p_0^{ATE} > p_0^{LOE}$ for ρ close to zero

We start by establishing a preliminary result. Since all the equilibrium variables are defined as roots of continuously differentiable functions that are well-defined in $\rho = 0$, with non-zero partial derivatives, we can apply the Implicit Function Theorem to show that:

Lemma A.18. *Every equilibrium object has a well-defined limit when $\rho \to 0$ and is continuous in ρ over $\rho \in [0, +\infty)$. Moreover, those limits satisfy the same equations as in the case $\rho > 0$ after letting $\rho = 0$.*

The proof is in Section A.11.17. We denote all these limits with a hat, " ˆ ". We now show that $\hat{p}_0^{ATE} < \hat{p}_0^{LOE}$ and the third point of Proposition 9 will follow by continuity.

In the ATE, after integrating the ODE for the price over $u \in [0, \phi_0]$ and taking the limit $\rho \to 0$, we obtain:

$$\hat{p}_0^{ATE} = e^{-r\hat{\phi}_0} \hat{p}_{\hat{\phi}_0}^{ATE} + \int_0^{\hat{\phi}_0} e^{-ru} \left[1 - \delta(1 - \mu_{hu}) (\hat{Q}_u^{ATE})^\sigma \right] du.$$

Inspecting the proof of Proposition 3, one notes that it does not use the fact that $\rho > 0$. Therefore it still holds when $\rho = 0$ and $u \mapsto \hat{Q}_u^{ATE}$ is hump-shaped. We also have $\hat{Q}_0^{ATE} = s$, and:

$$\hat{Q}_{\hat{\phi}_0}^{ATE} \geq \frac{\int_0^{\hat{\phi}_0} (s - \mu_{ht})\, dt}{\int_0^{\hat{\phi}_0} (1 - \mu_{ht})^{1+1/\sigma}\, dt} = s,$$

where the inequality follows from the market clearing condition (17) and the last equality comes from the definition of $\hat{\phi}_0$, equation (A.19). Taken together, these facts imply that:

$$\hat{p}_0^{ATE} < e^{-r\hat{\phi}_0} \hat{p}_{\hat{\phi}_0}^{ATE} + \int_0^{\hat{\phi}_0} e^{-ru} \left[1 - \delta(1 - \mu_{hu}) s^\sigma \right] du.$$

For the LOE, we integrate the ODEs (A.20) and (A.21) over $u \in [0, \phi_0]$ and we take the limit as $\rho \to 0$. It follows that:

$$\hat{p}_0^{LOE} = e^{-r\hat{\phi}_0}\hat{p}_{\hat{\phi}_0}^{LOE} + \int_0^{\hat{T}_\phi} e^{-ru}\left[1 - \delta_u + \frac{\partial}{\partial u}\left[\frac{\delta_u}{1-\mu_{hu}}\right]\int_u^{\hat{\phi}_u} e^{-r(z-u)}(1-\mu_{hz})\,dz\right] du$$
$$+ \int_{\hat{T}_\phi}^{\hat{\phi}_0} e^{-ru}\left[1 - \delta_{\hat{\phi}_u^{-1}}\frac{1-\mu_{hu}}{1-\mu_{h\hat{\phi}_u^{-1}}}\right] du,$$

where $\delta_u \equiv \delta((s-\mu_{hu})/(1-\mu_{hu}))^\sigma$. In the first line, we can compute the double integral by switching the order of integration as in the proof of Lemma A.15 (see Figure 9):

$$\int_0^{\hat{T}_\phi}\int_u^{\hat{\phi}_u} e^{-ru}\frac{\partial}{\partial u}\left[\frac{\delta_u}{1-\mu_{hu}}\right] e^{-r(z-u)}(1-\mu_{hz})\,dz\,du$$
$$= \int_0^{\hat{T}_\phi}\int_0^{z} e^{-rz}\frac{\partial}{\partial u}\left[\frac{\delta_u}{1-\mu_{hu}}\right](1-\mu_{hz})\,du\,dz + \int_{\hat{T}_\phi}^{\hat{\phi}_0}\int_0^{\hat{\phi}_z^{-1}} e^{-rz}\frac{\partial}{\partial u}\left[\frac{\delta_u}{1-\mu_{hu}}\right](1-\mu_{hz})\,du\,dz$$
$$= \int_0^{\hat{T}_\phi} e^{-rz}(1-\mu_{hz})\left[\frac{\delta_z}{1-\mu_{hz}} - \delta_0\right] dz + \int_{\hat{T}_\phi}^{\hat{\phi}_0} e^{-rz}(1-\mu_{hz})\left[\frac{\delta_{\hat{\phi}_z^{-1}}}{1-\mu_{h\hat{\phi}_z^{-1}}} - \delta_0\right] dz.$$

Plugging this back into the expression of $\hat{p}_0^{LOE}$ and substituting $\delta_0 = \delta s^\sigma$, we obtain:

$$\hat{p}_0^{LOE} = e^{-r\hat{\phi}_0}\hat{p}_{\hat{\phi}_0}^{LOE} + \int_0^{\hat{\phi}_0} e^{-ru}\left(1 - \delta(1-\mu_{hu})s^\sigma\right) du.$$

Since $\hat{p}_{\hat{\phi}_0}^{ATE} = \hat{p}_{\hat{\phi}_0}$, we obtain that $\hat{p}_0^{ATE} < \hat{p}_0^{LOE}$.

A.11 Proofs of preliminary and intermediate results

A.11.1 Proof of Lemma A.1

The left-hand side of (17) is continuous, strictly increasing for $Q_u < (1-\mu_{hu})^{-1/\sigma}$ and constant for $Q_u \geq (1-\mu_{hu})^{-1/\sigma}$. It is zero when $Q_u = 0$, and, when $Q_u = (1-\mu_{hu})^{-1/\sigma}$, it is equal to:

$$\int_0^u (1-\mu_{ht})\,\rho e^{-\rho(u-t)}\,dt > S_u = \int_0^u (s-\mu_{ht})\,\rho e^{-\rho(u-t)}\,dt,$$

since $s < 1$. Therefore, equation (17) has a unique solution, Q_u, and the solution satisfies $0 \leq Q_u < (1-\mu_{hu})^{-1/\sigma}$. To prove that Q_u is continuously differentiable we apply the Implicit Function Theorem (see, e.g., Theorem 13.7 Apostol, 1974). We note that (17) writes $K(u, Q_u) = 0$, where

$$K(u,Q) \equiv \int_0^u e^{\rho t}(1-\mu_{ht})\min\{(1-\mu_{ht})^{1/\sigma}Q, 1\}\,dt - \int_0^u e^{\rho t}\left(s-\mu_{ht}\right)\,dt. \tag{A.33}$$

Since we know that $Q_u < (1-\mu_{hu})^{-1/\sigma}$, we restrict attention to the domain $\{(u,Q) \in \mathbb{R}^2_+ \,:\, u > 0 \text{ and } Q < (1-\mu_{hu})^{-1/\sigma}\}$. In this domain, $\Psi(Q) < u$, and so equation (A.33) can be written, using the definition of $\Psi(Q)$:

$$\begin{aligned} K(u,Q) &= \int_0^{\Psi(Q)} e^{\rho t}(1-\mu_{ht})\,dt + \int_{\Psi(Q)}^{u} e^{\rho t}\,(1-\mu_{ht})^{1+1/\sigma}\,Q\,dt - \int_0^{u} e^{\rho t}\,(s-\mu_{ht})\,dt \\ &= \int_0^{u} e^{\rho t}(1-s)\,dt - \int_{\Psi(Q)}^{u} e^{\rho t}\,(1-\mu_{ht})\left[1-(1-\mu_{ht})^{1/\sigma}\,Q\right]\,dt, \end{aligned} \tag{A.34}$$

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, using (A.33):

$$\frac{\partial K}{\partial u} = e^{\rho u}(1-\mu_{hu})\min\{(1-\mu_{hu})^{1/\sigma}Q, 1\} - e^{\rho u}(s-\mu_{hu}).$$

and is clearly continuous. To calculate the partial derivative with respect to Q, we consider two cases. When $Q \in [0,1]$, then $\Psi(Q) = 0$, and so, using (A.34):

$$\frac{\partial K}{\partial Q} = \int_0^{u} e^{\rho t}\,(1-\mu_{ht})^{1+1/\sigma}\,dt = \int_{\Psi(Q)}^{u} e^{\rho t}\,(1-\mu_{ht})^{1+1/\sigma}\,dt,$$

When, on the other hand, $Q \in [1,(1-\mu_{hu})^{-1/\sigma})$, on the other hand, $\Psi(Q)$ solves $(1-\mu_{h\Psi(Q)})^{-1/\sigma} = Q$ and hence is continuously differentiable. Bearing this in mind when differentiating (A.34), we obtain again that

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

Since $\Psi(Q)$ is continuous, the above calculations show that $\partial K/\partial Q$ is continuous for all (u,Q) in its domain. Next, note that because $(1-\mu_{hu})^{1/\sigma}Q_u < 1$, we have $\Psi(Q_u) < u$ and therefore $\partial K/\partial Q > 0$ at (u,Q_u). Taken together, these observations allow to apply the Implicit Function Theorem and state that

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

where we used that $\psi_u \equiv \Psi(Q_u)$ and $Q_u(1-\mu_{hu})^{1/\sigma} < 1$.

