Market Transparency and Fragility*

Giovanni Cespa† and Xavier Vives‡

February, 2019

Abstract

We show that dealers' limited market participation, coupled with an informational friction resulting from lack of market transparency, can make liquidity demand upward sloping, inducing strategic complementarities: traders demand more liquidity when the market becomes less liquid, fostering market illiquidity. This can generate instability with an initial dearth of liquidity degenerating into a liquidity rout (as in a flash crash). In a fully transparent market, liquidity is increasing in the proportion of dealers continuously present in the market; however, in a less transparent market, liquidity can be U-shaped in this proportion and in the degree of transparency.

Keywords: Market fragmentation, liquidity fragility, flash crash, asymmetric information.
JEL Classification Numbers: G10, G12, G14

*A previous version of this paper circulated with the tile "High Frequency Trading and fragility." For helpful comments we thank Bruno Biais, Evangelos Benos, Thierry Foucault, Denis Gromb, Albert Menkveld, Sophie Moinas, Andreas Park, Joël Peress, Laura Veldkamp, Liyan Yang, Bart Yueshen, and seminar participants at INSEAD, HEC (Paris), Rotterdam School of Management, Stockholm School of Economics, the 9th Annual Central Bank Workshop on Microstructure (Frankfurt, 9/13), the conference on High Frequency Trading at Imperial College, Brevan Howard Centre (London, 12/14), the Workshop on Microstructure Theory and Applications (Cambridge, 3/15), the third workshop on Information Frictions and Learning (Barcelona, 6/15), the Bank of England, the AFA (San Francisco, 1/16), and the Wharton Conference on Liquidity and Financial Crises with Pete Kyle as discussant (Philadelphia, 10/16). Cespa acknowledges financial support from the Bank of England (grant no. RDC151391). The first version of this paper was prepared by Vives under the Wim Duisenberg Research Fellowship Program sponsored by the ECB. Any views expressed are only those of the authors and do not necessarily represent the views of the ECB or the Eurosystem.

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

‡IESE Business School, Avinguda Pearson, 21 08034 Barcelona, Spain. e-mail: xvives@iese.edu

"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.[1] However, episodes of extreme market turbulence where liquidity seems to inexplicably disappear, are by far not a staple of modern markets. On the contrary, as the stock market crash of October 19, 1987 makes clear, (apparently) fundamentals-unrelated crashes have been a worrying, regular feature of financial markets.[2]

A unifying feature of these episodes seems to be the jamming of the "rationing" function of market illiquidity. In normal market conditions, traders perceive a lack of liquidity as a cost, while arbitrageurs regard it as an opportunity. This, in turn, leads the former to limit their demand for immediacy, and the latter to increase their supply of liquidity (i.e., traders' demand for liquidity, and arbitrageurs' "liquidity supply" are respectively decreasing and increasing in the illiquidity of the market). In normal conditions, then, an illiquidity hike leads traders to contain their demand, and arbitrageurs to take advantage of profitable opportunities, easing the pressure on liquidity suppliers, thereby producing a stabilizing effect on the market. However, on occasions, a bout of illiquidity, which can hardly be construed as fundamentals driven, has a destabilizing impact, and fosters a disorderly "run for the exit" that is conducive to a rout. In these cases, traders attempt to place orders *despite* the liquidity shortage, and arbitrageurs flee the market, foregoing profitable (but risky) opportunities. In such conditions, liquidity is fragile. What can account for such a dualistic feature of market illiquidity?

In this paper, we argue that liquidity supply fragmentation, *and* lack of transparency about relevant market conditions, are important ingredients in the answer to this question. In current markets, trading automation arguably boosts liquidity supply fragmentation by limiting the market participation of some liquidity suppliers, or making it endogenous (Duffie (2010) and SEC (2010));[3] at the same time, computerized trading creates informational frictions by

[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] See https://en.wikipedia.org/wiki/List_of_stock_market_crashes_and_bear_markets.

[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

hampering some traders' access to reliable and timely market information (Ding et al. (2014)).[4] In less automated markets, liquidity supply fragmentation, and impaired access to market information arose because of different reasons. For example, in the 80s, access to the NYSE trading floor was crucial to have a good view of market conditions, but obviously constrained by physical limitations. Importantly, such frictions seem to have a bearing on episodes of liquidity crashes. Several accounts of the August 24, 2015 "flash-crash," point to the fact that uncertainty over the price of ETF constituents contributed to a huge investors' sellout, and sidelined the actions of arbitrageurs, exacerbating the liquidity dry-up in some ETFs.[5] In a somewhat similar fashion, Amihud et al. (1990), in their analysis of the 1987 "Black Monday," argue that a number of operational issues affected the opening trade session on the day of the event "[O]rders could not be executed, and information on market conditions, and on order execution was delayed." This, impaired the ability of traders outside of the market to provide liquidity, restricting total liquidity supply.

We study a model where two classes of risk-averse dealers provide liquidity to two cohorts of risk-averse, short-term traders, in a two-period market. Traders enter the market to partially hedge their exposure to a risky endowment that is correlated to the risky asset traded in the 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 one class of dealers, named 'full,' is able to participate. Full dealers 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 an informational friction, some liquidity traders are 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

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. "[n]ot 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."

[5] In the morning of August 24, 2015, the Dow dropped roughly 1,100 points in the first five minutes of trading, and trading in several stocks was halted due to unusual market turbulence. The ensuing lack of reliable price information allowed profitable, but risky, arbitrage opportunities to go unexploited, leading to a widening of spreads and a thinning of market depth. For example, during the event, the spread between the SPDR S&P500 (SPY) and the Guggenheim S&P 500 Equal Weight ETF (RSP), two very similar ETFs whose prices are normally in sync, at one point reached $21 (see What The E-T-F Happened On August 24?). In a similar vein, 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."

is because when first period traders load their positions–for example selling to hedge a positive endowment shock–a part of their orders is absorbed by standard dealers. These agents, however, are not in the market in the second period, when first period traders unwind. As a consequence, an order imbalance (induced by first period traders' unwinding orders and) affecting the second period price, arises. This has three important implications. First, as full dealers are in the market at both rounds, they exploit the predictable future price pressure, by leaning against traders' orders at the first round. This activity, which amounts to taking a directional bet, contributes to liquidity provision. Second, as standard dealers are unable to rebalance in the second period, they require a larger price concession to absorb traders' orders. Thus, when liquidity dries up, standard dealers absorb a larger imbalance, magnifying the propagation effect. Third, as second period traders observe a signal about the first period trading imbalance, they trade against the induced price pressure. This speculative activity, potentially adds to full dealers' second period liquidity provision, thereby easing the market impact of the propagated imbalance. The workings of the model crucially 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, and their speculative activity contributes to offset the propagated price pressure, inducing a stabilizing impact on the market. 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), and full dealers' directional bets are instead an increasing function of illiquidity: the less liquid is the market, the higher is the cost traders incur to reduce exposure, but the more profitable it is for full dealers to take a directional bet. Furthermore, illiquidity is increasing in traders' hedging aggressiveness (the inverse supply for liquidity is upward sloping). This is because lower hedging aggressiveness, and larger directional bets, shrink the order imbalance dealers need to absorb, allowing for cheaper liquidity provision. Thus, illiquidity in this case has a direct, "rationing" effect on traders' liquidity consumption, and a unique equilibrium arises. Additionally, 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 first period 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.[6] This can lead first period traders to consume more

[6]Endowed with a noisy signal, second period traders can end up selling (buying) when first period traders unwind a long (short) position, magnifying market impact.

liquidity, and full dealers to scale down the size of their directional bets (as holding exposure to the asset becomes riskier). In turn, this magnifies the imbalance dealers need to absorb (as their inventory of the risky asset increases), leading to an illiquidity hike, 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 lack of transparency 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 some degree of 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 lack of market market transparency 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 (corresponding to a technological 'glitch,' in modern markets, or the denial of responsibility that many dealers who refused to answer clients' call during the 1987 crash seemed to display).[7] 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, and an increase in the uncertainty first period traders face on their endowment value. In both cases, *small* parameter shocks produce *large* liquidity withdrawals. Interestingly, the former example is consistent with some narratives of the flash crash (see e.g. Easley et al. (2011)). The latter, appears instead to be in line with the events that occurred on August 24, 2015.[8]

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 lack of market transparency, 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

[7] According to a Fortune 500 CFO: "When the sell orders poured in, our specialist ran into a cave." See http://fortune.com/2012/10/21/the-great-crash-what-happened-and-whats-next-fortune-1987/

[8] See https://www.investopedia.com/articles/investing/011116/two-biggest-flash-crashes-2015.asp

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.

The rest of the paper is organized as follows. In the next section we review the literature related to our paper. We then 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. In Section 5, 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 Literature review

This paper is related to five 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 overlapping generations (OLG) model in which investors live for two periods. Our setup, in contrast, considers an economy with a finite number of trading rounds. Interestingly, our fragility result bears some resemblance with Gennotte and Leland (1990), in that fragility arises because of the change in slope of a demand curve. Note, however, that Gennotte and Leland (1990) analyze a market where non-linearity of the demand from portfolio insurers and sufficient uncertainty about its magnitude, may make the aggregate demand for the security upward sloping in its *price*. In this situation, *fundamentals-informed* traders may mistakenly attribute a price decline to fundamentals, and start a selling pressure that triggers a price crash. In our setup, in contrast, no trader has information on the *fundamentals*, and liquidity can rapidly deteriorate

when traders' demand becomes upward sloping in *illiquidity*.

Second, the paper is related to the literature on market transparency. A central tenet of this literature is that of the different impact of pre- and post-trade transparency on market liquidity. In this respect, it is generally accepted that visibility of competing quotes or incoming orders–two distinct dimensions of pre-trade transparency–improves market liquidity; conversely, order flow visibility–which instead relates to post-trade transparency–can have a mixed effect on the time series of market liquidity (see, Foucault et al. (2013)). Our paper adds to the debate on the effects of post-trade transparency showing that when traders' information on past order flows is noisy, the market can become fragile, but also arguing that transparency can have a non-monotone effect on market liquidity.[9]

Third, the paper relates to the literature that assesses the impact of limited market participation. Heston et al. (2010) and Bogousslavsky (2016) 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, 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 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.

