The Welfare Impact of High Frequency Trading*

Preliminary Draft

Giovanni Cespa† and Xavier Vives‡

March 2015

Abstract

We study a dynamic market where competitive, risk-averse fast dealers or high frequency traders (HFTs) together with standard (slow) dealers provide liquidity to a population of hedgers with heterogeneous investment horizons. Long term hedgers hold their position until the liquidation date, whereas short term ones unwind at interim. This setup generates endogenous persistence of liquidity shocks, liquidity complementarities, and potential multiple equilibria with different levels of liquidity. This can explain liquidity fragility and “flash crashes.” We use our model to assess the impact of HFT on market quality and traders’ welfare. Liquidity complementarities can make liquidity decreasing in the mass of HFTs, notwithstanding that an increase in HFT enhances the risk-bearing capacity of the market. Furthermore, equilibria with high liquidity are not necessarily associated with higher welfare. Finally, we endogenize the number of HFTs, and show that the solution chosen by a monopolistic exchange is generally at odds with the one advocated by a social planner.

Keywords: High frequency trading, flash crash, welfare, asymmetric information, hedgers, endogenous market structure.

JEL Classification Numbers: G10, G12, G14

---

*For helpful comments we thank Sophie Moinas, and seminar participants at the 9th Annual Central Bank Workshop on Microstructure (Frankfurt, 9/13), the conference on High Frequency Trading at Imperial College, Brevan Howard Centre (London, 12/14), and the Workshop on Microstructure Theory and Applications (Cambridge, 3/15). Cespa acknowledges financial support from the Bank of England (grant no. RDC151391).

†Cass Business School, City University London, CSEF, and CEPR. 106, Bunhill Row, London EC1Y 8TZ, UK. e-mail: giovanni.cespa@gmail.com

‡IESE Business School, Avinguda Pearson, 21 08034 Barcelona, Spain.
1 Introduction

The report on the causes of the “Flash-Crash” issued by the staffs of the CFTC-SEC (2010) highlights the role of High Frequency Traders (HFTs; HFT denotes High Frequency Trading)—a class of market players who engage in extremely short-term strategies—in exacerbating the sharp price drop that characterized the crash.\(^1\) Public perception about HFT is mixed (see, e.g. Harris (2013)): according to some, HFTs have replaced traditional market makers in the business of liquidity supply, taking advantage of the huge technological improvements that have swept modern markets. As such, they have come to perform in a more efficient way a vital service for the orderly workings of the market, generically leading to improved terms of trade. However, at the same time concern has been voiced both among academics and practitioners that the action of HFTs may undermine market stability and be detrimental to market participants. What is the impact of HFTs on the functioning of the market? How do they affect the welfare of market participants? Is their claimed liquidity enhancing activity a synonym for an overall improved market quality?

In this paper we address these issues, shedding light on the ways in which HFT can contribute to market instability and assessing its impact on liquidity and on the welfare of market participants. In contrast to most of the literature which considers risk neutral agents, we assume risk averse traders. This allows us to account for insurance and risk issues. We see HFTs as modern floor traders, who can act when others cannot, exploiting information on non-fundamental demand shocks to anticipate future prices.\(^2\) Trader heterogeneity, risk aversion, and a dynamic model where all traders have a well-defined utility function are essential ingredients of our analysis.

We find that the combination of hedgers with different horizons and the presence of HFTs leads to endogenous persistence of endowment shocks, liquidity complementarities, and potential multiple equilibria.\(^3\) Liquidity complementarities imply that liquidity can be hump-shaped in the number of HFTs. Thus, an increase in HFT can make the market thinner. In our setup, equilibria with high liquidity are not necessarily associated with high welfare. Furthermore, more HFTs can lower hedgers’ welfare. Thus, when the number of HFTs is determined by a monopoly exchange (that extracts the surplus of HFTs over normal dealers), which does not internalize the welfare of hedgers, too much or too little entry of HFTs can occur.

\(^1\)O’Hara (2014) offers a comprehensive overview of the way in which, due to technological advances, the organization of trading has changed, facilitating the advent of HFT as major market players. Biais and Woolley (2011) surveys the literature on HFT.

\(^2\)Evidence on the relevance of such shocks for asset prices is pervasive. Empirically, Peress and Schmidt (2015) find that the trades executed by individual investors are positively autocorrelated at weekly horizons, and the standard deviation of noise trades accounts for at least 44% of weekly volume. At a lower frequency, Coval and Stafford (2007) show that mutual funds faced with aggregate redemptions curtail their positions, generating a temporary price pressure. Theoretically, persistent liquidity trading demand affects price efficiency and, with asymmetric information, can generate multiple equilibria, in particular with short-term traders (Cespa and Vives (2012), and Cespa and Vives (2015)). With persistent, feedback noise trading, and short-term investors, an increase in transparency can disconnect prices from fundamentals (Banerjee, Davis, and Gondhi (2014)).

\(^3\)Other authors obtain multiple equilibria when agents are informed about non-fundamentals demand shocks and have a signal on fundamentals (Ganguli and Yang (2009), and Manzano and Vives (2011)). In this paper, multiplicity arises through a loop that only relies on non-fundamentals demand shock information.
More in detail, we study a 2-period market in which competitive, risk-averse high frequency traders (which we term “Fast Dealers,” FD) and standard dealers (D) trade a risky asset and provide liquidity to a population of hedgers who receive a random shock to their endowment. FD are in the market at all trading dates, whereas D only trade once. Hedgers have heterogeneous horizons, in that they load a position at the first trading round, and unwind it at different points in time. “Long term” hedgers (HL) hold their position until the liquidation date, whereas “short term” ones (HS) unwind at interim, exit the market, and are replaced by a new generation of hedgers who hedge a new random shock, holding their position until the liquidation date. This setup generates endogenous persistence of liquidity shocks, liquidity complementarities, and potential multiple equilibria with different levels of liquidity. The intuition for these results is as follows.

Suppose that first period hedgers receive a positive endowment of the risky asset, and thus want to reduce their exposure. Assume that the sell orders of HL overpower the demand of D for a given liquidity of the market. At equilibrium, this implies that FDs must step in to supply the missing liquidity. In the second trading round, when HSs unwind their position, the inventory held by FDs will consequently more than offset HSs’ demand, inducing a negative order imbalance that continues the downward price pressure originated by the first period endowment shock. Thus, this endowment shock propagation mechanism, endogenously delivers positive autocorrelation in the demand of uninformed investors, a property of noise trades that is often assumed in the literature (see e.g., He and Wang (1995), Cespa and Vives (2012), and Cespa and Vives (2015)), and that finds empirical support at short horizons (Peress and Schmidt (2015)). To see how liquidity complementarities can arise, go back to our example, and suppose that the first period market becomes less liquid. This leads HLs to curtail their sales (as liquidity consumption becomes more expensive), which reduces the imbalance carried over to the second period, attenuating the impact of the first period endowment shock on the second period price. As a consequence, second period hedgers face a lower return uncertainty, which leads them to hedge more aggressively, consuming more liquidity, and heightening the volatility of the second period price. The increased uncertainty on the liquidation price leads FDs to scale down their first period liquidity supply. Importantly, it also increases the endowment risk faced by HS and can lead them to hedge more aggressively their endowment shock, consuming more liquidity. We show that this self-sustaining loop can become strong enough to generate multiple equilibria, with different levels of liquidity, for the same deep parameters of the market.

We use the model to assess the impact of HFT on market quality. An important implication of liquidity complementarities is that an increase in the number of FDs can make the market less liquid. This is because an increase in FDs, augmenting the risk-bearing capacity of the market, has two contrasting effects on market liquidity. On the one hand, it lowers the risk borne by liquidity suppliers, thereby improving liquidity. On the other hand, it can spur second period

Liquidity complementarities also arise in other contexts, see e.g. Cespa and Foucault (2014), and Cespa and Vives (2015).

This is an immediate implication of market clearing. At equilibrium the aggregate supply of the asset from hedgers must exactly offset the aggregate demand from FDs and Ds. If HLs’ supply is larger than Ds’ demand, at equilibrium HFTs will absorb part of HLs’ supply.
hedgers to hedge more aggressively their endowment, triggering the loop described above, and leading to more liquidity consumption in the first period. As a result, liquidity can be hump-shaped in the number of FDs. This effect implies that a change in the number of FDs can tilt the market towards the region with multiple equilibria in which liquidity changes abruptly. Thus, because of liquidity complementarities a small reduction in HFT can lead to a discrete change in liquidity (a “flash-crash”).

The debate on the effects of HFT has predominantly focused on liquidity. However, little is known about welfare effects. As agents in the model have a well-defined utility function, we can compute the welfare of market participants and address this issue. We find that hedgers benefit from trading—as in this way they can take advantage of the insurance features of the market. However, except for HL, this does not imply that a more liquid market unequivocally makes them all better off. Indeed, HLs load a position at the first date to unwind it at the liquidation date. Thus, when the market is deeper they pay a lower cost to hedge. However, a deeper market affects the shock propagation mechanism described above and this can impair the welfare of HS and second period hedgers. To see this, consider that as HS unwind in the second period, they benefit from a liquidation price that is positively related to their endowment. This is because in general they retain part of their endowment and profit from a potential capital gain. Thus, when a more liquid market lowers the positive relationship between the second period price and the first period endowment shock, their welfare suffers. Similarly, second period hedgers’ take advantage of the price impact of the first period endowment shock, for example selling the asset when a positive price pressure arises. However, this implies that their strategies serve as poor hedge against the endowment shock they receive. Thus, when improved liquidity magnifies the impact of the first period endowment shock on the second period price, their welfare suffers. Overall, this implies that equilibria with higher liquidity are not necessarily associated with higher welfare.

The presence of the above, liquidity-related, externalities implies that in general a social planner chooses a number of HFTs that differs from the one that obtains with either a free-entry regime, or when the number of HFTs is determined by a monopolistic exchange. Thus, even though HFTs are able to supply liquidity when other agents can’t, the market solution is typically at odds with the one that maximizes social welfare.

The rest of the paper is organized as follows. After discussing the related literature, we present the model and solve for the equilibrium, studying its properties and relating multiplicity to the endowment shock propagation mechanism and to fragility. We then turn to the welfare analysis of the market participants, assess the welfare impact of HFT, and study entry decisions of HFTs in different market scenarios. The paper final section summarizes our results. Most formal proofs are relegated to the Appendix.

A notable exception in this respect is Biais, Foucault, and Moinas (2015). However, the focus of the analysis there is on the welfare maximizing level of HFT investment, rather than on the impact of HFT for the relationship between liquidity and welfare.
2 Related literature

There is by now a growing empirical literature which studies the automation of trading activity and HFT, focusing on its impact on standard measures of market performance. Hasbrouck and Saar (2013) study “low-latency” activity of HFTs, that is the set of strategies that respond to market events occurring within millisecond intervals (without necessarily leading to actual trades). Their findings point to a positive effect of such activity on standard measures of market performance, such as short-term volatility, quoted spreads, price impact, and the depth of the book. Brogaard, Hendershott, and Riordan (2012) find that HFT enhances the informational efficiency of the market by trading in the direction of permanent price changes (that is buying when the market buys in view of upcoming, confirming fundamentals information), and against the market when demand shocks are instead transitory (that is, basically supplying liquidity when orders are unlikely to originate from informed investors). The horizon over which HFT predicts price changes is however very short (less than 3-4 seconds), which greatly limits its contribution to informational efficiency. Using data from the Deutsche Borse allowing the identification of algorithmic generated trades, Hendershott and Riordan (2012) study liquidity provision by Algorithmic Traders (ATs, a class of algorithms which subsumes HFTs). Their finding points to the fact that ATs tend to consume liquidity when this is cheap, and supply it when it is expensive, stabilizing liquidity provision. Automation and increase in speed execution are no panacea, however. For instance, Hendershott and Moulton (2011) find that the increase in speed execution prompted by the introduction of the NYSE’s Hybrid Market led to an increase in price efficiency, but also in trading costs for market orders due to a worsening of adverse selection.