A.11.2 Proof of Lemma A.3

The continuity of $\overline{Q}_u$ is obvious. That $\overline{Q}_{0+} = s$ follows from an application of l'Hôpital rule, 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:

$$F_u \equiv (s-\mu_{hu})\int_0^{u} e^{\rho t}\,(1-\mu_{ht})^{1+1/\sigma}\,dt - (1-\mu_{hu})^{1+1/\sigma}\int_0^{u} e^{\rho t}\,(s-\mu_{ht})\,dt, \tag{A.35}$$

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

$$G_u \equiv \left(1+\frac{1}{\sigma}\right)(1-\mu_{hu})^{1/\sigma}\int_0^{u} e^{\rho t}\,(s-\mu_{ht})\,dt - \int_0^{u} e^{\rho t}\,(1-\mu_{ht})^{1+1/\sigma}\,dt, \tag{A.36}$$

is continuously differentiable. Now suppose that $\overline{Q}'_u = 0$. Then $F_u = 0$ and, after rearranging (A.35):

$$(1-\mu_{hu})^{1/\sigma} \int_0^u e^{\rho t} (s - \mu_{ht})\, dt = \frac{s - \mu_{hu}}{1 - \mu_{hu}} \int_0^u e^{\rho t} (1-\mu_{ht})^{1+1/\sigma}\, dt.$$

Plugging this back into G_u we find that:

RA.1. *Suppose that $F_u = 0$ for some $u > 0$. Then sign $[F'_u]$ = sign $\left[s\left(1+\frac{1}{\sigma}\right) - 1 - \frac{\mu_{hu}}{\sigma}\right]$.*

Now note that $G_0 = 0$ and $G'_0 = s(1 + 1/\sigma) - 1$. Thus,

RA.2. *If $s \leq \sigma/(1+\sigma)$, then $F_u < 0$ for all $u > 0$.*

To see this, first note that, from repeated 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 sign $[F'_v]$ = sign $[G_v]$. Now, since $G_0 = 0$, $G_v = vG'_w$ for some $w \in (0, v)$. But $G'_0 < 0$ so G'_w 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 RA.1 and the assumption that $s \leq \sigma/(1+\sigma)$.

RA.3. *If $s > \sigma/(1+\sigma)$, then $F_u > 0$ for small u, and $F_u < 0$ for $u \in [T_s, T_f]$.*

The first part follows from applying the same reasoning as in the above paragraph, since when $s > \sigma/(1+\sigma)$ we have $G'_0 > 0$. The second part follows from noting that, when $u \in [T_s, T_f]$, the first term of F_u is negative, and strictly negative when $u \in (T_s, T_f]$, while the second term is negative, and strictly negative when $u \in [T_s, T_f)$. So F_u changes sign in the interval $(0, T_s)$. We now show that:

RA.4. *If $s > \sigma/(1+\sigma)$, F_u changes sign only once in the interval $(0, T_s)$.*

Consider some u_0 such that $F_{u_0} = 0$. We can rewrite this equation as:

$$0 = -\int_0^{u_0} g(\mu_{ht}, \mu_{hu_0}) e^{\rho t}\, dt, \quad \text{where} \quad g(x, y) \equiv (s - x)(1-y)^{1+1/\sigma} - (1-x)^{1+1/\sigma}(s - y).$$

The function $x \mapsto g(x, y)$ is strictly concave, and it is such that $g(y, y) = 0$. Note that, for the above equation to hold, the function $x \mapsto g(x, \mu_{hu_0})$ has to change sign in the interval $(0, \mu_{hu_0})$. In particular, it must be the case that $\partial g/\partial x(\mu_{hu_0}, \mu_{hu_0}) < 0$. Otherwise, suppose that $\partial g/\partial x(\mu_{hu_0}, \mu_{hu_0}) \geq 0$. Then, by strict concavity, $g(x, \mu_{hu_0})$ lies strictly below its tangent at $x = \mu_{hu_0}$. But since $g(x, \mu_{hu_0}) = 0$ and is increasing when $x = \mu_{hu_0}$, the tangent is negative for $x \leq \mu_{hu_0}$, and so $g(x, \mu_{hu_0}) < 0$ for all $x \in (0, \mu_{hu_0})$, a contradiction. After calculating the partial derivative, we find:

$$\frac{\partial g}{\partial x}(\mu_{hu_0}, \mu_{hu_0}) < 0 \Leftrightarrow s\left(1 + \frac{1}{\sigma}\right) - 1 - \frac{\mu_{hu}}{\sigma} < 0.$$

Together with Result RA.1 this shows that if $F_{u_0} = 0$ for some $u_0 \in (0, T_s)$, then $F'_{u_0} < 0$, implying Result RA.4. We conclude that, over $(0, T_f]$, F_u is first strictly positive and then strictly negative, which shows that $\overline{Q}_u$ is hump-shaped.

A.11.3 Proof of Lemma A.4

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

$$\Delta_u' = -\frac{1}{\sigma}\frac{\mu_{hu}'}{1-\mu_{hu}}(1-\mu_{hu})^{1/\sigma} Q_u + (1-\mu_{hu})^{1/\sigma} Q_u'.$$

Using the formula (A.3) for Q_u', in Lemma A.1, we obtain:

$$\begin{aligned}\operatorname{sign}[\Delta_u'] &= \operatorname{sign}\left[-\frac{1}{\sigma}\frac{\mu_{hu}'}{1-\mu_{hu}} Q_u + Q_u'\right] \\ &= \operatorname{sign}\left[-\frac{e^{-\rho u}}{\sigma}\frac{\mu_{hu}'}{1-\mu_{hu}} Q_u \int_{\psi_u}^{u} e^{\rho t}(1-\mu_{ht})^{1+1/\sigma}\,dt + s - \mu_{hu} - (1-\mu_{hu})^{1+1/\sigma} Q_u\right]. \qquad \text{(A.37)}\end{aligned}$$

We first show:

RA.5. $\Delta_u' < 0$ *for* u *close to zero.*

To show this result, first note that when u is close to zero, $\overline{Q}_u \simeq s < 1$. Therefore $\psi_u = 0$ and, by Lemma A.2, $Q_u = \overline{Q}_u$. Plugging in $\psi_u = 0$ and the the expression (A.4) for $\overline{Q}_u$ in (A.37), one obtains:

$$\begin{aligned}\operatorname{sign}[\Delta_u'] = \operatorname{sign}\Bigg\{ &-\frac{e^{-\rho u}}{\sigma}\frac{\mu_{hu}'}{1-\mu_{hu}}\int_0^u e^{\rho t}(s-\mu_{ht})\,dt \int_0^u e^{\rho u}(1-\mu_{ht})^{1+1/\sigma}\,dt \\ &+ (s-\mu_{hu})\int_0^u e^{\rho t}(1-\mu_{ht})^{1+1/\sigma}\,dt - (1-\mu_{hu})^{1+1/\sigma}\int_0^u e^{\rho t}(s-\mu_{ht})\,dt\Bigg\}. \qquad \text{(A.38)}\end{aligned}$$

Now let $\gamma \equiv \mu_{h0}'$. Now, for the various functions appearing in the above formula, we calculate the first- and second derivatives at $u = 0$, and we obtain the following Taylor expansions:

$$\begin{aligned}&\frac{e^{-\rho u}}{\sigma}\frac{\mu_{hu}'}{1-\mu_{hu}} = \frac{\gamma}{\sigma}(1+o(1)) \\ &\int_0^u e^{\rho t}(s-\mu_{ht})\,dt = u\left(s + [\rho s - \gamma]\frac{u}{2}\right) = u\,(s+o(1)) \\ &\int_0^u e^{\rho t}(1-\mu_{ht})^{1+1/\sigma}\,dt = u\left(1 + \left[\rho - \gamma\left(1+\frac{1}{\sigma}\right)\right]\frac{u}{2}\right) + o(u) = u\,(1+o(1)) \\ &s - \mu_{hu} = s - \gamma u + o(u) \\ &(1-\mu_{ht})^{1+1/\sigma} = 1 - \gamma\left(1+\frac{1}{\sigma}\right) + o(u).\end{aligned}$$

Plugging these into (A.38) we obtain:

$$\begin{aligned}\operatorname{sign}[\Delta_u'] = \operatorname{sign}\Bigg[&-\frac{\gamma}{\sigma}u^2\,(1+o(1))\,(s+o(1)) \\ &+ u\,(s-\gamma u+o(u))\left(1+\left[\rho-\gamma\left(1+\frac{1}{\sigma}\right)\right]\frac{u}{2}\right) - u\left(1-\gamma\left(1+\frac{1}{\sigma}\right)+o(u)\right)\left(s+[\rho s-\gamma]\frac{u}{2}\right)\Bigg].\end{aligned}$$

After developing and rearranging, we obtain

$$\operatorname{sign}[\Delta_u'] = -\frac{\gamma u^2}{2} \times \operatorname{sign}\left[(1-s) + \frac{s}{\sigma}\right] < 0$$

establishing Result RA.5. Next, we show:

RA.6. *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 (A.37) as follows:

$$\begin{aligned}
\text{sign}\,[\Delta'_u] &= \text{sign}\left[-\frac{e^{-\rho u}}{\sigma}\frac{\mu'_{hu}}{1-\mu_{hu}}\frac{\Delta_u}{(1-\mu_{hu})^{1/\sigma}}\int_{\psi_u}^{u} e^{\rho t}(1-\mu_{ht})^{1+1/\sigma}\,dt + s - \mu_{hu} - (1-\mu_{hu})\Delta_u\right] \\
&= \text{sign}\left[-\frac{1}{\sigma}\frac{\mu'_{hu}}{1-\mu_{hu}}\Delta_u\int_{\psi_u}^{u} e^{-\rho(u-t)}(\frac{1-\mu_{ht}}{1-\mu_{hu}})^{1+1/\sigma}\,dt + \frac{s-\mu_{hu}}{1-\mu_{hu}} - \Delta_u\right] \\
&= \text{sign}\left[-\frac{\gamma}{\sigma}\Delta_u\int_{\psi_u} e^{\left[\gamma(1+\frac{1}{\sigma})-\rho\right](u-t)}\,dt + 1 - (1-s)e^{\gamma u} - \Delta_u\right] \\
&= \text{sign}\left[-\frac{\gamma}{\sigma}\Delta_u G(u-\psi_u) + 1 - (1-s)e^{-\gamma u} - \Delta_u\right], \quad \text{where} \quad G(y) \equiv \int_0^y e^{\left[\gamma(1+\frac{1}{\sigma})-\rho\right]x}\,dx,
\end{aligned}$$

and where we obtain the first equality after substituting in the expression for Q_u; the second equality after dividing by $1-\mu_{hu}$ and bringing $e^{-\rho u}$ inside the first integral; the third equality by using the functional form $\mu_{ht} = 1 - e^{-\gamma t}$; 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:

$$H_{u_0} \equiv -\frac{\gamma}{\sigma}\Delta_{u_0}G(u_0-\psi_{u_0}) + 1 - (1-s)e^{-\gamma u_0} - \Delta_{u_0} = 0.$$

If $Q_{u_0} < 1$ then $\psi_{u_0} = 0$ and $\psi'_{u_0} = 0$. Together with the fact that $\Delta'_{u_0} = 0$, this implies that

$$H'_{u_0} = -\frac{\gamma}{\sigma}\Delta_{u_0}G'(u_0) - (1-s)\gamma e^{-\gamma u_0} < 0,$$

since $G(u)$ is, clearly, a strictly increasing function. If $Q_u = 1$, then $\psi_{u_0} = 0$ and the left-derivative $\psi'_{u_0^-} = 0$, so the same calculation implies that $H'_{u_0^-} < 0$. If $Q_{u_0} > 1$ we first note that, around u_0,

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

So if $\Delta'_{u_0} = 0$, we must have that $\psi'_u = 1$. Plugging this back into H'_u we obtain that $H'_u = (1-s)e^{-\gamma u_0} < 0$. Lastly, if $Q_{u_0} = 1$, then the same calculation leads to $\psi_{u_0^+} = 1$ and so $H_{u_0^+} < 0$. In all cases, we find that H_u 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 RA.6 in mind, we then obtain:

RA.7. *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 RA.6.