Finally, the paper also adds to the theoretical literature on the impact of high frequency trading (HFT) on market performance. Indeed, our full dealers can be seen as HFT who can constantly monitor the market. In this respect, we show that an informational friction arising from liquidity provision fragmentation can be responsible for liquidity fragility, reversing the positive association between computerized trading and liquid markets (see, e.g., Hendershott et al. (2010)).[10] Additionally, our paper provides an alternative interpretation of some narrative

[9]Importantly, as none of the traders in our model has fundamentals information, transparency and opaqueness relate to information about liquidity-motivated order imbalances, differently from e.g., Banerjee et al. (2018).

[10]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 (2017), Bongaerts and Van Achter (2015), Foucault

of the May 10, 2010 "flash crash." 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 differential information on imbalances. 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 HFTs 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.

3 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, "Full Dealers" (denoted by FD) in the interval $(0, \mu)$, are active at both dates. Second, competitive, risk-averse "Standard Dealers" (SD) 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.

3.1 Liquidity providers

A FD has CARA preferences with risk-tolerance γ, and submits price-contingent orders x_t^{FD}, $t = 1, 2$, to maximize the expected utility of his final wealth: $W^{FD} = (v - p_2)x_2^{FD} + (p_2 - p_1)x_1^{FD}$.[11] A SD 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^{SD} to maximize the expected utility of his wealth $W^{SD} = (v - p_1)x_1^{SD}$. The inability of SD to trade in the second period captures some liquidity suppliers' limited market participation. This friction could be due to technological reasons as in the case of dealers with impaired access to a technology that allows trading at high frequencies. Alternatively, it could arise from limited access to the trading venue, as in the case of those liquidity suppliers who in the 80s could not access the NYSE trading floor.

et al. (2016), and Menkveld and Zoican (2017); see O'Hara (2015) and Menkveld (2016) for literature surveys)

[11] 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.

3.2 Short-term traders

Short-term traders represent the liquidity demand side of the model. In the first period, a unit mass of short-term, CARA traders with risk-tolerance γ_1 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 takes a position (a market order) x_1, anticipating that it will unwind it in the following period, and leave the market. Formally, a trader maximizes the expected utility of his wealth $\pi_1 = u_1 p_2 + (p_2 - p_1)x_1$:

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

where $\Omega_1 \equiv \{u_1\}$ 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, and observing a noisy signal of the previous period endowment shock $s_{u_1} = u_1 + \eta$. A second period trader has a CARA utility function with risk-tolerance γ_2, and submits a market order to maximize the expected utility of his wealth $\pi_2 = u_2 v + (v - p_2)x_2$:

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

where $\Omega_2 \equiv \{u_2, s_{u_1}\}$ denotes his information set.[12] We assume $u_t \sim N(0, \tau_{u_t}^{-1})$, $\eta \sim N(0, \tau_\eta^{-1})$ and $\text{Cov}[u_t, v] = \text{Cov}[u_t, \eta] = \text{Cov}[u_1, u_2] = 0$, $t = 1, 2$.

The assumption of a random endowment whose value is perfectly correlated with that of the risky asset traded in the market is standard in the literature (see, e.g., Huang and Wang (2009) and Vayanos and Wang (2012)).[13] Practically, it captures instances in which a trader uses a security that replicates his endowment to hedge his risk exposure. For example, the endowment could be a portfolio of stocks that tracks the market, say a fund, and the trader could hedge his exposure using a market-tracking ETF; alternatively, the endowment could represent a position on a S&P500 ETF, like the SPY, and the hedging instrument would be the Emini (while the former trades from 6am to 8pm, including extended trading hours, the latter trades 24/7, thus allowing overnight hedging).[14] Importantly, in our setup the price at which first period traders unwind their position depends on the market conditions prevailing at the second trading round, and is therefore endogenous. This means that when unwinding the ETF position at $t = 2$, the trader cashes $p_2 x_1$ while valuing the endowment at the ETF price p_2.

To simplify notation, in the following we denote by $E_t^k[Y]$, and $\text{Var}_t^k[Y]$, $k \in \{FD, SD\}$, the

[12] 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.

[13] These authors 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.

[14] For an example involving SPY, see https://money.stackexchange.com/questions/54373/why-dont-spy-spx-and-the-e-mini-sp-500-track-perfectly-with-each-other, and http://tastytradenetwork.squarespace.com/tt/blog/equating-futures-to-etfs, and for other ETF related examples, see https://investorplace.com/2017/10/portfolio-hedge-fund-consider-etfs/.

conditional expectation and variance that a k-type dealer forms about Y, in period $t = 1, 2$. Similarly, $E_1[Y]$, and $\text{Var}_1[Y]$ denote the conditional expectation and variance that a period-t trader forms about Y.

3.3 Market clearing conditions

We restrict attention to linear equilibria where prices reflect traders' information and endowment shocks:

$$p_1 = -\Lambda_1 u_1 \tag{1a}$$

$$p_2 = -\Lambda_2 u_2 - \Lambda_{21} u_1 - \Lambda_{22} \eta, \tag{1b}$$

and Λ_1, Λ_2, Λ_{21}, Λ_{22} are coefficients which will be pinned down at equilibrium.

The intuition for (1a) and (1b) is as follows. At equilibrium the orders of first period traders are absorbed by both FDs and SDs:

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

Traders know u_1, while, at equilibrium, dealers infer it from the price, which justifies (1a).

Consider now the second period equilibrium condition. When $\mu < 1$, FDs' aggregate position falls short of x_1: $\mu x_1^{FD} + x_1 \neq 0$. Thus, FDs second period inventory 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

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

Substituting (2) in the latter and rearranging yields:

$$\mu x_2^{FD} + x_2 + (1-\mu) x_1^{SD} = 0. \tag{3}$$

At equilibrium, dealers' and traders' strategies are a function of their information sets–$\{p_1, p_2\}$ for FDs, p_1 for SDs and Ω_2 for second period traders. As a consequence, the price will load on $\{u_1, u_2, \eta\}$, justifying (1b). Figure 1 displays the timeline of the model.

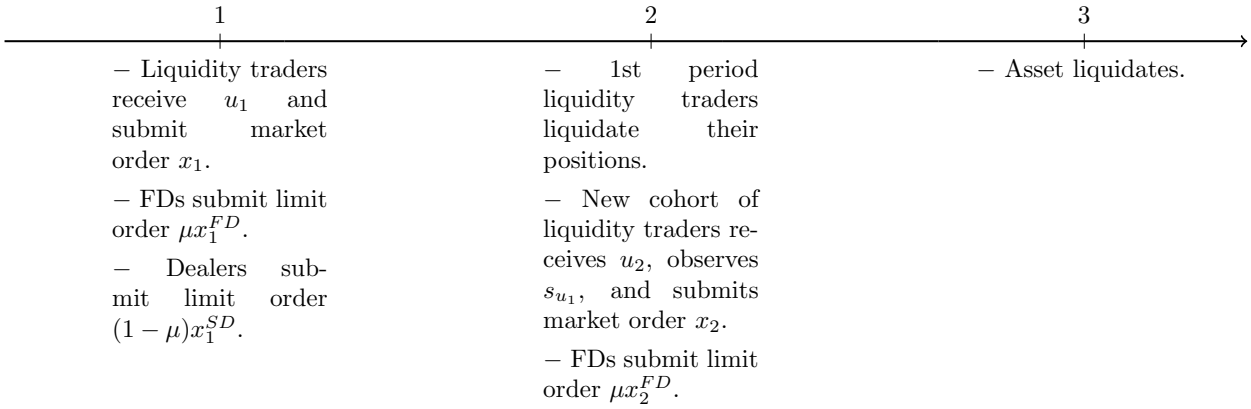


Figure 1: The timeline.

4 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, and derive the unique equilibrium of the model (Proposition 1). We then characterize the trading strategies of liquidity suppliers and demanders (Corollary 1, and 2), and derive the liquidity supply and demand functions (Corollary 3, and 4).

Proposition 1. *When the market is fully transparent there exists a unique linear equilibrium, where prices are given by* (1a) *and* (1b), $\Lambda_{22} = 0$, *and*

$$\Lambda_1^* = \frac{1}{(\gamma + (\gamma+\gamma_2)(\mu\gamma+\gamma_1)(\mu\gamma+\gamma_2)\tau_v\tau_{u_2})\tau_v} \tag{4a}$$

$$\Lambda_2^* = \frac{1}{(\mu\gamma+\gamma_2)\tau_v} \tag{4b}$$

$$\Lambda_{21}^* = -(1-\mu)\gamma\Lambda_1^*\Lambda_2^*\tau_v. \tag{4c}$$

The coefficient Λ_t, i.e. the period t endowment shock's negative price impact, is our measure of illiquidity:

$$\Lambda_t = -\frac{\partial p_t}{\partial u_t}. \tag{5}$$

According to the above result, when second period traders have a perfect signal on the first period endowment shock, the equilibrium is unique, and prices only reflect endowment shocks. To understand the logic of the equilibrium, the next two results describe the strategies of market participants:

Corollary 1. *When the market is fully transparent, liquidity providers' strategies are given by*

$$x_1^{FD} = \gamma \frac{E_1^{FD}[p_2 - p_1]}{\mathrm{Var}_1^{FD}[p_2]} - \gamma \tau_v p_1 \tag{6a}$$

$$x_1^{SD} = -\gamma \tau_v p_1 \tag{6b}$$

$$x_2^{FD} = -\gamma \tau_v p_2, \tag{6c}$$

where $\mathrm{Var}_1^{FD}[p_2] = \Lambda_2^2 \tau_{u_2}^{-1}$, $E_1^{FD}[p_2] = -\Lambda_{21} u_1$.

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 (a higher Λ_2) increases the price adjustment induced by the endowment shock, making liquidity provision more profitable, and increasing 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 which, substituting (1a) and the expressions for expectation and conditional variance in (6a), are given by:

$$x_1^{FD} = \underbrace{\gamma \tau_{u_2} \frac{\Lambda_1 - \Lambda_{21}}{\Lambda_2^2} u_1}_{\text{Speculation}} - \underbrace{\gamma \tau_v p_1}_{\text{Market making}}, \tag{7}$$

According to (7), 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 $u_1 > 0$, since due to (4a) and (4c) $\Lambda_1 - \Lambda_{21} > 0$), and selling otherwise. This is because he anticipates the future price impact of first period liquidity traders' reversion.[15] In this respect, FDs supply liquidity by placing "directional bets."[16] The larger is the wedge between the impact of u_1 on p_2 and p_1, and the smaller is the conditional return volatility, the larger is the size of the FD's directional bet, in absolute value terms. Additionally, a FD places a price contingent order to absorb the residual imbalance, like a standard dealer.