The welfare impact of HFT is however less well understood. Cartea and Penalva (2012) in a Grossman and Miller (1988) setup, model HFTs as parasitic agents who add a further inter-mediation layer between traders demanding liquidity and dealers who for technological reasons are not able to match that demand. Menkveld and Yueshen (2012) model HFT as middlemen that have no intrinsic value from holding an asset but are better informed about its liquidation value than the seller from which they buy the asset. Once they have acquired a position, HFTs can either find an investor interested in the asset, or engage in “swap trades” to unwind it. This generates price pressure, which impairs the inference of the seller when he returns to the market to dispose of a further unit of his asset endowment. In some cases this implies that HFTs have a negative effect on market liquidity due to the adverse selection problem they generate, which in turn impairs social welfare. Jovanovic and Menkveld (2010) argue that HFT can ameliorate the intertemporal adverse selection problem faced by a dealer, fostering liquidity provision. In their setup, a dealer anticipates that submitting a limit order exposes him to the risk of being adversely selected by traders that enter the market later during the trading day, and thus with potentially better information. In this situation, the anticipated adverse

\[\text{In their analysis, HFT heightens microstructure noise. Other authors emphasize HFTs' superior information processing ability looking at the consequence this has for the level of adverse selection that pervades the market and the welfare of market participants.}\]
selection risk may impair liquidity provision. However, HFTs’ ability to continuously monitor the book and thus adjust their quotes to reflect new information, reduces the magnitude of the problem, ameliorating liquidity provision, and improving welfare. Biais, Foucault, and Moinas (2015) study the welfare implications of investment in the acquisition of HFT technology. In their model HFT have a superior ability to match orders, and possess superior information compared to human (slow) traders. Therefore, an increase in HFT enhances liquidity provision while increasing adverse selection costs for slow traders. When investment decisions to acquire HFT are considered, the authors show that such decisions are strategic complements. Thus in equilibrium this can lead to excessive investment in HFT technology which, in view of the negative externality generated by HFT, can be welfare reducing. Hoffmann (2014) analyzes a market where agents can acquire a technology that speeds up their ability to revise posted limit orders in the light of new information. In this setup the welfare impact of technology acquisition depends on whether its introduction enhances or not the efficiency of the trading process. Budish, Cramton, and Shim (2014) argue that the investment arms’ race characterizing HFT is a byproduct of the predominant trading model centered on the continuous limit order book. With this form of market organization, even tiny speed advantages enabling to exploit publicly available information can give an important edge, that justifies massive (and potentially socially wasteful) investment in technology. It is this type of “toxic arbitrage” that Foucault, Kozhan, and Tam (2014) document using FX-market data, and show to be related to an increase in bid-ask spreads. Finally, Du and Zhu (2014) addresses the question of the optimal trading speed in a market in which traders can have heterogeneous speeds. The optimal solution when all traders act at the same frequency is to slow trades down when information releases are prescheduled, and speed them up when instead the arrival time of information is random. When trading speeds differ fast (slow) traders prefer high (low) trading frequencies.

3 The model

Consider a two-period stock market where a single risky asset with liquidation value \(v\), and a risk-less asset with unitary return are traded by a continuum of competitive, risk-averse, High Frequency Traders (which we refer to as “Fast Dealers” and denote by FD) in the interval \([0,\mu]\), competitive, risk-averse dealers (D) in the interval \([\mu,1]\), and hedgers (H). We now illustrate the preferences, and orders of the different players.

3.1 Liquidity providers

A FD \(i\) has CARA preferences (we denote by \(\gamma\) his risk-tolerance coefficient) and submits price-contingent orders \(x_{it}^{FD}\), \(t = 1,2\), to maximize the expected utility of his final wealth:

\[
W_{it}^{FD} = (v - p_2)x_{it}^{FD} + (p_2 - p_1)x_{i1}^{FD}
\]

A Dealer \(i\) also has CARA preferences with risk-tolerance

\(^8\)We assume, without loss of generality with CARA preferences, that the non-random endowment of FDs and dealers is zero.
\[\gamma\], but is in the market only in the first period. He thus submits a price-contingent order \(x_i^D\) to maximize the expected utility of his wealth \(W_i^D = (v - p_1)x_i^D\).

### 3.2 Hedgers

Hedgers differ depending on the date at which they enter the market. In the first period, each hedger \(i\) receives a random endowment of the risky asset \(\theta_{i1} = \theta_1 + \xi_{i1}\), correlated with the aggregate endowment \(\theta_1\), with \(\theta_1 \sim N(0, \tau_u^{-1})\), \(\xi_{i1} \sim N(0, \tau_\xi^{-1})\), and \(\text{Cov}[\theta_1, \theta_1] = \text{Cov}[\theta_1, \xi_{i1}] = \text{Cov}[\theta_2, \xi_{i1}] = \text{Cov}[\xi_{i1}, \xi_{j1}] = 0\) for all \(i, j\). There are two types of first period hedgers who differ in terms of their trading horizon. A fraction \(\beta\) (which we denote by HL) in the first period posts a market order \(x_{i1}^{HL}\) and holds the position until the liquidation date (period 3). The complementary fraction \((1 - \beta)\) (which we denote by HS) posts a market order \(x_{i1}^{HS}\) anticipating that it will unwind its holdings in the following period. Independently from their trading horizon, all hedgers have identical CARA preferences (we denote by \(\gamma_H^1\) the common risk-tolerance coefficient). Formally, a hedger \(i \in [0, \beta]\) maximizes the expected utility of his “long term” profit \(\pi_{i1}^{HL} = \theta_{i1}v + (v - p_1)x_{i1}^{HL}\):

\[
E[-\exp\{-\pi_{i1}^{HL}/\gamma\} | \Omega_{i1}^{HL}] .
\]

A hedger \(i \in [\beta, 1]\) instead maximizes the expected utility of his “short term” profit \(\pi_{i1}^{HS} = \theta_{i1}p_2 + (p_2 - p_1)x_{i1}^{HS}\):

\[
E[-\exp\{-\pi_{i1}^{HS}/\gamma\} | \Omega_{i1}^{HS}] .
\]

In the second period, a new generation of \((1 - \beta)\) hedgers enters the market, receiving a random endowment of the risky asset \(u_2 = u_2 + \xi_{i2}\), correlated with the aggregate endowment \(u_2 \sim N(0, \tau_u^{-1})\), \(\xi_{i2} \sim N(0, \tau_\xi^{-1})\), and \(\text{Cov}[u_2, \theta_1] = \text{Cov}[u_2, \xi_{i2}] = \text{Cov}[\xi_{i2}, \xi_{j2}] = 0\), for all \(i, j\). Also, we assume that \(u_2 \perp \theta_1\) and that the error terms in the endowment shocks are uncorrelated across trading dates. A second period hedger \(i\) has CARA utility function with risk-tolerance \(\gamma_2^H\), and submits a market order to maximize the expected utility of his profit \(\pi_{i2}^{H} = u_2v + (v - p_2)x_{i2}^{H}\):

\[
E[-\exp\{-\pi_{i2}^{H}/\gamma\} | \Omega_{i2}^{H}] .
\]

Finally, we make the convention that the SLLN holds:

\[
\frac{1}{\beta} \int_0^\beta \xi_{i1} di = 0, \quad \frac{1}{1 - \beta} \int_0^1 \xi_{i1} di = 0, \quad \frac{1}{1 - \beta} \int_0^1 \xi_{i2} di = 0. \]

We now turn to describe the information sets of the different market participants. At equilibrium, we conjecture that in period 1 a hedger \(i\) of type \(k \in \{L, S\}\) submits a market order

\[x_{i1}^{Hk} = b_{i1}^{Hk} \theta_{i1},\]  

(1)

with \(b_{i}^{Hk}\) (the “hedging aggressiveness”) to be determined in equilibrium, while a FD and a dealer respectively post a limit order \(x_{i1}^{FD} = \varphi_1^{FD}(p_1), x_{i1}^{D} = \varphi_1^{D}(p_1)\) where \(\varphi_1^{FD}(\cdot), \varphi_1^{D}(\cdot)\) are
linear functions of $p_1$. In the second period, a FD $i$ can observe a signal on the liquidation value $s_v = v + \epsilon$, with $\epsilon \sim N(0, \tau_\epsilon^{-1})$, and $\epsilon$ independent of all the other random variables in the model. We assume that a FD submits a limit order $x_{i2}^{FD} = a s_v - \varphi_2(p_1, p_2)$, where $a$ denotes the response to the private signal. A hedger in the second period observes a signal of the first period endowment shock $s_{\theta_1} = \theta_1 + \eta$, with $\eta \sim N(0, \tau_\eta^{-1})$, and independent from all the other random variables in the model, and submits a market order $x_{i2}^H = b_{21}^H u_2 + b_{22}^H s_{\theta_1}$.

- In the first period FDs have no information on the risky asset. As they submit limit orders, at a linear equilibrium they can retrieve the aggregate endowment shock $\theta_1$ from the equilibrium price. The argument is as follows. Denote by $\mu x_1^{FD} = \int_0^\mu x_{i1}^{FD} di = \mu \varphi_1^{FD}(p_1)$, $(1 - \mu) x_1^D = \int_1^\mu x_{i1}^D di = (1 - \mu) \varphi^D(p_1)$, and by $\beta x_1^{HL} + (1 - \beta) x_1^{HS} = \int_0^\beta x_{i1}^{HL} di + \int_1^\beta x_{i1}^{HS} di = (\beta b_1^{HL} + (1 - \beta) b_1^{HS}) \theta_1$ respectively the aggregate position of FDs, dealers and hedgers in the first period. Imposing market clearing yields:

$$
\mu x_1^{FD} + (1 - \mu) x_1^D + \beta x_1^{HL} + (1 - \beta) x_1^{HS} = 0 \iff \\
\mu \varphi_1^{FD}(p_1) + (1 - \mu) \varphi^D(p_1) + (\beta b_1^{HL} + (1 - \beta) b_1^{HS}) \theta_1 = 0.
$$

At equilibrium the coefficients of traders’ strategies are known, which implies that $p_1$ is observationally equivalent to $\theta_1$ and that both FDs and dealers can retrieve $\theta_1$ from the price. Therefore, the information set of a FD and a dealer in the first period coincide and are given by $\Omega_i^{FD} = \Omega_i^D = \{\theta_1\}$. In the second period, denote by $\mu x_2^{FD} = \int_0^\mu x_{i2}^{FD} di = \mu (a s_v - \varphi_2(p_1, p_2))$ and by $(1 - \beta) x_2^H = \int_1^\beta x_{i2}^H di = (1 - \beta) (b_2^H u_2 + b_2^H s_{\theta_1})$, respectively the aggregate position of FDs and second period hedgers. Impose market clearing:

$$
\mu (x_2^{FD} - x_1^{FD}) + (1 - \beta) (b_2^H u_2 + b_2^H s_{\theta_1} - x_1^{HS}) = 0,
$$

and rearrange the first period market clearing condition as follows

$$(1 - \mu) x_1^D + \beta x_1^{HL} = -\left(\mu x_1^{FD} + (1 - \beta) x_1^{HS}\right).
$$

Substitute the latter in the second period clearing equation to obtain

$$
\mu x_2^{FD} + (1 - \beta) (b_2^H u_2 + b_2^H s_{\theta_1}) + (1 - \mu) x_1^D + \beta x_1^{HL} = 0.
$$

Once again, at a linear equilibrium the coefficient of traders’ strategies are known, which implies that the price sequence $\{p_1, p_2\}$ is observationally equivalent to $\{\theta_1, u_2\}$, so that FDs can retrieve the aggregate endowment shock from the price. Thus, the second period information set of a FD $i$ is given by $\Omega_i^{FD} = \{s_v, \theta_1, u_2\}$.

