Working Paper Series

Giovanni Cespa, Xavier Vives

High frequency trading and fragility

Abstract

We show that limited dealer participation in the market, coupled with an informational friction resulting from high frequency trading, can induce demand for liquidity to be upward sloping and strategic complementarities in traders' liquidity consumption decisions: traders demand more liquidity when the market becomes less liquid, which in turn makes the market more illiquid, fostering the initial demand hike. This can generate market instability, where an initial dearth of liquidity degenerates into a liquidity rout (as in a flash crash). While in a transparent market, liquidity is increasing in the proportion of high frequency traders, in an opaque market strategic complementarities can make liquidity U-shaped in this proportion as well as in the degree of transparency.

Keywords: Market fragmentation, high frequency trading, flash crash, asymmetric information.
JEL Classification Numbers: G10, G12, G14

Nontechnical abstract

High Frequency Trading and Fragility

G. Cespa and X. Vives

"Flash events" have become pervasive in financial markets, ranging from stocks to commodities and including Treasury bonds (as for example the October 15, 2014 Treasury Bond crash, where the yield on the benchmark 10-year U.S. government bond dipped 33 basis points to 1.86%, and reversed to 2.13% by the end of the trading day). For US future markets, in the 5-year period from 2010, more than one hundred flash events have been documented.

These episodes seem to have in common the fact that illiquidity brings more illiquidity, contrary to the dynamic that in normal market conditions illiquidity, by moderating the demand for immediacy, produces a self-stabilizing effect. In a crash, in contrast, an increase in illiquidity fosters a disorderly "run for the exit," accentuating the price movement.

In the paper we trace this source of fragility to the fact that modern markets' pervasive presence of automated trading hampers market participants' access to (i) trade, and (ii) quote information. That is, two frictions shape trading activity nowadays: a participation friction (not all liquidity suppliers are continuously present in the market), and an informational friction (not all traders have access to timely market information).

We model these two computerized-trading related frictions in a two-period dynamic noisy rational expectations model. We consider risk-averse dealers who provide liquidity to two cohorts of risk-averse, short-term traders who receive a common endowment shock. Dealers can be "full", with a continuous market presence as stylized high frequency traders (HFTs), or only be present in the first period market. In the first round of trade both dealers' types absorb the (market) orders of the first traders' cohort. In the second trading round, only full dealers participate. Full dealers can therefore accommodate the reverting orders of the first traders' cohort, as well as those of the incoming second cohort who only observe an imperfect signal about the first period order imbalance.

We show that dealers' limited market participation favors the propagation of the first period traders' endowment shock across time, inducing a predictable price pattern. This, in turn, leads second period traders to speculate against the propagated endowment shock. The effect of speculation crucially depends on the degree of transparency of the market. When the market is transparent and there is no informational friction (with second period traders having perfect information about the first period market developments), speculation exerts a stabilizing effect and a higher proportion of full dealers always improves liquidity. When the market is opaque, with second period traders having imprecise information about the first period, speculation can augment first period traders' uncertainty, leading them to demand more immediacy when the market is less liquid. This can offset the rationing impact of illiquidity, and trigger a liquidity feedback loop in which a liquidity dry-up breeds a further, larger liquidity withdrawal. We show that in this scenario there is fragility: multiple equilibria, that can be ranked in terms of liquidity, price volatility, and demand for immediacy, can arise, and a small shock to market parameters can prompt a switch from the high liquidity equilibrium to an

equilibrium with low liquidity; furthermore, when a unique equilibrium arises, we show that a small reduction in the participation of full dealers (HFTs) can generate a large spike in illiquidity. Liquidity can be U-shaped both in full dealer participation and in the degree of transparency of the market. This implies that an increase in dealer market participation or market transparency may be counterproductive for liquidity.

From a methodological point of view, our work shows that fragility can arise in a context where prices are driven by a non-payoff related shock. When trading at high frequencies, the chances that payoff fundamentals drive prices are negligible. Thus our results offer an explanation for how the buildup of a large imbalance can precipitate the market into a crash, which does not rely on the effect of order flow toxicity. From a policy perspective, our paper has two important implications. First, our analysis shows that favoring HFTs' entry (i.e., reducing the participation friction) doesn't necessarily enhance liquidity. Second, the presence of liquidity providers acting with different trading frequencies may make liquidity fragile, either because a shock to parameters can prompt a switch across equilibria; or because a sudden reduction of these dealers' participation can lead to a large liquidity withdrawal. This supports regulatory concerns about the potential drawbacks of automated trading due to operational and transmission risks.

Finally, our analysis has empirical implications: (i) the liquidity impact of HFTs' entry should be assessed taking into account the effect of frictions in the access to market information; and (ii) an exclusive focus on changes in liquidity consumption can miss a potentially important explanatory factor in the changes of HFTs' strategies in the wake of crashes.

"The report describes how on October 15, some algos pulled back by widening their spreads and other reduced the size of their trading interest. Whether such dynamic can further increase volatility in an already volatile period is a question worth asking, but a difficult one to answer." (Remarks Before the Conference on the Evolving Structure of the U.S. Treasury Market (Oct. 21, 2015), Timothy Massad, Chairman, CFTC.)

1 Introduction

Concern for crashes has recently revived, in the wake of the sizeable number of "flash events" that have affected different markets. For futures, in the 5-year period from 2010, more than a 100 flash events have occurred (see Figure 1). For other contracts, the list of events where markets suddenly crash and recover is by now quite extensive.[1]
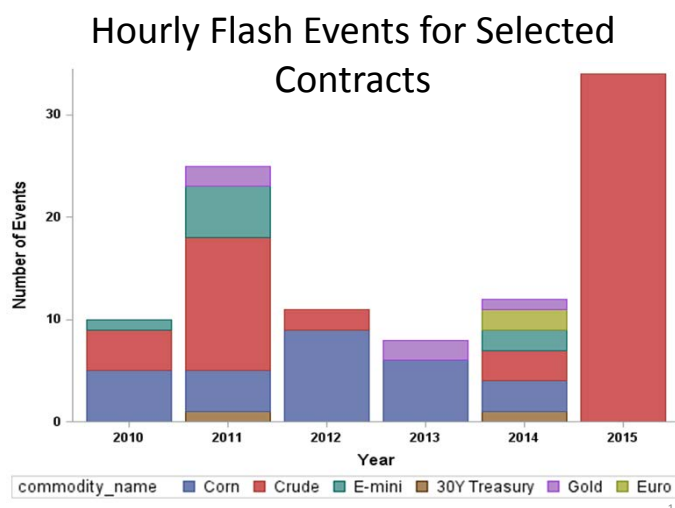

Figure 1: Number of *flash events* in futures contracts from 2010 to 2015. A flash event is an episode in which "the price of a contract moved at least 200 basis points within a trading hour but returned to within 75 basis points of the original or starting price within that same hour." (Source: Remarks Before the Conference on the Evolving Structure of the U.S. Treasury Market (Oct. 21, 2015), Timothy Massad, Chairman, CFTC.).

A common trait of these episodes seems to be the apparent jamming of the "rationing" function of market illiquidity. Indeed, in normal market conditions, traders perceive a lack of liquidity as a cost, which in turn leads them to limit their demand for immediacy (i.e., the demand for liquidity is downward sloping in the illiquidity of the market).[2] This eases the

---

[1]Starting with the May 6, 2010 U.S. "flash-crash" where U.S. equity indices dropped by 5-6% and recovered within half an hour; moving to the October 15, 2014 Treasury Bond crash, where the yield on the benchmark 10-year U.S. government bond, dipped 33 basis points to 1.86% and reversed to 2.13% by the end of the trading day; to end with the August 25, 2015 ETF market freeze, during which more than a fifth of all U.S.-listed exchange traded funds and products were forced to stop trading. More evidence of flash events is provided by NANEX.

[2]To minimize market impact and the associated trading costs they incur (e.g., by using algorithms that parcel out their orders).

pressure on liquidity suppliers, thereby producing a stabilizing effect on the market. However, on occasions a bout of illiquidity can have a destabilizing impact, and foster a disorderly "run for the exit" that is conducive to a rout. In this case traders attempt to place orders *despite* the liquidity shortage, and demand for liquidity may be upward sloping. In such conditions, liquidity is fragile. What can account for such a dualistic feature of market illiquidity?

In this paper, we argue that the fragmentation of liquidity supply *and* the informational frictions induced by computerized trading, are important ingredients in the answer to this question.[3] Indeed, trading automation boosts liquidity supply fragmentation by limiting the market participation of some liquidity suppliers (Duffie (2010) and SEC (2010)); at the same time, computerized trading creates informational frictions by hampering some traders' access to reliable and timely market information (Ding et al. (2014)).[4] Such frictions seem to have a bearing on episodes of sudden liquidity crashes. For example, in their account of the May 10, 2010 "Flash Crash" Easley et al. (2011) state: "This generalized severe mismatch in liquidity was exacerbated by the withdrawal of liquidity by some electronic market makers and by uncertainty about, or delays in, market data affecting the actions of market participants."

We analyze a model in which two classes of risk-averse dealers provide liquidity to two cohorts of risk-averse, short-term traders who receive a common endowment shock, in a two-period market. Traders enter the market to partially hedge their exposure to the risky asset. In the first round of trade both dealers' types absorb the (market) orders of the first traders' cohort. In the second trading round, only one class of dealers, named 'full,' is able to participate. Full dealers, like stylized high frequency traders (HFTs), are continuously in the market and can therefore accommodate the reverting orders of the first traders' cohort, as well as those of the incoming second cohort who observe an imperfect signal about the first period order imbalance.

In a nutshell, the message of the paper is as follows. When all market participants share the same information, the market is stable, and increasing the proportion of dealers with full market participation is good for liquidity. Suppose now that, due to a technological development, trading can be made more frequent but at the cost of introducing an informational friction (with some of the liquidity traders not able to know precisely the state of market imbalances in previous periods). Then the market may be unstable, and increasing the proportion of full dealers, or the degree of market transparency may be bad for liquidity (liquidity can be U-shaped in either variable).

More in detail, we start by showing that dealers' limited market participation favors the propagation of the endowment shock across time, inducing a predictable price pressure. This is because when first period traders load their positions, a part of their orders is absorbed by standard dealers. These agents, however, are not in the market in the second period, when first

---

[3]Automated trading is by now pervasive across different markets. For financial futures, automated trading accounts for about two-thirds of the activity in Eurodollars and Treasury contracts (Source: Keynote Address of CFTC Commissioner J. Christopher Giancarlo before the 2015 ISDA Annual Asia Pacific Conference).

[4]Ding et al. (2014) argue that in the U.S. "...not all market participants have equal access to trade and quote information. Both physical proximity to the exchange and the technology of the trading system contribute to the latency."

period traders unwind. As a consequence, an order imbalance (induced by first period traders' unwinding orders and) affecting the second period price, arises. As standard dealers are unable to rebalance in the second period, they require a larger price concession to absorb traders' orders. Thus, if liquidity dries up, standard dealers absorb a larger imbalance, magnifying the propagation effect. Importantly, second period traders speculate on the induced price pressure, and the effects of such activity depend on the transparency regime governing the market.

In our transparent market benchmark, second period traders have a perfect signal on the first period imbalance, a situation which is likely to arise at low trading frequencies (e.g., intradaily), or in a transparent setup where all market participants have access to the same type of feed, even at high frequencies. In this case, second period traders' speculation has a stabilizing impact on the market as it offsets the propagated price pressure. As a consequence, first period traders' demand for liquidity is a decreasing function of illiquidity (i.e., the compensation that dealers demand to hold the asset inventory in equilibrium): the less liquid is the market, the higher is the cost these traders incur to reduce exposure, and the less aggressive is their liquidity consumption (the lower is their hedging aggressiveness). Furthermore, illiquidity is increasing in traders' hedging aggressiveness (the inverse supply for liquidity is upward sloping). This is because lower aggressiveness limits liquidity consumption, which in turn shrinks dealers' inventory, allowing for cheaper liquidity provision. Thus, illiquidity in this case has a direct, "rationing" effect on traders' liquidity consumption, and a unique equilibrium arises. Furthermore, along this equilibrium, small shocks to the model's parameters have a minimal impact on market liquidity.

In contrast, when access to imbalance information is impaired, second period traders' speculation can boost first period traders' uncertainty, introducing a feedback, liquidity consumption "expanding" effect of illiquidity. This can create a self-sustaining loop that turns the demand for liquidity into an *increasing* function of illiquidity, fostering stronger liquidity consumption, and leading to multiple equilibria. To see this, note that as a higher illiquidity strengthens the endowment shock propagation (because standard dealers intermediate more of the outstanding imbalance), it also heightens second period traders' speculative activity. However, as information on the first period imbalance is noisy, speculation increases the first period uncertainty about the second period price. This can lead first period traders to consume more liquidity (as holding exposure to the asset becomes riskier), and liquidity suppliers to charge more to absorb the order imbalance (as their inventory of the risky asset increases), eventually reinforcing the initial, negative shock to market liquidity.

Equilibrium multiplicity induces three levels of liquidity that can be ranked in an increasing order (low, intermediate, and high). At the low (respectively, intermediate, and high) liquidity equilibrium, volatility and liquidity consumption are high (respectively, intermediate, and low). Thus, our paper highlights a channel through which the *combined* effect of a heightened demand for liquidity, and a reduced liquidity provision conjure to increase market volatility.

The liquidity consumption ranking across equilibria is a further manifestation of the fact that opaqueness jams the direct, rationing effect of illiquidity, while it strengthens its feedback,

liquidity consumption enhancing effect. The end result is that traders' demand for liquidity peaks at the equilibrium where the cost of trading is at its highest, consistently with the pattern observed in many crash events. Importantly, we also find that along a unique equilibrium with market opaqueness, illiquidity can be hump-shaped in the proportion of fast dealers, or in the degree of market transparency.

The strategic complementarity loop arising with market opaqueness implies that liquidity can be "fragile" in our setup. We show this with two types of examples. In the first one, we exploit equilibrium multiplicity and illustrate how a small shock to some parameter values can produce a switch from the high liquidity equilibrium to an equilibrium with low liquidity. In particular, we focus on the consequence of a shock that disconnects a small mass of full dealers from the market (a technological 'glitch'). We then analyze the effect of a positive shock to the volatility of first and second period traders' demand. These are meant to capture, respectively, an increase in the probability of a large order hitting the first period market (which is consistent with some narratives of the flash crash, see e.g. Easley et al. (2011)), and an increase in the uncertainty first period traders face on their endowment value. In all these examples *small* parameter shocks produce *large* liquidity withdrawals.

In the second type of example we review the impact of the glitch, but in this case taking account of the result that along a unique equilibrium with market opaqueness, illiquidity can be hump-shaped in the proportion of fast dealers. Based on this finding, we show that a high level of liquidity can suddenly evaporate because of a reduction in full dealers' participation *along the same* equilibrium. Furthermore, the evaporation of liquidity is related to a large increase in volatility. This provides an insight into the question of this paper's opening quotation.

It is also the case that illiquidity can be hump-shaped in the degree of market transparency. The reason is that first period illiquidity is positively associated to the return uncertainty faced by first period traders. For low transparency, a more informative signal for second period traders makes the market less liquid, as those traders speculate more aggressively on the propagated imbalance, increasing first period traders' uncertainty. However, as second period traders' signal precision increases, these traders' speculation increasingly reduces the propagated imbalance, lowering first period traders' uncertainty.

This paper is related to four strands of the literature. First, equilibrium multiplicity, complementarities, and liquidity fragility are phenomena known to obtain in economies where asset prices are driven by fundamentals information and noise trading (see, e.g., Cespa and Foucault (2014), Cespa and Vives (2015), Goldstein et al. (2014), and Goldstein and Yang (2015)). In this setup, in contrast, asset prices are exclusively driven by non-fundamentals information. However, the demand of all traders is responsive to the volatility of the price at which agents unwind their positions. In turn, such volatility depends on traders' demand. It is precisely this two-sided loop—which in a noise traders' economy cannot possibly arise—that is responsible for our multiplicity result. Other authors obtain multiple equilibria in setups where order flows are driven by only one type of shock (see, e.g., Spiegel (1998)). However, multiplicity there arises from the bootstrap nature of expectations in the steady-state equilibrium of an overlap-

ping generations (OLG) model in which investors live for two periods. Our setup, in contrast, considers an economy with a finite number of trading rounds.

Second, the paper adds to the theoretical literature on the impact of high frequency trading (HFT) on market performance, by showing that an informational friction arising from liquidity provision fragmentation can be responsible for liquidity fragility, and reverses the common wisdom that associates an increase in computerized trading with more liquid markets. Differently from our setup, the HFT literature has mostly concentrated on modeling risk neutral agents (e.g., Budish et al. (2015), Hoffmann (2014), Du and Zhu (2014), Bongaerts and Van Achter (2015), Foucault et al. (2015), and Menkveld and Zoican (2015); see O'Hara (2015) and Menkveld (2016) for literature surveys).[5] Easley et al. (2011, 2012), find that in the hours preceding the flash crash, signed order imbalance for the E-mini S&P500 futures contract was unusually high. They interpret this evidence as supportive of a high order flow "toxicity," which led HFTs to flee the market, eventually precipitating the crash. As argued above, our model also predicts that large imbalances can lead to a huge liquidity withdrawal. However, the channel we highlight is not related to adverse selection, but emphasizes the multiplier effect of illiquidity on the demand for immediacy that can arise when some traders have access to opaque information on imbalances. Menkveld and Yueshen (2012) argue that market *spatial fragmentation* can be detrimental to stability. In their model, HFTs have access to a private reselling opportunity which, due to impaired intermarket connectivity, can break down. When this happens, HFTs trade among themselves, providing an 'illusion' of liquidity to traders who observe volume, which in turn fosters further liquidity demand. Our focus is on the liquidity provision fragmentation induced by an informational friction in a single, *concentrated* market, a feature that is consistent with the futures markets flash events discussed above. Finally, Han et al. (2014), in a Glosten and Milgrom (1985) setup with HFT and low frequency market makers, also find that illiquidity is hump-shaped in the proportion of HFT. However, in their model this result arises from an adverse selection problem that HFT's ability to cancel quotes upon the arrival of a negative signal creates for low frequency dealers. In our model, instead, dealers face no adverse selection risk, and hump-shaped illiquidity arises because of strategic complementarities between traders' liquidity consumption decisions.