Corollary 2. *When the market is fully transparent, traders' strategies are given by* $x_1 = a_1 u_1, x_2 = a_2 u_2 + b u_1$, *where*

$$a_1 = \gamma_1 \frac{\Lambda_1 - \Lambda_{21}}{\mathrm{Var}_1[p_2]} - 1 \in \left(-1, -\frac{\mu\gamma}{\mu\gamma + \gamma_1} \right) \tag{8a}$$

$$a_2 = \gamma_2 \tau_v \Lambda_2 - 1 \in (-1, 0) \tag{8b}$$

$$b = -(1-\mu)(1+a_2)\gamma \tau_v \Lambda_1 < 0, \tag{8c}$$

where $\mathrm{Var}_1[p_2] = \Lambda_2^2 \tau_{u_2}^{-1}$.

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

[16] This is consisten with the literature linking liquidity supply to sophisticated traders' contrarian behavior (Nagel (2012), Brogaard et al. (2014), and Biais et al. (2015)).

According to (8a) and (8b), first and second period traders demand liquidity to hedge a fraction of their endowment shock (the hedging coefficient a_t, is negative). In the first period, for given u_1 a trader's order is in absolute value larger, the smaller is the expected short-term return (that corresponds to the wedge between the impact of u_1 on p_2 and p_1: $E_1[p_2 - p_1] = (\Lambda_1 - \Lambda_{21})u_1$), and the larger is the conditional return volatility. Thus, FDs and first period traders react in opposite ways to a change in the security short-term return expectation and variance.

This is because a FD has no initial exposure to the risky security and seeks to profit from placing a directional bet. The larger the expected appreciation, and the smaller the conditional return volatility, the more profitable and less risky the bet is, and the larger its size. Conversely, a trader holds an endowment u_1 whose value is correlated with that of the risky asset. He thus seeks protection against the hedging instrument's future price uncertainty, while banking (effectively, speculating) on its potential appreciation. Such protection is more valuable (and thus x_1 is larger in absolute value), the smaller the hedging instrument expected appreciation, and the higher its future price volatility.

According to (8c), second period traders also speculate on the propagated order imbalance by putting a negative weight on their signal ($b < 0$). Note that the "speculative aggressiveness" $|b|$ increases in first period illiquidity. This is because, for $u_1 > 0$, first period traders' reversion has a positive impact on p_2 (see (1b) and (4c)):

$$\text{Cov}[p_2, u_1] = -\Lambda_{21}\tau_{u_1}^{-1} > 0,$$

which prompts second period traders to short the asset. A less liquid first period market increases the position held by standard dealers (see (6b)), increasing the size of the propagated imbalance in the second period equilibrium condition (see (3)), and leading second period traders to step up their speculative aggressiveness.

We now derive the liquidity supply and demand schedules, which we employ to illustrate equilibrium determination in the paper.

Corollary 3. *The illiquidity coefficient Λ_t reflects dealers' risk-related compensation to absorb the outstanding imbalance in their inventory, i.e., the cost of supplying liquidity:*

$$\Lambda_1 = \left(1 - \frac{(\mu\gamma + \gamma_1)(1 + a_1)}{\gamma_1}\right)\frac{1}{\gamma\tau_v} \tag{9a}$$

$$\Lambda_2 = -\frac{a_2}{\mu\gamma\tau_v}. \tag{9b}$$

Λ_t is decreasing in a_t.

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 (9b)), this is immediate, since a_2 measures the fraction of the second period endowment shock that traders hedge

(see (8b)), the risk of the asset payoff is $\text{Var}[v] = 1/\tau_v$, and second period dealers' aggregate risk bearing capacity is given by $\mu\gamma$.[17]

For Λ_1, the argument is as follows. In view of (8a), at equilibrium first period traders hold a fraction

$$1 + a_1, \tag{10}$$

of their endowment shock. At the same time, according to (7) and (8a), FDs aggregate directional bet per unit of endowment shock is given by:

$$FD\text{s aggregate directional bet per unit of endowment shock } = \mu\gamma\frac{1+a_1}{\gamma_1}. \tag{11}$$

Thus, summing (10) and (11) yields the fraction of the endowment shock that does not dent liquidity suppliers' inventory bearing capacity:

$$1 + a_1 + \mu\gamma\frac{1+a_1}{\gamma_1} = \frac{(\mu\gamma+\gamma_1)(1+a_1)}{\gamma_1}, \tag{12}$$

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

$$\text{Dealer's inventory per unit of endowment shock } = 1 - \frac{(\mu\gamma+\gamma_1)(1+a_1)}{\gamma_1}. \tag{13}$$

At date 1, FDs know that they will be able to unwind their inventory in the second trading round, when x_1 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 (9a), the less liquid is the market.

According to (9a) and (9b), the liquidity supply curve is a decreasing function of the hedging coefficient a_t. As a_t moves towards its upper bound, first period trader's liquidity demand decreases, dealers' inventory shrinks (the cost of supplying liquidity declines), and dealers supply more liquidity.

Replacing (4b) and (4c) in (8a) yields the first period hedging coefficient as a function of Λ_1:

Corollary 4. *In a fully transparent market, the first period demand for liquidity is a decreasing function of* Λ_1*:*

$$a_1 = \gamma_1(\gamma+\gamma_2)(\mu\gamma+\gamma_2)\tau_v^2\tau_{u_2}\Lambda_1 - 1. \tag{14}$$

Other things equal, a higher Λ_1 increases the expected short term return on the endowment u_1, without affecting the conditional volatility of short-term returns. This, in turn, leads traders to lower their liquidity demand (the hedging coefficient a_1 approaches its upper bound):

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

illiquidity has a rationing effect. In Figure 2 we plot the liquidity supply and demand curves for a given set of parameter values (respectively, in blue and green). The intersection between the two curves identifies the equilibrium illiquidity and hedging coefficient in a transparent market.

In the figure, we also analyze the effect of an increase in the mass of FDs. A larger μ increases the aggregate speculative position of FDs (see (11)), lowering dealers' inventory. This reduces the cost of liquidity supply for all levels of a_1. As a result, when μ increases, the new function Λ_1 shifts downwards. Consider now a_1. A larger FD mass lowers traders' uncertainty on p_2, lowering their demand for liquidity.[18] Thus, when μ increases, a_1 shifts leftwards. Thus,

Corollary 5. *In a fully transparent market, an increase in the mass of FDs makes the market more liquid, and increases the first period liquidity demand.*

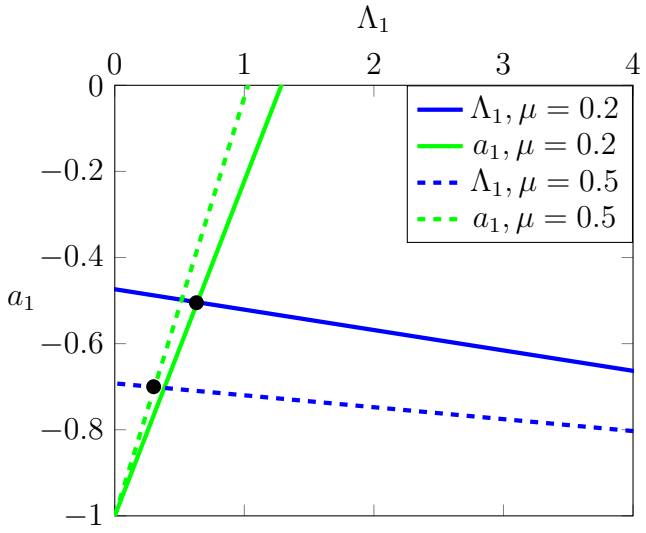


Figure 2: Liquidity consumption and illiquidity with transparency. Parameter values: $\gamma = \gamma_2 = 0.9$, $\gamma_1 = 0.2$, $\tau_{u_1} = 2$, $\tau_{u_2} = 200$, and $\tau_v = 0.1$.

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.

[18] To be sure, for given Λ_1, μ has a weaker impact on the expected return $E_1[p_2 - p_1] = \Lambda_1(\gamma+\gamma_2)u_1/(\mu\gamma+\gamma_2)$, than on the volatility of the expected return $\text{Var}_1[p_2] = 1/(\mu\gamma + \gamma_2)^2\tau_v^2\tau_{u_2}$.

5 Imperfect transparency 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). 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 the following, we first state the equilibrium existence result when the market is not fully transparent (Proposition 2). We then analyze how lack of transparency affects traders' strategies (Corollary 6), and can make liquidity demand upward sloping in illiquidity and FDs' directional bet decline with illiquidity (Corollary 8). We then concentrate on the analysis of equilibrium multiplicity and liquidity fragility, providing numerical examples.

In terms of notation, we use the superscript T to denote the equilibrium coefficient of the fully transparent market. The following result characterizes the equilibrium price coefficients:

Proposition 2. *When $0 < \tau_\eta < \infty$, an equilibrium exists where prices are given by (1a) and (1b), the second period price coefficients are as follows:*

$$\Lambda_2^* = \frac{1}{(\mu\gamma\kappa + \gamma_2)\tau_v} \tag{15a}$$

$$\Lambda_{21}^* = -\frac{1-\mu}{\mu}\Lambda_1\Lambda_2^*\tau_v(\mu\gamma\kappa + \alpha\gamma_2) \tag{15b}$$

$$\Lambda_{22}^* = \frac{1-\mu}{\mu}\Lambda_1\Lambda_2^*\tau_v(1-\alpha)\gamma_2, \tag{15c}$$

where

$$\kappa \equiv \tau_v \mathrm{Var}_2[v - p_2] = 1 + \left(\frac{1-\mu}{\mu}\right)^2 \frac{\Lambda_1^2\tau_v}{\tau_{u_1} + \tau_\eta} > 1, \ \alpha \equiv \frac{\tau_{u_1}}{\tau_{u_1} + \tau_\eta}, \tag{16}$$

and first period illiquidity obtains as a fixed point of the following map:

$$\psi(\Lambda_1) = \frac{(\Lambda_{22}^*)^2\tau_{u_2} + ((\Lambda_2^*)^2 + (\mu\gamma + \gamma_1)\tau_{u_2}\Lambda_{21}^*)\tau_\eta}{\gamma_1\tau_{u_2}\tau_\eta + \gamma(((\Lambda_{22}^*)^2\tau_{u_2} + (\Lambda_2^*)^2\tau_\eta)\tau_v + \mu\tau_\eta\tau_{u_2})}. \tag{17}$$

At equilibrium, $\Lambda_1^* \in (0, 1/\gamma\tau_v)$, and $\Lambda_2^* \in (0, \Lambda_2^T)$.