- A hedger in the first period only observes his endowment shock, independently from his trading horizon. Therefore, a first period hedger’s information set is given by $\Omega_i^{HL} = \{\theta_{1i}\}$, $k \in \{L, S\}$. A second period hedger receives an endowment shock $u_2$, and can observe a signal $s_{\theta_1}$. Summarizing, a second period hedger $i$’s information set is given by
\[ \Omega^H_{i+2} = \{ u_{i+2}, s_{\theta_1} \} . \]

Figure 1 displays the timeline of the model.
3.3 The propagation mechanism of endowment shocks

In our model, due to the different trading horizon of market participants, the first period endowment shock propagates to the second trading round. To see this, consider the market clearing equations (2) and (3). First period agents (hedgers, standard dealers, and FDs) load a position in the first period, but unless the position of HL (HS) exactly offsets the one of standard dealers (FDs):

\[(1 - \mu)x^D_1 + \beta x^{HL}_1 = 0 \iff -(\mu x^{FD}_1 + (1 - \beta)x^{HS}_1) = 0, \quad (4)\]

the impact of their orders also affects the second period price, i.e., \(p_2\) reflects \(u_2\) and \(\theta_1\). To fix ideas suppose that \(\beta x^{HL}_1 > -(1 - \mu)x^D_1\). In this case HL want to hedge less of their endowment than what D are ready to absorb. As a consequence, part of the liquidity supplied by the latter matches the orders of HS. At date 2, when HS unwind their orders, the inventory FDs carry over from the first period is insufficient to clear HS’s demand. This induces an imbalance (on top of the one due to second period hedgers) that affects the second period price, creating additional price volatility. In this context, having information on \(\theta_1\) is valuable to trade in period 2, and we assume that second period hedgers observe a noisy signal of \(\theta_1\). As the latter submit a market order, this implies that they face some residual uncertainty over their returns, which impacts their strategies.

4 The equilibrium

In this section we assume that (i) hedgers’ endowment shocks are perfectly correlated with the aggregate endowment shock in every period \(t = 1, 2\) (i.e., that \(\tau_{\xi} \to \infty\)), and that (ii) FDs’ signal in the second period is infinitely noisy: \(\tau_\epsilon = 0\). Formally, we have

\[\Omega^H_k = \{\theta_1\}, \quad k \in \{L, S\}, \quad (5)\]
\[\Omega^H = \{u_2, s_{\theta_1}\}. \quad (6)\]

These assumptions ensure that in the first period FDs, dealers, and hedgers have access to the same information \(\theta_1\); in the second period, FDs retain an informational advantage over second period hedgers in that they perfectly observe \(\theta_1\). We now discuss the strategies of the different market participants and then provide a characterization of the equilibrium.

We start from FDs. In the second period, these traders act like in a static market:

\[X^{FD}_2(p_1, p_2) = -\gamma \tau_\epsilon p_2.\]

Therefore, they speculate on the asset payoff (recall that \(E[v] = 0\)), and supply liquidity to second period hedgers, demanding a compensation that is inversely related to the risk they face.\(^9\)

\(^9\)At equilibrium it will be the case that \(b^{Hk}_1 < 0\), for \(k \in \{L, S\}\), so that if \(\theta_1 > 0\) HL and HS sell part of their endowment.
bear. In the first period, as we show in the appendix, we have

$$X_1^{FD}(p_1) = \frac{\gamma E[p_2|\theta_1] - p_1}{\text{Var}[p_2|\theta_1]} - \frac{\gamma}{\text{Var}[p_2|\theta_1]} p_1.$$  \hspace{1cm} (7)

The above expression implies that FDs once again perform two activities: they consume liquidity, speculating on the anticipation of the second period price, and they provide liquidity, demanding a compensation that is inversely related to the overall risk they bear. Note that an increase in the conditional volatility of $p_2$ lowers both FDs’ market making and speculative activities, since it increases the risk of the price at which they can unwind their positions. A traditional dealer in the first period trades according to

$$X_1^D(p_1) = -\gamma \tau v p_1.$$  \hspace{1cm} (7)

Importantly, the price change needed by FDs to accommodate the aggregate demand is smaller than the one demanded by traditional dealers:

$$\frac{1}{\gamma} \left( \frac{1}{\text{Var}[p_2|\theta_1]} + \frac{1}{\text{Var}[v]} \right)^{-1} \frac{\text{Var}[v]}{\gamma} < \frac{\text{Var}[v]}{\gamma}.$$  \hspace{1cm} (8)

This is because FDs can rebalance their position at interim, and thus manage more efficiently their asset inventory.

Consider now hedgers. In the appendix we show that a second period hedger trades according to

$$X_2^H(u_2, s_\theta) = \frac{\gamma H E[v - p_2|\Omega_2^H]}{\text{Var}[v - p_2|\Omega_2^H]} - \frac{\text{Cov}[v - p_2, v|\Omega_2^H]}{\text{Var}[v - p_2|\Omega_2^H]} u_2.$$  \hspace{1cm} (9)

Thus, a hedger’s strategy has a speculative and a hedging component. A hedger speculates on value change the more, the less liquid is the market (see the first term on the r.h.s. in (9)), while lowering his exposure to the asset risk the more, the higher is the covariance between the return on his position (i.e. $v - p_2$) and the final liquidation value ($v$), given his information. In this way he reduces the risk that his speculative strategy goes sour precisely when the value of his endowment collapses. Note that second period hedgers’ inability to perfectly infer $\theta_1$, affects the uncertainty over the price at which their order is executed. This is because, as observed in Section 3 due to the presence of HL and standard dealers, $p_2$ loads on $\theta_1$. Other things equal, a more volatile $\theta_1$ augments the effect of the first period endowment shock on $p_2$ and, as second period hedgers cannot perfectly infer $\theta_1$, this increases the execution risk they face.

In the first period, hedgers’ strategies are similar to (9):

$$X_1^{HL}(\theta_1) = \frac{\gamma H E[v - p_1|\theta_1]}{\text{Var}[v - p_1|\theta_1]} - \frac{\text{Cov}[v - p_1, v|\theta_1]}{\text{Var}[v - p_1|\theta_1]} \theta_1.$$  \hspace{1cm} (10a)

$$X_1^{HS}(\theta_1) = \frac{\gamma H E[p_2 - p_1|\theta_1]}{\text{Var}[p_2|\theta_1]} - \frac{\text{Cov}[p_2 - p_1, p_2|\theta_1]}{\text{Var}[p_2|\theta_1]} \theta_1.$$  \hspace{1cm} (10b)

Note that HS can partially anticipate the second period price, and thus speculate on it, e.g. by holding more of their endowment when $\theta_1 > 0$ (see the numerator of the first term on the right
hand side of (10b)). At the same time, due to the impact of second period hedgers’ demand on $p_2$, HS face uncertainty on the liquidation price, which is reflected in the conditional variance at the denominator of their strategies (10b).

In the appendix we prove the following result:

**Proposition 1.** There exists an equilibrium in linear strategies where

$$p_2 = \lambda_2 (1-\beta)(b_{21}^H u_2 + b_{22}^H s_{\theta_1}) + \lambda_2 \left( \beta b_1^{HL} + (1-\mu)\gamma \tau_v \Lambda_1 \right) \theta_1 \quad \text{(11a)}$$

$$p_1 = -\Lambda_1 \theta_1, \quad \text{(11b)}$$

$$\lambda_2 = 1/(\mu \gamma \tau_v) > 0,$$

$$b_1^{HL} = -(1-\gamma^H \tau_v \Lambda_1) \quad \text{(12a)}$$

$$b_1^{HS} = -\left( 1 - \gamma^H_1 \lambda_2 (\beta b_1^{HL} + (1-\beta)b_{22}^H + \gamma \tau_v \Lambda_1) \right) \quad \text{(12b)}$$

$$b_{21}^H = -\frac{1}{\gamma^H_2 \tau_v \lambda_2 (1-\beta) + \text{Var}[v-p_2|\Omega_2^H]} \quad \text{(12c)}$$

$$b_{22}^H = -\frac{\gamma^H_2 \tau_u \lambda_2 \left( \beta b_1^{HL} + (1-\mu)\gamma \tau_v \Lambda_1 \right)}{(\tau_u + \tau_v)(\gamma^H_2 \lambda_2 (1-\beta) + \text{Var}[v-p_2|\Omega_2^H])} \quad \text{(12d)}$$

$$\text{Var}[p_2|\theta_1] = \lambda_2^2 (1-\beta)^2 \left( \frac{(b_{21}^H)^2}{\tau_u} + \frac{(b_{22}^H)^2}{\tau_u} \right), \text{Var}[v-p_2|\Omega_2^H] = \text{Var}[v] + \lambda_2^2 \left( \beta b_1^{HL} + (1-\mu)\gamma \tau_v \Lambda_1 \right)^2 \times \text{Var}[\theta_1|s_{\theta_1}], \text{and } \Lambda_1 \text{ obtains as a fixed point of the following mapping:}$$

$$\psi(\Lambda_1) = \left( \mu \gamma \left( \frac{1}{\text{Var}[p_2|\theta_1]} + \frac{1}{\text{Var}[v]} \right) + (1-\mu)\gamma \frac{1}{\text{Var}[v]} \right)^{-1} \left( \mu \frac{\gamma \text{Cov}[p_2, \theta_1]}{\text{Var}[p_2|\theta_1] \text{Var}[\theta_1]} + b_1 \right), \quad \text{(13)}$$

where $b_1 \equiv \beta b_1^{HL} + (1-\beta)b_1^{HS}$.

According to (11b), $\Lambda_1$ captures the price impact of $\theta_1$, a measure of the (reciprocal of) first period liquidity:

$$\Lambda_1 \equiv -\frac{\partial p_1}{\partial \theta_1}.$$

We can interpret (13) as an aggregate best reply function yielding the liquidity level arising from first period traders’ optimal responses to a shock to $\Lambda_1$. Agents in the market trade to (i) hedge a shock (HL and HS), (ii) to speculate on the second period price (HS and FDs), or (iii) to accommodate the imbalance between demand and supply of the asset (FDs and D). Motives (i) and (ii) correspond to liquidity demand, whereas motive (iii) corresponds to liquidity supply.

The expressions for the strategy coefficients (12a)–(12d) show that a shock to $\Lambda_1$ affects all traders’ strategies. Thus, the above distinction leads to interpret $\psi(\Lambda_1)$ as the ratio of liquidity demand and liquidity supply cost responses to a shock to $\Lambda_1$. Numerical simulations show that $\Lambda_1 > 0$, and $b_1^{Hk} < 0$, $k \in \{L, S\}$. 

12
A similar reasoning implies that in the second period there are two measures of liquidity:

\[
\lambda_{21} \equiv \frac{\partial p_2}{\partial u_2} = \lambda_2(1 - \beta)b^H_{21} \quad (14a)
\]
\[
\lambda_{22} \equiv \frac{\partial p_2}{\partial \theta_1} = \lambda_2(1 - \beta)b^H_{22} \quad (14b)
\]

This is because the demand of liquidity from second period hedgers depends on two signals: the endowment shock \(u_2\) and the information about \(\theta_1\). The cost of supplying liquidity is instead exogenous, since the asset is liquidated in period 3 and FDs only face the uncertainty on \(v\).