Third, the paper relates to the literature that assesses the impact of limited market participation. Heston et al. (2010) and Bogousslavsky (2014) find that some liquidity providers' limited market participation can have implications for return predictability. Chien et al. (2012) focus instead on the time-series properties of risk premium volatility. Hendershott et al. (2014) concentrate on the effect of limited market participation for price departures from semi-strong efficiency. Our focus is, instead, on the destabilizing dynamics that is generated by bouts of illiquidity. In this respect, our paper is also related to Huang and Wang (2009) who show that with costly market participation, idiosyncratic endowment shocks can yield crashes. Note,

---

[5]Biais et al. (2015) study the welfare implications of investment in the acquisition of HFT technology. In their model HFTs have a superior ability to match orders, and possess superior information compared to human (slow) traders. They find excessive incentives to invest in HFT technology, which, in view of the negative externality generated by HFT, can be welfare reducing.

however, that in our setup traders are exposed to the same shock, which yields a different mechanism for market instability.

Fourth, by highlighting the first order asset pricing impact of uninformed traders' imbalance predictability, this paper shares some features of our previous work (Cespa and Vives (2012), and Cespa and Vives (2015)). In that setup, however, predictability obtained because of the assumed statistical properties of noise traders' demands, whereas in this paper it arises endogenously, because of a participation friction. A growing literature investigates the asset pricing implications of noise trading predictability. Collin-Dufresne and Vos (2015) argue that informed traders time their entry to the presence of noise traders in the market. This, in turn, implies that standard measures of liquidity (e.g., Kyle's lambda), may fail to pick up the presence of such traders. Peress and Schmidt (2015) estimate the statistical properties of a noise trading process, finding support for the presence of serial correlation in demand shocks.

The rest of the paper is organized as follows. In the next section we introduce the model, and show that with limited market participation, endowment shocks propagate across trading dates. Next, we analyze the benchmark with a transparent market. We then illustrate how the presence of an informational friction can generate strategic complementarities between traders' liquidity consumption decisions. We show that such complementarities are at the root of the loop responsible for equilibrium multiplicity and liquidity fragility. A final section contains concluding remarks. All proofs are in the appendix.

2   The model

A single risky asset with liquidation value $v \sim N(0, \tau_v^{-1})$, and a risk-less asset with unit return are exchanged in a market during two trading rounds. Three classes of traders are in the market. First, a continuum of competitive, risk-averse, High Frequency Traders (which we refer to as "Full Dealers" and denote by FD) in the interval $(0, \mu)$, are active at both dates. Second, competitive, risk-averse dealers (D) in the interval $[\mu, 1]$, are active only in the first period. Finally, a unit mass of short-term traders enters the market at date 1. At date 2, these traders unwind their position, and are replaced by a new cohort of short-term traders (of unit mass). The asset is liquidated at date 3. We now illustrate the preferences and orders of the different players.

2.1   Liquidity providers

A FD has CARA preferences (we denote by γ his risk-tolerance coefficient) and submits price-contingent orders x_t^{FD}, $t = 1, 2$, to maximize the expected utility of his final wealth: $W^{FD} = (v - p_2)x_2^{FD} + (p_2 - p_1)x_1^{FD}$.[6] A Dealer also has CARA preferences with risk-tolerance γ, but is in the market only in the first period. He thus submits a price-contingent order x_1^D to maximize the expected utility of his wealth $W^D = (v - p_1)x_1^D$. The inability of D to trade in the second

---

[6] We assume, without loss of generality with CARA preferences, that the non-random endowment of FDs and dealers is zero. Also, as equilibrium strategies will be symmetric, we drop the subindex i.

period captures some liquidity suppliers' limited market participation. This friction could be due to technological reasons (as, e.g. in the case of standard dealers with impaired access to a technology that allows trading at high frequencies).

2.2 Short-term traders

In the first period a unit mass of short-term traders is in the market. A short-term trader receives a random endowment of a risky asset u_1 whose payoff is perfectly correlated with the one of the asset traded in the market, and posts a market order x_1^L anticipating that it will unwind its holdings in the following period, and leave the market. We assume $u_1 \sim N(0, \tau_{u_1}^{-1})$, and $\text{Cov}[u_1, v] = 0$.[7] First period traders have identical CARA preferences (we denote by γ_1^L the common risk-tolerance coefficient). Formally, a trader maximizes the expected utility of his wealth $\pi_1^L = u_1 p_2 + (p_2 - p_1) x_1^L$:

$$E\left[-\exp\{-\pi_1^L/\gamma_1^L\}|\Omega_1^L\right],$$

where Ω_1^L denotes his information set. In period 2, first period traders are replaced by a new (unit) mass of traders receiving a random endowment of the same risky asset as their previous period peers u_2, where $u_2 \sim N(0, \tau_{u_2}^{-1})$ and $\text{Cov}[u_2, v] = \text{Cov}[u_2, u_1] = 0$. A second period trader has CARA utility function with risk-tolerance γ_2^L, and submits a market order to maximize the expected utility of his wealth $\pi_2^L = u_2 v + (v - p_2) x_2^L$:

$$E\left[-\exp\{-\pi_2^L/\gamma_2^L\}|\Omega_2^L\right],$$

where Ω_2^L denotes his information set.[8] We can interpret the second period traders as the proprietary desk of investment banks that trade to hedge their exposure to an asset whose payoff is perfectly correlated with the one of the asset traded in the market.

2.3 Information sets

We restrict attention to linear equilibria and conjecture that at equilibrium a period 1 trader submits an order $x_1^L = b_1^L u_1$, where b_1^L is the portion of the endowment shock a trader hedges, to be determined in equilibrium, while a FD and a dealer respectively post a limit order $x_1^{FD} = \varphi_1^{FD}(p_1)$, $x_1^D = \varphi_1^D(p_1)$ where $\varphi_1^{FD}(\cdot)$, $\varphi_1^D(\cdot)$ are linear functions of p_1. In the second period, a FD submits a limit order $x_2^{FD} = \varphi_2(p_1, p_2)$, where $\varphi_2(\cdot)$ is a linear function of prices. A second period trader observes a signal of the first period endowment shock $s_{u_1} = u_1 + \eta$, with $\eta \sim N(0, \tau_\eta^{-1})$, and independent from all the other random variables in the model, and submits a market order $x_2^L = b_{21}^L u_2 + b_{22}^L s_{u_1}$, where b_{21}^L and b_{22}^L are respectively the fraction of

---

[7] The assumption of a random endowment in the risky asset is akin to Huang and Wang (2009), and Vayanos and Wang (2012) who instead posit that traders receive an endowment in a consumption good that is perfectly correlated with the value of the risky asset at the terminal date.

[8] Our results are robust to the case in which the first period market is populated by a mass β of short-term traders, that unwind at date 2, and a mass $(1 - \beta)$ of long-term ones that hold their position until liquidation.

the endowment shock hedged by second period traders, and the response to the second period signal. With these assumptions, we obtain

Lemma 1. *At equilibrium, p_1 is observationally equivalent to u_1, and the sequence $\{p_1, p_2\}$ is observationally equivalent to $\{u_1, x_2^L\}$.*

A first period trader observes the endowment shock u_1. Therefore, his information set coincides with the one of Ds and FDs: $\Omega_1^L = \Omega_1^{FD} = \Omega_1^D = \{u_1\}$. A second period trader receives an endowment shock u_2, and can observe a signal s_{u_1}. Thus, his information set is $\Omega_2^L = \{u_2, s_{u_1}\}$. Finally, a FD in period 2 observes the sequence of prices: $\Omega_2^{FD} = \{p_1, p_2\}$ from which he retrieves $\{u_1, x_2^L\}$.

Thus, according to our model, liquidity provision is fragmented because (i) only one class of dealers is able to participate in the second period and (ii) some traders (the second cohort of short-term traders) have access to opaque information on the first period price. This assumption is consistent with the evidence that exchanges sell fuller access to their matching engine, as well as direct feeds of their market information at a premium (see, e.g., O'Hara (2015)).[9] Figure 2 displays the timeline of the model.

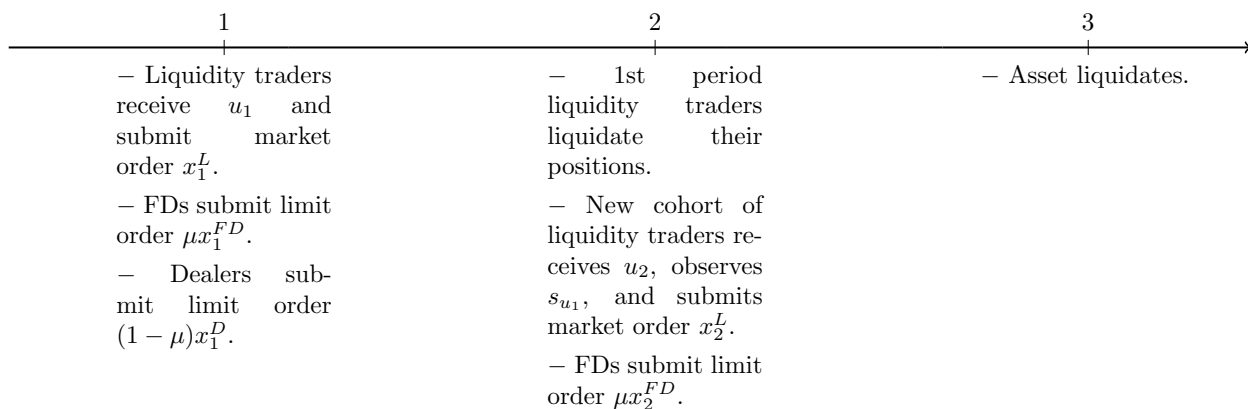| 1	2	3
− Liquidity traders receive u_1 and submit market order x_1^L.	− 1st period liquidity traders liquidate their positions.	− Asset liquidates.
− FDs submit limit order μx_1^{FD}.	− New cohort of liquidity traders receives u_2, observes s_{u_1}, and submits market order x_2^L.	
− Dealers submit limit order $(1-\mu)x_1^D$.	− FDs submit limit order μx_2^{FD}.	

Figure 2: The timeline.

2.4 Limited market participation and the propagation of endowment shocks

Due to limited market participation, the first period endowment shock propagates to the second trading round, thereby affecting p_2. To see this, consider the first period market clearing equation

$$\mu x_1^{FD} + (1-\mu)x_1^D + x_1^L = 0. \tag{1}$$

---

[9] This assumption is also similar to Foucault et al. (2015) who posit that HFTs receive market information slightly ahead of the rest of the market. Ding et al. (2014) compare the NBBO (National Best Bid and Offer, which is the price feed computed by the Security Industry Processors in the US) to the fuller feeds market participants obtain via a direct access to different trading platforms. Their findings point to sizeable price differences that can yield substantial profits to HFTs. Latency in the reporting of market data can also be profitably exploited for securities with centralized trading, see "High-speed traders exploit loophole," *Wall Street Journal*, May 1, 2013.

At equilibrium the orders of first period traders are absorbed by both FDs and Ds. Thus, when $\mu < 1$, FDs' aggregate position falls short of x_1^L: $\mu x_1^{FD} + x_1^L \neq 0$. As a consequence, the inventory FDs carry over from the first period is insufficient to absorb the reverting orders that first period traders post in period 2. This creates an order imbalance driven by the first period endowment shock u_1 that adds to the one originating from second period trades, and affects the second period price. Formally, from the second period market clearing equation we have

$$\mu(x_2^{FD} - x_1^{FD}) + (x_2^L - x_1^L) = 0.$$

Substituting (1) in the latter and rearranging yields:

$$\mu x_2^{FD} + x_2^L + (1 - \mu)x_1^D = 0. \tag{2}$$

According to Lemma 1, at equilibrium x_1^D depends on u_1. Thus, when $\mu < 1$, p_2 also reflects the first period endowment shock.

3 Market transparency and the rationing effect of illiquidity

In this section, we assume that second period traders have a perfect signal on the first period endowment shock: $\tau_\eta \to \infty$. This captures a scenario in which information on the first period imbalance is public, as is the case in a low frequency trade environment (e.g., intradaily). Alternatively, it represents a setup in which second period traders have access to the same information as FDs, or the same technology that parses demand information from prices. In this case, we obtain the following result:

Proposition 1. *When the market is transparent there exists a unique equilibrium in linear strategies, where* $x_1^D = -\gamma\tau_v p_1$, $x_1^{FD} = -(\gamma/\gamma_1^L \Lambda_1^*)(1+b_1^L)p_1 - \gamma\tau_v p_1$, $x_2^{FD} = -\gamma\tau_v p_2$, $x_1^L = b_1^L u_1$, $x_2^L = b_{21}^L u_2 + b_{22}^L u_1$,

$$p_2 = -\Lambda_2^*\left(u_2 - (1 - \mu)\gamma\tau_v\Lambda_1^* u_1\right) \tag{3a}$$

$$p_1 = -\Lambda_1^* u_1, \tag{3b}$$

$$\Lambda_2^* = -\frac{1}{\mu\gamma\tau_v}b_{21}^L > 0 \tag{4a}$$

$$\Lambda_1^* = \frac{1}{\gamma\tau_v}\left(1 - \frac{(\mu\gamma + \gamma_1^L)(1 + b_1^L)}{\gamma_1^L}\right) > 0, \tag{4b}$$

$$b_1^L = \gamma_1^L \frac{\text{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1^*}{\text{Var}[p_2|u_1]} - 1 \in \left(-1, -\frac{\mu\gamma}{\mu\gamma + \gamma_1^L}\right) \tag{5a}$$

$$b_{21}^L = \gamma_2^L \frac{\Lambda_2^*}{\text{Var}[v - p_2|\Omega_2^L]} - 1 \in (-1, 0) \tag{5b}$$

$$b_{22}^L = -(1 + b_{21}^L)(1 - \mu)\gamma\tau_v\Lambda_1^* < 0, \tag{5c}$$

$$\text{Cov}[p_2, u_1] = \frac{(1 - \mu)\gamma\Lambda_1^*}{\tau_{u_1}(\mu\gamma + \gamma_2^L)} \tag{6a}$$

$$\text{Var}[p_2|u_1] = \frac{1}{(\mu\gamma + \gamma_2^L)^2\tau_v^2\tau_{u_2}}, \tag{6b}$$

$\text{Var}[v - p_2|\Omega_2^L] = 1/\tau_v$.

The coefficient Λ_t^* in (3a) and (3b), i.e. the period t endowment shock's negative price impact, is our measure of illiquidity:

$$\Lambda_t^* = -\frac{\partial p_t}{\partial u_t}. \tag{7}$$

In the second period, FDs supply liquidity by posting a price contingent order, with an aggressiveness that is inversely proportional to their risk aversion and the risk of the asset payoff. Other things equal, a lower liquidity increases the size of their position. A similar behavior is displayed by standard dealers in the first period. The first period strategy of a FD has two components:

$$X_1^{FD}(p_1) = \underbrace{-\gamma\frac{1 + b_1^L}{\gamma_1^L\Lambda_1^*}p_1}_{\text{Speculation}} - \underbrace{\gamma\tau_v p_1}_{\text{Market making}}, \tag{8}$$

where $p_1 = -\Lambda_1^*u_1$, and it is worth remarking that $1 + b_1^L > 0$. According to the above expression, a FD provides liquidity in *two* distinct ways. First, for given u_1 he speculates on short-term returns, buying in the face of a price drop (i.e., when $u_1 > 0$), and selling otherwise. This is because he anticipates the future price impact of first period liquidity traders' reversion.[10] In this respect, FDs supply liquidity by placing "directional bets," consistently with the literature linking liquidity supply to sophisticated traders' contrarian behavior (Nagel (2012), Brogaard et al. (2014), and Biais et al. (2015)). Additionally, a FD places a price contingent order to absorb the residual imbalance, like a standard dealer.

According to (5a) and (5b), first and second period traders demand liquidity to hedge a fraction of their endowment shock. Other things equal, the more liquid is the market, the more volatile is the short term return, and the less risk tolerant traders are, the higher is the hedging "aggressiveness" ($|b_1^L|$ and $|b_{21}^L|$) for traders at both dates. Additionally, first period traders hedge more aggressively when the second period price is less positively associated with the first period endowment shock, and when the second period endowment shock is more dispersed.

---

[10]This is consistent with Hirschey (2016) who finds that HFTs trade ahead of other investors' order flow.

For the first effect, note that holding a fraction of the endowment shock, allows first period traders to speculate on the future price impact of their reversion. Since a lower $\text{Cov}[p_2, u_1]$ reduces the predictable impact of the first period endowment shock on p_2 (see (6a)), it also lowers first period traders' return from speculation. For the second effect, note that second period traders' hedging activity creates price volatility which heightens first period traders' uncertainty (see (6b)), leading to a higher liquidity demand in the first period.

According to (5c), second period traders also speculate on the propagated order imbalance by putting a negative weight on their signal ($b_{22}^L < 0$), which is decreasing in the first period illiquidity Λ_1^*, that is, the "speculative aggressiveness" $|b_{22}^L|$ increases in first period illiquidity. This is because, for $u_1 > 0$, first period traders' reversion has a positive impact on p_2, which prompts second period traders to short the asset. A less liquid first period market increases the position held by standard dealers, strengthening the positive dependence between p_2, and u_1 (see (6a)), and leading second period traders to step up their speculative aggressiveness.

Using equation (8) to compute FDs' aggregate position shows that first period traders' liquidity demand and FDs' aggregate directional bets are negatively related:

$$\mu X_1^{FD}(p_1) = \underbrace{-\mu\gamma\frac{1 + b_1^L}{\gamma_1^L \Lambda_1^*}p_1}_{\text{Speculation}} - \underbrace{\mu\gamma\tau_v p_1}_{\text{Market making}} . \tag{9}$$

Intuitively, for given endowment shock, first period traders scale up their hedging when facing higher uncertainty over p_2 ($|b_1^L|$ increases, see (5a)). But these conditions also increase the risk implied by FDs' speculation, and discourage these traders from placing directional bets.