Second period traders cannot perfectly observe u_1. This affects second period price coefficients via κ and α (see (15a)–(15c), and (16)), implies that the second period price loads also on the error term η (see (15c)), and impacts traders' behavior at both rounds:

Corollary 6. *When the market is not fully transparent, traders' strategies are given by $x_1 =$*

$a_1 u_1, x_2 = a_2 u_2 + b s_{u_1}$, *where*

$$a_1 = \gamma_1 \frac{\Lambda_1 - \Lambda_{21}}{\text{Var}_1[p_2]} - 1 < 0 \tag{18a}$$

$$a_2 = \frac{\gamma_2 \tau_v \Lambda_2 - 1}{\kappa} \in (a_2^T, 0) \tag{18b}$$

$$b = -(1-\mu)(1+\kappa a_2)(1-\alpha)\gamma\tau_v\Lambda_1 < 0, \tag{18c}$$

where $\text{Var}_1[p_2] = \Lambda_2^2 \tau_{u_2}^{-1} + \Lambda_{22}^2 \tau_\eta^{-1}$.

Traders at the second round face execution risk, and hedge a smaller fraction of their endowment shock (see (18b)), which explains why $\Lambda_2^* < \Lambda_2^T$. The impact of lack of transparency on second period traders' speculation and first period traders' hedging is instead more complex.

To see why, suppose that first period illiquidity increases. As in the transparent market benchmark, this increases the propagation of the first period endowment shock, since

$$\text{Cov}[p_2, u_1] = -\Lambda_{21}^* \tau_{u_1}^{-1} > 0,$$

is increasing in Λ_1. However, differently from the transparent market benchmark, as second period traders cannot observe u_1, this also increases these traders' return uncertainty $\text{Var}_2[v - p_2]$ (see (16)), in turn affecting their hedging and speculative aggressiveness:

Corollary 7. *In a not fully transparent market an increase in first period illiquidity, increases both the impact of the first period endowment shock on the second period price and second period traders' return uncertainty, while decreasing second period traders' hedging aggressiveness. The impact on second period speculative aggressiveness is inconclusive:*

$$\frac{\partial \text{Cov}[p_2, u_1]}{\partial \Lambda_1} > 0, \ \frac{\partial \text{Var}_2[v - p_2]}{\partial \Lambda_1} > 0, \ \frac{\partial |a_2|}{\partial \Lambda_1} < 0, \ \frac{\partial |b|}{\partial \Lambda_1} \lessgtr 0. \tag{19}$$

Based on (16), an increase in Λ_1 increases second period traders' execution risk, which heightens $\text{Var}_2[v - p_2]$, depressing $|a_2|$. The impact on second period traders' speculative aggressiveness instead obtains as the sum of two contrasting effects:

$$\frac{\partial |b|}{\partial \Lambda_1} = (1-\mu)(1-\alpha)\gamma\tau_v \Bigg| \underbrace{1 + \kappa a_2}_{\text{Speculation effect } (+)} + \Bigg(\underbrace{\frac{\partial \kappa}{\partial \Lambda_1} a_2 + \frac{\partial a_2}{\partial \Lambda_1} \kappa}_{\text{Uncertainty effect } (-)} \Bigg) \Lambda_1 \Bigg|. \tag{20}$$

As noted above, due to lack of transparency, a stronger endowment shock propagation, heightens execution risk, taming traders' reaction to their signal. However, as it also opens more profitable opportunities to second period traders, it pushes them to speculate more aggressively, injecting more noise in the second period price.[19] If the latter effect is sufficiently strong, it can in turn

[19] For example, trying to sell (buy) when $u_1 < 0$ ($u_1 > 0$), because $s_{u_1} > 0$ ($s_{u_1} < 0$).

lead to an increase in first period traders' return uncertainty:

$$\frac{\partial \text{Var}_1[p_2]}{\partial \Lambda_1} = 2\Bigg(\underbrace{\frac{\partial \Lambda_2^*}{\partial \Lambda_1}\tau_{u_2}^{-1}}_{\text{Uncertainty effect } (-)} + \underbrace{\frac{\partial \Lambda_{21}^*}{\partial \Lambda_1}\tau_{\eta}^{-1}}_{\text{Speculation effect } (\pm)} \Bigg), \tag{21}$$

which can offset the positive impact that a higher illiquidity has on the first period expected return (per unit of endowment shock):

$$\frac{\partial E_1[p_2 - p_1]/u_1}{\partial \Lambda_1} = \frac{1-\mu}{\mu}\tau_v \left(\Lambda_2^*(\mu\gamma\kappa + \alpha\gamma_2) + \Lambda_1 \frac{\mu\gamma\gamma_2(1-\alpha)\kappa'}{(\mu\gamma\kappa + \gamma_2)^2} \right) + 1 > 0, \tag{22}$$

where $\kappa' \equiv \partial\kappa/\partial\Lambda_1$.

As a result, we have that, differently from the transparent market benchmark:

Corollary 8. *Without full transparency, in the fist period an increase in illiquidity may lead to an increase in the demand for liquidity and a decrease in the size of FDs' directional bets.*

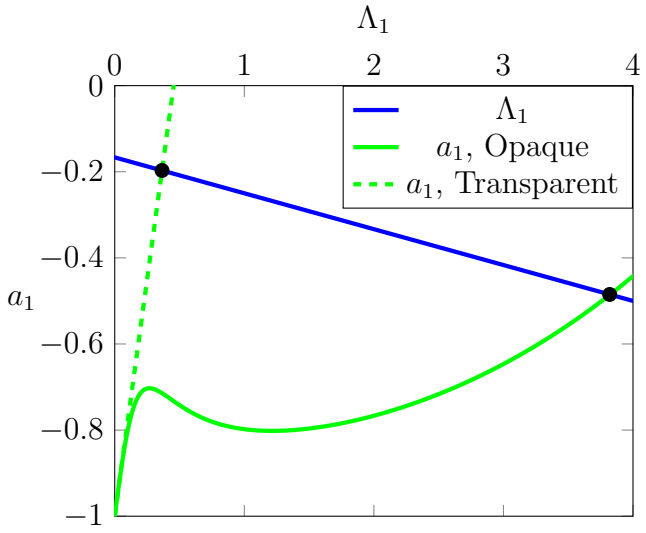


Figure 3: Liquidity consumption and illiquidity with opaqueness. Parameter values: $\gamma = \gamma_2 = 1$, $\gamma_1 = 0.5$, $\tau_{u_1} = 0.5$, $\tau_{u_2} = 200$, $\tau_\eta = 10$, $\mu = 0.1$, and $\tau_v = 0.1$.

In Figure 3 we show an example of the impact of alck of transparency for the equilibrium level of illiquidity.[20] In the figure, when the market is not fully transparent ("Opaque"), first period traders' liquidity demand has a backward bending segment (illiquidity becomes a Giffen good). This, in turn, magnifies the equilibrium level of illiquidity, as well as traders' liquidity demand:

$$\Lambda_1^* \begin{cases} 0.37 & \text{Transparent market} \\ 3.82 & \text{Opaque market} \end{cases} \qquad a_1^* \begin{cases} -0.2 & \text{Transparent market} \\ -0.49 & \text{Opaque market} \end{cases}$$

[20] The plot for Λ_1 as a function of a_1 is obtained using (9a), since lack of transparency does not change the logic behind the construction of the liquidity supply curve.

Note that in the example, lack of transparency leads to a 10-fold illiquidity increase, while traders' liquidity demand increases by a 2.5-factor. How can such a disproportionate effect arise?

To understand the reason for this effect, recall the discussion following Corollary 2. There we argued that FDs and first period traders react in opposite ways to a change in the risky security's short-term return expectation and variance. Due to market opaqueness, an illiquidity spike can generate an increase in return uncertainty that swamps the positive impact it has on expected returns. This leads first period traders to demand more liquidity (solid green curve in Figure 3), and FDs to reduce the size of their directional bets (solid green curve in Figure 4), exacerbating the liquidity rout.

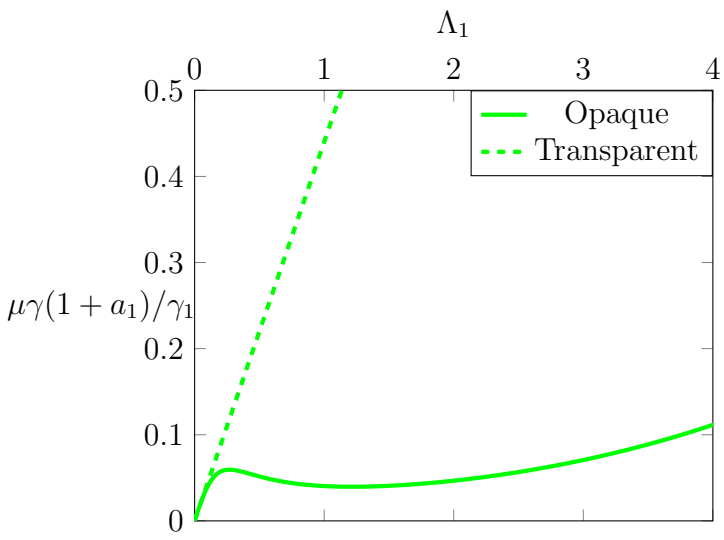


Figure 4: The effect of market opaqueness on FDs directional bets. Parameter values: $\gamma = \gamma_2 = 1$, $\gamma_1 = 0.5$, $\tau_{u_1} = 0.5$, $\tau_{u_2} = 200$, $\tau_\eta = 10$, $\mu = 0.1$, and $\tau_v = 0.1$.