A.11.4 Proof of Lemma A.6

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

$$Q_u = (1-\mu_{h\psi_u})^{-1/\sigma}. \tag{A.39}$$

Replacing into equation (A.3) for Q'_u of Lemma A.1 , one obtains that:

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

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 (A.39), $\psi'_{u_0} = 0$. Taking the derivative of X_u at such u_0, and using $X_{u_0} = 0$, leads:

$$\text{sign}\left[X'_{u_0}\right] = \text{sign}\left[-1 + \left(1+\frac{1}{\sigma}\right)\left(\frac{1-\mu_{hu_0}}{1-\mu_{h\psi_{u_0}}}\right)^{1/\sigma}\right] = \text{sign}\left[Y_{u_0}\right], \quad \text{where } Y_u \equiv -1 + \left(1+\frac{1}{\sigma}\right)\frac{s-\mu_{hu}}{1-\mu_{hu}},$$

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 Y_u is strictly decreasing.

A.11.5 Proof of Lemma A.7

Proof of the limit of $T_f(\rho)$, in equation (A.7). The defining equation for $T_f(\rho)$ is

$$H(\rho, T_f(\rho)) = 0 \quad \text{where } H(\rho, u) \equiv \int_0^u e^{\rho t}\left(s - \mu_{ht}\right) dt = 0.$$

Since $T_f > T_s$, we have

$$\frac{\partial H}{\partial u}(\rho, T_f(\rho)) = e^{\rho T_f(\rho)}\left(s - \mu_{hT_f(\rho)}\right) < 0.$$

Turning to the partial derivative with respect to ρ we note that since $\mu_{ht} - s$ changes sign at T_s:

$$\begin{aligned}\frac{\partial H}{\partial \rho}(\rho, T_f(\rho)) &= \int_0^{T_f(\rho)} t \times e^{\rho t}\left(\mu_{ht} - s\right) dt \\ &< \int_0^{T_s} T_s e^{\rho t}\left(s - \mu_{ht}\right) dt + \int_0^{T_s} T_s e^{\rho t}\left(s - \mu_{ht}\right) dt = T_s H(\rho, T_f) = 0.\end{aligned}$$

Taken together, $\partial H/\partial u < 0$ and $\partial H/\partial \rho < 0$ imply that $T_f(\rho)$ is strictly decreasing in ρ. In particular, it has a limit, $T_f(\infty)$, as ρ goes to infinity. To determine the limit, we integrate by part $H(\rho, T_f)$:

$$0 = H(\rho, T_f(\rho)) = s - \mu_{hT_f(\rho)} - se^{-\rho T_f} + \int_0^{\infty} \mathbb{I}_{\{t \in [0, T_f(\rho)]\}} \mu'_{ht} e^{-\rho(T_f - t)}\, dt.$$

Because $T_f(\rho)$ is bounded below by T_s, the second term goes to zero as $\rho \to \infty$. The integrand of the third term is bounded and goes to zero for all t except perhaps at $t = T_f(\infty)$. Thus, but dominated convergence, the third term goes to zero as $\rho \to \infty$. We conclude that $\mu_{hT_f(\infty)} = s$ and hence that $T_f(\infty) = T_s$.

Proof of the first–order expansion, in equation (A.8). Let

$$f(t, \rho) \equiv (1 - \mu_{ht}) \min\left\{(1-\mu_{ht})^{1/\sigma} Q_u(\rho), 1\right\} + \mu_{ht} - s. \tag{A.40}$$

By its definition, $Q_u(\rho)$ solves: $\int_0^u \rho e^{-\rho(u-t)} f(t,\rho)\,dt = 0$. 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:

$$0 = \int_0^u \rho e^{-\rho(u-t)} f(t,\rho)\,dt = f(u,\rho) - e^{-\rho u} f(0,\rho) - \int_0^u e^{-\rho(u-t)} f_t(t,\rho)\,dt, \tag{A.41}$$

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 $(1-\mu_{hu})^{-1/\sigma}$, this sequence has at least one accumulation point $Q_u(\infty)$. Taking the limit in (A.41) along a subsequence converging to this accumulation point, we obtain that $Q_u(\infty)$ solves the equation

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

whose unique solution is $Q_u(\infty) = (s-\mu_{hu})/(1-\mu_{hu})^{1+1/\sigma}$. 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 (A.41):

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

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:

$$0 = f(u,\rho) - \frac{1}{\rho} f_t(u,\rho) + o\left(\frac{1}{\rho}\right). \tag{A.42}$$

Equation (A.8) obtains after substituting in the expressions for $f(u,\rho)$ and $f_t(u,\rho)$, using that $\mu'_{ht} = \gamma(1-\mu_{ht})$.

Proof of the convergence of the argmax, in equation (A.9). 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 $s(1+1/\sigma) \leq 1$, then both $Q_u(\rho)$ and $Q_u(\infty)$ are strictly increasing, achieve their maximum at $u=0$, and the result follows. Otherwise, if $s(1+1/\sigma) > 1$, 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 (A.3) to write the corresponding first–order condition, $Q'_{T_\psi(\rho)} = 0$, we obtain after rearranging that

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

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.

A.11.6 Proof of Lemma A.8

In the perfect cognition case:

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

Since $T_s = -\gamma \log(1-s)$ goes to $+\infty$ when s goes to 1, then the integrand of $I(s)$ converges pointwise towards e^{-ru}. Moreover, the integrand is bounded by e^{-ru}. Therefore, by an application of the Dominated Convergence Theorem, $I(s)$ goes to $\int_0^{+\infty} e^{-ru} \, du = 1/r$ when $s \to 1$.

In the market with limited cognition, for $u > 0$, we note that $Q_u(s)$ is an increasing function of s and is bounded above by $(1-\mu_{hu})^{-1/\sigma}$. Letting $s \to 1$ in the market clearing condition (17) then shows that $Q_u \to (1-\mu_{hu})^{-1/\sigma} > 1$. 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 $1/r$.

A.11.7 Proof of Lemma A.9

In the market with perfect cognition, we can compute:

$$I'(s) = \int_0^{T_s} e^{-ru} \sigma e^{\gamma u} \left(1-(1-s)e^{\gamma u}\right)^{\sigma-1} du + \frac{\partial T_s}{\partial s} \left(1-(1-s)e^{\gamma T_s}\right)^\sigma . \tag{A.43}$$

The second term is equal to 0 since $e^{\gamma T_s} = (1-\mu_{hT_s})^{-1} = (1-s)^{-1}$. We then compute an approximation of the first term when s goes to 1.

Consider first the case when $r > \gamma$. Equation (A.43) rewrites:

$$I'(s) = \sigma \int_0^{+\infty} \mathbb{I}_{\{u<T_s\}} e^{-(r-\gamma)u} \left(1-(1-s)e^{\gamma u}\right)^{\sigma-1} du.$$

Since T_s goes to infinity when s goes to 1, the integrand goes to, and is bounded by, $e^{-(r-\gamma)u}$. Therefore, by dominated convergence, $I'(s)$ goes to $\sigma/(r-\gamma)$.

When $r = \gamma$, equation (A.43) becomes:

$$I'(s) = \sigma \int_0^{T_s} \left(1-(1-s)e^{\gamma u}\right)^{\sigma-1} du = \sigma \int_0^{T_s} \left(1-(1-s)e^{\gamma(T_s-z)}\right)^{\sigma-1} dz = \sigma \int_0^{T_s} \left(1-e^{-\gamma z}\right)^{\sigma-1} dz.$$

where we make the change of variable $z \equiv T_s - u$ to obtain the second equality, and we use that $1-s = e^{-\gamma T_s}$ to obtain the third equality. The integrand goes to 1 when T_s goes to infinity. Thus, the Cesàro mean $I'(s)/T_s$

converges to σ, i.e.:

$$I'(s) \sim \sigma T_s = -\sigma\gamma \log(1-s).$$

Consider now that $r < \gamma$. We make the change of variable $z \equiv T_s - u$ in equation (A.43):

$$\begin{aligned} I'(s) &= \sigma \int_0^{T_s} e^{(\gamma-r)(T_s-z)} \left(1-(1-s)e^{\gamma(T_s-z)}\right)^{\sigma-1} dz \;= \sigma e^{(\gamma-r)T_s} \int_0^{T_s} e^{-(\gamma-r)z} \left(1-e^{-\gamma z}\right)^{\sigma-1} dz \\ &= \sigma e^{(\gamma-r)T_s} \int_0^{+\infty} \mathbb{I}_{\{z<T_s\}} e^{-(\gamma-r)z} \left(1-e^{-\gamma z}\right)^{\sigma-1} dz, \end{aligned}$$

where we use that $1-s = e^{-\gamma T_s}$ to obtain the second equality in the first line. 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 $e^{-\gamma T_s} = 1-s$, we obtain:

$$I'(s) \sim \sigma \left(\frac{1}{1-s}\right)^{1-r/\gamma} \int_0^{+\infty} e^{-(\gamma-r)z} \left(1-e^{-\gamma z}\right)^{\sigma-1} dz.$$

A.11.8 Proof of Lemma A.10

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

$$\gamma + \gamma/\sigma - \rho > 0 \iff \gamma + \sigma(\gamma - \rho) > 0, \tag{A.44}$$