As is standard in economies with risk-averse liquidity suppliers, Λ_t^* reflects dealers' risk-related compensation to absorb the outstanding imbalance in their inventory: the cost of supplying liquidity. However, differently from a noise trader economy, in this model dealers' inventory depends on the equilibrium trading decisions of FDs *and* liquidity traders. For Λ_2^* (see (4a)), this is immediate, since b_{21}^L measures the fraction of the second period endowment shock that traders hedge (see (5b)), the risk of the asset payoff is $\text{Var}[v] = 1/\tau_v$, and second period dealers' aggregate risk bearing capacity is given by $\mu\gamma$.[11]

For Λ_1^* (see (4b)), the argument is as follows. In view of (5a), at equilibrium first period traders hold a fraction

$$1 + b_1^L = \gamma_1^L \frac{\text{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1^*}{\text{Var}[p_2|u_1]}, \tag{10}$$

of their endowment shock. At the same time, according to (9), FDs aggregate speculative position per unit of endowment shock is given by

$$\mu\gamma\frac{1 + b_1^L}{\gamma_1^L}. \tag{11}$$

Thus, summing (10) and (11) yields the total speculative exposure of FDs and first period

---

[11]Limited market participation implies that only a proportion μ of FDs is in the market at date 2.

traders per unit of u_1, i.e., the fraction of the endowment shock that does not dent liquidity suppliers' "inventory bearing" capacity:

$$1 + b_1^L + \mu\gamma\frac{1 + b_1^L}{\gamma_1^L} = \frac{(\mu\gamma + \gamma_1^L)(1 + b_1^L)}{\gamma_1^L}, \tag{12}$$

while the complement to one of (12) captures dealers' inventory (per unit of endowment shock):

$$\text{Dealer's inventory per unit of endowment shock} = 1 - \frac{(\mu\gamma + \gamma_1^L)(1 + b_1^L)}{\gamma_1^L}. \tag{13}$$

At date 1, FDs know that they will be able to unwind their inventory in the second trading round, when x_1^L reverts. However, at that point in time, a new generation of traders enters the market. These traders hedge a new endowment shock, exposing FDs to the risk of holding their initial inventory until the liquidation date. Thus, for given inventory (13), the riskier is the asset, and the more risk averse FDs are, the higher is the risk borne by liquidity suppliers, and, according to (4b), the less liquid is the market.

When second period traders' endowment shock is null, liquidity suppliers' risk vanishes, b_1^L reaches its upper bound, and the market is infinitely liquid:

Corollary 1. *In a transparent market, when the second period endowment shock is null ($\tau_{u_2} \to \infty$), first period traders' liquidity demand matches FDs' relative risk-bearing capacity, and the market is infinitely liquid ($b_1^L \to -(\mu\gamma + \gamma_1^L)^{-1}\mu\gamma$, and $\Lambda_1^* \to 0$).*

The equilibrium (Λ_1^*, b_1^L) obtains as the unique solution to the system (4b)–(5a):

$$\Lambda_1 = \frac{1}{\gamma\tau_v}\left(1 - \frac{(\mu\gamma + \gamma_1^L)(1 + b_1^L)}{\gamma_1^L}\right) \tag{14a}$$

$$b_1^L = \gamma_1^L(\gamma + \gamma_2^L)(\mu\gamma + \gamma_2^L)\tau_v^2\tau_{u_2}\Lambda_1 - 1, \tag{14b}$$

where (14b) obtains by replacing (6a) and (6b) in (5a) (and evaluating the resulting expression at Λ_1). Such solution can be understood as the intersection between the "inverse supply" and "demand" of liquidity (respectively, (14a) and (14b)). This is so because b_1^L measures the fraction of the endowment shock that first period traders hedge in the market, while Λ_1 captures the price adjustment dealers require to accommodate the order imbalance. A less liquid first period market increases the cost of scaling down traders' exposure, and leads the latter to hedge less of their endowment. Thus, in this case a drop in liquidity has a "rationing" effect on liquidity consumption, and the demand for liquidity is a *decreasing* function of Λ_1.[12] Conversely, a lower hedging aggressiveness implies a larger speculative position for FDs, which shrinks the imbalance that liquidity suppliers have to clear in the first period, and leads to a more liquid market. Hence, the (inverse) supply of liquidity is *decreasing* in b_1^L. Figure 3, Panel (a), provides a graphical illustration of the equilibrium determination. In Panel (b) we

---

[12] As $b_1^L < 0$, and positively sloped in Λ_1, a higher illiquidity implies that traders shed a lower fraction of their endowment, or that their liquidity demand subsides.

plot FDs' aggregate directional bets per unit of endowment shock, as a function of Λ_1 which, consistently with (11), are increasing in Λ_1. Thus, in the symmetric information benchmark, an increase in illiquidity increases the expected return from speculation, without affecting its implied risk (see (6b)).
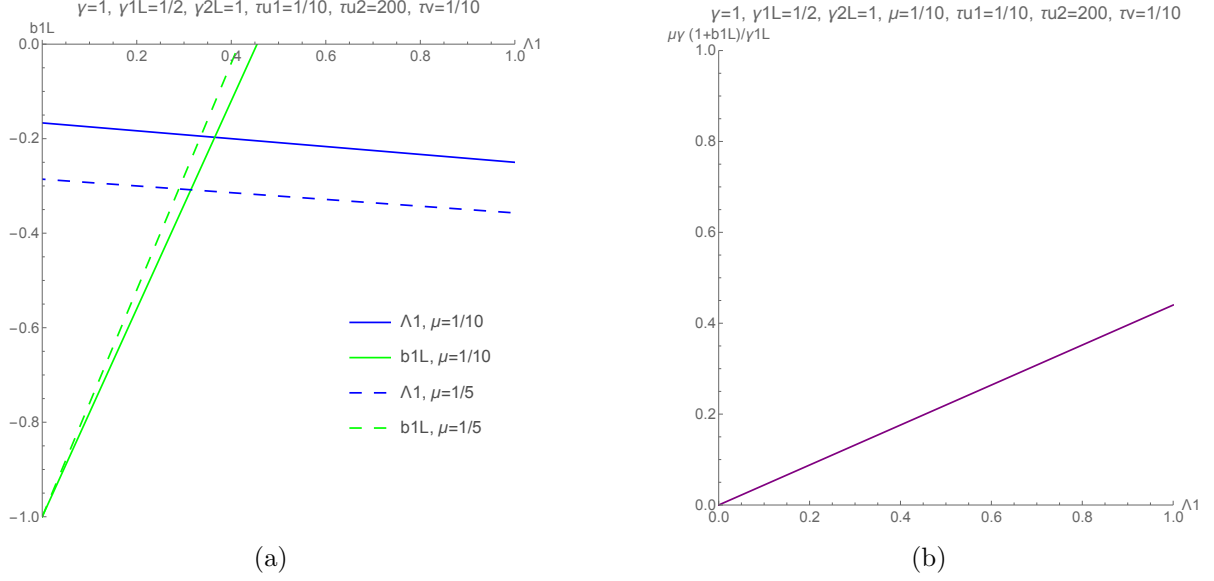

(a)

(b)

Figure 3: Transparency and equilibrium uniqueness. For Panel (a): the green (blue) curve represents first period traders' liquidity demand (inverse liquidity supply); the solid (dashed) curves are drawn assuming $\mu = 1/10$ ($\mu = 1/5$). When $\mu = 1/10$, $\{\Lambda_1^*, b_1^L\} = \{.4, -.2\}$, while when $\mu = 1/5$, $\{\Lambda_1^*, b_1^L\} = \{.3, -.3\}$. For Panel (b): the purple curve represents FDs aggregate directional bets.

In Panel (a) of Figure 3 we also graphically analyze the effect of an increase in the mass of FDs on Λ_1^*. The solid (dashed) curves in the figure are drawn for $\mu = 1/10$ ($\mu = 1/5$). A larger μ has a positive effect on the cost of trading for all levels of b_1^L, since, according to (11), the aggregate speculative position of FDs increases, lowering dealers' inventory. As a result, when μ increases, the new function Λ_1 shifts downwards. Consider now b_1^L. Based on (10), a larger μ has two contrasting effects on first period traders' hedging aggressiveness: on the one hand, as implied by (6b), first period return uncertainty is decreasing in μ. Therefore, a larger μ lowers first period traders' uncertainty about p_2, and makes them consume less liquidity. However, according to (6a),

$$\frac{\partial \mathrm{Cov}[p_2, u_1]}{\partial \mu} < 0 \tag{15}$$

and a higher μ lowers the positive association between the second period price and the first period endowment shock, making speculation less profitable. This pushes first period traders to shed a larger fraction of their endowment, increasing dealers' inventory, and consuming more liquidity. When the market is transparent, this latter effect is never strong enough to offset the former two and we obtain:

Corollary 2. *In a transparent market, liquidity increases in the proportion of fast dealers* $(\partial \Lambda_1^* / \partial \mu < 0)$.

We concentrate our analysis on the liquidity of the first period market. However, note that as the volatility of the first period price is given by

$$\text{Var}[p_1] = (\Lambda_1^*)^2 \tau_{u_1}^{-1},$$

our liquidity results can also be interpreted in terms of price volatility.

4 Opaqueness and the feedback effect of illiquidity

Suppose now that second period traders' signal on u_1 has a bounded precision ($\tau_\eta < \infty$). This setup characterizes a scenario where some traders (FDs, in our setup) have access to better market information (for example on order imbalances) compared to others (the second cohort of traders), and given our previous discussion, is likely to hold at a high trading frequency. Alternatively, this assumption captures a situation where all traders can observe past prices in real time but in which only FDs have the ability to exactly parse u_1 out of p_1. In this case, we obtain the following result:

Proposition 2. *When $0 < \tau_\eta < \infty$, an equilibrium exists. At equilibrium: first period strategies and p_1 are as in Proposition 1, $x_2^L = b_{21}^L u_2 + b_{22}^L s_{u_1}$,*

$$p_2 = -\Lambda_2^* u_2 - \frac{1-\mu}{\mu} \Lambda_1^* \left(\left(1 + b_{21}^L \kappa\right) \beta_{u_1|s_{u_1}} \eta + \left(\left(1 + b_{21}^L \kappa\right) \beta_{u_1|s_{u_1}} - 1 \right) u_1 \right), \qquad (16)$$

$$b_{21}^L = \gamma_2^L \frac{\Lambda_2^*}{\text{Var}[v - p_2|\Omega_2^L]} - \frac{1}{\kappa} \in (-1, 0) \qquad (17a)$$

$$b_{22}^L = -(1-\mu)\gamma\tau_v\Lambda_1^*\beta_{u_1|s_{u_1}} \left(1 + b_{21}^L \kappa\right) < 0, \qquad (17b)$$

where

$$\kappa \equiv \tau_v \text{Var}[v - p_2|\Omega_2^L] > 1 \qquad (18a)$$

$$\beta_{u_1|s_{u_1}} = \frac{\text{Cov}[s_{u_1}, u_1]}{\text{Var}[s_{u_1}]}, \qquad (18b)$$

the first and second period return uncertainty are respectively given by $\text{Var}[p_2|u_1] = \lambda_2^2((b_{21}^L)^2/\tau_{u_2} + (b_{22}^L)^2/\tau_\eta)$, $\text{Var}[v - p_2|\Omega_2^L] = \text{Var}[v] + (\lambda_2(1-\mu)\gamma\tau_v\Lambda_1^*)^2\text{Var}[u_1|s_{u_1}]$, *and the covariance between the second period price and the first period endowment shock is given by*

$$\text{Cov}[p_2, u_1] = -\frac{1-\mu}{\mu}\Lambda_1^* \left(\left(1 + b_{21}^L \kappa\right) \beta_{u_1|s_{u_1}} - 1 \right) \tau_{u_1}^{-1} > 0. \qquad (19)$$

Differently from the transparent market benchmark, the second period price now loads also on the error term η (see (16)), and second period traders face uncertainty on the price at which their order is executed, besides that on the liquidation value (this additional source of

uncertainty is captured by the coefficient κ, see (18a)). As a consequence, they hedge a lower fraction of their endowment shock:

Corollary 3. *When the market is opaque, second period traders hedge a lower portion of their endowment shock compared to the transparent market case:*

$$|b_{21}^L|_{\tau_\eta < \infty} < |b_{21}^L|_{\tau_\eta \to \infty}. \tag{20}$$

The next result shows that, as in the transparent market benchmark, a less liquid first period market strengthens the propagation of the first period imbalance to the second trading date, increasing the positive association between p_2 and u_1. Additionally, because of opaqueness, a lower first period liquidity also heightens the return uncertainty of second period traders, lowering their hedging aggressiveness:

Corollary 4. *At equilibrium, the impact of the first period endowment shock on the second period price, second period traders' return uncertainty are increasing in illiquidity, while hedging aggressiveness decreases in illiquidity:*

$$\frac{\partial \mathrm{Cov}[p_2, u_1]}{\partial \Lambda_1} > 0, \ \frac{\partial \mathrm{Var}[v - p_2 | \Omega_2^L]}{\partial \Lambda_1} > 0, \ \frac{\partial |b_{21}^L|}{\partial \Lambda_1} < 0. \tag{21}$$

Thus, first period illiquidity impacts second period traders' strategies (see (21) and (17a), (17b)). In turn, second period traders' strategies impact the return uncertainty faced by first period traders ($\mathrm{Var}[p_2|u_1]$—see Proposition 2), and thus these agents' liquidity demand. Hence, while the equilibrium still obtains as a solution of the system (4b)–(5a), b_1^L becomes a non-linear function of Λ_1. More in detail, substituting the expressions for $\mathrm{Cov}[p_2, u_1]$ and $\mathrm{Var}[p_2|u_1]$ that appear in Proposition 2 in (5a), and rearranging, yields:

$$b_1^L = \gamma_1^L \tau_{u_2} \tau_\eta \frac{\lambda_2(b_{22}^L + (1-\mu)\gamma\tau_v\Lambda_1) + \Lambda_1}{\lambda_2^2((b_{21}^L)^2\tau_\eta + (b_{22}^L)^2\tau_{u_2})} - 1. \tag{22}$$

To assess the effect of a shock to first period illiquidity on the first period liquidity demand, we need to pin down how b_{21}^L and $\mathrm{Var}[p_2|u_1]$ react to an increase in Λ_1. First, a higher illiquidity has two contrasting effects on the speculative aggressiveness of second period traders ($|b_{22}^L|$). Substituting (17a) in (17b), rearranging, and considering the absolute value of b_{22}^L, yields

$$|b_{22}^L| = \frac{\gamma_2^L(1-\mu)\tau_v\beta_{u_1|s_{u_1}}}{\mu}\Lambda_1^*|b_{21}^L|.$$

Differentiating the above expression, and determining its sign:

$$\mathrm{sign}\left(\frac{\partial |b_{22}^L|}{\partial \Lambda_1}\right) = \mathrm{sign}\left(\underbrace{|b_{21}^L|}_{\text{Speculation effect } (+)} + \underbrace{\Lambda_1 \frac{|\partial b_{21}^L|}{\partial \Lambda_1}}_{\text{Uncertainty effect } (-)}\right). \tag{23}$$

On the one hand, like in the transparent market benchmark, an increase in Λ_1 augments second period traders' speculative opportunities, and drives them to trade more against the u_1-led imbalance (the first term in the parenthesis in (23)). On the other hand, a higher Λ_1 augments second period traders return uncertainty, and makes them speculate less (the second term in the parenthesis in (23)).

Consider now the effect of an increase in Λ_1 on $\mathrm{Var}[p_2|u_1]$:

$$
\frac{\partial \mathrm{Var}[p_2|u_1]}{\partial \Lambda_1} = 2\lambda_2^2 \left(\underbrace{\frac{b_{21}^L}{\tau_{u_2}} \frac{\partial b_{21}^L}{\partial \Lambda_1}}_{(-)} + \underbrace{\frac{b_{22}^L}{\tau_\eta} \frac{\partial b_{22}^L}{\partial \Lambda_1}}_{(\pm)} \right).
\tag{24}
$$

In the transparent market benchmark, a higher illiquidity has *no impact* on first period traders' uncertainty over p_2 (see (6b)). In contrast, according to (21), opaqueness introduces two channels through which a shock to liquidity feeds back to first period traders' uncertainty. First, an increase in Λ_1 lowers second period traders' hedging activity, lowering $\mathrm{Var}[p_2|u_1]$.[13] However, as we argued above, a less liquid first period market can spur more speculation by second period traders. As traders' information is imprecise, this yields a second feedback channel that can instead magnify first period traders' uncertainty. Thus, according to (24), the ultimate impact of a shock to Λ_1 on first period traders' uncertainty depends on the strength of the speculation effect.

Finally, because of opaqueness, an increase in Λ_1 introduces an additional effect on b_1^L. Differentiating (5a):

$$
\frac{\partial b_1^L}{\partial \Lambda_1} = \frac{\gamma_1^L}{\mathrm{Var}[p_2|u_1]^2} \times
\tag{25}
$$

$$
\left(\underbrace{\left(\frac{\partial \mathrm{Cov}[p_2, u_1]}{\partial \Lambda_1} \tau_{u_1} + 1 \right) \mathrm{Var}[p_2|u_1]}_{\text{Direct effect } (+)} - \underbrace{\frac{\partial \mathrm{Var}[p_2|u_1]}{\partial \Lambda_1} (\mathrm{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1)}_{\text{Feedback effect } (\pm)} \right).
$$

For given $\mathrm{Var}[p_2|u_1]$, as p_2 is more positively associated with u_1, a larger Λ_1 leads first period traders to speculate more (and hedge less, that is b_1^L is closer to zero), as per the *direct* liquidity consumption "rationing" effect of the transparent market benchmark. However, when the speculation effect leads $\mathrm{Var}[p_2|u_1]$ to increase in Λ_1, a less liquid market now also has a *feedback* liquidity consumption "expanding" effect on b_1^L. As a higher Λ_1 increases the risk to which first period traders are exposed, a less liquid market can lead them to hedge more (that is, a more negative b_1^L). As a result, first period traders' demand for liquidity can become *increasing* in Λ_1, as shown in Panel (a) of Figure 4. In this case, an increase in the cost of liquidity provision incites more liquidity consumption. Mirroring the discussion around Figure 3, Panel (b) in the figure illustrates the effect of opaqueness on FDs' directional bets. Differently from the

---

[13] According to Corollary 4, $\partial|b_{21}^L|/\partial\Lambda_1 < 0$, or equivalently $\partial b_{21}^L/\partial\Lambda_1 > 0$, and since $b_{21}^L < 0$, we have $b_{21}^L \times (\partial b_{21}^L/\partial\Lambda_1) < 0$.

transparent market benchmark, an increase in Λ_1 leads FDs to scale *down* their speculative activity, exacerbating the price impact of the liquidity demand hike. This is because a higher Λ_1 has a positive impact on traders' expected returns and risk from speculation. However, as the figure shows, the impact on the latter can overcome the one on the former.