5.1 Equilibrium multiplicity

A second effect of alck of transparency is the possibility of multiple, self-fulfilling equilibria which arise out of strategic complementarities in *liquidity demand*. According to Corollary 7, 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 (20)), which can increase the uncertainty faced by first period traders on p_2 (see (21)). As a consequence, first period traders can decide to hedge more (Corollary 8), and FDs to speculate less. This chain of effects turns out to be particularly strong when the risk bearing capacity of FDs is sufficiently high, 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. *Without full market transparency, there exists a set of parameter values, such that for, $\tau_v < \bar{\tau}_v$, $\mu < \bar{\mu}$, and $\gamma_1 < \bar{\gamma}_1$, three equilibrium levels of liquidity $(\Lambda_1^*)^H, (\Lambda_1^*)^I, (\Lambda_1^*)^L$ arise, where*

$$0 < (\Lambda_1^*)^H < \frac{\mu}{1-\mu} < (\Lambda_1^*)^I < \frac{1}{1-\mu} < (\Lambda_1^*)^L < \frac{1}{\gamma\tau_v}. \tag{23}$$

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 the expression for $\Lambda_1(a_1)$ coincides with the one obtained in Section 4. Since $\Lambda_1(a_1)$ is decreasing in a_1 (Corollary 3), 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.

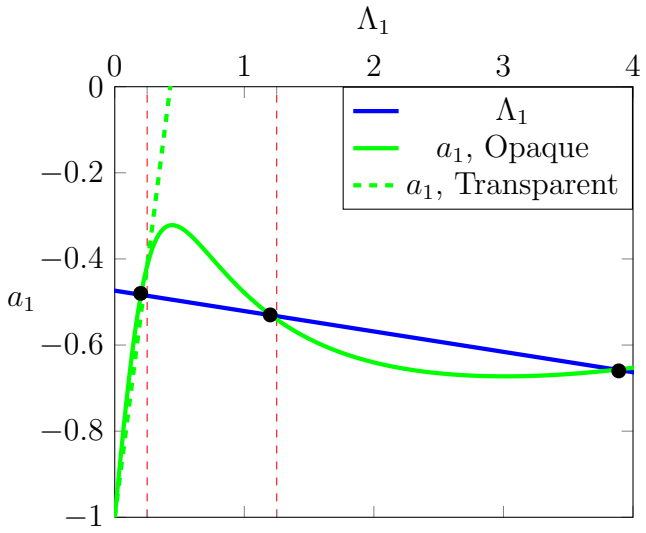


Figure 5: Liquidity consumption and illiquidity with opaqueness. Parameter values: $\gamma = \gamma_2 = 0.9$, $\gamma_1 = 0.2$, $\tau_{u_1} = 2$, $\tau_{u_2} = 600$, $\tau_\eta = 10$, $\mu = 0.2$, and $\tau_v = 0.1$. The first (second) gridline is drawn at $\mu/(1-\mu) = 0.25$ ($1/(1-\mu) = 1.25$), which according to (23) define the intervals where the HLE (ILE) are located.

The following corollary follows from Proposition 3:

Corollary 9. *Without full market transparency, when the volatility of the second period endowment shock vanishes ($\tau_{u_2} \to \infty$), $(\Lambda_1^*)^H \to 0$. When the market is fully transparent, this is the unique equilibrium of the model.*

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 $b = 0$ (see (18c)). When second period traders' signal on u_1 is noiseless, this equilibrium is unique. 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, we plot the function $\psi(\Lambda_1)$ (see (17)), 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.[21] Figure 6 provides a graphical representation of the best response, for a set of parameters yielding multiple equilibria.

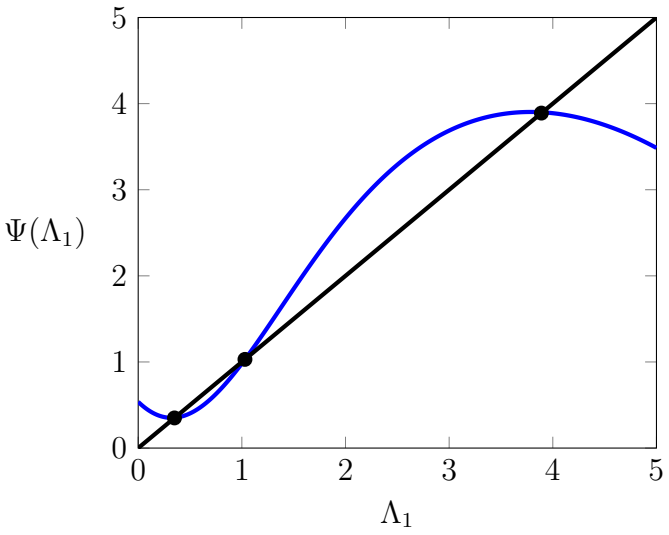


Figure 6: The best response mapping $\Psi(\Lambda_1)$. Parameter values: $\gamma = \gamma_2 = 0.9$, $\gamma_1 = 0.2$, $\tau_{u_1} = 2$, $\tau_{u_2} = 400$, $\tau_\eta = 10$, $\mu = 0.2$, and $\tau_v = 0.1$.

Higher values of τ_{u_2}, γ_2, 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, 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 11 and 12 (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 11, Panel (a), (b), and (c)).

[21] See the proof of Proposition 2, and in particular the discussion around (A.22).

2. An increase in γ_1 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 11, 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, $\text{Var}_1[p_2]$ starts decreasing (see the expression in Proposition 2), leading first period traders to hold more of their 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 $\text{Var}_1[p_2]$ disappears, and (ii) second period traders' uncertainty no longer depends on Λ_1^* (see the expression for $\text{Var}_2[v - p_2]$ 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 4). Figure 12, 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 increases the strength of strategic complementarity. Other things equal, according to (18b) and (18c), 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 12, 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 12, Panel (c) in the Appendix illustrates this effect.[22] In the limit case where $\tau_{u_2} \to 0$, we obtain the following result:

Corollary 10. *When* $\tau_{u_2} \to 0$, $\Lambda_1^* = 1/\gamma\tau_v$, *and* $a_1 = -1$.

[22] According to Corollary 9 in the extreme case in which the second period endowment shock is null (almost surely), there always exist an equilibrium where $\Lambda_1^* = 0$, $x_1^{SD} = 0$, so that first period traders' orders are absorbed by FDs' speculative trades, which implies that no imbalance arises in the second period, and $b = 0$ (see (18c)). 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: $a_1 \to -\mu\gamma/(\mu\gamma+\gamma_1)$.

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.

5.2 Fragility

The feedback loop induced by lack of market transparency 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, 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.

5.2.1 Equilibrium switch

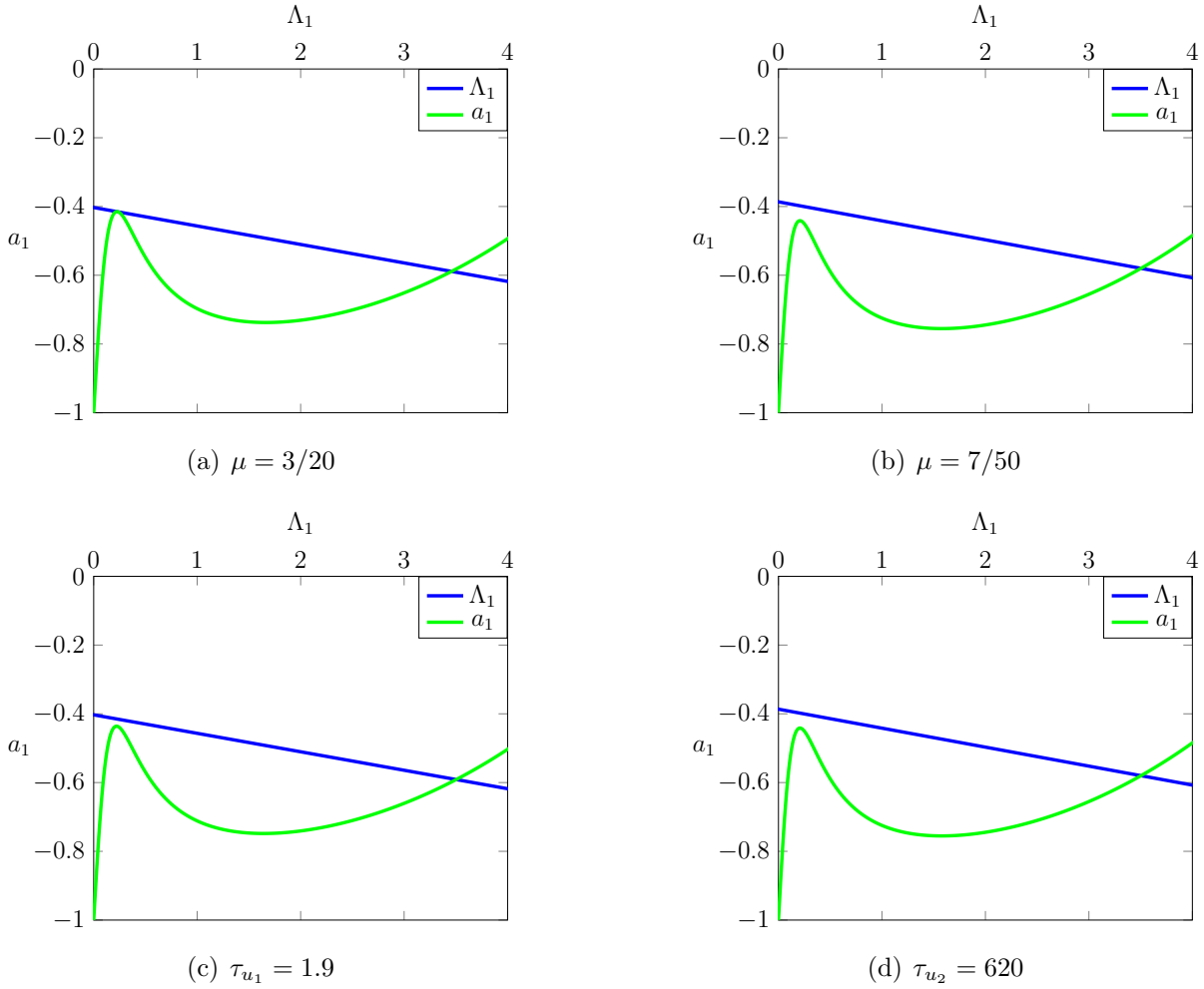
Consider Panel (a) in Figure 7, and suppose that initially the market is at the high liquidity equilibrium, where $\Lambda_1^* = 0.22$. Suppose that a technical "glitch" disconnects 6% of the FDs. In this new situation, the plot for $\Lambda_1(a_1)$ shifts upwards, while the one for $a_1(\Lambda_1)$ 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 a_1 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 $a_1 = -0.6$. 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 $a_1(\Lambda_1^*)$ moves downwards, while the one for $\Lambda_1^*(a_1)$ 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.