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 (17); and $0 < T_1 < T_2 < T_f$ are defined as in the proof of Proposition 3, 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 Section A.3: $Q_u = (1-\mu_{h\psi_u})^{-1/\sigma}$. By construction, $\psi_u \in (0, u)$, and, as shown in Section A.11.9:

$$\frac{\partial \psi_u}{\partial s} = \frac{\gamma + \sigma(\gamma-\rho)}{\gamma\rho} \frac{(1-e^{-\rho u})\, e^{\gamma u}}{e^{-(\rho-\gamma)(u-\psi_u)} - e^{-(\gamma/\sigma)(u-\psi_u)}}. \tag{A.45}$$

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

$$J'(s) > \frac{\gamma + \sigma(\gamma-\rho)}{\rho} \int_{T_1}^{T_2} e^{-ru} \frac{(1-e^{-\rho u})\, e^{\gamma\psi_u}}{e^{-(\rho-\gamma)(u-\psi_u)} - e^{-(\gamma/\sigma)(u-\psi_u)}}\, du. \tag{A.46}$$

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 A.11.6, for all $u > 0$, $Q_u \to (1-\mu_{hu})^{-1/\sigma}$ as $s \to 1$. Since the integrand in

(A.46) is strictly positive, we have:

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

where the second line follows from the fact, proven is Section A.11.9, that $u-\psi_u$ is strictly increasing in u when $\psi_u>0$. In Section A.11.9 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 $r\le\gamma$. In this case we make the change of variable $z\equiv T_s-u$ in equation (A.46) and we use that $e^{-\gamma T_s}=(1-s)$:

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

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 A.11.9 that, if $\psi_{T_s-z}>0$, then:

$$
e^{\gamma\psi_{T_s-z}} > \begin{cases} \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}}(1-s)^{-1}e^{-\gamma z} & \text{if } \rho\neq\gamma, \\ e^{-(1+\sigma)}(1-s)^{-1}e^{-\gamma z} & \text{if } \rho=\gamma, \end{cases} \tag{A.47}
$$

and:

$$
\left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})}-e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}\right)^{-1} > \frac{\gamma}{\gamma+\sigma(\gamma-\rho)}\frac{\rho}{\max\{2\rho-\gamma,\gamma\}}. \tag{A.48}
$$

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

$$
J'(s) > \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}}\frac{\gamma}{\max\{2\rho-\gamma,\gamma\}}(1-s)^{-1+\frac{r}{\gamma}}\int_0^{+\infty}\mathbb{I}_{\{\max\{T_s-T_2,0\}<z<T_s-T_1\}}e^{-(\gamma-r)z}\left(1-e^{-\rho(T_s-z)}\right)dz. \tag{A.49}
$$

In Section A.11.9 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 (A.49) rewrites:

$$
\begin{aligned}
J'(s) &> \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}}\frac{\gamma}{\max\{2\rho-\gamma,\gamma\}}\int_{\max\{T_s-T_2,0\}}^{T_s-T_1}\left(1-e^{-\rho(T_s-z)}\right)dz \\
&= \left(\frac{\gamma+\sigma(\gamma-\rho)}{\rho}\right)^{\frac{\gamma}{\rho-\gamma}}\frac{\gamma}{\max\{2\rho-\gamma,\gamma\}}\left(T_s-T_1-\max\{T_s-T_2,0\}-\frac{e^{-\rho T_1}-e^{-\rho\min\{T_2,T_s\}}}{\rho}\right).
\end{aligned}
$$

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 $\gamma = \rho$.

A.11.9 Intermediate results for the proofs of Lemma A.8, A.9 and A.10

Derivative of the ψ_u function when $\psi_u > 0$. When $\psi_u > 0$, time–t low–valuation traders hold $q_{t,u} = 1$ if $t < \psi_u$, and $q_{t,u} = (1-\mu_{ht})^{1/\sigma}(1-\mu_{h\psi_u})^{-1/\sigma}$ if $t > \psi_u$. The market clearing condition (17) rewrites:

$$\int_0^{\psi_u} e^{\rho t}(1-\mu_{ht})\,dt + \int_{\psi_u}^{u} e^{\rho t}(1-\mu_{ht})^{1+1/\sigma}(1-\mu_{h\psi_u})^{-1/\sigma}\,dt = \int_0^u e^{\rho t}(s-\mu_{ht})\,dt. \tag{A.50}$$

We differentiate this equation with respect to s:

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

After computing the integrals and rearranging the terms we obtain equation (A.45).

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 (17) and using that $e^{-\gamma T_s} = 1-s$, we obtain, when $\gamma \neq \rho$:

$$\overline{Q}_{T_s} = \left(\frac{1}{\gamma-\rho} + \frac{\gamma(1-s)^{1-\frac{\rho}{\gamma}}}{\rho(\rho-\gamma)} + \frac{1-s}{\rho}\right)\frac{\gamma+\gamma/\sigma-\rho}{1-(1-s)^{\frac{\gamma+\gamma/\sigma-\rho}{\gamma}}}.$$

When $\gamma < \rho$, $\overline{Q}_{T_s}$ goes to infinity when s goes to 1. When $\gamma > \rho$, $\overline{Q}_{T_s}$ goes to $(\gamma+\gamma/\sigma-\rho)/(\gamma-\rho) > 1$. When $\gamma = \rho$, a similar computation shows that $\overline{Q}_{T_s} \sim (\gamma+\gamma/\sigma-\rho)T_s$, which goes to infinity.

Proof that $u - \psi_u$ is strictly increasing in u when $\psi_u > 0$. When $\rho \neq \gamma$, after computing the integrals in (A.50) and rearranging, we obtain:

$$\left(\frac{1}{\rho-\gamma} + \frac{1}{\gamma+\gamma/\sigma-\rho}\right)\left(1-e^{-(\rho-\gamma)(u-\psi_u)}\right) - \frac{1}{\gamma+\gamma/\sigma-\rho}\left(1-e^{-(\gamma/\sigma)(u-\psi_u)}\right) = (1-s)\frac{e^{\gamma u}-e^{-(\rho-\gamma)u}}{\rho}. \tag{A.51}$$

Taking the derivative of the left-hand side with respect to $u - \psi_u$ we easily obtain that the left-hand side of that equation is strictly increasing in $u-\psi_u$. 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 same computation leads to:

$$\int_0^{u-\psi_u}\left(e^{-\gamma/\sigma z}-1\right)dz + \int_0^u e^{-\rho t}(1-s)\,dt = 0$$

which is strictly decreasing in $u-\psi_u$ and strictly increasing in u, 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 A.11.6, for any u, $Q_u \to (1-\mu_{hu})^{-1/\sigma}$ 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 $s \to$.

Proof of equation (A.47). When $\gamma \neq \rho$, we make the change of variable $z \equiv T_s - u$ in the market clearing condition (A.51):

$$\begin{aligned}
&\left(\frac{1}{\rho-\gamma} + \frac{1}{\gamma+\gamma/\sigma-\rho}\right) e^{(\rho-\gamma)\psi_{T_s-z}} - \frac{(1-s)^{\frac{\gamma+\gamma/\sigma-\rho}{\gamma}} e^{(\gamma+\gamma/\sigma-\rho)z}}{\gamma+\gamma/\sigma-\rho} e^{(\gamma/\sigma)\psi_{T_s-z}} \\
&= (1-s)^{-\frac{\rho-\gamma}{\gamma}} \left(\frac{e^{-(\rho-\gamma)z}}{\rho-\gamma} - \frac{e^{-\rho z}}{\rho}\right) + \frac{1-s}{\rho},
\end{aligned} \tag{A.52}$$

where we have used that $e^{-\gamma T_s} = (1-s)$. This implies that:

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

where, to move from the first to the second line, we have collected terms on the left–hand side and used $e^{-\rho z} < e^{-(\rho-\gamma)z}$ on the right-hand side. Equation (A.47) then follows after applying to both sides the increasing transformation:

$$x \mapsto \left(\frac{(\rho-\gamma)[\gamma+\sigma(\gamma-\rho)]}{\gamma} x\right)^{\frac{\gamma}{\rho-\gamma}}$$

Finally, when $\gamma = \rho$, after computing the integrals in the market clearing condition (A.50), making the change of variable $z \equiv u - T_s$, and using that $e^{-\gamma T_s} = (1-s)$, we obtain:

$$T_s - z - \psi_{T_s-z} - \frac{1 - e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}}{\gamma/\sigma} = \frac{e^{-\gamma z}}{\gamma} - \frac{1-s}{\gamma}.$$

This implies:

$$\psi_{T_s-z} > T_s - z - \frac{\sigma}{\gamma} - \frac{e^{-\gamma z}}{\gamma} > T_s - z - \frac{\sigma+1}{\gamma},$$

from which equation (A.47) follows by multiplying by γ, taking the exponent of the expression, and using that $e^{-\gamma T_s} = (1-s)$.

Proof of equation (A.48). When $\gamma \neq \rho$, we make the change of variable $z \equiv T_s - u$ in equation (A.51):

$$\begin{aligned}
&\left(\frac{1}{\rho-\gamma} + \frac{1}{\gamma+\gamma/\sigma-\rho}\right) e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})} - \frac{1}{\gamma+\gamma/\sigma-\rho} e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})} \\
&= \frac{1}{\rho-\gamma} - \frac{e^{-\gamma z}}{\rho} + \frac{e^{-\gamma z-\rho(T_s-z)}}{\rho}.
\end{aligned} \tag{A.53}$$

When $\rho > \gamma$, we add $-1/(\rho-\gamma) \times e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}$, which is negative, to the left–hand side of (A.53) and obtain:

$$\begin{aligned}
&\frac{\gamma}{(\rho-\gamma)[\gamma+\sigma(\gamma-\rho)]}\left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})}-e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}\right)<\frac{1}{\rho-\gamma}-\frac{e^{-\gamma z}}{\rho}+\frac{e^{-\gamma z-\rho(T_s-z)}}{\rho}\\
\Longrightarrow\quad&\left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})}-e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}\right)>\frac{(\rho-\gamma)[\gamma+\sigma(\gamma-\rho)]}{\gamma}\left(\frac{1}{\rho-\gamma}+\frac{1}{\rho}\right)\\
\Longrightarrow\quad&\left(e^{-(\rho-\gamma)(T_s-z-\psi_{T_s-z})}-e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}\right)>\frac{(2\rho-\gamma)[\gamma+\sigma(\gamma-\rho)]}{\rho\gamma}, \qquad \text{(A.54)}
\end{aligned}$$

where we move from the first to the second line by multiplying both sides by $(\rho-\gamma)[\gamma+\sigma(\gamma-\rho)]/\gamma$, which is a positive number since $\rho>\gamma$. Equation (A.48) then follows.