The expanding effect of illiquidity can be responsible for a destabilizing dynamic whereby to a sizeable evaporation of liquidity, first period traders respond with an even more aggressive liquidity consumption, while FDs scale down their directional bets. In the figure we use the same parameter values of Figure 3, but assume that $\tau_\eta = 10$ (instead of $\tau_\eta \to \infty$). As a result, at equilibrium we obtain $b_1^L|_{\tau_\eta=10} = -0.5$, and $\Lambda_1^*|_{\tau_\eta=10} = 3.8$. Compared to the values of the example of Figure 3, these results correspond to a more than two- and an almost ten-fold increase in liquidity consumption and illiquidity.
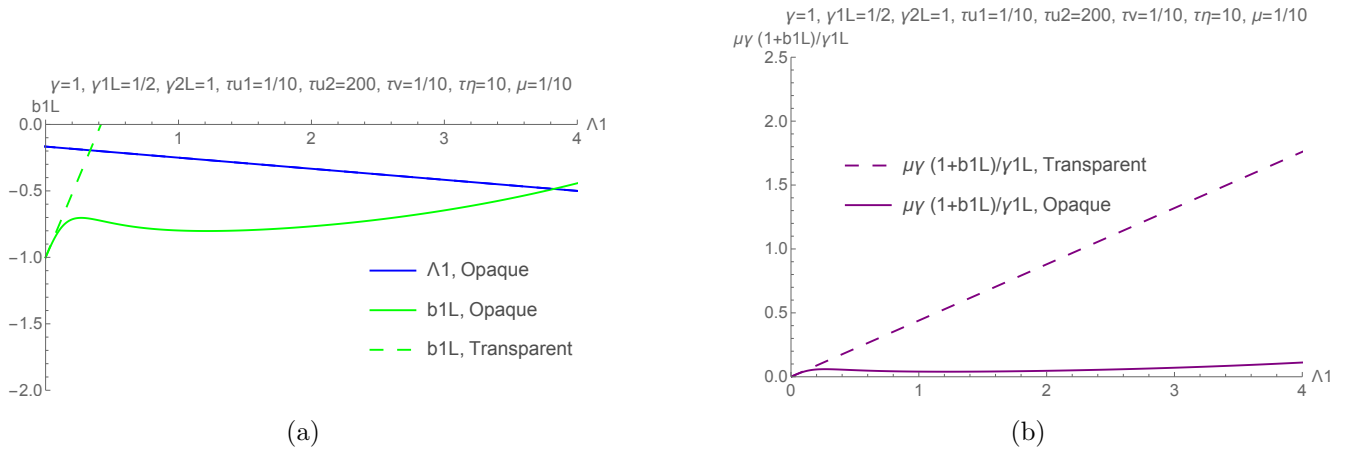
Figure 4: When the market is opaque, first period traders' demand for liquidity can turn increasing in Λ_1 (Panel (a)) and FDs' directional bets decreasing in Λ_1 (Panel (b)). The dashed curves in the two panels corresponds respectively to b_1^L, and FDs' aggregate directional bets in the transparent market case.

4.1   Equilibrium multiplicity

A second effect of opaqueness is the possibility of multiple, self-fulfilling equilibria which arise out of strategic complementarities in *liquidity demand*. According to Corollary 4, a less liquid first period market heightens the time-propagation of the first period shock. This, in turn, can lead second period traders to speculate more aggressively on the u_1-led imbalance (see (23)), which can increase the uncertainty faced by first period traders on p_2 (see (24)). As a consequence, first period traders can decide to hedge more, and FDs to speculate less (see (25)). This chain of effects turns out to be particularly strong when the risk bearing capacity of FDs is not too low, first period traders are sufficiently risk averse, second period traders have a sufficiently informative signal, and face low endowment risk, and the risk of the asset payoff is large. In these conditions, an initial dearth of liquidity escalates into a loop that sustains three equilibrium levels of liquidity:

Proposition 3. *There exists a set of parameter values, such that for, $\tau_v < \bar{\tau}_v$, $\mu < \bar{\mu}$, and $\gamma_1^L < \bar{\gamma}_1^L$, three equilibrium levels of liquidity $(\Lambda_1^*)^H, (\Lambda_1^*)^I, (\Lambda_1^*)^L$ arise, where*

$$0 < (\Lambda_1^*)^H < \frac{\mu}{1-\mu} < (\Lambda_1^*)^I < \frac{1}{1-\mu} < (\Lambda_1^*)^L < \frac{1}{\gamma\tau_v} \ . \qquad (26)$$

We will refer to the equilibrium where Λ_1^* is low (resp., intermediate, and high) as the High, (resp., Intermediate, and Low) liquidity equilibrium (HLE, ILE, and LLE). Note that since the function $\Lambda_1^*(b_1^L)$ is decreasing in b_1^L (see (14a)), the hedging activity of first period traders is respectively high, intermediate, and low along $(\Lambda_1^*)^L$, $(\Lambda_1^*)^I$, and $(\Lambda_1^*)^H$. This is a further manifestation of the fact that the feedback effect of liquidity jams the stabilizing impact of an increase in illiquidity on traders' hedging demand. Figure 5 provides a numerical example of the proposition.
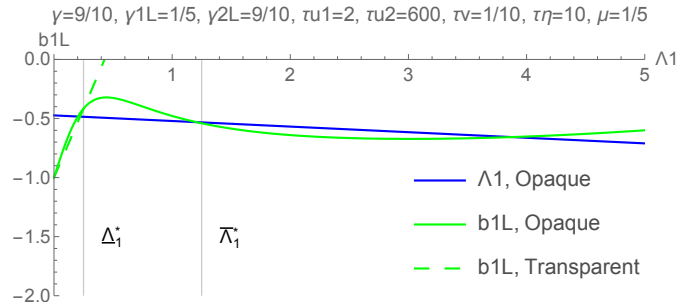
Figure 5: Market opaqueness and equilibrium multiplicity. At equilibrium $\{\Lambda_1^*, b_1^L\} \in \{\{0.4, -0.5\}, \{1, -0.5\}, \{3.9, -0.7\}\}$.

The following corollary follows from Proposition 3:

Corollary 5. *When the volatility of the second period endowment shock vanishes ($\tau_{u_2} \to \infty$) and the following parameter restriction applies: $\tau_v < \bar{\tau}_v$, $\tau_\eta > \underline{\tau}_\eta$, $\mu < \bar{\mu}$, $\gamma_1^L < \bar{\gamma}_1^L$, then $(\Lambda_1^*)^H = 0$.*

When $\tau_{u_2} \to \infty$, second period traders have no endowment to hedge, and only trade to speculate on the u_1-induced imbalance. In the equilibrium where $\Lambda_1^* = 0$, $x_1^D = 0$, so that first period traders' orders are absorbed by FDs' speculative trades, no imbalance arises in the second period, and $b_{22}^L = 0$ (see (17b)). When second period traders' signal on u_1 is fully revealing, this equilibrium is unique (Corollary 1). For τ_η finite, however, first period traders cannot rule out the possibility that second period traders speculate on a certain realization of s_{u_1} that gives an incorrect signal about u_1 (e.g., $s_{u_1} > 0$, while $u_1 < 0$). This increases the uncertainty they face, and triggers the loop that can lead to the appearance of two further equilibria.

To study comparative statics, it is convenient to introduce the "best response" function $\psi(\Lambda_1)$, which collapses the demand and supply of liquidity equations (respectively (22), and (14a)) into a single one that can be interpreted as an aggregate best response of first period traders to a change in first period illiquidity. As we explain in the appendix, the fixed points of $\psi(\Lambda_1)$

correspond to the equilibria of the market.[14] Figure 6 provides a graphical representation of the best response, for a set of parameters yielding multiple equilibria.
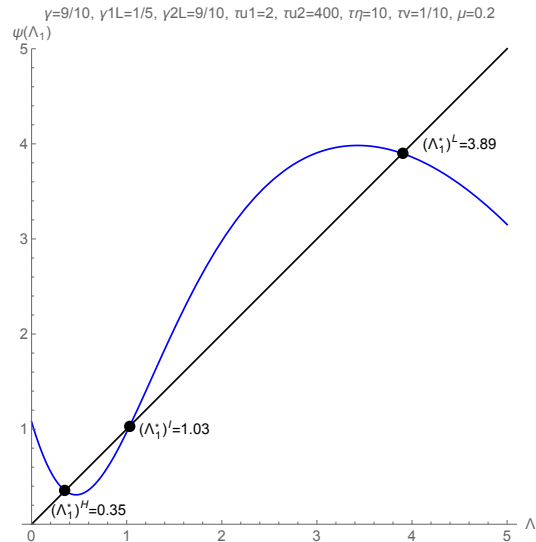

γ=9/10, γ1L=1/5, γ2L=9/10, τu1=2, τu2=400, $\tau\eta$=10, τv=1/10, μ=0.2

Figure 6: The best response mapping $\psi(\Lambda_1)$ for $\gamma = \gamma_2^L = 9/10$, $\gamma_1^L = 1/5$, $\tau_{u_1} = 2$, $\tau_{u_2} = 400$, $\tau_\eta = 10$, $\tau_v = 1/10$, and $\mu = 1/5$.

Higher values of τ_{u_2}, γ_2^L, and intermediate values of τ_η, strengthen the degree of strategic complementarity, facilitating equilibrium multiplicity. Intuitively, under these conditions second period traders are relatively better positioned to speculate on the propagated imbalance, which reinforces the impact of their orders on first period traders' uncertainty. A positive shock to μ, γ, γ_1^L, τ_{u_1}, and τ_v moves the best response down and also mildly decreases the degree of strategic complementarity. These shocks increase the risk bearing capacity of the market, and reduce the risk to which first period traders are exposed. As a consequence, they weaken the impact of second period traders' orders on first period traders' uncertainty

More in detail, as shown in Figures 12 and 13 (in the Appendix):

1. An increase in μ increases the risk bearing capacity of the market, and lowers the second period imbalance due to u_1. This shifts the best response mapping downwards, implying a higher (lower) liquidity at $(\Lambda_1^*)^H$, $(\Lambda_1^*)^L$ $((\Lambda_1^*)^I)$. A similar effect obtains when γ or τ_v increase (see Figure 12, Panel (a), (b), and (c)).

2. An increase in γ_1^L or in τ_{u_1} works to lower the supply shock that Ds and FDs absorb in the first period, and thus the second period imbalance due to u_1, shifting the best response mapping downwards (see Figure 12, Panel (d) and (e)).

3. An increase in τ_η has two contrasting effects on the strength of strategic complementarity. For τ_η small, a more precise signal on u_1 boosts second period traders' speculation on the u_1-induced imbalance, heightening first period traders' uncertainty on p_2, increasing Λ_1^*, and strengthening complementarity. As τ_η increases further, $\text{Var}[p_2|u_1]$ starts decreasing (see the expression in Proposition 2), leading first period traders to hold more of their

---

[14]See the proof of Proposition 2, and in particular the discussion around (A.23).

endowment shock, increasing the liquidity of the first period market, reducing the size of the u_1-led imbalance, and weakening complementarity. When $\tau_\eta \to \infty$, first period short term traders anticipate the trades of second period short term traders. The consequence is that (i) the impact of second period traders' speculation on $\text{Var}[p_2|u_1]$ disappears, and (ii) second period traders' uncertainty no longer depends on Λ_1^* (see the expression for $\text{Var}[v - p_2|\Omega_2^L]$ in Proposition 2). This severs the link between trading decisions at the two dates, yielding a unique equilibrium (as we know from the analysis of Section 3). Figure 13, panel (a), illustrates this effect. Note that as τ_η increases, the intermediate and low liquidity equilibria eventually disappear, but liquidity at the high liquidity equilibrium diminishes (compared with the case with low signal precision).

4. An increase in γ_2^L increases the strength of strategic complementarity. Other things equal, according to (17a) and (17b), when second period traders become more risk tolerant they hedge less of their endowment shock, and speculate more aggressively against the u_1-led imbalance. The first of the above effects works to weaken strategic complementarities, because it reduces the impact of the endowment shock on first period traders' uncertainty. Conversely, the second effect strengthens them, and always prevails in our simulations as illustrated in Figure 13, Panel (b).

5. Finally, for $\tau_{u_2} < \infty$, a lower τ_{u_2} magnifies the risk to which first period traders are exposed, and works to shift the best response upwards. The smaller τ_{u_2} becomes, the harder it is to sustain the equilibrium with high liquidity. When τ_{u_2} becomes very small, our simulations show that only the low liquidity equilibrium survives. Figure 13, Panel (c) in the Appendix illustrates this effect.[15] In the limit case where $\tau_{u_2} \to 0$, we obtain the following result:

Corollary 6. *When $\tau_{u_2} \to 0$, $\Lambda_1^* = 1/\gamma\tau_v$, and $b_1^L = -1$.*

Therefore, when second period traders' endowment shock volatility explodes, first period traders cannot anticipate p_2, and hedge all of their endowment, like noise traders.

4.2 Fragility

The feedback loop induced by market opaqueness implies that liquidity can be "fragile," in the sense that a relatively small shock to one of the model's parameters can lead to a disproportionately large change in liquidity. We show this with two examples. In the first one, we consider the effect of a parameter shock yielding a switch from the HLE to an equilibrium with low liquidity. In the second example, we show that when a unique equilibrium obtains, with opaqueness Λ_1^* can be hump-shaped in μ. This implies that a sudden reduction in the mass of FDs can lead to a large drop in liquidity.

---

[15] According to Corollary 5 in the extreme case in which the second period endowment shock is null (almost surely), there always exist an equilibrium where $\Lambda_1^* = 0$, $x_1^D = 0$, so that first period traders' orders are absorbed by FDs' speculative trades, which implies that no imbalance arises in the second period, and $b_{22}^L = 0$ (see (17b)). A full analytical characterization of this equilibrium is complex. Numerically, it can be seen that first period liquidity traders hedge the smallest possible fraction of their endowment shock: $b_1^L \to -\mu\gamma/(\mu\gamma + \gamma_1^L)$.

4.2.1 Equilibrium switch

Consider Panel (a) in Figure 7, and suppose that initially the market is at the high liquidity equilibrium, where $\Lambda_1^* = 0.22$. Suppose that a technical "glitch" disconnects 6% of the FDs. In this new situation, the plot for $\Lambda_1(b_1^L)$ shifts upwards, while the one for $b_1^L(\Lambda_1)$ moves downwards, as illustrated in panel (b) of the figure. As a result, a new, unique equilibrium obtains with $\Lambda_1^* = 3.5$, which corresponds to a 16-fold liquidity decrease.

A similar effect also arises if we shock the volatility of first or second period traders' endowment. To see this, suppose now that starting from Panel (a) in the figure, we introduce a 5% negative shock to τ_{u_1} (i.e., we move τ_{u_1} from 2 to 1.9), which increases the likelihood that an order of an unusual magnitude hits the first period market. As shown in Panel (c) of Figure 7, this leads to a downward shift in the plot for b_1^L which is large enough to eliminate the HLE and move the market towards a new equilibrium with low liquidity in which $\Lambda_1^* = 3.5$ and $b_1^L = -0.6$. Finally, suppose that we increase the volatility of the second period endowment shock, introducing a 7% negative shock to τ_{u_2} (lowering it to 620). In this new situation, the plot for the function $b_1^L(\Lambda_1^*)$ moves downwards, while the one for $\Lambda_1^*(b_1^L)$ is unchanged (see Panel (d) in Figure 7). A unique equilibrium obtains, where $\Lambda_1^* = 3.5$, implying liquidity dry-up comparable to the one of the previous examples.

Table 1 summarizes the results of these exercises and compares them with the effects that obtain in the transparent market case.
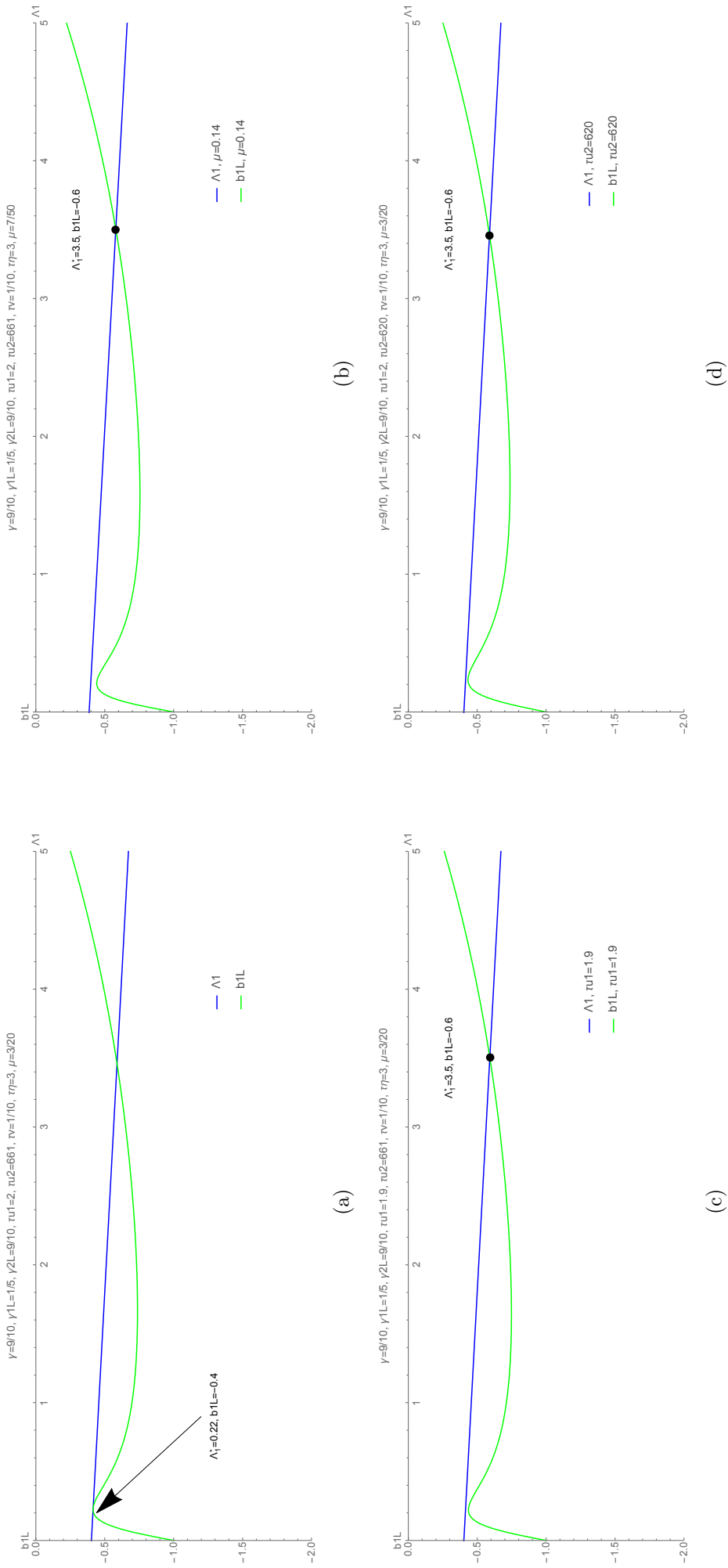
Figure 7: Liquidity fragility. Comparing panel (a) with (b) illustrates the effect of a decrease in the mass of FDs. Comparing panel (a) with (c) and (d) illustrates the effect of an increase in the volatility of first and second period liquidity traders' demand.