The dynamics of the latter comparative statics exercise appear to be consistent with the events that occurred on August 24, 2015, when the S&P500 lost 5% in a matter of minutes at the open, eventually recovering the loss along the day. Importantly, the narrative of the event points to the fact that due to the lack of bids, the opening of many NYSE stocks was delayed on that day. This, in turn, complicated establishing the fair value of ETFs, causing traders to sell more and bid less at the market opening.[23]

Table 1 summarizes the results of these exercises and compares them with the effects that obtain in the fully transparent market case.

[23] See https://www.investopedia.com/articles/investing/011116/two-biggest-flash-crashes-2015.asp



(a) $\mu = 3/20$

(b) $\mu = 7/50$

(c) $\tau_{u_1} = 1.9$

(d) $\tau_{u_2} = 620$

Figure 7: Liquidity fragility. In panel (a) we set $\gamma = \gamma_2 = 0.9$, $\gamma_1 = 0.2$, $\tau_{u_1} = 2$, $\tau_{u_2} = 661$, $\tau_\eta = 3$, $\mu = 0.15$, and $\tau_v = 0.1$ 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^*	$\Delta\Lambda_1^*/\Lambda_1^*$
Fully transparent market ($\tau_\eta \to \infty$)	0.237	$\Delta\mu/\mu = -6\%$	0.245	3.2%
		$\Delta\tau_{u_1}/\tau_{u_1} = -5\%$	0.237	0%
		$\Delta\tau_{u_2}/\tau_{u_2} = -7\%$	0.252	6.3%
Not fully transparent market ($\tau_\eta < \infty$)	0.22	$\Delta\mu/\mu = -6\%$	3.5	1470%
		$\Delta\tau_{u_1}/\tau_{u_1} = -5\%$	3.5	1470%
		$\Delta\tau_{u_2}/\tau_{u_2} = -7\%$	3.5	1470%

Table 1: Illiquidity impact of an equilibrium switch according to transparency.

5.2.2 Fragility along a unique equilibrium

Liquidity fragility can also arise when the equilibrium is unique because, with some 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:

$$\frac{\partial \Lambda_1^*}{\partial \mu} = -\frac{1}{\gamma \tau_v} \times \Bigg(\underbrace{\frac{\gamma(1+a_1)}{\gamma_1}}_{\text{Direct effect}_{(+)}} + \underbrace{\frac{\partial a_1}{\partial \mu}\left(1+\frac{\mu\gamma}{\gamma_1}\right)}_{\text{Indirect effect}_{(\pm)}} \Bigg). \tag{24}$$

For given a_1, 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 ($\partial a_1/\partial \mu$), and on each FDs speculative position ($(\mu\gamma/\gamma_1)(\partial a_1/\partial \mu)$). 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 $\text{Var}_1[p_2]$ and leading first period traders to hold more of their endowment, and each FD to speculate more on the short term capital gain;[24] 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.[25] When the market is not fully transparent, second period traders face execution risk, which tames their hedging aggressiveness (see (18b)), and lowers first period traders' return uncertainty. In this situation, the uncertainty reduction effect of μ on a_1 can be dwarfed by the one due to reduced predictability. As a consequence, when μ increases, first period traders' demand for immediacy can increase 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.

[24] 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.

[25] In a fully transparent market, this latter effect is never strong enough to overcome the previous two, and liquidity increases in the mass of FDs.

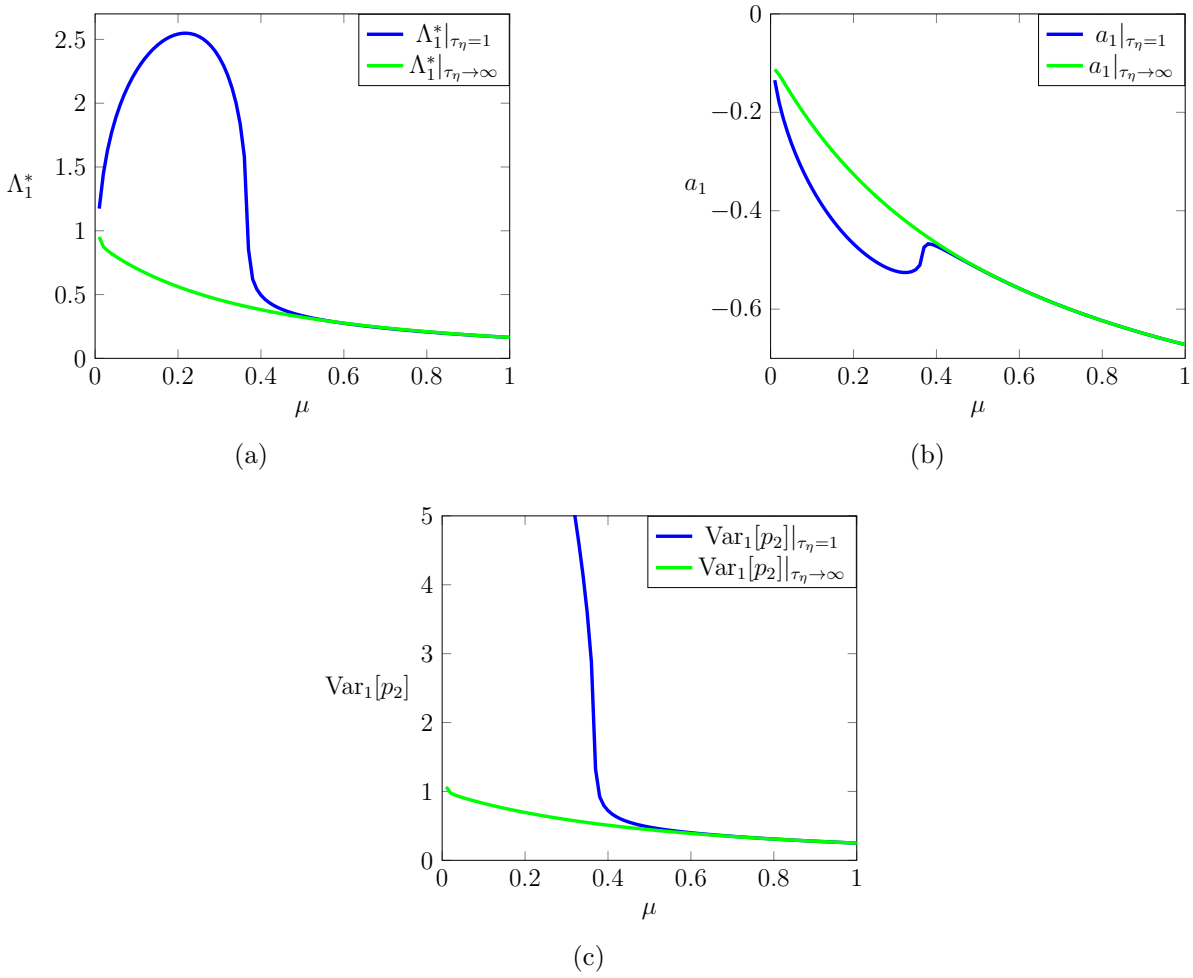


Figure 8: Fragility at an equilibrium with low liquidity. In panel (a), (b), and (c) we plot Λ_1^*, a_1, and $\mathrm{Var}_1[p_2]$, for $\tau_v = \tau_{u_1} = 0.1$, $\tau_{u_2} = 100$, $\tau_\eta = \gamma = \gamma_2 = 1$, $\gamma_1 = 0.5$, and $\mu \in \{0.01, .02, \ldots, 1\}$; the blue (green) plot 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^*, a_1, and $\mathrm{Var}_1[p_2]$.

This implies that illiquidity can be hump-shaped in the proportion of FDs, as shown in Figure 8 (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 $a_1 = -0.47$ (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 $\mu = 0.36$, $a_1 =$

-0.51 (see Figure 8, Panel (b)), which, compared to the status quo liquidity demand, corresponds to a 9% increase in liquidity consumption by first period traders (i.e., $a_1 = -0.47$ when $\mu = 0.4$). How can such a comparatively small increase in liquidity consumption generate an illiquidity spike of this magnitude? The explanation is once again 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, $\text{Var}_1[p_2]$ 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 (24) to break down the different effects yields:

$$\underbrace{\frac{\Delta\Lambda_1^*}{\Delta\mu}}_{\approx -27} \approx -\underbrace{\frac{1}{\gamma\tau_v}}_{=-10} \times \tag{25}$$

$$\underbrace{\left(\frac{\gamma(1+a_1)}{\gamma_1}\right.}_{\substack{\Delta \text{ in } FD\text{s' aggregate speculative position} \approx 1}} + \underbrace{\frac{\mu\gamma}{\gamma_1}}_{=4/5} \times \underbrace{\frac{\Delta a_1}{\Delta\mu}}_{\substack{\Delta \text{ in each FD speculative position} \approx 1}} + \underbrace{\left.\frac{\Delta a_1}{\Delta\mu}\right)}_{\substack{\Delta \text{ in liquidity demand} \approx 1}}$$

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.[26] This suggests that in the wake of a liquidity crash, we should observe FDs unwinding their directional bets.[27]

[26] According to (25), 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 $\Delta a_1/\Delta\mu$ accounts for 1.

[27] 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.