When $\rho<\gamma$, we also add $-1/(\rho-\gamma)\times e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}$ to the left–hand side of (A.53). But, since $\rho<\gamma$ this term is negative so we obtain the opposite inequality. This inequality is reversed when we multiply both sides of the equation by $(\rho-\gamma)[\gamma+\sigma(\gamma-\rho)]/\gamma$, which is a negative number since $\rho<\gamma$. Thus, we end up with the same inequality, (A.54), and equation (A.48) follows.

Finally, when $\gamma=\rho$, equation (A.48) follows since $1-e^{-(\gamma/\sigma)(T_s-z-\psi_{T_s-z})}<1$.

A.11.10 Proof of Lemma A.11

One sees easily that, since the left–hand side of (A.18) is strictly increasing and $s>\sigma/(1+\sigma)$, equation (A.18) has a unique solution. Moreover, since $s<1$, the left–hand side of (A.18) is greater than the right–hand side when evaluated at μ_{hT_s}, implying that $T_\phi<T_s$. Next, define

$$\begin{aligned}
&H(t,\phi)\equiv\int_t^\phi e^{\rho u}g(\mu_{ht},\mu_{hu})\,du,\\
\text{where}\quad &g(x,y)\equiv(1-y)^{1+1/\sigma}(s-x)-(1-x)^{1+1/\sigma}(s-y).
\end{aligned}$$

Let $t<T_\phi$ and $x=\mu_{ht}$. Since $x=\mu_{ht}<\mu_{hT_\phi}<\mu_{hT_s}=s$, the function $y\mapsto g(x,y)$ is strictly convex. Moreover, $g(x,x)=0$, and

$$\frac{\partial g}{\partial y}(x,x)=(1-x)^{1/\sigma}\left[\frac{x}{\sigma}-\left(1+\frac{1}{\sigma}\right)s+1\right].$$

This partial derivative is strictly negative since $x<\mu_{hT_\phi}$. Therefore $g(x,y)$ is strictly negative for y just above x. Since $g(x,1)=(1-x)^{1+1/\sigma}(1-s)>0$, this implies that $y\mapsto g(x,y)$ has a root in $(x,1)$. Because of strict convexity it is the only root; we denote it by $\mu_{ht'}$ with $t'>t$. It follows that $\phi\mapsto H(t,\phi)$ is strictly decreasing over $[t,t']$ and strictly increasing over $[t',+\infty)$. Now we note that $H(t,t)=0$. Moreover, $H(t,\phi)$ goes to $+\infty$ when $\phi\to+\infty$: indeed, μ_{hu} converges to 1 when $u\to+\infty$ and $g(x,1)>0$. Taken together, this means that $\phi\mapsto H(t,\phi)$ has a unique root $\phi_t>t'>t$.

We now establish that $t\mapsto\phi_t$ is a strictly decreasing function. First we note that $\partial H/\partial\phi(t,\phi_t)=e^{\rho\phi_t}g(\mu_{ht},\mu_{h\phi_t})>0$ since $\phi_t>t'$. Then,

$$\frac{\partial H}{\partial t}(t,\phi_t)=-e^{\rho t}g(\mu_{ht},\mu_{ht})+\mu'_{ht}\int_t^{\phi_t}e^{\rho u}\frac{\partial g}{\partial x}(\mu_{ht},\mu_{hu})\,du.$$

The first term is equal to zero because $g(x,x)=0$ for all $x\in[0,1]$. To evaluate the sign of the second term, we note that

$$\int_t^{\phi_t} e^{\rho u} \frac{\partial g}{\partial x}(\mu_{ht}, \mu_{hu})\, du = \int_t^{\phi_t} e^{\rho u} \left[-(1-\mu_{hu})^{1+1/\sigma} + \left(1 + \frac{1}{\sigma}\right)(1-\mu_{ht})^{1/\sigma}(s - \mu_{hu}) \right] du.$$

But, since $H(t, \phi_t) = 0$, we have

$$\int_t^{\phi_t} e^{\rho u}(1-\mu_{hu})^{1+1/\sigma}\, du = \frac{(1-\mu_{ht})^{1+1/\sigma}}{s - \mu_{ht}} \int_t^{\phi_t} e^{\rho u}(s - \mu_{hu})\, du.$$

Plugging this into the equation just before, we obtain that $\partial H/\partial t(t, \phi_t)$ has the same sign as

$$-(1-\mu_{ht}) + \left(1 + \frac{1}{\sigma}\right)(s - \mu_{ht}) = -\frac{\mu_{ht}}{\sigma} + \left(1 + \frac{1}{\sigma}\right) s - 1,$$

which is strictly positive since $t < T_\phi$. An application of the Implicit Function Theorem shows, then, that ϕ_t is strictly decreasing and continuously differentiable.

It remains to show that $\lim_{t \to T_\phi} \phi_t = T_\phi$ and that $\phi_0 < T_f$. First, note that since ϕ_t is strictly decreasing for $t \in [0, T_\phi)$, it has a well defined limit as $t \to T_\phi$. Moreover, it must be that $\phi_t \geq T_\phi$. Indeed, if $\phi_{t_1} < T_\phi$ for some t_1, then for all $t_2 \in (\phi_{t_1}, T_\phi)$ we have $\phi_{t_2} > t_2 > \phi_{t_1}$, which is impossible since ϕ_t is strictly decreasing. In particular, we must have that $\lim_{t \to T_\phi} \phi_t \geq T_\phi$. Now, towards a contradiction, assume that $\lim_{t \to T_\phi} \phi_t > T_\phi$. Note that, given $\partial g/\partial y(\mu_{hT_\phi}, \mu_{hu}) > 0$ for all $u > T_\phi$ and $g(\mu_{hT_\phi}, \mu_{hT_\phi}) = 0$, we have $g(\mu_{hT_\phi}, \mu_{hu}) > 0$ for all $u > T_\phi$. Therefore

$$0 < \int_{T_\phi}^{\lim_{t \to T_\phi} \phi_t} e^{\rho u} g(\mu_{hT_\phi}, \mu_{hu})\, du = H(T_\phi, \lim_{t \to T_\phi} \phi_t) = \lim_{t \to T_\phi} H(t, \phi_t) = 0,$$

by continuity of H, which is a contradiction. Therefore, we conclude that $\lim_{t \to T_\phi} \phi_t = T_\phi$.

In order to show that $\phi_0 < T_f$ we only need to show that $H(0, T_f) > 0$ because $H(0, \phi) \leq 0$ for all $\phi \leq \phi_0$. But we have

$$H(0, T_f) = \int_0^{T_f} e^{\rho u}(1-\mu_{hu})^{1+1/\sigma} s\, du - \int_0^{T_f} e^{\rho u}(s - \mu_{hu})\, du > 0$$

since the first integral is strictly positive and the second integral is equal to zero by definition of T_f.

A.11.11 Proof of Lemma A.12

Direct calculations show that

$$\frac{d}{dt}\left[\frac{s - \mu_{ht}}{(1-\mu_{ht})^{1+1/\sigma}}\right] = \frac{\mu'_{ht}}{(1-\mu_{ht})^{2+1/\sigma}} \left[s\left(1 + \frac{1}{\sigma}\right) - 1 - \frac{\mu_{ht}}{\sigma} \right] \tag{A.55}$$

The first multiplicative term on the right–hand side is always strictly positive because $\mu'_{ht} > 0$. When $t = T_\phi$, the second multiplicative term is zero by definition of T_ϕ. When $t < T_\phi$, it is strictly positive given that μ_{ht} is strictly increasing.

A.11.12 Proof of Lemma A.13

Let us now turn to the proof of Lemma A.13. For $t < T_\phi$, $rp_t - \dot{p}_t$ is given by equation (A.20). Since $\delta < 1$, $(s - \mu_{ht})/(1 - \mu_{ht}) < 1$, and using Lemma A.12, we obtain that $rp_t - \dot{p}_t > 0$. In order to show that it is strictly below 1, we need some additional computations. First, note that equation (A.55) implies that:

$$\frac{d}{dt}\left[\left(\frac{s-\mu_{ht}}{(1-\mu_{ht})^{1+1/\sigma}}\right)^{\sigma}\right] = \sigma\left(\frac{s-\mu_{ht}}{(1-\mu_{ht})^{1+1/\sigma}}\right)^{\sigma-1}\frac{\mu'_{ht}}{(1-\mu_{ht})^{2+1/\sigma}}\left[s\left(1+\frac{1}{\sigma}\right)-1-\frac{\mu_{ht}}{\sigma}\right].$$

Next, we plug $\mu'_{ht} = \gamma(1 - \mu_{ht})$ in the above, and then we plug the resulting expression in equation (A.20). After some algebraic manipulations, we obtain:

$$rp_t - \dot{p}_t = 1 - \delta\left(\frac{s-\mu_{ht}}{1-\mu_{ht}}\right)^{\sigma}\left\{1 - \left[1+\sigma-\sigma\frac{1-\mu_{ht}}{s-\mu_{ht}}\right]\frac{\gamma}{1-\mu_{ht}}\int_t^{\phi_t} e^{-(r+\rho)(u-t)}(1-\mu_{hu})\,du\right\}. \quad \text{(A.56)}$$

One easily check that the term in brackets, $1 + \sigma - \sigma(1 - \mu_{ht})/(s - \mu_{ht})$, is strictly smaller than 1 because $s < 1$, and greater than zero because $t \leq T_\phi$. On the other hand, after multiplying the integral term by $\gamma/(1 - \mu_{ht}) = \gamma e^{\gamma t}$ we find:

$$\frac{\gamma}{r+\rho+\gamma}\left(1 - e^{-(r+\rho+\gamma)(\phi_t - t)}\right) < 1. \quad \text{(A.57)}$$

Taken together, these inequalities imply that $rp_t - \dot{p}_t \in (0, 1)$.