### 4.1 Endowment shock propagation and the time series properties of supply

As anticipated in Section 3.3, the second period price loads on \(u_2\) and, provided \(\beta b^{HL}_1 + (1 - \mu)\gamma \tau v A_1 \neq 0\), also on \(\theta_1\). More in detail, suppose that \(\theta_1 > 0\), and that on average HL sell less than what D are ready to absorb:

\[
\beta b^{HL}_1 > -(1 - \mu)\gamma \tau v A_1 \iff \frac{\beta x^{HL}_1}{\theta_1} > -\frac{(1 - \mu)x^D_1}{\theta_1}. \quad (15)
\]

Equivalently, due to market clearing, HS want to hedge a larger quantity of their endowment compared to what FDs are ready to absorb:

\[
\frac{(1 - \beta)x^{HS}_1}{\theta_1} < -\frac{\mu x^D_1}{\theta_1}.
\]

As a consequence, part of D’s liquidity supply is matched by the orders of HS. When, at date 2 HS unwind their position, D are not in the market and the inventory FDs carried over from the first period is smaller than HS’ order. This imbalance propagates the impact of \(\theta_1\) to the second period, exerting a positive pressure on \(p_2\). Note that in this case if \(\tau_x > 0\), according to (12d) second period hedgers put a negative weight on \(s_{\theta_1}\): \(b^H_{22} < 0\). Indeed, as \(\theta_1 > 0\) has a positive impact on \(p_2\), second period hedgers have the opportunity to sell the asset at a higher price. Thus, when second period hedgers are informed about \(\theta_1\), their strategy contains a speculation component. Finally, from (12c), \(b^H_{21} < 0\).

Based on the above discussion,

\[
z_2 \equiv (1 - \beta)x^H_2 + \beta x^{HL}_1 + (1 - \mu)x^D_1,
\]

captures the second period aggregate supply of the asset. Defining \(z_1 \equiv \beta x^{HL}_1 + (1 - \beta)x^{HS}_1\), a natural measure of supply (first period endowment shock) persistence (propagation) is given
by

\[
\text{Cov}[z_1, z_2] = \left(\beta b^{HL}_1 + (1 - \beta) b^{HS}_1\right) \left(\frac{(1 - \beta) b^{H}_2 + \beta b^{HL}_1 + (1 - \mu) \gamma \tau v \Lambda_1}{\tau_u (\tau_u + \tau_1)(\gamma_2 \lambda_2 (1 - \beta) + \text{Var}[v - p_2 | \Omega^H_2])}\right)
\]

Based on the numerical simulations showing \( b^{HL}_1 < 0 \), this implies the following result:

**Corollary 2.** The following conditions are equivalent:

1. **HLs’ demand (supply) falls short of Ds’ supply (demand):** \( \beta b^{HL}_1 > -(1 - \mu) \gamma \tau v \Lambda_1 \).

2. **The second period price is positively correlated with the first period endowment shock:** \( \text{Cov}[p_2, \theta_1] > 0 \).

3. **The asset supply is negatively autocorrelated:** \( \text{Cov}[z_1, z_2] < 0 \).

Condition (15) depends on the magnitude of \( \Lambda_1 \), which is determined at equilibrium. Thus, in this model the time-series properties of the asset supply are endogenous. For instance, at an equilibrium with \( \beta x^{HL}_1 < -(1 - \mu) x^{D}_1 \), we have \( \text{Cov}[z_1, z_2] > 0 \) which corresponds to the assumption of (exogenously) positively autocorrelated noise trading that is often made in the literature.

Furthermore, other things equal, an increase in \( \Lambda_1 \) has a non-monotone impact on \( \text{Cov}[p_2, \theta_1] \), and on the autocorrelation of the asset supply. For equilibria at which the opposite of (15) holds, a less liquid first period market attenuates the negative (positive) effect of \( \theta_1 \) (the first period supply) on \( p_2 \) (the second period supply). This is because in this case \( \Lambda_1 \) is small, and HL’s supply is partly absorbed by FDs who thus have a large inventory when entering the market in the second period. This depresses \( p_2 \). As \( \Lambda_1 \) increases, HL hedge less and more of their orders are absorbed by D, which attenuates the impact of \( \theta_1 \) on \( p_2 \). Conversely, when (15) holds, a less liquid first period market strengthens the positive (negative) effect of \( \theta_1 \) (the first period supply) on \( p_2 \) (the second period supply).

**4.2 Endowment shock propagation and liquidity: Unique equilibrium**

Because endowment shocks propagate across trading dates, in general a shock to the first period liquidity affects the second period price, second period hedgers’ return uncertainty (\( \text{Var}[v - p_2 | \Omega^H_2] \)), and thus the latter response to their endowment shock. This, in turn, impacts the uncertainty faced by the first period investors that re-trade in the second period (\( \text{Var}[p_2 | \theta_1] \)), as well as the trading decisions of long term traders (HL and D), ultimately affecting first period liquidity. Thus, first period liquidity arises from a loop that ties the first and second period traders’ strategies. The following 3 corollaries provide parameters’ restrictions
ensuring that second period hedgers’ uncertainty is exogenous, in which case the loop breaks down, and a unique equilibrium obtains.

We start by pinning down the conditions ensuring that \( \theta_1 \) does not affect \( p_2 \). From our previous discussion, this happens when \( \beta b^{HL}_1 + (1 - \mu) \gamma \tau_v \Lambda_1 = 0 \), in which case the aggregate supplies of the asset at dates 1 and 2 are uncorrelated, and \( \text{Cov}[p_2, \theta_1] = 0 \). This condition arises either when all first period hedgers are HS and there are only FDs in the market (i.e., \( \mu = 1 \) and \( \beta = 0 \)), or, assuming \( \mu, \beta \in (0, 1) \), when the aggregate demand of HL is exclusively cleared by standard dealers:

\[
\beta \frac{x^{HL}_1}{\theta_1} = -(1 - \mu) \frac{x^D_1}{\theta_1}, \tag{17}
\]

In these cases, the uncertainty faced by second period hedgers is exogenous:

\[
\text{Var}[v - p_2|\Omega^H_2] = \tau_v^{-1} \tag{18}
\]

and the equilibrium is unique, as we argue in the next two results.

**Corollary 3.** If \( \beta = 0 \) and \( \mu = 1 \), there exists a unique equilibrium where

\[
\Lambda_1 = \frac{1}{\tau_v(\gamma + (\gamma + \gamma^H_1)(\gamma + \gamma^H_2)^2 \tau_u \tau_v)}. \tag{19}
\]

If we impose \( \mu, \beta \in (0, 1) \), there are two sets of parameter restrictions for which (17) holds:

**Corollary 4.** If \( \mu, \beta \in (0, 1) \), and either

\[
\tau_u \tau_v = \beta \frac{1}{\beta(1 - \beta) + \gamma \mu} \tag{20a}
\]

\[
\beta = \frac{1 - \mu}{1 + \mu}, \tag{20b}
\]

or

\[
\tau_u \tau_v = \frac{(1 - \beta)^2((1 - \beta)(\gamma + \beta \gamma^H_1) - \gamma \mu)}{\beta(\gamma^H_1(1 - \beta) + \gamma \mu)(\gamma^H_2(1 - \beta) + \gamma \mu)^2} \tag{21a}
\]

\[
(\gamma^H_2(1 - \beta) + \gamma \mu)^2 \tau_u \tau_v < (1 - \beta)^2 \tag{21b}
\]

\[
\beta \geq 1 - \mu, \tag{21c}
\]

there exists a unique equilibrium where \( \beta x^{HL}_1 = -(1 - \mu)x^D_1 \), and

\[
\Lambda^*_1 = \frac{\beta}{\tau_v(\beta \gamma^H_1 + (1 - \mu) \gamma \Lambda_1)}. \tag{22}
\]

When parameters’ values satisfy (20a)–(20b), we obtain \( b^{HL}_1 = b^HS_1 \) and \( \text{Var}[p_2|\theta_1] = \text{Var}[v] \). When they satisfy (21a)–(21c), we have instead \( b^{HL}_1 < b^HS_1 \) and \( \text{Var}[p_2|\theta_1] > \text{Var}[v] \).

---

10This is because \( \text{Var}[v - p_2|\Omega^H_2] = \text{Var}[v] + \text{Var}[p_2|\Omega^H_2] = \text{Var}[v] + \lambda^2_2 \left( (\beta b^{HL}_1 + (1 - \mu) \gamma \tau_v \Lambda_1)^2 \text{Var}[\theta_1|s_{\theta_1}] \right) \).
Remark 1. A further implication of Corollary 4 is that \( \beta x^H_{HL} > -(1 - \mu) x^D_1 \iff \Lambda_1 > \Lambda^*_1 \), and \( \text{Var}[v - p_2|\Omega^H_2] \) is a convex function of \( \Lambda_1 \), with a minimum at \( \Lambda_1 = \Lambda^*_1 \). When \( \Lambda_1 < \Lambda^*_1 \), part of HL’s orders are absorbed by FDs. As \( \Lambda_1 \) increases, HL hedge less and more of their orders are matched by D. This reduces the dependence of \( p_2 \) on \( \theta_1 \), attenuating the uncertainty faced by second period hedgers. Conversely, at equilibria where \( \Lambda_1 > \Lambda^*_1 \), a more illiquid first period market increases second period hedgers’ uncertainty. Therefore, through the endowment shock propagation mechanism, an increase in \( \Lambda_1 \) can either dampen or heighten the uncertainty faced by second period hedgers.

Second period hedgers’ uncertainty is also exogenous when these traders have a perfect signal on \( \theta_1 \). In this case, the following result holds:

**Corollary 5.** If second period hedgers observe perfectly \( \theta_1 \), there is a unique equilibrium in linear strategies where

\[
\Lambda_1 = \frac{(1 - \beta)^2 + \beta \tau_u \tau_v ((1 - \beta) \gamma^H_1 + \gamma \mu)((1 - \beta) \gamma^H_2 + \gamma \mu)}{\tau_v ((1 - \beta)^2(\gamma + \beta \gamma^H_1) + (\gamma + \beta \gamma^H_1 + \gamma^H_2 (1 - \beta))((\gamma + \gamma^H_1 (1 - \beta))(\gamma \mu + \gamma^H_2 (1 - \beta)) \tau_u \tau_v)},
\]

with \( \psi'(\Lambda_1) < 0, \partial \Lambda_1 / \partial \mu < 0 \). When second period hedgers also become infinitely risk averse, they act like noise traders:

\[
\lim_{\gamma^H_2 \to 0} \left( \lim_{\tau_\eta \to \infty} b_{21}^H \right) = -1, \lim_{\gamma^H_2 \to 0} \left( \lim_{\tau_\eta \to \infty} b_{22}^H \right) = 0.
\]

If second period hedgers’ signal on \( \theta_1 \) is not perfect, convergence to the noise traders’ case does not obtain. This is because in this case second period hedgers still face execution risk, which tames their hedging aggressiveness.