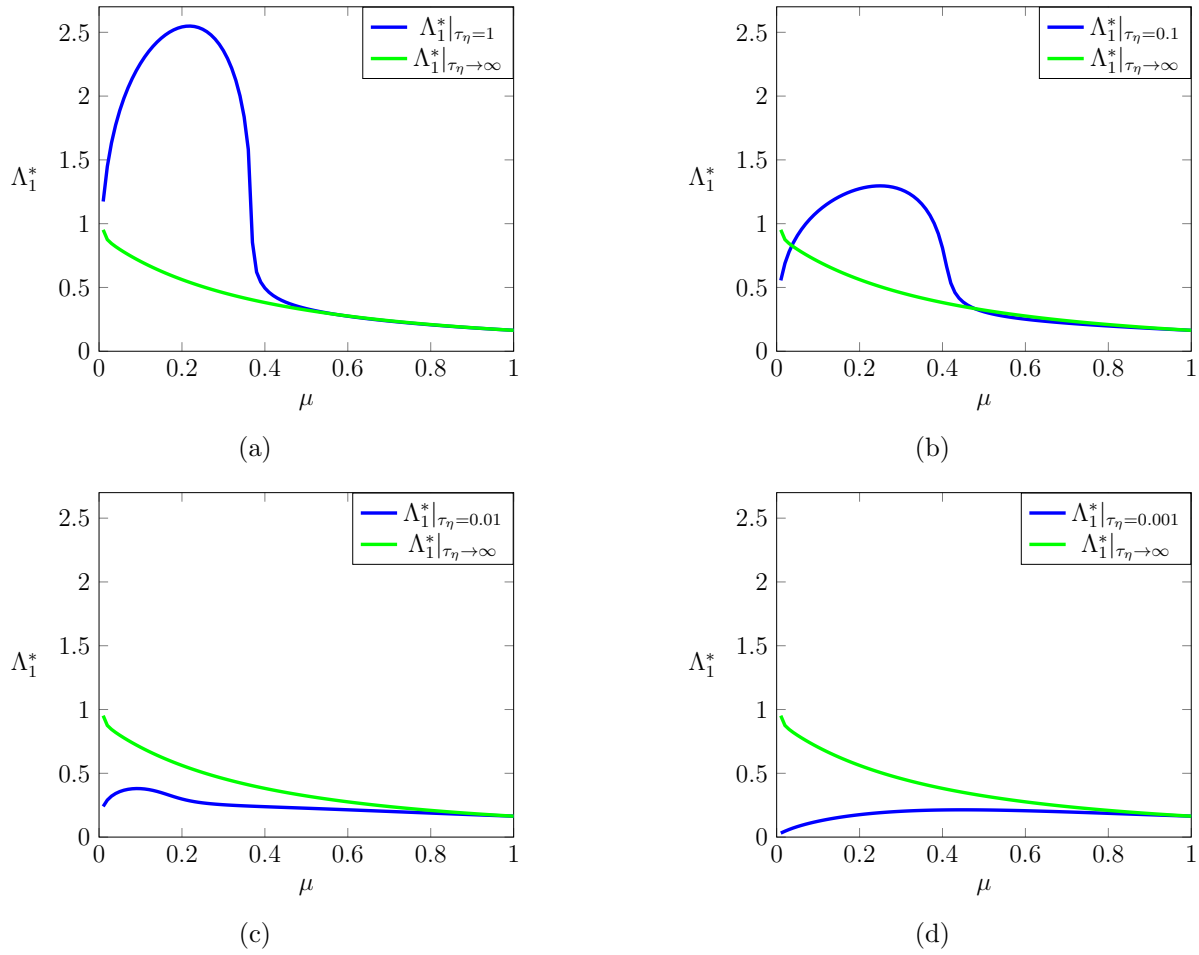


Figure 9: Liquidity comparison between the opaque and transparent market case. In panel (a), we plot Λ_1^*, for $\tau_v = \tau_{u_1} = 0.1$, $\tau_{u_2} = 100$, $\tau_\eta = \gamma = \gamma_2 = 1$, $\gamma_1 = 0.5$, and $\mu \in \{0.01, .02, \ldots, 1\}$; the blue (green) plot relates to the transparent (opaque) case. In panel (b), (c), and (d) we progressively reduce τ_η.

A final implication of this example is that when information on prices and/or order imbalances is opaque, an increase in the mass of FDs (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.[28]

5.2.3 Transparency and liquidity

An important feature of the example in Figure 8 is that for all values of μ the market is at least as liquid with full transparency as it is without it. This result is not general. In Figure 9 we reproduce Panel (a) of Figure 8, and in panels (b), (c), and (d), investigate the effect of progressively lowering the informativeness of the signal observed by second period traders. The

[28] 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 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 11. *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:*

$$\Lambda_1^*|_{\tau_\eta \to 0} \leq \Lambda_1^*|_{\tau_\eta \to \infty}. \tag{26}$$

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 a_2, 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 9 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 10, where we plot Λ_1^* 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 5.1). 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 11, a very precise signal (τ_η large) fosters second period traders' hedging aggressiveness, which heightens first period traders' uncertainty compared to the case in which the signal is uninformative ($\tau_\eta \approx 0$).

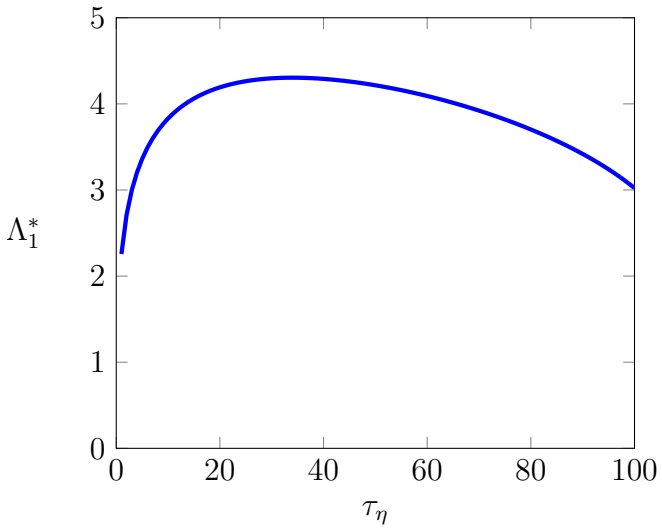


Figure 10: Illiquidity as a function of transparency, for $\tau_v = \tau_{u_1} = 0.1$, $\tau_{u_2} = 100$, $\gamma = \gamma_2 = 1$, $\gamma_1 = 0.5$, $\mu = 0.1$.

5.2.4 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, without full market transparency, the self-stabilizing mechanism whereby an illiquidity spike depresses liquidity consumption, and fosters' FDs directional bets can jam, and instead be replaced by a vicious cycle that creates a liquidity rout.

Transparency	Full ($\tau_\eta \to \infty$)	Partial ($0 < \tau_\eta < \infty$)
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
FDs speculative demand	Increasing in Λ_1	Can be decreasing in Λ_1

Table 2: A summary of our results.

6 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. Without it, 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 somewhat 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 lack of market transparency, 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 not fully transparent 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 FDs entry should be assessed taking into account the effect of frictions in the access to market information. Second, with noisy market information, fragmented liquidity supply 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.[29]

Finally, our analysis of fragility along a unique equilibrium highlights the role that changes in FDs' strategies in the wake of crashes have to explain huge illiquidity spikes. As we argued in our numerical examples, the unwinding of FDs' 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, FDs 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)).

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

References

Amihud, Y., H. Mendelson, and R. A. Wood (1990). Liquidity and the 1987 stock market crash. *Journal of Portfolio Management 16*(3), 65–69.

Banerjee, S., J. Davis, and N. Gondhi (2018). When transparency improves, must prices reflect fundamentals better? *Review of Financial Studies 31*(1), 2377–2414.

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

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

Bogousslavsky, V. (2016). Infrequent rebalancing, return autocorrelation, and seasonality. *Journal of Finance 71*(6), 2967–3006.

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 (2017). What is the optimal trading frequency in financial markets? *Review of Economic Studies 84*(4), 1606–1651.

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 (2016). News trading and speed. *Journal of Finance 71*(1), 335–382.

Foucault, T., M. Pagano, and A. Röell (2013). *Market Liquidity*. Oxford University Press.

Gennotte, G. and H. Leland (1990). Market liquidity, hedging, and crashes. *American Economic Review 80*(5), 999–1021.

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., C. M. Jones, and A. Menkveld (2010). Does algorithmic trading improve liquidity? *Journal of Finance 66*(1), 1–33.

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 8*, 1–24.

Menkveld, A. J. and M. A. Zoican (2017). Need for speed? Exchange latency and liquidity. *Review of Financial Studies 30*(4), 1188–1228.

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 1. *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 Proposition 1

From the proof of Proposition 2 (see (A.29)) we know that when $\tau_\eta \to \infty$, a unique equilibrium obtains where

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

implying that the second period price coefficients (A.14a) and (A.14b) have the closed form solution ((4b), and (4c)).

Replacing (4b) and (4c) in (8a) yields

$$a_1 = \gamma_1(\gamma_2 + \gamma)(\gamma_2 + \mu\gamma)\Lambda_1\tau_{u_2}\tau_v^2 - 1.$$

Furthermore, $\lim_{\tau_\eta \to \infty} \mathrm{Var}_2[v - p_2] = \tau_v^{-1}$, so that a_2 and b in (A.5) have the closed form solution displayed in (8b) and (8c).

□

Proof of Proposition 2

We work by backward induction. In the second period, CARA and normality assumptions imply that the objective function of a liquidity trader is given by

$$E_2[-\exp\{-\pi_2/\gamma_2\}] = -\exp\left\{-\frac{1}{\gamma_2}\left(E_2[\pi_2] - \frac{1}{2\gamma_2}\mathrm{Var}_2[\pi_2]\right)\right\}, \tag{A.2}$$

where $\pi_2 \equiv (v - p_2)x_2 + u_2 v$. Maximizing (A.2) with respect to x_2, and solving for the optimal strategy yields:

$$x_2 = \gamma_2 \frac{E_2[v - p_2]}{\mathrm{Var}_2[v - p_2]} - \frac{\mathrm{Cov}_2[v - p_2, v]}{\mathrm{Var}_2[v - p_2]} u_2. \tag{A.3}$$

Using (1b) we can compute:

$$E_2[v-p_2] = \Lambda_2 u_2 + \frac{\Lambda_{21}\tau_\eta + \Lambda_{22}\tau_{u_1}}{\tau_\eta + \tau_{u_1}} s_{u_1} \tag{A.4a}$$

$$\mathrm{Var}_2[v-p_2] = \frac{1}{\tau_v} + \frac{(\Lambda_{21}-\Lambda_{22})^2}{\tau_\eta + \tau_{u_1}} \tag{A.4b}$$

$$\mathrm{Cov}_2[v-p_2, v] = \frac{1}{\tau_v}. \tag{A.4c}$$

Substituting (A.4a)–(A.4c) in (A.3), and rearranging yields:

$$x_2 = \underbrace{\frac{\gamma_2\tau_v\Lambda_2 - 1}{\tau_v \mathrm{Var}_2[v-p_2]}}_{a_2} u_2 + \underbrace{\gamma_2 \frac{\Lambda_{21}\tau_\eta + \Lambda_{22}\tau_{u_1}}{(\tau_\eta + \tau_{u_1})\mathrm{Var}_2[v-p_2]}}_{b} s_{u_1}. \tag{A.5}$$