	Status quo Λ_1^*	Shock to parameter	New Λ_1^*	$\Delta\Lambda_1^*/\Lambda_1^*$
		$\Delta\mu/\mu = -6\%$	0.245	3.2%
Transparent market ($\tau_\eta \to \infty$)	0.237	$\Delta\tau_{u_1}/\tau_{u_1} = -5\%$	0.237	0%
		$\Delta\tau_{u_2}/\tau_{u_2} = -7\%$	0.252	6.3%
		$\Delta\mu/\mu = -6\%$	3.5	1470%
Opaque market ($\tau_\eta < \infty$)	0.22	$\Delta\tau_{u_1}/\tau_{u_1} = -5\%$	3.5	1470%
		$\Delta\tau_{u_2}/\tau_{u_2} = -7\%$	3.5	1470%

Table 1: Equilibrium switch: The impact of a shock to μ, τ_{u_1}, and τ_{u_2}, in the transparent, and opaque market case. Other parameters' values are as in Figure 7.

4.2.2 Fragility along a unique equilibrium

Liquidity fragility can also arise when the equilibrium is unique. This is because, with opaqueness, parameter shocks can have a non monotone effect on market liquidity. For example, an increase in μ triggers two potentially contrasting effects on liquidity:

$$\frac{\partial \Lambda_1^*}{\partial \mu} = -\frac{1}{\gamma\tau_v} \times \left(\underbrace{\frac{\gamma(1+b_1^L)}{\gamma_1^L}}_{\text{Direct effect }(+)} + \underbrace{\frac{\partial b_1^L}{\partial \mu}\left(1 + \frac{\mu\gamma}{\gamma_1^L}\right)}_{\text{Indirect effect }(\pm)} \right). \tag{27}$$

For given b_1^L, the direct effect captures the increase in FDs' aggregate speculative position, which works to lower dealers' inventory, and make the market more liquid. The indirect effect reflects the impact of the change in μ on first period traders' demand for liquidity ($\partial b_1^L/\partial\mu$), and on each FD's speculative position (($\mu\gamma/\gamma_1^L$)($\partial b_1^L/\partial\mu$)). The sign of this effect is, instead, ambiguous. Indeed, an increase in the mass of FDs can lower the impact of second period traders' orders on p_2, thereby lowering $\text{Var}[p_2|u_1]$ and leading first period traders to hold more of their endowment, and each FD to speculate more on the short term capital gain;[16] at the same time, however, it can also lower the propagation of u_1 to the second period, impairing the predictability of p_2, and inducing traders to shed more of their endowment, and each FD to speculate less.[17] When the market is opaque, second period traders face execution risk, which tames their hedging aggressiveness (see (17a)), and lowers first period traders' return uncertainty. In this situation, the uncertainty reduction effect of μ on b_1^L can be dwarfed by the one due to reduced predictability. As a consequence, when μ increases, first period

---

[16] The volatility reduction can happen for two different reasons. As μ increases, (i) less of the first period endowment shock propagates to the second period, and (ii) more FDs absorb second period liquidity traders' orders, enhancing risk sharing.

[17] In a transparent market, this latter effect is never strong enough to overcome the previous two, and liquidity increases in the mass of FDs (see Corollary 2).

traders' demand for immediacy can increase (see Figure 11 in the Appendix) and the individual speculative activity of each FD can abate, offsetting the direct positive effect of FDs' aggregate speculative trades. Hence, a wider FDs' participation can impair liquidity.

This implies that illiquidity can be hump-shaped in the proportion of FDs, as shown in Figure 9 (Panel (a)). The non-monotone relationship between Λ_1^* and μ illustrates an additional channel through which liquidity fragility can arise. To see this, suppose that the proportion of FDs in the market is initially $\mu = 0.4$. According to the figure, for this fraction of FDs' participation, in the opaque market case we have $\Lambda_1^* = 0.5$, and $b_1^L = -0.47$ (see Panel (a) and Panel (b)). Suppose now that a glitch disconnects 10% of FDs, implying that a proportion $\mu = 0.36$ of FDs supplies liquidity. In the opaque market case, this implies a new illiquidity level $\Lambda_1^* = 1.58$, corresponding to a 216% liquidity withdrawal. Conversely, in the case with transparent markets, when $\mu = 0.4$, $\Lambda_1^* = 0.38$, while when $\mu = 0.36$, $\Lambda_1^* = 0.41$, corresponding to an 8% liquidity decrease. This shows that along an equilibrium with low liquidity, following a reduction in FD participation, liquidity can dry up quite dramatically.

The example highlights an additional implication of our analysis. When $\mu = 0.36$, $b_1^L = -0.51$ (see Figure 9, Panel (b)), which, compared to the status quo liquidity demand, corresponds to a 9% increase in liquidity consumption by first period traders (i.e., $b_1^L = -0.47$ when $\mu = 0.4$). How can such a comparatively small increase in liquidity consumption generate an illiquidity spike of this magnitude? To understand this, note that increased liquidity consumption only accounts for *part* of the total effect, as it occurs *jointly* with a steep increase in first period return uncertainty (according to Panel (c) in the figure, $\text{Var}[p_2|u_1]$ experiences a 310% increase across the two equilibrium outcomes). This, leads each FD to scale down his speculative position, thereby adding to the aggregate effect of a reduced liquidity supply. Using expression (27) to break down the different effects yields:

$$
\underbrace{\frac{\Delta \Lambda_1^*}{\Delta \mu}}_{\approx -27} \approx - \underbrace{\frac{1}{\gamma \tau_v}}_{= -10} \times \tag{28}
$$

$$
\left(\underbrace{\frac{\gamma(1 + b_1^L)}{\gamma_1^L}}_{\substack{\Delta \text{ in FDs' aggregate specula-} \\ \text{tive position} \approx 1}} + \underbrace{\frac{\mu \gamma}{\gamma_1^L}}_{= 4/5} \times \underbrace{\frac{\Delta b_1^L}{\Delta \mu}}_{\substack{\Delta \text{ in each FD speculative po-} \\ \text{sition} \approx 1}} + \underbrace{\frac{\Delta b_1^L}{\Delta \mu}}_{\Delta \text{ in liquidity demand} \approx 1} \right)
$$

Thus, the increase in liquidity consumption accounts for roughly 36% of the drought, whereas the lion share of it (about 64%) is due to the combined effect of the aggregate and individual reduction in FDs' speculative activity.[18] This suggests that in the wake of a liquidity crash, we should observe HFTs unwinding their directional bets.[19]

A final implication of this example is that when information on prices and/or order im-

---

[18] According to (28), the sum of Δ in FDs' aggregate speculative position, Δ in each FD speculative position, and Δ in traders' liquidity demand amounts to 2.8, of which $\Delta b_1^L / \Delta \mu$ accounts for 1.

[19] In this example too, for μ small, the stabilizing effect of illiquidity is jammed and first period traders demand more immediacy *precisely* when the cost of liquidity supply is increasing.

balances is opaque, an increase in the mass of HFTs (promoting full participation), can lower market liquidity. This finding is consistent with Boehmer et al. (2015) who show that greater algorithmic trading intensity is associated with more liquidity for average firm size, the same is not true for small market cap firms. For these firms, when algorithmic trading increases, liquidity declines.[20]

4.2.3   Transparency and liquidity

An important feature of the example in Figure 9 is that for all values of μ the market is at least as liquid with full transparency as it is with opaqueness. This result is not general. In Figure 10 we reproduce Panel (a) of Figure 9, and in panels (b), (c), and (d), investigate the effect of progressively lowering the informativeness of the signal observed by second period traders. The figure shows that while the hump-shaped feature of the illiquidity plot is preserved, the liquidity ranking is progressively *reversed*. In fact, we can prove the following result:

Corollary 7. *When the market is strongly opaque ($\tau_\eta \to 0$), there exists a unique equilibrium along which liquidity is higher than in the transparent market benchmark:*

$$\Lambda_1^*|_{\tau_\eta \to 0} \leq \Lambda_1^*|_{\tau_\eta \to \infty}. \tag{29}$$

Intuitively, if second period traders receive an uninformative signal, they stop speculating on the propagated imbalance, killing the loop that is responsible for equilibrium multiplicity. Thus, in this case, first period traders' uncertainty only depends on b_{21}^L, similarly to the transparent benchmark where first period traders can perfectly anticipate the impact of second period speculation. Furthermore, due to execution risk, second period traders also hedge less of their endowment shock when the signal is uninformative (since $\kappa > 1$) than when it is perfect (in which case, instead, $\kappa = 1$). As a consequence, with strong opaqueness, first period traders face lower uncertainty than in the transparent market benchmark. Thus, in the latter case they hedge more, and FDs speculate less, lowering market liquidity.

Figure 10 also suggests that for $\mu < 1$, increasing transparency won't necessarily enhance market liquidity, precisely because a more informative signal leads second period traders to trade more aggressively, increasing first period traders' uncertainty. This insight is confirmed in Figure 8, where we plot Λ_1^* in the opaque market case, as a function of τ_η. For low values of transparency, a more informative signal makes the market less liquid. This is because, second period traders speculate more aggressively on the propagated imbalance, increasing first period traders' uncertainty (as we observed in Section 4.1). As second period traders' signal precision increases, first period traders become increasingly better able to anticipate their strategy, and face lower return uncertainty. This explains the declining branch of the plot. As argued in Corollary 7, a very precise signal (τ_η large) fosters second period traders' hedging aggressiveness, which heightens first period traders' uncertainty compared to the case

---

[20]See also Breckenfelder (2014) for other evidence on the negative impact of an increase in HFT competition on market liquidity for a sample of stocks traded on the Stockholm Stock Exchange.
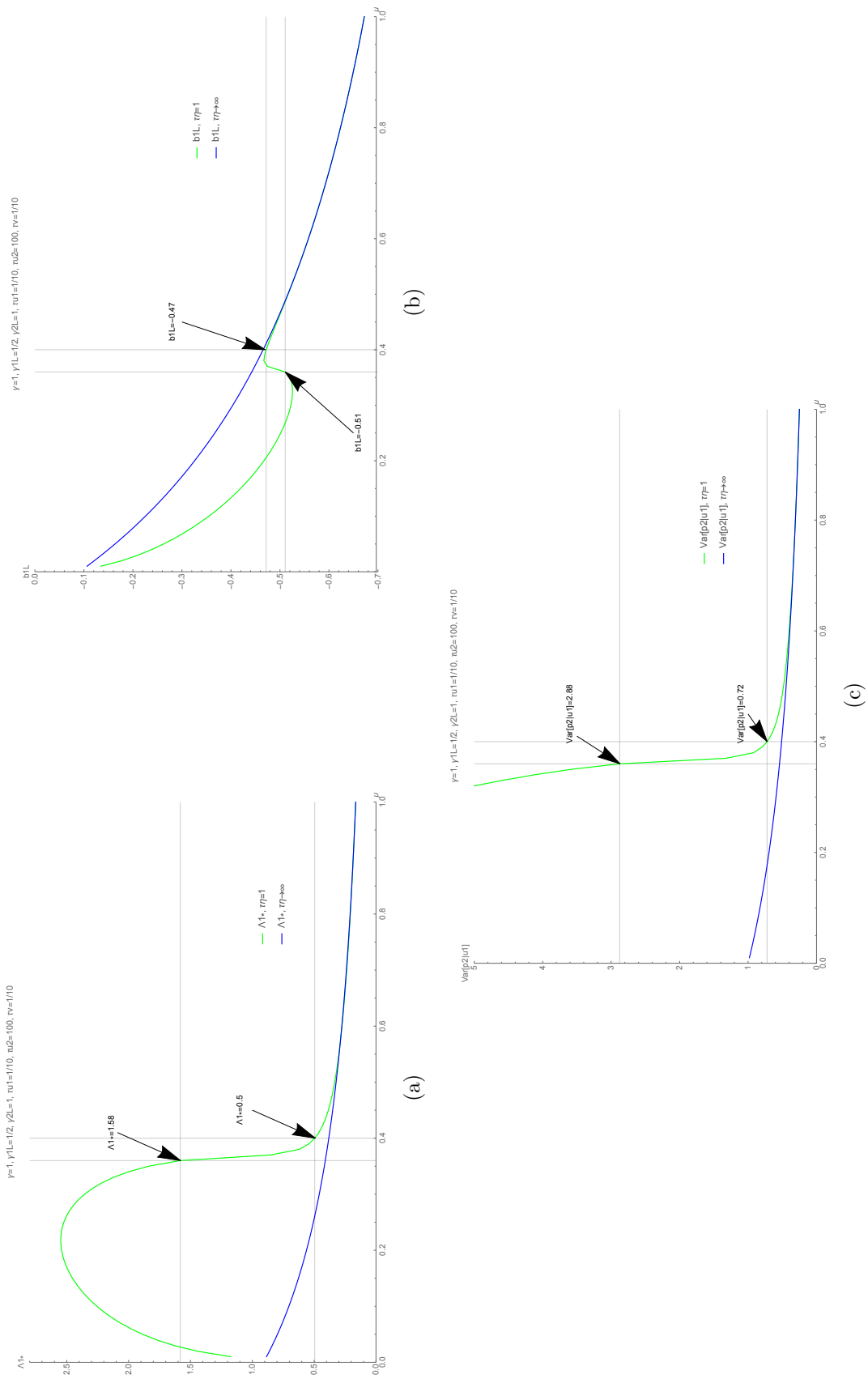
Figure 9: Fragility at an equilibrium with low liquidity. In panel (a), (b), and (c) we plot Λ_1^*, b_1^L, and $\mathrm{Var}[p_2|u_1]$, for $\tau_v = \tau_{u_1} = 0.1$, $\tau_{u_2} = 100$, $\tau_\eta = \tau = \gamma = \gamma_2^L = 1$, $\gamma_1^L = 0.5$, and $\mu \in \{0.01, 0.02, \ldots, 1\}$; the blue (green) plot relates to the transparent (opaque) case. A glitch disconnecting 10% of FDs yields a 9% increase in the demand for immediacy, a 310% increase in first period uncertainty and a 216% illiquidity spike. The gridlines are drawn at $\mu = \hat\mu \in \{0.36, 0.4\}$ and at the corresponding values for Λ_1^*, b_1^L, and $\mathrm{Var}[p_2|u_1]$.
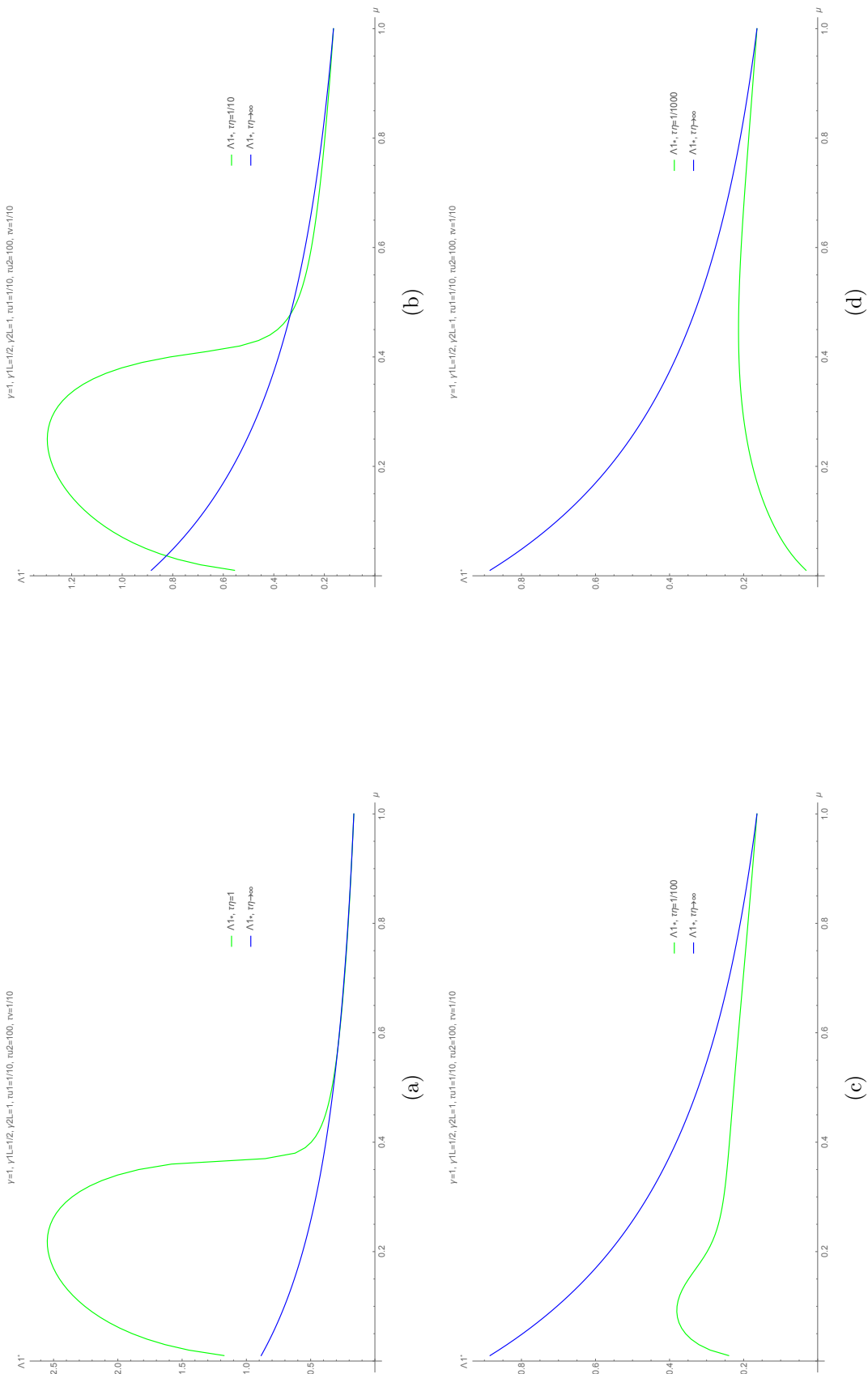
Figure 10: Liquidity comparison between the opaque and transparent market case. In panel (a), we plot Λ_1^*, for $\tau_v = \tau_{u_1} = 0.1$, $\tau_{u_2} = 100$, $\tau_\eta = \gamma = \gamma_2^L = 1$, $\gamma_1^L = 0.5$, and $\mu \in \{0.01, .02, \ldots, 1\}$; the blue (green) plot relates to the transparent (opaque) case. In panel (b), (c), and (d) we progressively reduce τ_η.

in which the signal is uninformative ($\tau_\eta \approx 0$).
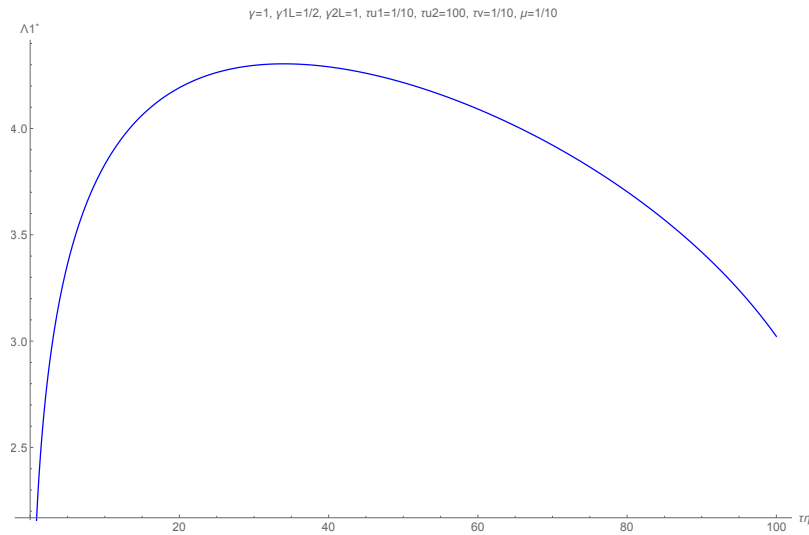
Figure 8: Illiquidity as a function of transparency, for $\tau_v = \tau_{u_1} = 0.1$, $\tau_{u_2} = 100$, $\gamma = \gamma_2^L = 1$, $\gamma_1^L = 0.5$, $\mu = 0.1$.