### 4.3 Endowment shock propagation and liquidity: Multiplicity

We now turn to investigate the effect of second period hedgers’ endogenous uncertainty for equilibrium determination. Suppose first that \( \tau_\eta = 0 \), so that \( b_{22}^H = 0 \) (see (12d)). Consider an equilibrium where \( \Lambda_1 < \Lambda^*_1 \), and assume that \( \Lambda_1 \) increases. This lowers the uncertainty faced by second period hedgers (Remark 1), leading them to lower their asset exposure (increase \( |b_{21}^H| \), see (12c)). Thus, both the price impact of trades in the second period (see (14a)), and the variance of hedgers’ demand in the second period \( ((b_{21}^H)^2 \tau_u^{-1}) \) increase, boosting the conditional uncertainty on the second period price (\( \text{Var}[p_2|\theta_1] \)). This effect increases the cost of supplying liquidity, which lowers the denominator in (13). Consider now the effect of the liquidity shock on first period liquidity consumption. On the one hand a larger \( \Lambda_1 \) lowers the demand of liquidity by HL (see (12a)) and FDs’ speculation activity. Both effects reduce the numerator

\[\text{This is because, according to Corollary 2 for } \Lambda_1 < \Lambda^*_1, \text{ Cov}[p_2, \theta_1] < 0 \text{ and as } \Lambda_1 \text{ increases, the latter approaches } 0, \text{ while Var}[p_2|\theta_1] \text{ increases.} \]
On the other hand, an increase in $\Lambda_1$ has an ambiguous impact on HS’s liquidity consumption. To see this, let us rewrite (12b) in the following way:

$$b_H^1 = \left(1 - \gamma^H \Lambda_1 \left( \frac{\lambda_2 (\beta b_{HL}^1 + (1 - \mu) \gamma^v \Lambda_1)}{\text{Var}[p_2|\theta_1]} \right) \right).$$

For $\Lambda_1 < \Lambda_1^*$, an increase in $\Lambda_1$ reduces the dependence of $p_2$ on $\theta_1$ and heightens $\text{Var}[p_2|\theta_1]$. This makes HS hold more of their endowment:

$$\frac{\partial}{\partial \Lambda_1} \left( \frac{\gamma^H \Lambda_1}{\text{Var}[p_2|\theta_1]} \right) < 0,$$

which reduces their liquidity consumption. However, as $\text{Var}[p_2|\theta_1]$ increases, HS find it increasingly risky to hold their endowment and even if the market becomes less liquid, may choose to hedge more. In fact, direct differentiation yields

$$\frac{\partial b_H^1}{\partial \Lambda_1} < 0$$

for $\Lambda_1^* < \Lambda_1 < \Lambda_1^*$.

As a consequence, the aggregate best response can display a positively sloped branch which can lead to multiple equilibria. Figure 2 (panel (a)) illustrates a case in which multiple equilibria arise.

Consider now the effect of an increase in $\mu$. A larger fraction of FDs corresponds to an increase in the risk-bearing capacity of the market since these traders’ supply is less price-elastic than the one of standard dealers (see (8)). This effect tends to lower $\Lambda_1$, thereby improving liquidity. However, due to the loop linking second period hedgers’ response to $\Lambda_1$, an increase in $\mu$ can lead to a reduction in the liquidity of the first period market. To see this, consider an equilibrium with $\Lambda_1 > \Lambda_1^*$. Differentiating (12c) yields

$$\Lambda_1 > \Lambda_1^* \implies \frac{\partial b_{21}^H}{\partial \mu} < 0.$$
Figure 2: In panel (a) we plot the equilibrium mapping \([13]\) against the 45-degree line (the vertical gray line is drawn at \(\Lambda_1^*\)). In panel (b) we plot the equilibrium first period liquidity as a function of \(\mu\). Parameters’ values are \(\tau_\eta = 0, \tau_u = \tau_v = 1/4, \gamma = 1/2, \gamma^H_1 = \gamma^H_2 = .3, \beta = .54,\) and \(\mu = .65\). In panel (c) and (d) we repeat the same plots with the following parameters’ values: \(\tau_\eta = 75, \tau_u = 3.5, \tau_v = 2.5, \gamma = 1, \gamma^H_1 = \gamma^H_2 = .8, \beta = 0,\) and \(\mu = .01\).

about \(\theta_1\) (Section 4.1), which can heighten liquidity complementarities. To see this, suppose that \(\beta = 0\), in which case \(\Lambda_1^* = 0\), and an increase in \(\Lambda_1\) always increases the uncertainty faced by 2nd period hedgers, so that \(\partial b_{21}^H / \partial \Lambda_1 > 0\), which works to reduce the intensity of the loop. However, when \(\beta = 0\), according to \([12d]\), \(b_{22}^H < 0\), and second period hedgers faced with \(s_{\theta_1} > 0\) sell to speculate on the anticipated demand due to HS unwinding at date 2. In this case, an increase in \(\Lambda_1\) leads D to supply more liquidity, potentially augmenting the
speculative opportunities for second period hedgers. This, in turn, can make \( \partial b_{22}^H / \partial \Lambda_1 < 0 \), which strengthens the liquidity consumption loop across dates. In Figure 2, panels (c) and (d) we show an example of equilibrium multiplicity when \( \tau_\eta > 0 \), together with the effects of complementarities for \( \Lambda_1(\mu) \).

**Remark 2.** As argued in Corollary 5, when \( \tau_\eta \to \infty \) a unique equilibrium arises. The intuition for this result can now be understood through the effect of a perfect \( \theta_1 \)-signal on liquidity complementarities. When second period hedgers observe \( \theta_1 \), their return uncertainty no longer depends on \( \Lambda_1 \) (Var[\( v - p_2 | \Omega^H_2 \)] = Var[\( v \)]). Thus, a shock to first period liquidity leaves second period hedgers strategies unaffected, and importantly also the uncertainty faced by first period investors (Var[\( p_2 | \theta_1 \)]). Thus, a perfect \( \theta_1 \)-signal effectively shuts down the two sources of the loop responsible for liquidity complementarities.

### 4.4 Liquidity fragility

The analysis of the previous section has shown that as the number of FDs changes, the market can hover in a region with multiple equilibria with different levels of liquidity. We now use this insight to show that this can explain liquidity “fragility,” understood as abrupt changes (in market liquidity) originating in small changes in the underlying parameters of the market.

Consider Figure 3, in which we use the same parameter values of Figure 2 (panel (a)) but set \( \mu = .66 \). In this situation, we have a unique equilibrium with \( \Lambda_1 = 4 \). Suppose now that a computer glitch disconnects a small number of FDs (a 3%-reduction, from \( \mu = .66 \) to \( \mu = .64 \)). As \( \mu \) decreases, the complementarity loop strengthens, and the market moves initially to the region with multiple equilibria depicted in Figure 2 (panel (a)). There are now three equilibria, and since the one associated with the highest liquidity is stable (\( \psi'(\Lambda_1) < 1 \)), we have \( \Lambda_1 = 4.2 \). Thus, the initial liquidity drop is rather contained (5%). Note, however, that the other two equilibria are associated with much lower levels of liquidity. The effect of the final adjustment is illustrated in panel (c). As the market moves towards the state with \( \mu = .64 \), the high liquidity equilibrium disappears, and \( \Lambda_1 = 5.7 \). Because of the liquidity complementarity loop, a 3%-reduction in the number of FDs has triggered a 43%-liquidity withdrawal, inducing a collapse of liquidity (a “Flash Crash”).

### 5 Welfare analysis

In this section we compute and analyze the welfare functions of market participants. We have the following result

14 This is because if \( \beta = 0 \), only HS are in the market at date 1, and when D absorb a larger portion of the aggregate demand, the propagation of \( \theta_1 \) to the second period strengthens.

15 It is possible to show that for \( \Lambda_1 > 1 + \sqrt{1 - \mu \text{Var}[\theta_1 | s_{01}](\gamma_2^H + \mu \gamma)/(1 - \mu)^2 \gamma \tau_v}, or 0 < \Lambda_1 < 1 - \sqrt{1 - \mu \text{Var}[\theta_1 | s_{01}](\gamma_2^H + \mu \gamma)/(1 - \mu)^2 \gamma \tau_v}, \partial b_{22}^H / \partial \Lambda_1 < 0.\)
Figure 3: In panel (a) we plot the equilibrium mapping \([13]\) against the 45-degree line (the vertical gray line is drawn at \(\Lambda^*_1\)). Parameters’ values are \(\tau_{\eta} = 0\), \(\tau_u = \tau_v = 1/4\), \(\gamma = 1/2\), \(\gamma_H^1 = \gamma_H^2 = .3\), \(\beta = .54\), and \(\mu = .66\). In panel (b) and (c) we repeat the same plots, lowering \(\mu\) to \(.65\), and \(.64\).

**Proposition 6.** At a linear equilibrium the unconditional expected utilities of market participants are given by

\[
EU^{HL} = - \left( 1 + \frac{\text{Var}[x^H_L] \text{Var}[v]}{(\gamma_H^1)^2} - \frac{\text{Var}[\theta_1] \text{Var}[v]}{(\gamma_H^1)^2} \right)^{-1/2}
\]

\[
EU^{HS} = - \left( 1 + \frac{\text{Var}[x^H_S] \text{Var}[p_2|\theta_1]}{(\gamma_H^1)^2} + \frac{2 \text{Cov}[p_2, \theta_1]}{\gamma_H^1} - \frac{\text{Var}[\theta_1] \text{Var}[p_2|\theta_1]}{(\gamma_H^1)^2} \right)^{-1/2},
\]

(24a) (24b)
respectively for first period long and short term hedgers,

\[
EU^H = -\left(1 + \frac{\text{Var}[x_t^H]\text{Var}[v - p_t^2|\Omega_t^H]}{(\gamma_t^H)^2} - \frac{\text{Var}[u_2]\text{Var}[v]}{(\gamma_t^H)^2}\right) + \\
\left(p_{x_t^H,u_2}^2 - 1\right)\frac{\text{Var}[v]\text{Var}[u_2]\text{Var}[x_t^H]\text{Var}[v - p_t^2|\Omega_t^H]}{(\gamma_t^H)^4}\right)^{-1/2},
\]

for second period hedgers, and

\[
EU^D = -\left(1 + \frac{\text{Var}[p_1]}{\text{Var}[v]}\right)^{-1/2}
\]

\[
EU^{FD} = -\left(1 + \frac{\text{Var}[p_2|p_1]}{\text{Var}[v]} + \frac{\text{Var}[E[p_2|p_1] - p_1]}{\text{Var}[p_2|p_1]}\right)^{-1/2},
\]

respectively for first period dealers and FDs.

Expressions (24a) and (24b) show that first period hedgers are negatively affected by the uncertainty over the value of their endowment (due to the effect of \(\text{Var}[\theta_1]\text{Var}[v]\) and \(\text{Var}[\theta_1]\text{Var}[p_2|\theta_1]\) for HL and HS) and benefit from trading (due to the effect of \(\text{Var}[x_t^H], k \in \{L,S\}\)). Note that for HL, this implies the following result:

**Corollary 7.** An increase in liquidity improves HLs’ welfare.

This result follows immediately from the fact that \(\text{Var}[x_t^{HL}] = (b_t^{HL})^2\tau_u^{-1}\), and from (12a), according to which a more liquid market makes HL trade more aggressively, as it reduces hedging costs. Coupled with Corollary 5, the above result also implies that an increase in FDs always improves HLs’ welfare.

For HS, (24b) shows that they also benefit when \(p_2\) is positively correlated with their endowment. To see why, suppose that \(\theta_1 > 0\). Then, short term hedgers reduce their exposure to the asset (as \(b_t^{HS} < 0\)), selling part of their endowment in the first period, and keeping \((1 - |b_t^{HS}|)\) of it in their portfolio. In this case, the higher is \(p_2\), the more they earn out of such position. Thus, HS’s utility contains a capital gains component.

According to (25) second period hedgers also benefit from trading and suffer from the uncertainty of their endowment value. Additionally, the lower is the correlation between their endowment shock and their strategy \((\rho_{x_t^H,u_2}^2)\), the lower is their utility. This is because in this case their strategy serves as a poor hedging instrument against the random exposure to the risky asset payoff. Note that this effect disappears if \(\tau_\eta = 0\), in which case their strategy only loads on \(u_2\).

Overall, we can say that (i) all types of hedgers benefit when they can trade more and suffer from the uncertainty of their endowment value, (ii) HS are adversely affected by short term price reversion, and (iii) second period hedgers suffer when their strategy acts as a poor hedging instrument against their endowment shock.