For $t \in (T_\phi, \phi_0)$, $rp_t - \dot{p}_t$ is given by equation (A.21). We have $(1 - \mu_{h\phi_t^{-1}})/(1 - \mu_{ht}) \in (0, 1)$ since $\phi_t^{-1} < T_\phi < t$, and $(s - \mu_{h\phi_t^{-1}})/(1 - \mu_{h\phi_t^{-1}}) \in (0, 1)$ since $\phi_t^{-1} < T_\phi < T_s$. Therefore $rp_t - \dot{p}_t \in (0, 1)$.

For $t \in (\phi_0, T_\phi)$, $rp_t - \dot{p}_t$ is the same in the ATE, thus it also lies in $(0, 1)$.

A.11.13 Proof of Lemma A.14

Plugging equation (A.56) into the definition of Q_u, we obtain that, for $u \in (0, T_\phi)$,

$$Q_u = \frac{s-\mu_{hu}}{(1-\mu_{hu})^{1+1/\sigma}}\left[1 - \frac{\gamma}{r+\rho+\gamma}\left(1-e^{-(r+\rho+\gamma)(\phi_u-u)}\right)\left(1+\sigma-\sigma\frac{1-\mu_{hu}}{s-\mu_{hu}}\right)\right]^{1/\sigma}.$$

From Lemma A.12 we know that $(s - \mu_{hu})/((1 - \mu_{hu})^{1+1/\sigma}$ is strictly increasing in u over $(0, T_\phi)$. In the term in brackets, the first term in parentheses is strictly positive and strictly decreasing in u since ϕ_u is strictly decreasing. The second term in parentheses is strictly positive because $u < T_\phi$ and it is strictly decreasing in u since $\mu_{hu} < s < 1$ when $u < T_\phi < T_s$.

For $t \in (T_\phi, \phi_0)$,

$$Q_u = \frac{s-\mu_{h\phi_u^{-1}}}{(1-\mu_{h\phi_u^{-1}})^{1+1/\sigma}}.$$

This is a strictly decreasing function of u because $u \mapsto \phi_u^{-1}$ is strictly decreasing and belongs to $(0, T_\phi)$, and $x \mapsto (s - x)/(1 - x)^{1+1/\sigma}$ is strictly increasing over $(0, \mu_{hT_\phi})$ by Lemma A.12.

For $t \in (\phi_0, T_f)$, $Q_u = \overline{Q}_u$. Since $u \mapsto \overline{Q}_u$ is hump-shaped by Lemma A.3 all we need to show is $\overline{Q}'_{\phi_0} < 0$.

To that end, note first that $H(0,\phi_0)=0$ writes as

$$\int_0^{\phi_0} e^{\rho t}(1-\mu_{ht})^{1+1/\sigma} s\,dt = \int_0^{\phi_0} e^{\rho t}(s-\mu_{ht})\,dt.$$

From the proof of Lemma A.3, in Section A.11.2, equation (A.35) we know that $\overline{Q}'_{\phi_0}$ has the same sign as

$$\int_0^{\phi_0} e^{\rho t}\left[(1-\mu_{ht})^{1+1/\sigma}(s-\mu_{h\phi_0}) - (1-\mu_{h\phi_0})^{1+1/\sigma}(s-\mu_{ht})\right] dt.$$

Replacing the first equation into the second, we find that $\overline{Q}'_{\phi_0}$ as the same sign as

$$s-\mu_{h\phi_0} - s(1-\mu_{h\phi_0})^{1+1/\sigma} = -g(0,\mu_{h\phi_0}),$$

where the function g is defined in the proof of Lemma A.11. But we already know from this proof that $g(0,\mu_{h\phi_0})>0$ has to hold for $H(0,\phi_0)=0$, which from the above imply that $\overline{Q}'_{\phi_0}<0$.

It remains to show that $u\mapsto Q_u$ is continuous at $u=T_\phi$ and $u=\phi_0$. Starting from equations (A.20) and (A.21), continuity at $u=T_\phi$ follows from $\phi^{-1}_{T_\phi}=T_\phi$ and Lemma A.12. Turning to continuity at $u=\phi_0$, (A.21) evaluated at $t=\phi_0$ yields $Q_{\phi_0^-}=s$. On the other hand, plugging $H(0,\phi_0)=0$ into the definition (A.4) of $\overline{Q}_u$ we obtain that $Q_{\phi_0^+}=\overline{Q}_{\phi_0}=s$.

A.11.14 Proof of Lemma A.15

By construction the price is continuously differentiable in all the open intervals $(0,T_\phi)$, (T_ϕ,ϕ_0), (ϕ_0,T_f), and (T_f,∞). Let us show that it is also continuously differentiable at the boundary points of these intervals. First note that, by definition of Q_t, in equation (A.24), it follows that

$$rp_u = 1-\delta(1-\mu_{hu})Q_u^\sigma + \dot{p}_u. \tag{A.58}$$

Since by Lemma A.14, Q_u is continuous over $(0,T_f)$, and since the price is continuous by construction, it follows that $\dot{p}_u$ is continuous over $(0,T_f)$ as well. Turning to $t=T_f$, we have $Q_{T_f}=\overline{Q}_{T_f}=0$ by definition of T_f and of $\overline{Q}_u$ in equation (A.4). Plugging $Q_{T_f}=0$ in (A.58), it follows that $rp_{T_f}=1+\dot{p}_{T_f^-}$. Since $p_{T_f}=1/r$ it follows that $\dot{p}_{T_f^-}=0$. Since p_t is constant for $t>T_f$, this shows that $\dot{p}_t$ is continuous at T_f.

Next, we show that the price is strictly increasing over $(0,T_f)$. We start with the time interval (T_ϕ,T_f). Since by Lemma A.14, Q_u is strictly decreasing over (T_ϕ,T_f), it follows that $\Delta_u=(1-\mu_{hu})^{1/\sigma}Q_u$ is strictly decreasing over (T_ϕ,T_f). Using the same argument as in the proof Proposition 1 in Section A.4, it follows that the price is strictly increasing over (T_ϕ,T_f).

The proof is more difficult for the initial time interval, $[0,T_\phi]$. We start by defining, for $t\in[0,T_\phi]$:

$$\delta_t \equiv \delta\left(\frac{s-\mu_{ht}}{1-\mu_{ht}}\right)^\sigma.$$

Clearly, since μ_{ht} is strictly increasing, we have $\delta'_t<0$. Also, using $\mu'_{ht}=\gamma(1-\mu_{ht})$ one easily sees after some algebra that:

$$\gamma\delta_t+\delta'_t = \gamma\sigma\delta\frac{(s-\mu_{ht})^{\sigma-1}}{(1-\mu_{ht})^\sigma}\left[s\left(1+\frac{1}{\sigma}\right)-1-\frac{\mu_{ht}}{\sigma}\right]\geq 0 \tag{A.59}$$

for $t \in [0, T_\phi]$, by definition of T_ϕ. With the definition of δ_t, and keeping in mind that $1 - \mu_{ht} = e^{-\gamma t}$, ODE (A.20) writes:

$$\begin{aligned} rp_t - \dot{p}_t &= 1 - \delta_t + \frac{d}{dt}\left[\delta_t e^{\gamma t}\right] e^{-\gamma t} \int_t^{\phi_t} e^{-(r+\rho+\gamma)(u-t)}\, du \\ &= 1 - \delta_t + \frac{\delta_t + \gamma \delta_t'}{r+\rho+\gamma} \left(1 - e^{-(r+\rho+\gamma)(\phi_t - t)}\right) \end{aligned} \tag{A.60}$$

And, ODE (A.21) writes:

$$rp_t - \dot{p}_t = 1 - e^{-\gamma(t - \phi_t^{-1})} \delta_{\phi_t^{-1}}. \tag{A.61}$$

Next, we differentiate equations (A.60) and (A.61) to find ODEs for $d_t \equiv \dot{p}_t$:

$$\begin{aligned} t \in (0, T_\phi): \quad rd_t - \dot{d}_t &= -\delta_t' + \left[\gamma \delta_t + \delta_t'\right] e^{-(r+\rho+\gamma)(\phi_t - t)} \left(\phi_t' - 1\right) + \frac{\gamma \delta_t' + \delta_t''}{r+\rho+\gamma} \left(1 - e^{-(r+\rho+\gamma)(\phi_t - t)}\right), \\ t \in (T_\phi, \phi_0): \quad rd_t - \dot{d}_t &= -\delta'_{\phi_t^{-1}} e^{-\gamma(t - \phi_t^{-1})} + \left[\gamma \delta_{\phi_t^{-1}} + \delta'_{\phi_t^{-1}}\right] \left(1 - \frac{1}{\phi'_{\phi_t^{-1}}}\right) e^{-\gamma(t - \phi_t^{-1})}. \end{aligned}$$

We already know that the price is continuously differentiable and hence that $d_t = \dot{p}_t$ is continuous over $[0, T_f]$. This allows to write:

$$d_t = \int_t^{\phi_t} e^{-r(u-t)} \left(rd_u - \dot{d}_u\right) du + e^{-r(\phi_t - t)} d_{\phi_t}.$$

Since $\phi_t \geq T_\phi$, we already know that $d_{\phi_t} \geq 0$. So, in order to show that $d_t \geq 0$, it suffices to show that the integral is positive. To that end, equipped with the above analytical expressions of $rd_t - \dot{d}_t$, the integral can be written as sum of five terms:

$$\begin{aligned} \text{Term (I)}: &- \int_t^{T_\phi} e^{-r(u-t)} \delta_u'\, du \\ \text{Term (II)}: &- \int_{T_\phi}^{\phi_t} e^{-r(u-t)} \delta'_{\phi_u^{-1}} e^{-\gamma(u - \phi_u^{-1})}\, du \\ \text{Term (III)}: &\int_t^{T_\phi} e^{-r(u-t)} \frac{\gamma \delta_u' + \delta_u''}{r+\rho+\gamma} \left(1 - e^{-(r+\rho+\gamma)(\phi_u - u)}\right) du \\ \text{Term (IV)}: &\int_t^{T_\phi} e^{-r(u-t)} \left[\gamma \delta_u + \delta_u'\right] \left(\phi_u' - 1\right) e^{-(r+\rho+\gamma)(\phi_u - u)}\, du \\ \text{Term (V)}: &\int_{T_\phi}^{\phi_t} e^{-r(u-t)} \left[\gamma \delta_{\phi_u^{-1}} + \delta'_{\phi_u^{-1}}\right] \left(1 - \frac{1}{\phi'_{\phi_u^{-1}}}\right) e^{-\gamma(u - \phi_u^{-1})}\, du. \end{aligned}$$

We make the change of variable $u = \phi_z$ in Term (V) and obtain that

$$\text{Term (V)} = \int_t^{T_\phi} e^{-r(z-t)} \left[\gamma \delta_z + \delta_z'\right] \left(1 - \phi_z'\right) e^{-(r+\gamma)(\phi_z - z)}\, dz.$$

Since, by equation (A.59), $\gamma \delta_z + \delta_z' \geq 0$ for $z \leq T_\phi$, and since by Lemma A.11, $\phi_z' \leq 0$, this implies that the sum of terms (IV) and (V) is positive.