4.2.4 Summary

This section highlights the role of informational frictions in generating a liquidity feedback loop that can have a destabilizing effect on the market. Second period traders, endowed with a noisy signal on the first period endowment shock, speculate against the propagated order imbalance, generating additional volatility. This can feed back on first period traders' strategies, leading them to consume more liquidity and FDs to retreat from speculation, thereby magnifying the outstanding imbalance liquidity suppliers need to absorb, and further lowering market liquidity. This self-sustaining loop can induce multiple equilibria and liquidity fragility. Equilibria can be ranked in terms of liquidity and first period traders' hedging activity, with the most (least) liquid equilibrium occurring with the least (highest) liquidity consumption. Thus, with market opaqueness, the self-stabilizing mechanism whereby an illiquidity spike depresses liquidity consumption can jam, and instead be replaced by a vicious cycle that creates a liquidity rout.

	Transparent ($\tau_\eta \to \infty$)	Opaque ($0 < \tau_\eta < \infty$)
Equilibrium	Unique	Possible ME
Illiquidity	Decreasing in μ	Can be 'fragile' and hump-shaped in μ and τ_η (Numerical)
First period traders' demand for immediacy	Decreasing in Λ_1	Can be increasing in Λ_1

Table 2: Summary of results.

5   Concluding remarks

We study a 2-period model in which two classes of dealers—full and standard—intermediate the orders of two successive cohorts of short-term traders, in a context where markets are fragmented due to both an informational and a participation friction. We show that dealers' limited market participation favors the propagation of first period traders' endowment shock across time, inducing a predictable price pressure. This, in turn, leads second period traders to speculate against the propagated endowment shock. The effect of speculation crucially depends on the transparency regime governing the market. With full transparency the market is stable and increasing the proportion of full dealers is good for liquidity. With opaqueness, the market may be unstable, and liquidity hump-shaped in the proportion of full dealers and in the degree of transparency. More in detail, our main findings can be summarized as follows (see Table 2):

1. When second period traders have perfect information about the first period endowment shock, speculation exerts a stabilizing effect. In this context, a unique equilibrium obtains, and a dearth of liquidity increases first period traders' cost of hedging, reducing their liquidity consumption. Furthermore, higher FDs participation always has a beneficial impact on liquidity.

2. When the market is opaque—in that second period traders' information is imprecise—speculation can *augment first period traders' uncertainty*, leading them to demand more immediacy when the market is less liquid. This can offset the rationing impact of illiquidity, and trigger a liquidity feedback loop in which a liquidity dry-up breeds a further, larger liquidity withdrawal. We show that in this scenario,

   (a) Multiple equilibria—that can be ranked in terms of liquidity, price volatility, and demand for immediacy—can arise.

   (b) Liquidity can be fragile, either because a shock to parameter values can prompt a switch from the high liquidity equilibrium to an equilibrium with low liquidity; or, with a unique equilibrium, because illiquidity may be hump-shaped and with steep slope in the mass of full dealers, implying that a reduction in FDs participation can generate a large spike in illiquidity.

3. Our results are robust to the case in which the first period market is populated by a proportion β of short-term traders, that unwind at date 2, and a complementary proportion $(1 - \beta)$ of long-term ones that hold their position until liquidation.

From a methodological point of view, our work shows that because of market opaqueness, uncertainty on future prices can become an increasing function of current illiquidity, implying that fragility can arise in a context where prices are driven by a non-payoff related shock. We view this as a realistic feature of trading at high frequencies since in those conditions, the chances that payoff fundamentals drive prices are negligible. This also allows us to offer an alternative explanation for how the buildup of a large imbalance can precipitate the market into a crash, which does not rely on the effect of order flow toxicity (Easley et al. (2011, 2012)).

From a policy perspective, our paper has two important implications. First, our analysis of the opaque market model shows that favoring FDs' entry (i.e., reducing the participation friction) doesn't necessarily enhance liquidity. Indeed, illiquidity can be hump-shaped in the proportion of FDs. This can also serve as a guide to empirical analysis, as it suggests that the liquidity impact of HFT entry should be assessed taking into account the effect of frictions in the access to market information. Second, with noisy market information, the presence of liquidity providers acting with different trading frequencies may make liquidity fragile, either because a shock to parameters can prompt a switch across equilibria; or because, due to the hump-shaped relationship between illiquidity and the proportion of FDs, a sudden reduction of these dealers' participation can lead to a large liquidity withdrawal. This supports regulatory concerns about the potential drawbacks of automated trading due to operational and transmission risks and implies that fragility can arise in the absence of order flow toxicity.[21]

Finally, our analysis of fragility along a unique equilibrium highlights the role that changes in HFTs' strategies in the wake of crashes have to explain huge illiquidity spikes. As we argued in our numerical example, the unwinding of HFTs' directional bets can act as a multiplier of the initial increase in traders' liquidity consumption. This can be of help in empirical analyses of these events, in that an exclusive focus on changes in liquidity consumption can miss an important explanatory factor and suggests the following empirical prediction. We expect that in the wake of a liquidity crash, HFTs should unwind their directional bets, thereby depleting the market of the liquidity that they otherwise supply with contrarian marketable orders (see the evidence in Brogaard et al. (2014), and Biais et al. (2015)).

---

[21]See Joint Staff Report: The U.S. Treasury Market on October 15, 2014.

References

Biais, B., F. Declerck, and S. Moinas (2015). Who supplies liquidity, how and when? *Working paper*.

Biais, B., T. Foucault, and S. Moinas (2015). Equilibrium fast trading. *Journal of Financial Economics 116*, 292–313.

Boehmer, E., K. Fong, and J. Wu (2015). International evidence on algorithmic trading. *Working paper*.

Bogousslavsky, V. (2014). Infrequent rebalancing and short-term return predictability. *Working Paper*.

Bongaerts, D. and M. Van Achter (2015). High frequency trading and market stability. *Working Paper, Erasmus University*.

Breckenfelder, H. J. (2014). Competition between high-frequency traders, and market quality. *Working Paper*.

Brogaard, J., T. Hendershott, and R. Riordan (2014). High frequency trading and price discovery. *Review of Financial Studies 27*, 2267–2306.

Budish, E., P. Cramton, and J. Shim (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics 139*, 1547–1621.

Cespa, G. and T. Foucault (2014). Illiquidity contagion and liquidity crashes. *Review of Financial Studies 27*(6), 1615–1660.

Cespa, G. and X. Vives (2012). Dynamic trading and asset prices: Keynes vs. Hayek. *Review of Economic Studies 79*, 539–580.

Cespa, G. and X. Vives (2015). The Beauty Contest and short-term trading. *Journal of Finance 70*(5), 2099–2154.

Chien, Y., H. Cole, and H. Lustig (2012). Is the volatility of the market price of risk due to intermittent portfolio rebalancing? *American Economic Review 102*(6), 2859–96.

Collin-Dufresne, P. and V. Vos (2015). Do prices reveal the presence of informed trading? *Journal of Finance 70*, 1555–1582.

Ding, S., J. Hanna, and T. Hendershott (2014). How slow is the NBBO? a comparison with direct exchange feeds. *Financial Review* (49), 313–332.

Du, S. and H. Zhu (2014). Welfare and optimal trading frequency in dynamic double auctions. *Working paper*.

Duffie, D. (2010). Presidential address: Asset price dynamics with slow-moving capital. *Journal of Finance 65*(4), 1237–1267.

Easley, D., M. O'Hara, and M. López de Prado (2011). The microstructure of the Flash-Crash: flow toxicity, liquidity crashes, and the probability of informed trading. *Journal of Portfolio Management 37*, 118–128.

Easley, D., M. O'Hara, and M. López de Prado (2012). Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies 25*, 1457–1493.

Foucault, T., J. Hombert, and I. Rosu (2015). News trading and speed. *Journal of Finance (Forthcoming)*.

Glosten, L. and P. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*, 71–100.

Goldstein, I., Y. Li, and L. Yang (2014). Speculation and hedging in segmented markets. *Review of Financial Studies 27*(3), 881–922.

Goldstein, I. and L. Yang (2015). Information diversity and complementarities in trading and information acquisition. *Journal of Finance 70*, 1723–1765.

Han, J., M. Khapko, and A. S. Kyle (2014). Liquidity with high-frequency market making. *Working Paper*.

Hendershott, T., S. Li, A. Menkveld, and M. Seasholes (2014). Asset price dynamics with limited attention. *Working Paper*.

Heston, S. L., R. A. Korajczyk, and R. Sadka (2010). Intraday patterns in the cross-section of stock returns. *The Journal of Finance 65*(4), 1369–1407.

Hirschey, N. H. (2016). Do high-frequency traders anticipate buying and selling pressure? *Working paper*.

Hoffmann, P. (2014). A dynamic limit order market with fast and slow traders. *Journal of Financial Economics 113*(1), 156–169.

Huang, J. and J. Wang (2009). Liquidity and market crashes. *Review of Financial Studies 22*(7), 2607–2643.

Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics (Forthcoming)*.

Menkveld, A. J. and B. Yueshen (2012). Middlemen interaction and its effect on market quality. *Working Paper*.

Menkveld, A. J. and M. A. Zoican (2015). Need for speed? exchange latency and liquidity. *Working Paper, Timbergen Institute*.

Nagel, S. (2012). Evaporating liquidity. *Review of Financial Studies 25*(7), 2005–2039.

O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics 116*, 257–270.

Peress, J. and D. Schmidt (2015). Noise traders incarnate: Describing a realistic noise trading process. *Working paper*.

SEC (2010). Concept release on equity market structure. *Federal Register*.

Spiegel, M. (1998). Stock price volatility in a multiple security overlapping generations model. *Review of Financial Studies 11*(2), 419–447.

Vayanos, D. and J. Wang (2012). Liquidity and asset returns under asymmetric information and imperfect competition. *Review of Financial Studies 25*(5), 1339–1365.

Vives, X. (2008). *Information and learning in markets: the impact of market microstructure*. Princeton University Press.

A Appendix

The following is a standard results (see, e.g. Vives (2008), Technical Appendix, pp. 382–383) that allows us to compute the unconditional expected utility of market participants.

Lemma 2. *Let the n-dimensional random vector $z \sim N(0, \Sigma)$, and $w = c + b'z + z'Az$, where $c \in \mathbb{R}$, $b \in \mathbb{R}^n$, and A is a $n \times n$ matrix. If the matrix $\Sigma^{-1} + 2\rho A$ is positive definite, and $\rho > 0$, then*

$$E[-\exp\{-\rho w\}] = -|I + 2\rho\Sigma A|^{-1/2} \exp\{-\rho(c - \rho b'(\Sigma + 2\rho A)^{-1}b)\}.$$

Proof of Lemma 1

Denote by $\mu x_1^{FD} = \int_0^\mu x_1^{FD} di$, $(1-\mu)x_1^D = \int_\mu^1 x_1^D di$, and by x_1^L respectively the aggregate position of FDs, dealers and liquidity traders in the first period. Imposing market clearing yields:

$$\mu x_1^{FD} + (1-\mu)x_1^D + x_1^L = 0 \iff \mu\varphi_1^{FD}(p_1) + (1-\mu)\varphi^D(p_1) + b_1^L u_1 = 0. \qquad (\text{A.1})$$

At equilibrium the coefficients of traders' strategies are known, which implies that, provided $b_1^L \neq 0$ (which is true in equilibrium), p_1 is observationally equivalent to u_1 and that both FDs and dealers can retrieve u_1 from the price. Therefore, the information set of a FD and a dealer in the first period coincide and are given by $\Omega_1^{FD} = \Omega_1^D = \{u_1\}$. In the second period, denote by $\mu x_2^{FD} = \int_0^\mu x_2^{FD} di$ and by x_2^L, respectively the aggregate position of FDs and second period liquidity traders. Impose market clearing:

$$\mu(x_2^{FD} - x_1^{FD}) + (x_2^L - x_1^L) = 0,$$

and rearrange the first period market clearing condition as follows

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

Substitute the latter in the second period clearing equation to obtain

$$\mu x_2^{FD} + x_2^L + (1-\mu)x_1^D = 0. \qquad (\text{A.2})$$

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 $\{u_1, x_2^L\}$. $\qquad \square$

Proof of Proposition 1

From the proof of Proposition 2 (see (A.30)) we know that when $\tau_\eta \to \infty$, a unique equilibrium obtains where

$$\Lambda_1^* = \frac{1}{\tau_v(\gamma + (\mu\gamma + \gamma_1^L)(\mu\gamma + \gamma_2^L)(\gamma_2^L + \gamma)\tau_{u_2}\tau_v)}. \qquad (\text{A.3})$$

Substituting (A.3) in (4b) and solving for b_1^L yields

$$b_1^L = \gamma_1^L(\gamma_2^L + \gamma)(\gamma_2^L + \mu\gamma)\Lambda_1^* \tau_{u_2}\tau_v^2 - 1.$$

Furthermore, $\lim_{\tau_\eta \to \infty} \text{Var}[v - p_2|\Omega_2^L] = \tau_v^{-1}$, so that $\lim_{\tau_\eta \to \infty} \kappa = 1$, implying that (A.14a) and (A.14b) have a closed form solution.

To see why the expression for the second period price simplifies to (3a), note that when $\tau_\eta \to \infty$, $\beta_{u_1|s_{u_1}} \to 1$, while $\eta \xrightarrow{P} 0$, implying that $p_2 = -\Lambda_2^*(u_2 - (1 - \mu)\gamma\tau_v\Lambda_1^*u_1)$. $\qquad\square$

Proof of Corollary 1

Taking the limit for $\tau_{u_2} \to \infty$ in (A.3) yields the desired result. $\qquad\square$

Proof of Corollary 2

Differentiating (A.3) with respect to μ yields the desired result. $\qquad\square$

Proof of Proposition 2

We prove our argument by backward induction. In the second period a new mass of liquidity traders with risk-tolerance coefficient $\gamma_2^L > 0$ enters the market. A date-2 liquidity trader submits a market order

$$X_2^L(u_2, s_{u_1}) = b_{21}^L u_2 + b_{22}^L s_{u_1}, \tag{A.4}$$

with $u_2 \sim N(0, \tau_{u_2}^{-1})$, and $s_{u_1} = u_1 + \eta$, with $\eta \sim N(0, \tau_\eta^{-1})$ and u_2, η independent of all the other random variables in the model. Consider the sequence of market clearing equations

$$\mu x_1^{FD} + (1 - \mu)x_1^D + x_1^L = 0 \tag{A.5a}$$
$$\mu(x_2^{FD} - x_1^{FD}) + (b_{21}^L u_2 + b_{22}^L s_{u_1} - x_1^L) = 0. \tag{A.5b}$$

Rearrange (A.5a) as follows: $(1 - \mu)x_1^D = -\left(\mu x_1^{FD} + x_1^L\right)$. Substituting the latter in (A.5b) yields:

$$\mu x_2^{FD} + b_{21}^L u_2 + b_{22}^L s_{u_1} + (1 - \mu)x_1^D = 0. \tag{A.6}$$

Consider first Full Dealers. A FD maximizes the expected utility of his second period wealth:

$$E\left[-\exp\left\{-\frac{1}{\gamma}\left((p_2 - p_1)x_1^{FD} + (v - p_2)x_2^{FD}\right)\right\}|p_1, p_2\right] = \tag{A.7}$$

$$= \exp\left\{-\frac{1}{\gamma}(p_2 - p_1)x_1^{FD}\right\}\left(-\exp\left\{-\frac{1}{\gamma}\left(E[v - p_2|p_1, p_2]x_2^{FD} - \frac{(x_2^{FD})^2}{2\gamma}\text{Var}[v - p_2|p_1, p_2]\right)\right\}\right),$$

where the last expression in (A.7) is due to CARA and normality. For given x_1^{FD} the above is a concave function of the second period strategy x_2^{FD}. Solving the first order condition, yields

that a second period FD's strategy is given by $X_2^{FD}(p_1, p_2) = -\gamma\tau_v p_2$. Similarly, due to CARA and normality, in the first period a traditional market maker maximizes

$$E\left[-\exp\left\{-\frac{1}{\gamma}(v-p_1)x_1^D\right\}|p_1\right] = -\exp\left\{-\frac{1}{\gamma}\left(E[v-p_1|p_1]x_1^D - \frac{(x_1^D)^2}{2\gamma}\mathrm{Var}[v-p_1|p_1]\right)\right\}. \quad (A.8)$$