Expression (26a) shows that dealers derive utility from liquidity provision to first period hedgers. More in detail, \(\text{Var}[p_1]\), and \(\text{Var}[v]\) are respectively the variance of the profit from
liquidity provision that is explained and unexplained by \( p_1 \) (that is \( \text{Var}[E[v - p_1|p_1]] = \text{Var}[p_1] \), and \( \text{Var}[v] = \text{Var}[v - p_1|p_1] \)). Therefore, the higher is \( \text{Var}[p_1] \) (\( \text{Var}[v] \)) the more (less) accurately dealers can anticipate their profit — namely \( (v - p_1) \) — from the knowledge of \( p_1 \).

From (26b) we can instead see that FDs derive utility from three trading activities: (i) liquidity supply to first period hedgers like traditional dealers, (ii) speculation on short term returns, and (iii) liquidity supply to second period hedgers. A change in the conditional volatility of returns has two contrasting effects on FDs’ utility: the higher is \( \text{Var}[p_2|p_1] \), the larger are the potential returns from second period liquidity provision (as this implies a larger liquidity demand from second period hedgers), but also the higher is the risk of speculating on short-term price changes (the short term return variance unexplained by \( p_1 \)).\(^{16}\)

Comparing (26a) with (26b) yields

\[
EU^{FD} > EU^D,
\]

implying that the ability to trade at a higher frequency creates value for FDs.

We will measure the welfare of market participants using certainty equivalents: the welfare of a type \( k \) first period hedger and the one of a second period hedger are respectively measured by

\[
CE^{Hk} \equiv -\gamma^H_1 \ln (-EU^{Hk}), \quad k \in \{L, S\},
\]

\[
CE^{H} \equiv -\gamma^H_2 \ln (-EU^{H}).
\]

Finally, for a FD and a traditional dealer we have

\[
CE^{FD} \equiv -\gamma \ln (-EU^{FD}),
\]

\[
CE^{D} \equiv -\gamma \ln (-EU^{D}).
\]

5.1 A numerical example

The market microstructure literature considers the liquidity of a market as an important quality parameter, intimately related to investors’ welfare. Intuitively, a more liquid market minimizes trading costs, thereby encouraging hedgers to trade. This is the intuition behind the effect of liquidity on HLs’ welfare in Corollary \(^7\). However, HS and second period hedgers’ welfare also depends on other factors, so that the ultimate welfare effect of liquidity is not clear.

\(^{16}\)At date 1, the short term profit of a FD is \( (p_2 - p_1) \). The variance of this profit that is explained (unexplained) by \( p_1 \) is \( \text{Var}[E[p_2|p_1] - p_1] \) (\( \text{Var}[p_2 - p_1|p_1] = \text{Var}[p_2|p_1] \)). Thus, the higher is \( \text{Var}[E[p_2|p_1] - p_1] \) (\( \text{Var}[p_2|p_1] \)), the more (less) accurately FDs can anticipate their short term profit based on \( p_1 \).
\( \tau_v = 2.5, \tau_u = 3.5, \tau_n = 75, \gamma = 1, \gamma_1^H = .3, \gamma_2^H = .75, \beta = .4 \)

<table>
<thead>
<tr>
<th>( \mu = .014 )</th>
<th>( \mu = .015 )</th>
<th>( \mu = .016 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Lambda_1 )</td>
<td>.23</td>
<td>.22</td>
</tr>
<tr>
<td>(</td>
<td>\lambda_{21} + \lambda_{22}</td>
<td>)</td>
</tr>
<tr>
<td>( \text{Var}[\nu - p_2</td>
<td>\Omega_2^H] )</td>
<td>1.03</td>
</tr>
<tr>
<td>( \text{Var}[p_2</td>
<td>\theta_1] )</td>
<td>.58</td>
</tr>
<tr>
<td>( \text{Cov}[p_2, \theta_1] )</td>
<td>.23</td>
<td>.15</td>
</tr>
<tr>
<td>( \text{Cov}[x^{NR}<em>1, s</em>{\theta_1}] )</td>
<td>2</td>
<td>1.52</td>
</tr>
</tbody>
</table>

**Table 1: Multiple equilibria and welfare.**

In Table 1 we compute welfare and standard market quality indicators for a given set of parameter values, and for \( \mu \in \{.014, .015, .016\} \).\(^{17}\) When \( \mu = .014 \), and \( \mu = .016 \) we obtain a unique equilibrium, whereas for \( \mu = .015 \) three equilibria arise. In the table we also compute Hedgers’ Welfare (HW) and Total Welfare (TW), defined in the following way:

\[
\begin{align*}
HW & \equiv \beta CE^{HL} + (1 - \beta) (CE^{HS} + CE^H) \\
TW & \equiv \mu CE^{FD} + (1 - \mu) CE^D + HW.
\end{align*}
\]

The table offers a number of insights. First, note that an increase in FD increases liquidity across both trading dates. Second, this liquidity increase benefits HL (consistent with Corollary \(^{17}\)), and affects negatively the welfare of all other market participants, implying that both HW and TW decrease as \( \mu \) increases. For FDs and D the result is intuitive as an increase in \( \mu \) heightens competition, thus eroding the profits from liquidity provision. For second period hedgers, the effect goes through the reduced trading due to lower reliance on \( s_{\theta_1} \). Finally, for HS, the increase in \( |b_1^{HL}| \) and the decrease in \( \text{Var}[p_2|\theta_1] \), increase their speculative activity, thereby reducing their

\(^{17}\)The other parameter values are \( \tau_v = 2.5, \tau_u = 3.5, \tau_n = 75, \gamma = 1, \gamma_1^H = .3, \gamma_2^H = .75, \beta = .4 \).
hedging aggressiveness. This effect is not compensated by the capitals gains component, as an increase in $\mu$ lowers the association between HS’s endowment and $p_2$.

6 Entry

We determine the number of FDs that enter the intermediation market in three different scenarios: when (i) there is “Free-Entry” in the industry and all FDs that can break even paying a cost $c > 0$ enter; (ii) the decision is made by a monopolistic exchange that can restrict access to its platform (e.g., by selling co-location space to a limited number of FDs’ computers—in this case the assumption is that the exchange bears the platform cost $c$ and passes it on to FDs via a fee); and (iii) when the decision is made by a social planner who maximizes a utilitarian welfare function (and thus internalizes the FD cost as well as the effects of a change in $\mu$ on all market participants).

We start by determining the value that a traditional dealer obtains by becoming a FD. In view of (27) this value is positive and is defined as the certainty equivalent $\phi$ of being a FD compared to being a traditional dealer:

$$\phi(\mu) \equiv CE^{FD} - CE^{D}.$$

Using (26a), (26b), (29a), (29b), and in view of (27) we obtain:

$$\phi(\mu) = \gamma \ln \frac{EU^{D}}{EU^{FD}} > 0.$$  \hspace{1cm} (30)

The free entry solution can now be defined as the value $\mu^{FE}$, such that

$$\phi(\mu^{FE}) = c.$$  \hspace{1cm} (31)

The monopolistic solution obtains at $\mu^{EX}$ such that

$$\mu^{EX} \in \arg \max_{\mu} \Pi^{EX} \equiv (\phi(\mu) - c)\mu.$$  \hspace{1cm} (32)

Finally, the planner solution obtains at $\mu^{P}$ such that

$$\mu^{P} \in \arg \max_{\mu} \Pi^{P}(\mu) \equiv TW(\mu) - c\mu = \Pi^{EX}(\mu) + CE^{D}(\mu) + HW(\mu).$$  \hspace{1cm} (33)

From (32) it is apparent that a monopolistic exchange does not internalize the effect of FDs on hedgers’ welfare. Numerically, it can be shown that in general $\mu^{EX} \neq \mu^{P}$. In Figure 4 we present the results of numerical simulations and in Table 2 we summarize them. An important implication of the figure is that total welfare can be convex in $\mu$, whereas the profit function of the exchange is typically concave in the number of FDs. This suggests that social welfare can be maximized by extreme solutions, while the market solution tends to be interior.
When second period hedgers’ signal is perfect, we can solve the model in closed form and obtain an analytical characterization of the solutions, for some parameter values:

**Corollary 8.** If second period hedgers observe perfectly \( \theta_1 \) and become infinitely risk-averse,
when $\beta = 1$, $\mu^P = \mu^{EX}$, and the monopolistic exchange implements the first best.

When $\beta = 1$ only HL matter for welfare maximization. However, as no interim trading by HS can be anticipated by FDs, the strategy of the latter coincides with the one of D. This implies that HLs’ strategies do not depend on $\mu$. When $\beta < 1$, our numerical simulations show that the monopolistic solution typically involves too much entry for $\beta$ small, and too little for $\beta$ large. We present this numerical result in Figure 5.

![Figure 5: Comparing the three solutions, case $\beta < 1$.](image)

7 Conclusions

When hedgers’ horizons are heterogeneous and dealers are risk averse endowment shocks become endogenously persistent, which can generate liquidity complementarities and possibly multiple equilibria. In our model HFTs have better inventory management and can therefore supply liquidity at a lower cost. However, being in the market all the time are also sensitive to anticipated bouts of volatility. We show that because of liquidity complementarities this can induce “fragility”: a small change in the mass of HFT can create discrete changes in liquidity (a “Flash-Crash”).

Our analysis shows that higher liquidity is not necessarily indicative of higher welfare and that total welfare can be convex in the number of HFTs, which can yield extreme solutions. Conversely, the monopolistic exchange (which does not internalize the external effects of fast traders) typically has an interior solution.
References


Appendix

Proof of Proposition 1

Assume that in the second period a fraction \((1 - \beta)\) of hedgers endowed with risk-tolerance coefficient \(\gamma^H > 0\) enter the market. A date-2 hedger submits a market order

\[ X^H_2(u_2, s_{\theta_1}) = b^H_{21}u_2 + b^H_{22}s_{\theta_1}, \tag{34} \]

with \(u_2 \sim N(0, \tau^{-1}_u)\), and \(s_{\theta_1} = \theta_1 + \eta\), with \(\eta \sim N(0, \tau^{-1}_\eta)\) and \(u_2, \eta\) independent of all the other random variables in the model.\(^{18}\)

Consider the sequence of market clearing equations

\[ \mu x^{FD}_1 + (1 - \mu) x^{D}_1 + \beta x^{HL}_1 + (1 - \beta) x^{HS}_1 = 0 \tag{35a} \]

\[ \mu(x^{FD}_2 - x^{FD}_1) + (1 - \beta)(b^H_{21}u_2 + b^H_{22}s_{\theta_1} - x^{HS}_1) = 0. \tag{35b} \]

Rearrange (35a) as follows:

\[ (1 - \mu)x^{D}_1 + \beta x^{HL}_1 = - (\mu x^{FD}_1 + (1 - \beta)x^{HS}_1). \]

Substitute the latter in (35b):

\[ \mu x^{FD}_2 + (1 - \beta)(b^H_{21}u_2 + b^H_{22}s_{\theta_1}) + (1 - \mu)x^{D}_1 + \beta x^{HL}_1 = 0. \]

Due to CARA and normality, in the second period a FD’s limit order is given by

\[ X^{FD}_2(p_1, p_2) = -\gamma \tau_v p_2. \]

Similarly, in the first period the limit order of a traditional market maker is given by

\[ X^{D}_1(p_1) = -\gamma \tau_v. \]