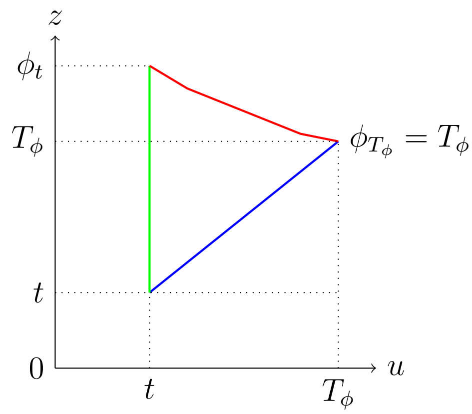

Figure 9: The green vertical line is the segment $u = t$, from $u = 0$ to $u = \phi_t$. The blue upward slopping line is the segment $z = u$, from $u = t$ to $u = T_\phi$. The red downward slopping curve is the function $z = \phi_u$, from $u = t$ to $u = T_\phi$. Thus, the domain of integration is the area enclosed in the "triangle-shaped" area between the green line, the blue line, and the red curve.

Next, since

$$\frac{1 - e^{-(r+\rho+\gamma)(\phi_u - u)}}{r + \rho + \gamma} = \int_u^{\phi_u} e^{-(r+\rho+\gamma)(z-u)}\, dz$$

we can rewrite Term (III) as

$$\int_t^{T_\phi} [\gamma\delta'_u + \delta''_u] e^{(\rho+\gamma)u} \int_u^{\phi_u} e^{-(\rho+\gamma)z - r(z-t)}\, dz\, du.$$

Using the graphical representation of the domain of integration shown in Figure 9, we switch round the two integrals and obtain

$$\begin{aligned}\text{Term (III)} = &\int_t^{T_\phi} \left[\int_t^z [\gamma\delta'_u + \delta''_u] e^{(\rho+\gamma)u}\, du\right] e^{-(\rho+\gamma)z - r(z-t)}\, dz \\ &+ \int_{T_\phi}^{\phi_t} \left[\int_t^{\phi_z^{-1}} [\gamma\delta'_u + \delta''_u] e^{(\rho+\gamma)u}\, du\right] e^{-(\rho+\gamma)z - r(z-t)}\, dz.\end{aligned}$$

Since $\delta'_t < 0$, it follows that

$$[\gamma\delta'_u + \delta''_u] e^{(\rho+\gamma)u} > [(\rho+\gamma)\delta'_u + \delta''_u] e^{(\rho+\gamma)u} = \frac{d}{du}\left[\delta'_u e^{(\rho+\gamma)u}\right].$$

Plugging this inequality into the last expression we found for Term (III) and explicitly integrating with respect to u, we find

$$\begin{aligned}
\text{Term (III)} > &\int_t^{T_\phi} \left[\delta'_z e^{(\rho+\gamma)z} - \delta'_t e^{(\rho+\gamma)t}\right] e^{-(\rho+\gamma)z-r(z-t)}\,dz \\
&+ \int_{T_\phi}^{\phi_t} \left[\delta'_{\phi_z^{-1}} e^{(\rho+\gamma)\phi_z^{-1}} - \delta'_t e^{(\rho+\gamma)t}\right] e^{-(\rho+\gamma)z-r(z-t)}\,dz \\
> &\int_t^{T_\phi} \delta'_z e^{-r(z-t)}\,dz + \int_{T_\phi}^{\phi_t} \delta'_{\phi_z^{-1}} e^{-(\rho+\gamma)(z-\phi_z^{-1})-r(z-t)}\,dz - \int_t^{\phi_t} \delta'_t e^{-(r+\rho+\gamma)(z-t)}\,dz.
\end{aligned}$$

The last term is positive since $\delta'_t < 0$. The first term cancels out with Term (I). Adding up the second term in the above equation with Term (II), we obtain

$$\int_{T_\phi}^{\phi_t} e^{-r(z-t)} \delta'_{\phi_z^{-1}} \left[e^{-(\rho+\gamma)(z-\phi_z^{-1})} - e^{-\gamma(z-\phi_z^{-1})}\right] dz,$$

which is positive given that $\delta'_u < 0$. Taken together, this all show that the sum of Terms (I) through (V) is positive, and hence that $d_t = \dot{p}_t > 0$ for all $t \in [0, T_\phi]$.

A.11.15 Proof that condition (i) and (ii) hold

From the definition of $M(u,q)$ in equation (A.29), and keeping in mind that $\xi_u = 1-\delta(1-\mu_{hu})Q_u^\sigma$ by definition of Q_u, we have

$$\frac{\partial M}{\partial q}(u,q) = 1 - \xi_u - \delta\frac{1-\mu_{h_u}}{1-\mu_{ht}} q^\sigma = \delta\frac{1-\mu_{hu}}{1-\mu_{ht}}\left[(1-\mu_{ht})Q_u^\sigma - q^\sigma\right]. \tag{A.62}$$

By the definition of $q_{t,u}$ in equation (A.24) and (A.26) this implies that

$$\frac{\partial M}{\partial q}(u, q_{t,u}) = 0 \quad \text{for all } t \in (0, T_\phi],\ u \in [\phi_t, T_f);\ \text{and for all } t \in [T_\phi, T_f),\ u \in [t, T_f]. \tag{A.63}$$

This clearly implies that condition (i) holds for $t \geq T_\phi$. For $t \in (0, T_\phi]$, we define

$$N_t = \int_t^{\phi_t} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u})\,du.$$

First, we note that $\phi_{T_\phi} = T_\phi$ implies $N_{T_\phi} = 0$. Then, we differentiate N_t with respect to t

$$\begin{aligned}
N'_t &= (r+\rho)N_t - \frac{\partial M}{\partial q}(t, q_{t,t}) + \int_t^{\phi_t} e^{-(r+\rho)(u-t)} \frac{\partial}{\partial t}\left[\frac{\partial M}{\partial q}(u, q_{t,t})\right] du \\
&= (r+\rho)N_t - \left(1-\xi_t - \delta q_{t,t}^\sigma\right) - \delta\int_t^{\phi_t} e^{-(r+\rho)(u-t)}(1-\mu_{hu})\frac{\partial}{\partial t}\left[\frac{q_{t,t}^\sigma}{1-\mu_{ht}}\right] du \\
&= (r+\rho)N_t,
\end{aligned}$$

where: for the first equality we used that $\partial M/\partial q(u, q_{t,\phi_t}) = 0$, and in the integral we have substituted $q_{t,u} = q_{t,t}$ since $u \in [t, \phi_t]$; we obtain the second equality by evaluating $\partial M/\partial q(u,q)$, in equation (A.62), at $(t, q_{t,t})$; we obtain the third equality from equation (A.20) and $q_{t,t} = (s-\mu_{ht})/(1-\mu_{ht})$. Therefore we have a differential equation for N_t. Given the boundary condition $N_{T_\phi} = 0$, we have

$$N_t = 0 \quad \text{for all } t \in [0, T_\phi]. \tag{A.64}$$

With this in mind, we turn to condition (i) for $t \in [0, T_\phi]$ and note that:

$$\begin{aligned}
\int_t^{T_f} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u}) q_{t,u}\, du =& q_{t,t} \int_t^{\phi_t} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u})\, du + \int_{\phi_t}^{T_f} \frac{\partial M}{\partial q}(u, q_{t,u}) q_{t,u}\, du \\
=& q_{t,t} N_t + \int_{\phi_t}^{T_f} \frac{\partial M}{\partial q}(u, q_{t,u}) q_{t,u}\, du = 0,
\end{aligned}$$

where: we obtain the first equality by breaking the interval of integration into $[t, \phi_t]$ and $[\phi_t, T_f]$, and noting that $q_{t,u} = q_{t,t}$ over $[t, \phi_t]$; we obtain the second equality by recognizing that the first integral is equal to N_t; the last equality follows from (A.63) and (A.64). This establishes condition (i) holds.