Hence, his strategy is given by $X_1^D(p_1) = -\gamma\tau_v p_1$. Substituting these strategies in (A.6) and solving for p_2 yields

$$p_2 = \lambda_2\left(b_{21}^L u_2 + b_{22}^L s_{u_1}\right) - \frac{1-\mu}{\mu}p_1, \quad (A.9)$$

where $\lambda_2 = 1/\mu\gamma\tau_v$. The assumption that first period liquidity traders' strategies are linear implies that $p_1 = -\Lambda_1 u_1$ (see below). As a consequence we can rewrite (A.9) as follows:

$$p_2 = \lambda_2(b_{21}^L u_2 + b_{22}^L s_{u_1}) + \lambda_2(1-\mu)\gamma\tau_v\Lambda_1 u_1. \quad (A.10)$$

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

$$E[-\exp\{-\pi_2^L/\gamma_2^L\}|\Omega_2^L] = -\exp\left\{-\frac{1}{\gamma}\left(E[\pi_2^L|\Omega_2^L] - \frac{1}{2\gamma}\mathrm{Var}[\pi_2^L|\Omega_2^L]\right)\right\}, \quad (A.11)$$

where $\Omega_2^L = \{u_2, s_{u_1}\}$, and $\pi_2^L \equiv (v-p_2)x_2^L + u_2 v$. Maximizing (A.11) with respect to x_2^L, and solving for the strategy yields:

$$X_2^L(u_2, s_{u_1}) = \gamma_2^L\frac{E[v-p_2|\Omega_2^L]}{\mathrm{Var}[v-p_2|\Omega_2^L]} - \frac{\mathrm{Cov}[v-p_2, v|\Omega_2^L]}{\mathrm{Var}[v-p_2|\Omega_2^L]}u_2, \quad (A.12)$$

where

$$E[v-p_2|\Omega_2^L] = -\left(\lambda_2(b_{21}^L u_2 + b_{22}^L s_{u_1}) + \frac{1-\mu}{\mu}\Lambda_1\frac{\tau_\eta}{\tau_\eta + \tau_{u_1}}s_{u_1}\right) \quad (A.13a)$$

$$\mathrm{Var}[v-p_2|\Omega_2^L] = \frac{\mu^2(\tau_{u_1}+\tau_\eta) + ((1-\mu)\Lambda_1)^2\tau_v}{\mu^2(\tau_{u_1}+\tau_\eta)\tau_v} \quad (A.13b)$$

$$\mathrm{Cov}[v-p_2, v|\Omega_2^L] = \frac{1}{\tau_v}. \quad (A.13c)$$

Substituting (A.13a), (A.13b), and (A.13c) in (A.12) and identifying coefficients yields $X_2^L(u_2, s_{u_1}) = b_{21}^L u_2 + b_{22}^L s_{u_1}$, where

$$b_{21}^L = -\frac{1}{\tau_v(\gamma_2^L\lambda_2 + \mathrm{Var}[v-p_2|\Omega_2^L])} = -\frac{\mu\gamma}{\gamma_2^L + \mu\gamma\kappa}, \quad (A.14a)$$

$$b_{22}^L = -\frac{\gamma_2^L\tau_\eta\lambda_2(1-\mu)\gamma\tau_v\Lambda_1}{(\tau_\eta + \tau_{u_1})(\gamma_2^L\lambda_2 + \mathrm{Var}[v-p_2|\Omega_2^L])}, \quad (A.14b)$$

and $\kappa \equiv \tau_v\mathrm{Var}[v-p_2|\Omega_2^L]$.

Consider now the first period. We start by characterizing the strategy of a FD. Substituting a FD's second period strategy in (A.7), rearranging and applying Lemma 2 yields the following

expression for the first period objective function of a FD:

$$E[U((p_2 - p_1)x_1^{FD} + (v - p_2)x_2^{FD})|u_1] = -\left(1 + \frac{\text{Var}[p_2|u_1]}{\text{Var}[v]}\right)^{-1/2} \times \tag{A.15}$$

$$\exp\left\{-\frac{1}{\gamma}\left(\frac{\gamma\tau_v}{2}\nu^2 + (\nu - p_1)x_1^{FD} - \frac{(x_1^{FD} + \gamma\tau_v\nu)^2}{2\gamma}\left(\frac{1}{\text{Var}[p_2|u_1]} + \frac{1}{\text{Var}[v]}\right)^{-1}\right)\right\},$$

where

$$\nu \equiv E[p_2|u_1] = \left(\lambda_2 b_{22}^L + \frac{1-\mu}{\mu}\Lambda_1\right)u_1 \tag{A.16a}$$

$$\text{Var}[p_2|u_1] = \lambda_2^2\left(\frac{(b_{21}^L)^2}{\tau_{u_2}} + \frac{(b_{22}^L)^2}{\tau_\eta}\right). \tag{A.16b}$$

Maximizing (A.15) with respect to x_1^{FD} and solving for the first period strategy yields

$$X_1^{FD}(p_1) = \frac{\gamma}{\text{Var}[p_2|u_1]}\nu - \gamma\left(\frac{1}{\text{Var}[p_2|u_1]} + \frac{1}{\text{Var}[v]}\right)p_1. \tag{A.17}$$

As we argued above, due to CARA and normality, for traditional market makers at date 1 we have $X_1^D(p_1) = -\gamma\tau_v p_1$. At equilibrium we then have

$$\mu\left(\frac{\gamma}{\text{Var}[p_2|u_1]}\nu - \gamma\left(\frac{1}{\text{Var}[p_2|u_1]} + \frac{1}{\text{Var}[v]}\right)p_1\right) + (-(1-\mu)\gamma\tau_v p_1) + b_1^L u_1 = 0,$$

implying that p_1 is linear in u_1: $p_1 = -\Lambda_1 u_1$, with Λ_1 to be determined.

We now turn to the characterization of first period liquidity traders' strategies. CARA and normality imply

$$E[-\exp\{-\pi_1^L/\gamma_1^L\}] = -\exp\left\{-\frac{1}{\gamma}\left(E[\pi_1^L|u_1] - \frac{1}{2\gamma_1^L}\text{Var}[\pi_1^L|u_1]\right)\right\}, \tag{A.18}$$

where $\pi_1^L \equiv (p_2 - p_1)x_1^L + u_1 p_2$. Maximizing (A.18) with respect to x_1^L, and solving for the optimal strategy, yields

$$X_1^L(u_1) = \gamma_1^L \frac{E[p_2 - p_1|u_1]}{\text{Var}[p_2 - p_1|u_1]} - \frac{\text{Cov}[p_2 - p_1, p_2|u_1]}{\text{Var}[p_2 - p_1|u_1]}u_1, \tag{A.19}$$

where

$$p_2 - p_1 = \left(\lambda_2 b_{22}^L + \frac{\Lambda_1}{\mu}\right)u_1 + \lambda_2\left(b_{21}^L u_2 + b_{22}^L \eta\right),$$

and

$$E[p_2 - p_1|u_1] = \left(\lambda_2 b_{22}^L + \frac{\Lambda_1}{\mu}\right)u_1 \tag{A.20a}$$

$$\text{Cov}[p_2 - p_1, p_2|u_1] = \text{Var}[p_2|u_1]. \tag{A.20b}$$

Substituting the above in (A.19) and identifying yields

$$X_1^L(u_1) = b_1^L u_1, \tag{A.21}$$

where

$$b_1^L = \gamma_1^L \frac{\mu\lambda_2 b_{22}^L + \Lambda_1}{\mu \text{Var}[p_2|u_1]} - 1. \tag{A.22}$$

Substituting (A.17), x_1^D, and (A.21) in the first period market clearing condition and solving for the price yields $p_1 = -\Lambda_1 u_1$, where

$$
\begin{aligned}
\Lambda_1 = \psi(\Lambda_1) &\equiv -\left(\mu\gamma\left(\frac{1}{\text{Var}[p_2|u_1]} + \frac{1}{\text{Var}[v]}\right) + (1-\mu)\gamma\frac{1}{\text{Var}[v]}\right)^{-1}\left(\mu\frac{\gamma\text{Cov}[p_2, u_1]}{\text{Var}[p_2|u_1]\text{Var}[u_1]} + b_1^L\right) \\
&= -\frac{\mu\gamma\text{Cov}[p_2, u_1]\tau_{u_1} + b_1^L\text{Var}[p_2|u_1]}{\gamma(\mu + \tau_v\text{Var}[p_2|u_1])}.
\end{aligned} \tag{A.23}
$$

According to (A.22), the equilibrium coefficient of a first period liquidity trader depends on b_{21}^L, and b_{22}^L, both of which in turn depend on Λ_1. Therefore, recursive substitution of the equilibrium strategies' coefficients in (A.23) shows that Λ_1 is pinned down by the solution of the following equation in Λ_1:

$$\psi(\Lambda_1) - \Lambda_1 = \frac{(\mu\gamma + \gamma_1^L)(\text{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1) + \text{Var}[p_2|u_1](\gamma\tau_v\Lambda_1 - 1)}{\gamma(\mu + \tau_v\text{Var}[p_2|u_1])} = 0. \tag{A.24}$$

For $\mu \in (0, 1]$ the denominator in the above expression is positive, which implies that equilibria are pinned down by solutions to the quintic at the numerator of (A.24), which is given by:

$$
\begin{aligned}
f(\Lambda_1) = {}& \gamma^2(1-\mu)^4\Lambda_1^5(\mu\gamma + \gamma_1^L)\tau_{u_2}\tau_v^4 \\
&+ \mu\gamma(1-\mu)^2\tau_{u_2}\tau_v^3\Lambda_1^3\big(((\mu\gamma + \gamma_1^L)(2\mu\gamma + \gamma_2^L(1+\mu)) + (\gamma_2^L)^2)\tau_\eta + 2\tau_{u_1}(\mu\gamma + \gamma_1^L)(\mu\gamma + \gamma_2^L)\big) \\
&- (\gamma_2^L)^2(1-\mu)^2\mu\tau_\eta\tau_{u_2}\tau_v^2\Lambda_1^2 \\
&+ \mu^2(\tau_\eta + \tau_{u_1})\tau_v\Lambda_1\big(\mu\gamma(\tau_\eta + \tau_{u_1}) + (\mu\gamma + \gamma_1^L)(\mu\gamma + \gamma_2^L)(\mu\gamma(\tau_\eta + \tau_{u_1}) + \gamma_2^L(\mu\tau_\eta + \tau_{u_1}))\tau_{u_2}\tau_v\big) \\
&- \mu^3(\tau_\eta + \tau_{u_1})^2 = 0.
\end{aligned} \tag{A.25}
$$

The above expression shows that there are three sign changes in the sequence formed by the quintic's coefficients. Therefore, by Descartes' rule of sign, there are up to three positive roots of the equation $f(\Lambda_1) = 0$.

Compute $\text{Cov}[p_2, u_1]$:

$$\text{Cov}[p_2, u_1] = \frac{(1-\mu)\Lambda_1(\tau_{u_1}\gamma_2^L\lambda_2 + (\tau_{u_1} + \tau_\eta)\text{Var}[v - p_2|\Omega_2^L])}{\mu\tau_{u_1}(\tau_{u_1} + \tau_\eta)(\gamma_2^L\lambda_2 + \text{Var}[v - p_2|\Omega_2^L])}, \tag{A.26}$$

which is positive if and only if $\Lambda_1 > 0$. Consider (A.24) and suppose that at equilibrium $\Lambda_1^* < 0$. From (A.26), $\text{Cov}[p_2, u_1] < 0$. Due to (A.24) this implies $f(\Lambda_1^*) < 0$, which is impossible. Thus, at equilibrium, $\Lambda_1^* > 0$, and $\text{Cov}[p_2, u_1] \geq 0$.

To sign the strategy coefficient of a first period liquidity trader, we use (A.22):

$$b_1^L = \gamma_1^L \frac{\text{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1}{\text{Var}[p_2|u_1]} - 1. \tag{A.27}$$

From (A.27) we obtain $\text{Var}[p_2|u_1](1+b_1^L)/\gamma_1^L = \text{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1$, which substituted in (A.24) yields

$$f(\Lambda_1) = \frac{\text{Var}[p_2|u_1]}{\gamma_1^L} \left((\mu\gamma + \gamma_1^L)(1 + b_1^L) + \gamma_1^L(\gamma\tau_v\Lambda_1 - 1) \right) = 0. \tag{A.28}$$

Solving the above for Λ_1 yields: $\Lambda_1^* = (-\mu\gamma - b_1^L(\mu\gamma + \gamma_1^L))/\gamma_1^L\gamma\tau_v$. Since $\Lambda_1^* > 0$, the previous equation implies that at equilibrium $b_1^L < 0$. Furthermore, using (A.27), $1 + b_1^L > 0$, which proves our result. Finally, defining $\Lambda_2^* \equiv -\lambda_2 b_{21}^L$ and rearranging (A.10) yields (16). Taking the limit for $\tau_\eta \to \infty$ in $\psi(\Lambda_1)$ yields:

$$\lim_{\tau_\eta \to \infty} \psi(\Lambda_1) = \frac{1 - \Lambda_1(\gamma_2^L + \mu\gamma)(\gamma_1^L(\gamma + \gamma_2^L) + \mu\gamma^2(1-\mu))\tau_{u_2}\tau_v^2}{\gamma\tau_v(1 + \mu(\gamma_2^L + \mu\gamma)^2\tau_{u_2}\tau_v)}. \tag{A.29}$$

Identifying Λ_1 yields

$$\Lambda_1^* = \frac{1}{\tau_v(\gamma + (\mu\gamma + \gamma_1^L)(\mu\gamma + \gamma_2^L)(\gamma_2^L + \gamma)\tau_{u_2}\tau_v)}. \tag{A.30}$$

$\square$

Proof of Corollary 3

With a transparent market, solving for b_{21}^L yields $b_{21}^L = -\mu\gamma/(\gamma_2^L + \mu\gamma)$. When the market is opaque, instead, replacing Λ_2^* in (17a) and solving for b_{21}^L yields $b_{21}^L = -\mu\gamma/(\gamma_2^L + \mu\gamma\kappa)$. Since $\kappa > 1$, the result follows. $\square$

Proof of Corollary 4

From Proposition 2 it is immediate that $\text{Var}[v - p_2|\Omega_2^L]$ is increasing in Λ_1. Differentiating $\text{Cov}[p_2, u_1]$ yields

$$\frac{\partial \text{Cov}[p_2, u_1]}{\partial \Lambda_1} = \frac{1 - \mu}{\mu(\tau_{u_1} + \tau_\eta)} +$$

$$\frac{(1-\mu)\tau_\eta(\text{Var}[v - p_2|\Omega_2^L](\gamma_2^L\lambda_2 + \text{Var}[v - p_2|\Omega_2^L]) + \gamma_2^L\lambda_2\Lambda_1\text{Var}[v - p_2|\Omega_2^L])}{\mu\tau_{u_1}(\tau_{u_1} + \tau_\eta)(\gamma_2^L\lambda_2 + \text{Var}[v - p_2|\Omega_2^L])^2} \geq 0,$$

for $\mu \leq 1$. Finally, from Proposition 2 we know that b_{21}^L is increasing in κ, and thus in $\text{Var}[v - p_2|\Omega_2^L]$. This implies that b_{21}^L is increasing in Λ_1. $\square$

The equilibrium quintic (A.25) can be expressed as the sum of two polynomials: a quintic in Λ_1^* that multiplies τ_{u_2}, and a first degree polynomial in Λ_1^* that does not depend on τ_{u_2}, as shown in the expression below:

$$f(\Lambda_1) = \Big[\Lambda_1(\tau_{u_1}(\gamma_2^L\lambda_2 + \mathrm{Var}[v - p_2|\Omega_2^L]) + \tau_\eta(\mu\gamma_2^L\lambda_2 + \mathrm{Var}[v - p_2|\Omega_2^L]))(\mu\gamma + \gamma_1^L)\times \tag{A.31}$$
$$(\tau_{u_1} + \tau_\eta)(\gamma_2^L\lambda_2 + \mathrm{Var}[v - p_2|\Omega_2^L])\tau_\eta + \lambda_2^2(\gamma_2^L\tau_\eta\lambda_2(1 - \mu)\gamma\tau_v\Lambda_1)^2(\gamma\tau_v\Lambda_1 - 1)\mu\Big]\tau_{u_2}\tau_v^2$$
$$+ \lambda_2^2(\tau_{u_1} + \tau_\eta)^2\mu\tau_\eta(\gamma\tau_v\Lambda_1 - 1).$$

Inspection of the equilibrium mapping $\psi(\Lambda_1)$ shows that if we let $\tau_{u_2} \to \infty$, the corresponding equilibrium quintic is proportional to the term in square brackets in (A.31) (i.e., the one that multiplies τ_{u_2}). We first concentrate on the analysis of this quintic:

$$\hat{f}(\Lambda_1) = \Lambda_1(\tau_{u_1}(\gamma_2^L\lambda_2 + \mathrm{Var}[v - p_2|\Omega_2^L]) + \tau_\eta(\mu\gamma_2^L\lambda_2 + \mathrm{Var}[v - p_2|\Omega_2^L]))(\mu\gamma + \gamma_1^L)\times \tag{A.32}$$
$$(\tau_{u_1} + \tau_\eta)(\gamma_2^L\lambda_2 + \mathrm{Var}[v - p_2|\Omega_2^L])\tau_\eta + \lambda_2^2(\gamma_2^L\tau_\eta\lambda_2(1 - \mu)\gamma\tau_v\Lambda_1)^2(\gamma\tau_v\Lambda_1 - 1)\mu.$$

First, note that $\hat{f}(0) = 0$, implying that when $\tau_{u_2} \to \infty$, $\Lambda_1^* = 0$ is an equilibrium of the model. Additionally, considering $h(\Lambda_1) \equiv \hat{f}(\Lambda_1)/\Lambda_1$, a quartic in Λ_1, we can pin down parameter restrictions that ensure the existence of two additional equilibria. To see this, we start by evaluating $h(\cdot)$ at $\Lambda_1 = 0$ obtaining:

$$h(0) = \frac{\tau_\eta(\mu\gamma + \gamma_1^L)(\mu\gamma + \gamma_2^L)(\tau_{u_1} + \tau_\eta)(\mu\gamma(\tau_{u_1} + \tau_\eta) + \gamma_2^L(\tau_{u_1} + \mu\tau_\eta))}{\gamma^2\mu^2} > 0. \tag{A.33}$$