Substituting these strategies in the above market clearing equation and solving for \(p_2\) yields

\[ p_2 = \lambda_2 \left( (1 - \beta)(b^H_{21}u_2 + b^H_{22}s_{\theta_1}) + \beta x^{HL}_1 \right) - \frac{1 - \mu}{\mu} p_1, \tag{36} \]

where \(\lambda_2 = 1/\mu \gamma \tau_v\). The assumption that first period hedgers’ strategies are linear implies that \(p_1 = -\Lambda_1 \theta_1\) (see below). As a consequence we can rewrite (36) as follows:

\[ p_2 = \lambda_2(1 - \beta)(b^H_{21}u_2 + b^H_{22}s_{\theta_1}) + \lambda_2 \left( \beta b^H_{11} + (1 - \mu) \gamma \tau_v \Lambda_1 \right) \theta_1. \tag{37} \]

CARA and normality assumptions imply that the objective function of a second period hedger is given by

\[ E[-\exp\{-\pi^H_2/\gamma^H_2\}|\Omega^H_2] = -\exp\left\{ -\frac{1}{\gamma} \left( E[\pi^H_2|\Omega^H_2] - \frac{1}{2\gamma} \text{Var}[\pi^H_2|\Omega^H_2] \right) \right\}, \tag{38} \]

where \(\Omega^H_2 = \{u_2, s_{\theta_1}\}\), and \(\pi^H_2 \equiv (v - p_2)x^H_2 + u_2v\). Maximizing (38) with respect to \(x^H_2\), the

\(^{18}\)Linear equilibria will be symmetric. Therefore we drop the \(i\) index from traders’ strategies.
strategy of a second period hedger is given by

\[ X^H_2(u_2, s_{\theta_1}) = \gamma^H_2 \left( u_2 \right) E[v - p_2|\Omega^H_2] - \frac{\text{Cov}[v - p_2, v|\Omega^H_2]}{\text{Var}[v - p_2|\Omega^H_2]} u_2. \]  

(39)

Computing

\[
E[v - p_2|\Omega^H_2] = -\left( \lambda_2(1 - \beta)(b^H_{21}u_2 + b^H_{22}s_{\theta_1}) + \left( \lambda_2\beta b^H_1 + \frac{1 - \mu}{\mu} - \frac{\tau_\eta}{\tau_\eta + \tau_u} \right) s_{\theta_1} \right) \]  

(40a)

\[
\text{Var}[v - p_2|\Omega^H_2] = \mu^2(\tau_u + \tau_\eta) + \frac{(\mu\lambda_2\beta b^H_1 + (1 - \mu)\Lambda_1)^2\tau_v}{\mu(\tau_\eta + \tau_u)} \]  

(40b)

\[
\text{Cov}[v - p_2, v|\Omega^H_2] = \frac{1}{\tau_v}. \]  

(40c)

Substituting (40a), (40b), and (40c) in (39) and identifying coefficients yields

\[ X^H_2(u_2, s_{\theta_1}) = b^H_{21}u_2 + b^H_{22}s_{\theta_1}, \]

where

\[
b^H_{21} = -\frac{1}{\tau_v\left( \gamma^H_2\lambda_2(1 - \beta) + \text{Var}[v - p_2|\Omega^H_2] \right)} \]  

(41a)

\[
b^H_{22} = -\frac{\gamma^H_2\tau_\eta(\mu\lambda_2\beta b^H_1 + (1 - \mu)\Lambda_1)}{\mu(\tau_\eta + \tau_u)(\gamma^H_2\lambda_2(1 - \beta) + \text{Var}[v - p_2|\Omega^H_2])}. \]  

(41b)

Note that

1. Both (41a) and (41b) depend on \( b^H_1 \) and \( \Lambda_1 \) via \( \text{Var}[v - p_2|\Omega^H_2] \). This is because by assumption second period hedgers can only observe \( s_{\theta_1} \), which exposes them to the volatility that the first period endowment shock creates on \( v - p_2 \).

2. According to (39) second period hedgers’ strategies react both to endowment and informational shocks. Thus, there are \textit{two} measures of the price impact of trades in the second period (see (37)):

\[
\lambda_{21} \equiv \frac{\partial p_2}{\partial u_2} = \lambda_2(1 - \beta)b^H_{21} \]  

(42a)

\[
\lambda_{22} \equiv \frac{\partial p_2}{\partial s_{\theta_1}} = \lambda_2(1 - \beta)b^H_{22}. \]  

(42b)

Expressions (42a) and (42b) respectively correspond to the price impact of a marginal increase in the endowment shock and in the realization of the signal about \( \theta_1 \) observed by second period hedgers.

Consider now the first period. We start by characterizing the strategy of a FD. Standard
arguments imply that the first period objective function of a FD is given by

\[ E[U((p_2 - p_1)x_1^{FD} + (v - p_2)x_2^{FD})|\theta_1] = -\left(1 + \frac{\text{Var}[p_2|\theta_1]}{\text{Var}[v]}\right)^{-1/2} \times \]

\[ \exp \left\{ -\frac{1}{\gamma} \left( \frac{\gamma\tau_v}{2} \nu^2 + (\nu - p_1)x_1^{FD} - \frac{(x_1^{FD} + \gamma\tau_v\nu)^2}{2\gamma} \right) \right\}, \]

where

\[ \nu \equiv E[p_2|\theta_1] = \left( \lambda_2(\beta b_{1H}^L + (1 - \beta)b_{22}^H) + \frac{1 - \mu}{\mu} \Lambda_1 \right) \theta_1 \]

(44a)

\[ \text{Var}[p_2|\theta_1] = \lambda_2^2(1 - \beta)^2 \left( \frac{b_{21}^H}{\tau_u} + \frac{b_{22}^H}{\tau_\eta} \right)^2. \]

(44b)

Maximizing (43) with respect to \( x_1^{FD} \) and solving for the first period strategy yields

\[ x_1^{FD}(p_1) = \frac{\gamma}{\text{Var}[p_2|\theta_1]} \nu - \gamma \left( \frac{1}{\text{Var}[p_2|\theta_1]} + \frac{1}{\text{Var}[v]} \right) p_1. \]

(45)

Due to CARA and normality, for traditional market makers at date 1 we have \( x_1^{FD}(p_1) = -\gamma\tau_v p_1 \). At equilibrium we then have

\[ \mu \left( \frac{\gamma}{\text{Var}[p_2|\theta_1]} \nu - \gamma \left( \frac{1}{\text{Var}[p_2|\theta_1]} + \frac{1}{\text{Var}[v]} \right) p_1 \right) + (-\mu)\gamma\tau_v p_1 + (\beta b_{1H}^L + (1 - \beta)b_{1H}^S) \theta_1 = 0, \]

implying that \( p_1 \) is linear in \( \theta_1 \): \( p_1 = -\Lambda_1 \theta_1 \), with \( \Lambda_1 \) to be determined.

We now turn to the characterization of first period hedgers' strategies. From CARA and normality a HL's objective function is given by

\[ E[-\exp\{-\pi_{11}^{HL}/\gamma_1^H\}] = -\exp \left\{ -\frac{1}{\gamma} \left( E[\pi_{11}^{HL}|\theta_1] - \frac{1}{2\gamma} \text{Var}[\pi_{11}^{HL}|\theta_1] \right) \right\}, \]

(46)

where \( \pi_{11}^{HL} \equiv (v - p_1)x_1^{HL} + \theta_1 v \). Maximizing (46) with respect to \( x_1^{HL} \), and solving for the optimal strategy, yields

\[ x_1^{HL}(\theta_1) = \gamma_1^H \left( \frac{E[v - p_1|\theta_1]}{\text{Var}[v - p_1|\theta_1]} - \frac{\text{Cov}[v - p_1, v|\theta_1]}{\text{Var}[v - p_1|\theta_1]} \right) \theta_1. \]

(47)

Computing

\[ E[v - p_1|\theta_1] = -\Lambda_1 \theta_1 \]

(48a)

\[ \text{Cov}[v - p_1, v|\theta_1] = \text{Var}[v - p_1|\theta_1] = \frac{1}{\tau_v}. \]

(48b)

Substituting the above in \( x_1^{HL} \) and rearranging yields

\[ x_1^{HL}(\theta_1) = b_{1H}^{HL} \theta_1, \]

(49)
where
\[ b_1^{\text{HL}} = -(1 - \gamma_1^\text{H} \Lambda_1 \tau_v). \] (50)

For a HS, CARA and normality imply
\[ E[- \exp\{-\pi_{i1}^{\text{HS}} / \gamma_1^\text{H}\}] = - \exp \left\{ - \frac{1}{\gamma} \left( E[\pi_{i1}^{\text{HS}} | \theta_1] - \frac{1}{2\gamma} \text{Var}[\pi_{i1}^{\text{HS}} | \theta_1] \right) \right\}, \] (51)

where \( \pi_{i1}^{\text{HS}} \equiv (p_2 - p_1)x_1^{\text{HS}} + \theta_1p_2 \). Maximizing (51) with respect to \( x_1^{\text{HS}} \), and solving for the optimal strategy, yields
\[ X_1^{\text{HS}}(\theta_1) = \gamma_1^\text{H} \frac{E[p_2 - p_1 | \theta_1]}{\text{Var}[p_2 - p_1 | \theta_1]} - \frac{\text{Cov}[p_2 - p_1, p_2 | \theta_1]}{\text{Var}[p_2 - p_1 | \theta_1]} \theta_1. \] (52)

Computing
\[ p_2 - p_1 = \left( \lambda_2 (\beta b_1^{\text{HL}} + (1 - \beta)b_{22}^\text{H}) + \frac{\Lambda_1}{\mu} \right) \theta_1 + \lambda_2 (1 - \beta) (b_{21}^\text{H}\mu_2 + b_{22}^\text{H}\eta), \]
and
\[ E[p_2 - p_1 | \theta_1] = \left( \lambda_2 (\beta b_1^{\text{HL}} + (1 - \beta)b_{22}^\text{H}) + \frac{\Lambda_1}{\mu} \right) \theta_1 \] (53a)
\[ \text{Cov}[p_2 - p_1, p_2 | \theta_1] = \text{Var}[p_2 - p_1 | \theta_1] = \lambda_2^2 (1 - \beta)^2 \text{Var}[x_2^\text{H}], \] (53b)

where
\[ \text{Var}[x_2^\text{H}] = (b_{21}^\text{H})^2 \tau_a^{-1} + (b_{22}^\text{H})^2 (\tau_a^{-1} + \tau_\eta^{-1}). \]

Substituting the above in the strategy of a HS and identifying yields
\[ X_1^{\text{HS}}(\theta_1) = b_1^{\text{HS}} \theta_1, \] (54)

where
\[ b_1^{\text{HS}} = \gamma_1^\text{H} \frac{\mu \lambda_2 (\beta b_1^{\text{HL}} + (1 - \beta)b_{22}^\text{H}) + \Lambda_1}{\mu \text{Var}[p_2 | \theta_1]} - 1. \] (55)

Substituting (45), \( x_1^D \), (49), and (54) in the first period market clearing condition and solving for the price yields
\[ p_1 = -\Lambda_1 \theta_1, \]
where
\[ \Lambda_1 = \left( \mu \gamma \left( \frac{1}{\text{Var}[p_2 | \theta_1]} + \frac{1}{\text{Var}[\nu]} \right) + (1 - \mu) \gamma \frac{1}{\text{Var}[\nu]} \right)^{-1} \left( \mu \gamma \text{Cov}[p_2, \theta_1] \frac{\text{Var}[p_2 | \theta_1]}{\text{Var}[\nu]} + b_1 \right), \] (56)

and \( b_1 = \beta b_1^{\text{HL}} + (1 - \beta)b_1^{\text{HS}} \). According to (41a) and (41b) the equilibrium coefficients of a second period hedger’s strategy depend on \( b_1^{\text{HL}} \). Also, according to (55), the equilibrium coefficient of a HS depends on \( b_{21}^\text{H}, b_{22}^\text{H} \) and \( b_1^{\text{HL}} \). Finally, from (50) we see that \( b_1^{\text{HL}} \) only depends on \( \Lambda_1 \). Therefore, recursive substitution of the equilibrium strategies’ coefficients in (56) shows that
Λ₁ is pinned down by the solution of the following quintic equation:

\[ f(Λ₁) ≡ ψ(Λ₁) − Λ₁ = (μγ + (1 − β)γ₁^H)(Cov[p₂, θ₁]τ_u + Λ₁) + Var[p₂|θ₁](γτ_vΛ₁ + βb₁^{HL} − (1 − β)) = 0, \]

which proves our claim.