Let us now turn to condition (ii). Note first that this condition holds for $t \in [T_\phi, T_f]$ since in that case $\partial M/\partial q(u, q_{t,u}) = 0$. To show that it also holds for $t \in (0, T_\phi)$, we show:

RA.8. *There exists $u_1 \in (t, \phi_t)$ such that $u \mapsto \partial M/\partial q(u, q_{t,u}) < 0$ for $u \in (t, u_1)$ and $\partial M/\partial q(u, q_{t,u}) > 0$ for $u \in (u_1, \phi_t)$.*

Indeed, by equation (A.62):

$$\text{sign}\left[\frac{\partial M}{\partial q}(u, q_{t,u})\right] = \text{sign}\left[F_u\right] \qquad \text{where } F_u \equiv (1 - \mu_{ht}) Q_u^\sigma - q_{t,u}^\sigma.$$

By Lemma A.14 we know that Q_u is strictly increasing before T_ϕ and strictly decreasing after T_ϕ. Also, by construction of the candidate LOE, $q_{t,u}$ is constant over $[t, \phi_t]$. Thus F_u is strictly increasing over $[t, T_\phi)$, and strictly decreasing over $(T_\phi, \phi_t]$. Second, when $u = t$, we have:

$$F_t = (1 - \mu_{ht}) Q_t^\sigma - q_{t,t}^\sigma = \frac{1 - \xi_t}{\delta} - q_{t,t}^\sigma = \frac{1}{\delta}\left[1 - \delta\left(\frac{s - \mu_{ht}}{1 - \mu_{ht}}\right)^\sigma - \xi_t\right],$$

where the first equality follows by definition of Q_u, in equation (A.24), and the second equality follows by definition of $q_{t,t}$ for $t \in (0, T_\phi)$, in equation (A.23). Now, by equation (A.20) and Lemma A.12, this last expression is strictly negative for $t < T_\phi$, i.e., $F_t < 0$. Third, the asset holding plan is continuous at $u = \phi_t$ so $F_{\phi_t} = (1 - \mu_{ht}) Q_{\phi_t}^\sigma - q_{t,t}^\sigma = (1 - \mu_{ht}) Q_{\phi_t}^\sigma - q_{t,\phi_t}^\sigma$. But this last expression is equal to zero by equation (A.24). Taken together, the above shows that F_u is increasing over $[t, T_\phi]$, decreasing over $[T_\phi, \phi_t]$, negative at $u = t$, and zero at $u = \phi_t$. This shows that there exists $u_1 \in (t, \phi_t)$ such that $F_u < 0$ for $u \in (t, u_1)$ and $F_u > 0$ for $u \in (u_1, \phi_t)$. Because $u \mapsto \partial M/\partial q(u, q_{t,u})$ has the same sign as F_u, result RA.8 follows.

Next, for any decreasing function $\tilde{q}_{t,u}$, we have

$$\begin{aligned}
&\int_t^{T_f} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u}) \tilde{q}_{t,u}\, du \\
\leq &\int_t^{u_1} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u}) \tilde{q}_{t,u_1}\, du + \int_{u_1}^{\phi_t} e^{-(r+\rho)(u-t)} \frac{\partial M}{\partial q}(u, q_{t,u}) \tilde{q}_{t,u_1}\, du = N_t \tilde{q}_{t,u_1},
\end{aligned}$$

but $N_t = 0$ because of (A.64), and therefore condition (ii) holds.

A.11.16 Proof of Lemma A.17

For $t \in (0, \phi_u^{-1})$. In the ATE, time–t low–valuation traders hold:

$$q_{t,u}^{ATE} = \min\{(1-\mu_{ht})^{1/\sigma} Q_u^{ATE}, 1\}. \tag{A.65}$$

In the LOE, on the other hand:

$$q_{t,u}^{LOE} = \frac{s-\mu_{ht}}{1-\mu_{ht}},$$

which is strictly less than 1. Thus, if $q_{t,u}^{ATE} = 1$, we have that $q_{t,u}^{LOE} < q_{t,u}^{ATE}$. Now, if $q_{t,u}^{ATE} < 1$, we write

$$\begin{aligned} q_{t,u}^{LOE} &= (1-\mu_{ht})^{1/\sigma} \frac{s-\mu_{ht}}{(1-\mu_{ht})^{1+1/\sigma}} \\ &< (1-\mu_{ht})^{1/\sigma} \frac{s-\mu_{h\phi_u^{-1}}}{(1-\mu_{h\phi_u^{-1}})^{1+1/\sigma}} \\ &= (1-\mu_{ht})^{1/\sigma} Q_u^{LOE} \end{aligned}$$

where the first line follows by multiplying and dividing by $(1-\mu_{ht})^{1/\sigma}$, the second line from Lemma A.12, and the third line by combining equations (A.21) and (A.24). Therefore, if $Q_u^{ATE} \geq Q_u^{LOE}$ implies that $q_{t,u}^{LOE} < (1-\mu_{ht})^{1/\sigma} Q_u^{ATE} = q_{t,u}^{ATE}$ given equation (A.65) and our assumption that $q_{t,u}^{ATE} < 1$.

For $t \in (\phi_u^{-1}, u)$. In the ATE, time–t low–valuation traders holdings are still determined by equation (A.65). In the LOE, their holdings are given by

$$q_{t,u}^{LOE} = (1-\mu_{ht})^{1/\sigma} Q_u^{LOE}.$$

Since $Q_u^{ATE} \geq Q_u^{LOE}$ and $q_{t,u}^{LOE} \leq 1$, it follows that $q_{t,u}^{ATE} \geq q_{t,u}^{LOE}$.

A.11.17 Proof of Lemma A.18

For $\rho > 0$, T_f is defined as the unique $u > 0$ such that:

$$K(u,\rho) \equiv \int_0^u e^{\rho t}(s-\mu_{ht})\, dt = 0.$$

This equation has a unique strictly positive solution when $\rho \geq 0$. When $\rho > 0$ it is equal to T_f. When $\rho = 0$ we denote it by $\hat{T}_f$. Moreover, $K(\,\cdot\,,\,\cdot\,)$ is continuously differentiable in u and ρ with $\partial K/\partial u(\hat{T}_f, 0) \neq 0$. Thus, by the Implicit Function Theorem, the unique strictly positive solution of $K(u,\rho) = 0$ is a continuous function of ρ in a neighborhood of $\rho = 0$. In particular, $T_f \to \hat{T}_f$ as $\rho \to 0$.

We now prove an analogous result for Q_u^{ATE}. The only subtlety is that the support of Q_u^{ATE}, which is $(0, T_f)$, implicitly depends on ρ. We first note that, since $\partial K/\partial u(T_f,\rho) < 0$ and $\partial K/\partial \rho(T_f,\rho) < 0$, then T_f is decreasing in ρ, i.e., T_f increases to $\hat{T}_f$ when ρ decreases to 0. Therefore, for all $u < \hat{T}_f$, Q_u^{ATE} is well defined for ρ close enough to zero. Using the same argument as for T_f, in this neighborhood of $\rho = 0$, Q_u^{ATE} goes to a well-defined limit $\hat{Q}_u^{ATE}$ satisfying:

$$\int_0^u (1-\mu_{ht}) \min\{(1-\mu_{ht})^{1/\sigma} \hat{Q}_u^{ATE}, 1\}\, dt = \int_0^u (s-\mu_{ht})\, dt.$$

The price p_u^{ATE} is obtained by integrating the ODE (18) over $t \in [u, T_f]$:

$$p_u^{ATE} = e^{-r(T_f - u)} \frac{1}{r} + \int_u^{T_f} e^{-r(t-u)} \left[1 - \delta(1 - \mu_{ht})(Q_t^{ATE})^\sigma\right] dt.$$

By continuity, p_u^{ATE} goes to:

$$\hat{p}_u^{ATE} = e^{-r(\hat{T}_f - u)} \frac{1}{r} + \int_u^{\hat{T}_f} e^{-r(t-u)} \left[1 - \delta(1 - \mu_{ht})(\hat{Q}_t^{ATE})^\sigma\right] dt.$$

Using similar arguments, we obtain analogous results in the LOE. Time T_ϕ does not depend on ρ, thus $\hat{T}_\phi = T_\phi$. For all $t \in [0, T_\phi)$, ϕ_t goes to the unique $\hat{\phi}_t \in (T_\phi, \hat{T}_f)$ such that:

$$\int_t^{\hat{\phi}_t} \left[(1 - \mu_{hu})^{1+1/\sigma} (s - \mu_{ht}) - (1 - \mu_{ht})^{1+1/\sigma} (s - \mu_{hu})\right] du = 0.$$

The price path also goes to a well-defined limit $\hat{p}_u^{LOE}$ satisfying the same ODE as when $\rho > 0$ after letting $\rho = 0$.

References

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

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.

Bruno Biais and Pierre-Olivier Weill. Liquidity shocks and order book dynamics. Working paper, TSE and UCLA, 2009.

Bruno Biais, Piere Hillion, and Chester Spatt. An empirical analysis of the limit order book and the order flow in the paris bourse. *Journal of Finance*, 50:1655–1689, 1995.

Bruno Biais, Johan Hombert, and Pierre-Olivier Weill. Supplementary material for "trading and liquidity with limited cognition". Working Paper, TSE, HEC, UCLA, 2010.

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

Jonathan Brogaard. High frequency trading and its impact on market quality. Working paper, Northwestern University, 2010.

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

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

René Descartes. *Meditations and other philosophical writings*. Penguin Classics, 1641.

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

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

Andrew Ellul, Graig W. Holden, Pankaj K. Jain, and Robert H. Jennings. Order dynamics: Recent evidence from the nyse. Working paper, Indiana University Bloomington, 2007.

Thierry Foucault. Order flow composition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, 2:99–134, 1999.

Thierry Foucault, Ohad Kadan, and Eugene Kandel. Limit order book as a market for liquidity. *Review of Financial Studies*, 18:1171–1217, 2005.

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

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

Ronald L. Goettler, Christine A. Parlour, and Uday Rajan. Equilibrium in a dynamic limit order market. *Journal of Finance*, 60:2149–2192, 2005.

Ronald L. Goettler, Christine A. Parlour, and Uday Rajan. Informed traders and limit order markets. *Journal of Financial Economics*, 2009. Forthcoming.

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.

Mark D. Griffiths, Brian F. Smith, D. Alasdair Turnbull, and Robert W. White. The costs and determinants of order aggressiveness. *Journal of Financial Economics*, 56:65–88, 2000.

Larry Harris. *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford University Press, New York, New York, 2003.

Terrence Hendershott and Albert J. Menkveld. Price pressures. Working paper, UC Berkeley and University of Amsterdam, 2010.

Terrence Hendershott and Ryan Riordan. Algorithmic trading and information. Working paper, UC Berkeley, 2010.

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

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

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

Ricardo Lagos, Guillaume Rocheteau, and Pierre-Olivier Weill. Crises and liquidity in otc markets. Working Paper, NYU, UCI, UCLA, 2007.

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.

Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, Oxford, 1995.

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

Emiliano Pagnotta. Information and liquidity trading at optimal frequencies. Working paper, NYU Stern, 2010.

Christine A. Parlour. Price dynamics in limit order markets. *Review of Financial Studies*, 11: 789–816, 1998.

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

Ioanid Rosu. A dynamic model of the limit-order book. *Review of Financial Studies*, pages 4601–4641, 2009.

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

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