Next, evaluating $h(\cdot)$ at the point $\bar{\Lambda}_1^* = 1/(1 - \mu)$, yields

$$h(\bar{\Lambda}_1^*) = \frac{\tau_\eta}{\gamma^2\tau_\eta^4}\times \tag{A.34}$$
$$\Big((\mu\gamma + \gamma_1^L)(\mu(\mu\gamma + \gamma_2^L)(\tau_{u_1} + \tau_\eta) + \gamma\tau_v)(\mu\gamma_2^L\mu(\tau_{u_1} + \mu\tau_\eta) + \gamma(\mu^2(\tau_\eta + \tau_{u_1}) + \tau_v))$$
$$- \mu(\gamma_2^L)^2\tau_\eta(1 - \mu - \gamma\tau_v)\Big),$$

which is negative when the following parameter restrictions are satisfied:

$$0 < \mu < \bar{\mu} \equiv \frac{\gamma_2^L(\sqrt{5\gamma^2 + \gamma_2^L(2\gamma + \gamma_2^L)} - (\gamma + \gamma_2^L))}{2\gamma^2} \tag{A.35a}$$

$$0 < \tau_v < \bar{\tau}_v \equiv \frac{(1 - \mu)(\gamma_2^L)^2 - \mu\gamma(\mu\gamma + \gamma_2^L)}{\gamma(\gamma_2^L)^2} \tag{A.35b}$$

$$\tau_\eta > \underline{\tau}_\eta \equiv \frac{\gamma(\mu(\mu\gamma + \gamma_2^L)\tau_{u_1} + \gamma\tau_v)}{(1 - \mu - \gamma\tau_v)(\gamma_2^L)^2 - \mu\gamma(\mu\gamma + \gamma_2^L)} \tag{A.35c}$$

$$0 < \gamma_1^L < \bar{\gamma}_1^L \equiv \frac{\mu((\gamma_2^L)^2\tau_\eta(1 - \mu - \gamma\tau_v) - \mu\gamma\gamma_2^L(\tau_{u_1} + \tau_\eta) - \gamma^2(\tau_v + \mu^2(\tau_{u_1} + \tau_\eta)))}{\gamma\tau_v + \mu(\mu\gamma + \gamma_2^L)(\tau_{u_1} + \tau_\eta)}. \tag{A.35d}$$

Therefore, when (A.35a)-(A.35d) hold, two additional equilibria exist $(\Lambda_1^*)^I \in (0, \bar{\Lambda}_1^*)$, and $(\Lambda_1^*)^L \in (\bar{\Lambda}_1^*, 1/\gamma\tau_v)$. This establishes that in the case $\tau_{u_2} \to \infty$, when (A.35a)-(A.35d) hold, three equilibria: $0 < (\Lambda_1^*)^I < \bar{\Lambda}_1^* < (\Lambda_1^*)^L < 1/\gamma\tau_v$, arise. Consider now the general quintic (A.31). First, note that

$$f(0) = -\frac{\tau_\eta(\tau_\eta + \tau_{u_1})^2}{\mu(\gamma\tau_v)^2} < 0. \tag{A.36}$$

Next, evaluating $f(\cdot)$ at $\underline{\Lambda}_1^* = \mu/(1-\mu) < \bar{\Lambda}_1^*$ yields

$$f(\underline{\Lambda}_1^*) = \frac{\tau_\eta}{\mu\gamma^2}\left\{\frac{(\tau_\eta + \tau_{u_1})^2(\mu\gamma\tau_v - (1-\mu))}{(1-\mu)\tau_v^2} + \frac{\tau_{u_2}}{1-\mu}\left((\gamma_2^L)^2\tau_\eta(\mu\gamma\tau_v - (1-\mu)) + \right.\right. \tag{A.37}$$

$$\left.\left. + (\mu\gamma + \gamma_1^L)(\gamma_2^L(\tau_{u_1} + \tau_\eta) + \mu\gamma(\tau_{u_1} + \tau_\eta + \tau_v))(\gamma_2^L(\tau_{u_1} + \mu\tau_\eta) + \mu\gamma(\tau_{u_1} + \tau_\eta + \tau_v))\right)\right\}.$$

The sign of (A.37) is determined by the sign of the expression inside the curly brackets. As $\underline{\Lambda}_1^* < 1/\gamma\tau_v$, the term $\mu\gamma\tau_v - (1-\mu) < 0$. Also, by inspection, the expression within parentheses is positive provided that

$$\gamma > \underline{\gamma} \equiv \frac{(1-\mu)\tau_\eta - \gamma_1^L(\tau_{u_1} + \tau_\eta)(\tau_{u_1} + \mu\tau_\eta)}{\mu(\tau_v\tau_\eta + (\tau_{u_1} + \tau_\eta)(\tau_{u_1} + \mu\tau_\eta))}. \tag{A.38}$$

Hence, if (A.38) holds, and

$$\tau_{u_2} > \underline{\tau}_{u_2} \equiv \left((\tau_v^2((\gamma_2^L)^2(\tau_\eta(\mu\gamma\tau_v - (1-\mu)) + (\mu\gamma + \gamma_1^L)(\tau_1 + \tau_\eta)(\tau_{u_1} + \mu\tau_\eta)) + \mu\gamma\gamma_2^L\times\right.$$

$$(\mu\gamma + \gamma_1^L)(\tau_1 + \tau_v + \tau_\eta)(2\tau_{u_1} + \tau_\eta(1+\mu)) + (\mu\gamma + \gamma_1^L)(\mu\gamma)^2(\tau_{u_1} + \tau_\eta + \tau_v)^2)))^{-1}\times$$

$$\left((\tau_{u_1} + \tau_\eta)^2(1 - \mu - \mu\gamma\tau_v)\right), \tag{A.39}$$

expression (A.37) is positive. This establishes the existence of an equilibrium $0 < (\Lambda_1^*)^H < \underline{\Lambda}_1^*$. Finally, provided (A.35a)-(A.35d) hold, $f(\bar{\Lambda}_1^*) < 0$, since $\hat{f}(\bar{\Lambda}_1^*) < 0$ and $\gamma\tau_v/(1-\mu) < 1$. Therefore, we can conclude that when $\tau_{u_2} < \infty$, if (A.35a)-(A.35d), and (A.38), (A.39) hold, the model displays three equilibria:

$$0 < (\Lambda_1^*)^H < \underline{\Lambda}_1^* < (\Lambda_1^*)^I < \bar{\Lambda}_1^* < (\Lambda_1^*)^L < \frac{1}{\gamma\tau_v}. \tag{A.40}$$

$\square$

Proof of Corollary 5

See the proof of Proposition 3. $\square$

Proof of Corollary 6

Taking the limit for $\tau_{u_2} \to 0$ in the equilibrium quintic (A.25), and collecting terms yields

$$\lim_{\tau_{u_2} \to 0} f(\Lambda_1) = \mu^3(\tau_{u_1} + \tau_\eta)^2(\gamma\tau_v\Lambda_1 - 1) = 0. \tag{A.41}$$

The above equation has a unique real solution $\Lambda_1^* = 1/\gamma\tau_v$. Substituting this solution in (4b) implies that at this equilibrium $b_1^L = -1$. $\qquad\square$

<div align="center"><i>Proof of Corollary 7</i></div>

If we let $\tau_\eta \to 0$, (A.25) reads as follows:

$$f(\Lambda_1) = -\mu^3\tau_{u_1}^2(1 - \gamma\tau_v\Lambda_1) + \Lambda_1(\gamma_1^L + \mu\gamma)\tau_{u_2}\tau_v^2(\gamma\tau_v\Lambda_1^2(1 - \mu)^2 + \mu\tau_{u_1}(\gamma_2^L + \mu\gamma))^2 = 0. \tag{A.42}$$

By inspection, $f(0) < 0$, $f(1/\gamma\tau_v) > 0$, and $f'(\Lambda_1) > 0$, which implies that in this case there exists a unique positive root. Evaluating (A.42) at (A.30) yields $f(\Lambda_1^*) > 0$ which implies that a strongly opaque market is more liquid than a fully transparent one. $\qquad\square$
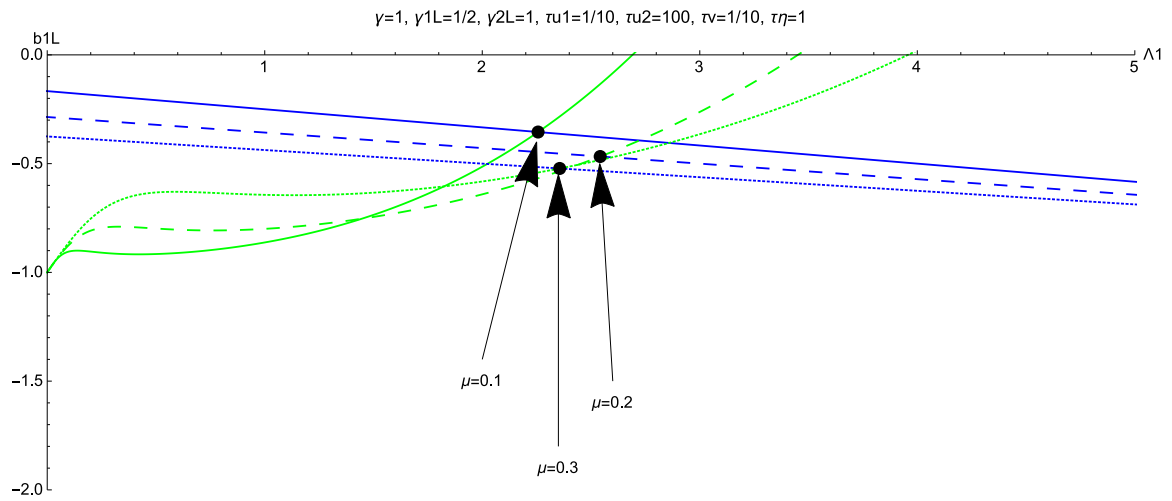
Figure 11: When the market is opaque, an increase in μ can have a non-monotone effect on market liquidity.
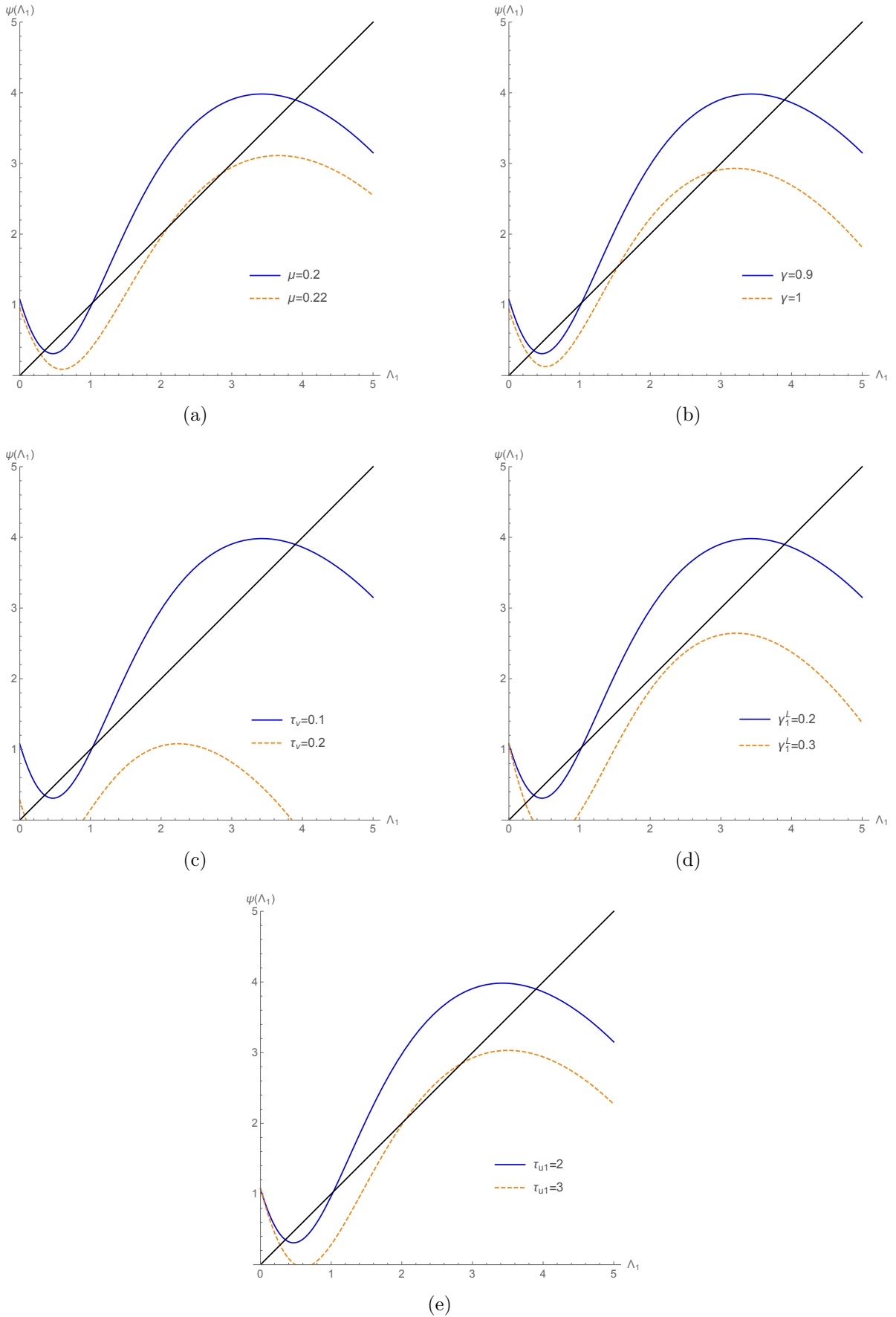
Figure 12: Comparative statics. The figure displays the effect of a change in the proportion of FD (Panel (a)), the risk tolerance of D and FD (Panel (b)), the volatility of the payoff (Panel (c)), the risk tolerance of first period traders (Panel (d)), and the volatility of the first period endowment shock (Panel (e)). The continuous blue curve represents the function $\psi(\Lambda_1)$ for the set of parameters' values used in Figure 6.
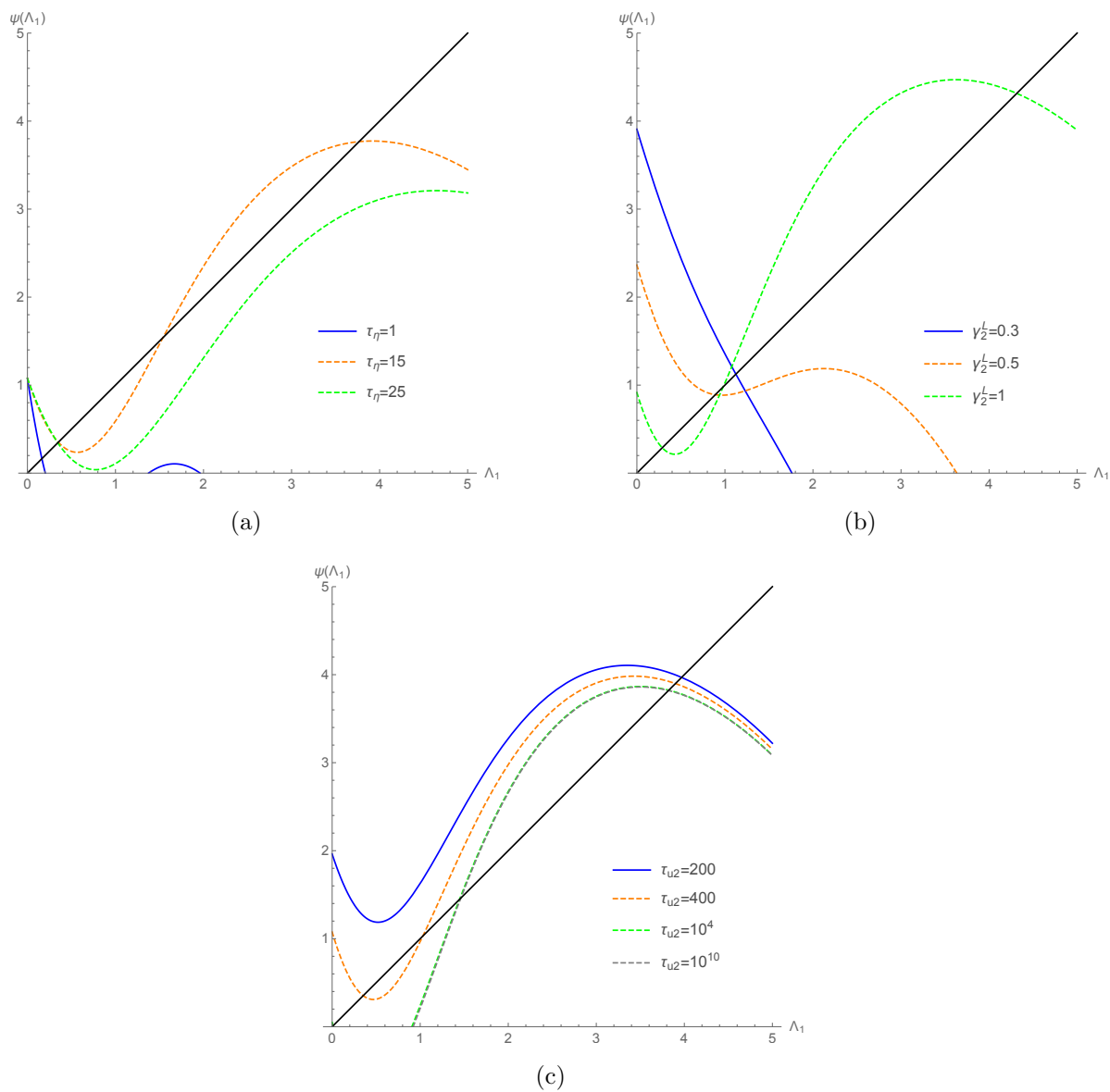
Figure 13: Comparative statics. The figure displays the effect of a change in the precision of second period liquidity traders' signal (Panel (a)), second period liquidity traders' risk tolerance (Panel (b)), and the precision of second period liquidity traders' endowment shock (Panel (c)). The other parameters' values are as in Figure 6.

Giovanni Cespa
Cass Business School, City, University of London, and CEPR; email: giovanni.cespa@gmail.com

Xavier Vives
IESE Business School Barcelona.