Proof of Corollary 3

Suppose that \( μ = 1 \) and \( β = 0 \). From (12d) these parameters imply that \( b₂^{H} = 0 \), which in turn implies that \( Cov[p₂, θ₁] = 0 \), and that \( Var[v − p₂|Ω₂^H] = Var[v] \). Thus, \( b₂^{H} = −γ/(γ + ψ₂^H) \), and \( Var[p₂|θ₁] = 1/(γ + ψ₂^H)^2τ_uτ_v^2 \). Substituting these expressions in (13) and rearranging yields (19).

Proof of Corollary 4

Suppose now that \( βx₁^{HL} = −(1 − μ)x₁^D \). Substituting the strategies of HL and D in this condition yields a linear equation in Λ₁:

\[ −β(1 − γ₁^Hτ_vΛ₁) + (1 − μ)γτ_vΛ₁ = 0, \]

which can be solved to obtain (22). To complete the proof of the equilibrium, we need to restrict the set of parameters’ values so that we ensure that \( βx₁^{HL} = −(1 − μ)x₁^D \). To do this we substitute \( Λ₁^* \) in (13) and impose

\[ f(Λ₁^*) ≡ ψ(Λ₁^*) − Λ₁^* = 0. \]

Manipulating the above equation yields

\[ \frac{(1 − β)^2((1 − β)(γ + βγ₁^H) − γμ) − β((1 − β)γ₁^H + γμ)((1 − β)γ₂^H + γμ)^2τ_uτ_v}{γτ_v(βγ₁^H + γ(1 − μ))(1 − β)^2 + (γ₂^H(1 − β) + γμ)^2τ_uτ_v} = 0. \]

There are two set of conditions ensuring that the numerator in the above expression is identically null. The first one is

\[ τ_uτ_v = \left( \frac{1 − β}{γ₂^H(1 − β) + γμ} \right)^2 \] (58a)

\[ β = \frac{1 − μ}{1 + μ}. \] (58b)

These conditions ensure that \( Var[p₂|θ₁] = Var[v] \), and thus that \( b₁^{HL} = b₁^{HS} \). The second set of
conditions requires instead that

\[ \tau_u \tau_v = \frac{(1 - \beta)^2((1 - \beta)(\gamma + \beta \gamma_1) - \gamma \mu)}{\beta(\gamma_1^H(1 - \beta) + \mu)(\gamma_2^H(1 - \beta) + \gamma \mu)^2} \] (59a)

\[ (\gamma_2^H(1 - \beta) + \gamma \mu)^2 \tau_u \tau_v < (1 - \beta)^2 \] (59b)

\[ \beta \geq 1 - \mu. \] (59c)

Under (59a)–(59c) we have instead \( \text{Var}[p_2|\theta_1] \geq \text{Var}[v] \) and \( b_1^{HS} < b_1^{HL} \).

\[ \Box \]

**Proof of Proposition 6**

We start from long term hedgers. Substituting (47) in (46) and rearranging yields

\[ -\exp \left\{ -\frac{1}{\gamma_1^H} \left( E[\pi_{i1}^{HL}|\theta_1] - \frac{1}{2\gamma_1^H} \text{Var}[\pi_{i1}^{HL}|\theta_1] \right) \right\} = -\exp \left\{ -\frac{\theta_1^2}{\gamma_1^H} \left( \frac{(b_1^{HL})^2 - 1}{2\gamma_1^H\tau_v} \right) \right\}, \]

where \( \theta_1 \sim N(0, \tau_u^{-1}) \). The argument at the exponential is a quadratic form of a normal random variable. To compute the unconditional expectation we can apply a well-known result from normal theory (see, e.g. Vives (2008), Technical Appendix, pp. 382–383), and obtain

\[ E[-\exp(\pi_{i1}^{HL}/\gamma_1^H)] = -\left( \frac{\gamma_1^H \tau_u}{\gamma_1^H \tau_u + 2C^{HL}} \right)^{1/2}, \] (60)

where

\[ C^{HL} = \frac{(b_1^{HL})^2 - 1}{2\gamma_1^H\tau_v}, \] (61)

so that (60) corresponds to (24a). For HS, a similar argument establishes

\[ E[-\exp(\pi_{i1}^{HS}/\gamma_1^H)] = -\left( \frac{\gamma_1^H \tau_u}{\gamma_1^H \tau_u + 2C^{HS}} \right)^{1/2}, \] (62)

where

\[ C^{HS} = \frac{((b_1^{HS})^2 - 1)\text{Var}[p_2|\theta_1] + \text{Cov}[p_2, \theta_1]}{2\gamma}, \] (63)

so that (62) corresponds to (24b). Finally, for second period hedgers, substituting the optimal strategy (39) in the objective function (38) yields

\[ E \left[ -\exp \left\{ -\pi_{i2}^{\gamma}/\gamma_2^H \right\} | \Omega_2^H \right] = -\exp \left\{ -\frac{1}{\gamma_2^H} \left( \frac{\text{Var}[v - p_2|\Omega_2^H](x_2^H)^2 - \text{Var}[v]u_2^2}{2\gamma_2^H} \right) \right\} \]

\[ = -\exp \left\{ -\frac{1}{\gamma_2^H} \left( x_2^H \ u_2 \right) \left( \frac{1}{2\gamma_2^H} \left( \begin{array}{cc} \text{Var}[v - p_2|\Omega_2^H] & 0 \\ 0 & -\text{Var}[v] \end{array} \right) \right) \left( \begin{array}{c} x_2^H \\ u_2 \end{array} \right) \right\}. \] (64)

The argument of the exponential is a quadratic form of the normally distributed random vector

\[ \left( \begin{array}{c} x_2^H \\ u_2 \end{array} \right) \sim N \left( \left( \begin{array}{c} 0 \\ 0 \end{array} \right), \Sigma \right), \]
where

$$
\Sigma \equiv \left( \begin{array}{cc} \text{Var}[x_H^2] & b_H^2 \text{Var}[u_2] \\
\tilde{b}_2^H \text{Var}[u_2] & \text{Var}[u_2] \end{array} \right).
$$

Therefore, we can again apply the usual result to (64), obtaining

$$
E \left[ - \exp \left\{ - \frac{\pi_H^2}{\gamma_H^2} \right\} | \Omega_H^2 \right] = - \left| I + 1/\left(\gamma_H^2\right)^2 \Sigma^{-1} A \right|^{-1/2},
$$

where

$$
A \equiv \frac{1}{2\gamma_2} \left( \begin{array}{cc} \text{Var}[v - p_2|\Omega_2^H] & 0 \\
0 & -\text{Var}[v] \end{array} \right).
$$

Rearranging (65) yields (25).

Consider now traditional dealers. Because of CARA and normality, their conditional expected utility evaluated at the optimal strategy is given by

$$
E[U((v - p_1)x_1^D)|p_1] = - \exp \left\{ - \frac{(E[v|p_1] - p_1)^2}{2\text{Var}[v]} \right\}
= - \exp \left\{ - \frac{\tau_v \Lambda_1^2}{2} \right\}.
$$

Thus, traditional dealers derive utility from the expected, long term capital gain obtained supplying liquidity to first period hedgers. A standard arguments yields

$$
EU^D \equiv E[U((v - p_1)x_1^D)] = - \left( 1 + \frac{\text{Var}[p_1]}{\text{Var}[v]} \right)^{-1/2}
= - \left( \frac{\tau_u}{\tau_u + \tau_v \Lambda_1^2} \right)^{1/2}.
$$

We now turn to the analysis of HFTs welfare. Replacing (45) in (43) and rearranging yields

$$
E[U((p_2 - p_1)x_1^{HF} + (v - p_2)x_2^{HF})|\theta_1] = - \left( \frac{\tau_u}{\tau_u + \tau_v \Lambda_1^2} \right)^{1/2} \times \exp \left\{ - \frac{g(\theta_1)}{\gamma} \right\},
$$

where

$$
g(\theta_1) \equiv \frac{\Gamma_1(\nu - p_1)^2}{2} - \frac{\gamma \tau_v \nu^2}{2} + \gamma \tau_v \nu p_1
= \frac{(\Gamma_1 - \gamma \tau_v)(\nu - p_1)^2 + \gamma \tau_v p_1^2}{2}
= \frac{\gamma}{2} \left( \frac{(E[p_2|p_1] - p_1)^2}{\text{Var}[p_2|p_1]} + \frac{(E[v|p_1] - p_1)^2}{\text{Var}[v]} \right).
$$

Comparing the latter with (66) shows that given their second period utility, at date 1 HFTs derive utility from two sources: the “long term” capital gain due to liquidity supply to first period hedgers and the “short term” capital gain due to the anticipation of $p_2$. We can now
use (44a) and compute

\[ \nu - p_1 = -\frac{\mu\beta\lambda_2 + (\beta + 1)\Lambda_1}{\mu}\theta_1. \]

Replacing the latter, (44a), and the expression for \( p_1 = \Lambda_1\theta_1 \) in \( g(\theta_1) \) we obtain

\[ g(\theta_1) = \frac{\gamma\tau_u}{2} \left( \frac{\text{Var}[E[p_2|p_1] - p_1]}{\text{Var}[p_2|p_1]} + \frac{\text{Var}[p_1]}{\text{Var}[v]} \right) \theta_1^2, \]

where an explicit expression for \( K \) is given by

\[ K = \frac{1}{2} \left( \frac{\gamma\tau_u}{\lambda_2^2} \left( \frac{\beta\mu\lambda_2 + (1 + \mu\gamma\tau_u\beta\lambda_2)\Lambda_1}{\mu} \right)^2 + \gamma\tau_u\Lambda_1^2 \right). \]

Note that

\[ \frac{\tau_u}{\tau_u + \lambda_2^2\tau_v} = \left( 1 + \frac{\text{Var}[p_2|p_1]}{\text{Var}[v]} \right)^{-1}. \]

Therefore, we have

\[ E[\{((p_2 - p_1)x_1^{HF} + (v - p_2)x_2^{HF})|\theta_1\}] = -\left( 1 + \frac{\text{Var}[p_2|p_1]}{\text{Var}[v]} \right)^{-1/2} \times \exp \left\{ -\frac{K}{\gamma} \theta_1^2 \right\}, \]

which is a quadratic form of \( \theta_1 \sim N(0, \tau_u^{-1}) \). We can therefore apply the usual transformation and obtain

\[ EU^{HF} \equiv E[\{((p_2 - p_1)x_1^{HF} + (v - p_2)x_2^{HF})\}] = -\left( 1 + \frac{\text{Var}[p_2|p_1]}{\text{Var}[v]} \right)^{-1/2} \left( 1 + \frac{\text{Var}[p_1]}{\text{Var}[v]} + \frac{\text{Var}[E[p_2|p_1] - p_1]}{\text{Var}[p_2|p_1]} \right)^{-1/2}. \quad (69) \]

Expression (69) shows that HFTs derive utility from two trading activities: (i) liquidity supply to first period hedgers like traditional dealers, and (ii) speculation on short price movements. The higher is the volatility of \( p_2 \) given \( p_1 \), and the lower is the volatility of the payoff, the more rewarding are both trading activities.

\[ \square \]