A FD maximizes the expected utility of his second period wealth:

$$\begin{aligned} &E_2^{FD}\left[-\exp\left\{ -\frac{1}{\gamma}\left((p_2-p_1)x_1^{FD} + (v-p_2)x_2^{FD}\right)\right\}\right] = \\ &= \exp\left\{ -\frac{1}{\gamma}(p_2-p_1)x_1^{FD}\right\}\left(-\exp\left\{ -\frac{1}{\gamma}\left(E_2^{FD}[v-p_2]x_2^{FD} - \frac{(x_2^{FD})^2}{2\gamma}\mathrm{Var}_2^{FD}[v-p_2]\right)\right\}\right), \end{aligned} \tag{A.6}$$

where the last expression in (A.6) 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:

$$x_2^{FD} = \gamma \frac{E_2^{FD}[v-p_2]}{\mathrm{Var}_2^{FD}[v-p_2]}. \tag{A.7}$$

Computing expectation and variance in the above expression:

$$E_2^{FD}[v-p_2] = -p_2 \tag{A.8a}$$

$$\mathrm{Var}_2^{FD}[v-p_2] = \frac{1}{\tau_v}, \tag{A.8b}$$

and substituting these in x_2^{FD} yields:

$$x_2^{FD} = -\gamma\tau_v p_2. \tag{A.9}$$

Similarly, due to CARA and normality, in the first period an SD maximizes

$$\begin{aligned} &E_1^{SD}\left[-\exp\left\{ -\frac{1}{\gamma}(v-p_1)x_1^{SD}\right\}\right] = \\ &\qquad -\exp\left\{ -\frac{1}{\gamma}\left(E_1^{SD}[v-p_1]x_1^{SD} - \frac{(x_1^{SD})^2}{2\gamma}\mathrm{Var}_1^{SD}[v-p_1]\right)\right\}. \end{aligned} \tag{A.10}$$

Maximizing the above and solving for x_1^{SD} yields:

$$x_1^{SD} = \gamma \frac{E_1^{SD}[v - p_1]}{\text{Var}_1^{SD}[v - p_1]}. \tag{A.11}$$

Computing the conditional expectation and variance:

$$E_1^{SD}[v - p_1] = -p_1 \tag{A.12a}$$

$$\text{Var}_1^{SD}[v - p_1] = \frac{1}{\tau_v}, \tag{A.12b}$$

so that

$$x_1^{SD} = -\gamma \tau_v p_1. \tag{A.13}$$

We can now substitute (A.5), (A.9), and (A.13) in (3), and identify the price coefficients, obtaining:

$$\Lambda_2 = \frac{(\tau_\eta + \tau_{u_1})\mu}{((1-\mu)^2 \gamma \tau_v \Lambda_1^2 + (\tau_\eta + \tau_{u_1})(\mu\gamma + \gamma_2)\mu)\tau_v} \tag{A.14a}$$

$$\Lambda_{21} = -\frac{(1-\mu)(\tau_{u_1} + (1 - \gamma_2 \tau_v \Lambda_2)\tau_\eta)\Lambda_1}{(\tau_\eta + \tau_{u_1})\mu} \tag{A.14b}$$

$$\Lambda_{22} = \frac{(1-\mu)\gamma_2 \Lambda_1 \Lambda_2 \tau_v \tau_\eta}{(\tau_\eta + \tau_{u_1})\mu}. \tag{A.14c}$$

Note that $\Lambda_2 > 0$, while the sign of Λ_{21} and Λ_{22} depend on that of Λ_1. Indeed,

$$1 - \gamma_2 \tau_v \Lambda_2 = \gamma \frac{(1-\mu)^2 \tau_v \Lambda_1^2 + (\tau_\eta + \tau_{u_1})\mu^2}{(1-\mu)^2 \gamma \tau_v \Lambda_1^2 + (\tau_\eta + \tau_{u_1})(\mu\gamma + \gamma_2)\mu} > 0.$$

Consider now the first period. We start by characterizing the strategy of a FD. Substituting a FD's second period strategy (A.7) in (A.6), rearranging and applying Lemma 1 yields the following expression for the first period objective function of a FD:

$$E_1^{FD}[U((p_2 - p_1)x_1^{FD} + (v - p_2)x_2^{FD})] = -\left(1 + \frac{\text{Var}_1^{FD}[p_2]}{\text{Var}[v]}\right)^{-1/2} \times \tag{A.15}$$
$$\exp\left\{-\frac{1}{\gamma}\left(\frac{\gamma\tau_v}{2}\nu^2 + (\nu - p_1)x_1^{FD} - \frac{(x_1^{FD} + \gamma\tau_v\nu)^2}{2\gamma}\left(\frac{1}{\text{Var}_1^{FD}[p_2]} + \frac{1}{\text{Var}[v]}\right)^{-1}\right)\right\},$$

where

$$E_1^{FD}[p_2] = -\Lambda_{21} u_1 \tag{A.16a}$$

$$\text{Var}_1^{FD}[p_2] = \frac{\Lambda_{21}^2}{\tau_{u_2}} + \frac{\Lambda_{22}^2}{\tau_\eta}. \tag{A.16b}$$

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

$$X_1^{FD}(p_1) = \frac{\gamma}{\text{Var}_1^{FD}[p_2]}\nu - \gamma\left(\frac{1}{\text{Var}_1^{FD}[p_2]} + \frac{1}{\text{Var}[v]}\right)p_1. \tag{A.17}$$

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

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

$$E[-\exp\{-\pi_1/\gamma_1\}] = -\exp\left\{-\frac{1}{\gamma}\left(E_1[\pi_1] - \frac{1}{2\gamma_1}\text{Var}_1[\pi_1]\right)\right\}, \tag{A.18}$$

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

$$X_1(u_1) = \gamma_1\frac{E_1[p_2 - p_1]}{\text{Var}_1[p_2]} - \frac{\text{Cov}_1[p_2 - p_1, p_2]}{\text{Var}_1[p_2]}u_1, \tag{A.19}$$

where

$$E_1[p_2 - p_1] = -(\Lambda_{21} - \Lambda_1)u_1 \tag{A.20a}$$

$$\text{Var}_1[p_2] = \text{Var}_1^{FD}[p_2] = \frac{\Lambda_2^2}{\tau_{u_2}} + \frac{\Lambda_{22}^2}{\tau_\eta} \tag{A.20b}$$

$$\text{Cov}_1[p_2 - p_1, p_2] = \text{Var}_1[p_2]. \tag{A.20c}$$

Substituting (A.20a)–(A.20c) in (A.19) and rearranging yields:

$$x_1 = \underbrace{\left(\gamma_1\frac{\Lambda_1 - \Lambda_{21}}{\text{Var}_1[p_2]} - 1\right)}_{a_1}u_1. \tag{A.21}$$

We can now substitute (A.13), (A.17), and (A.21) into (2) to identify the first period price coefficient Λ_1:

$$\begin{aligned}\Lambda_1 = \psi(\Lambda_1) &= \frac{\Lambda_{22}^2\tau_{u_2} + (\Lambda_2^2 + (\mu\gamma + \gamma_1)\tau_{u_2}\Lambda_{21})\tau_\eta}{\gamma_1\tau_{u_2}\tau_\eta + \gamma((\Lambda_{22}^2\tau_{u_2} + \Lambda_2^2\tau_\eta)\tau_v + \mu\tau_\eta\tau_{u_2})} \\ &= -\frac{\mu\gamma\text{Cov}[p_2, u_1]\tau_{u_1} + a_1\text{Var}_1[p_2]}{\gamma(\mu + \tau_v\text{Var}_1[p_2])}.\end{aligned} \tag{A.22}$$

According to (A.22), Λ_1 is pinned down by the solution of the following equation:

$$f(\Lambda_1) \equiv \psi(\Lambda_1) - \Lambda_1 = \frac{(\mu\gamma + \gamma_1)(\text{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1) + \text{Var}_1[p_2](\gamma\tau_v\Lambda_1 - 1)}{\gamma(\mu + \tau_v\text{Var}_1[p_2])} = 0. \tag{A.23}$$

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.23), which is given by:

$$\begin{aligned} f(\Lambda_1) = {} & \gamma^2(1-\mu)^4\Lambda_1^5(\mu\gamma+\gamma_1)\tau_{u_2}\tau_v^4 \\ & + \mu\gamma(1-\mu)^2\tau_{u_2}\tau_v^3\Lambda_1^3\big(((\mu\gamma+\gamma_1)(2\mu\gamma+\gamma_2(1+\mu))+\gamma_2^2)\tau_\eta + 2\tau_{u_1}(\mu\gamma+\gamma_1)(\mu\gamma+\gamma_2)\big) \\ & - \gamma_2^2(1-\mu)^2\mu\tau_\eta\tau_{u_2}\tau_v^2\Lambda_1^2 \\ & + \mu^2(\tau_\eta+\tau_{u_1})\tau_v\Lambda_1\big(\mu\gamma(\tau_\eta+\tau_{u_1}) + (\mu\gamma+\gamma_1)(\mu\gamma+\gamma_2)(\mu\gamma(\tau_\eta+\tau_{u_1})+\gamma_2(\mu\tau_\eta+\tau_{u_1}))\tau_{u_2}\tau_v\big) \\ & - \mu^3(\tau_\eta+\tau_{u_1})^2 = 0. \end{aligned} \tag{A.24}$$

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 $f(\Lambda_1) = 0$.

Computing $\mathrm{Cov}[p_2, u_1]$ yields:

$$\mathrm{Cov}[p_2, u_1] = -\Lambda_{21}\tau_{u_1}^{-1}, \tag{A.25}$$

which is positive if and only if $\Lambda_1 > 0$. Consider (A.23) and suppose that at equilibrium $\Lambda_1 < 0$. From (A.25), $\mathrm{Cov}[p_2, u_1] < 0$. Due to (A.23) this implies $f(\Lambda_1) < 0$, which is impossible. Thus, at equilibrium, $0 < \Lambda_1 < 1/\gamma\tau_v$, and $\mathrm{Cov}[p_2, u_1] \geq 0$.

To sign the strategy coefficient of a first period liquidity trader, we use (A.21):

$$a_1 = \gamma_1 \frac{\mathrm{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1}{\mathrm{Var}_1[p_2]} - 1. \tag{A.26}$$

From (A.26) we obtain $\mathrm{Var}_1[p_2](1+a_1)/\gamma_1 = \mathrm{Cov}[p_2, u_1]\tau_{u_1} + \Lambda_1$, which substituted in (A.23) at equilibrium yields

$$f(\Lambda_1) = \frac{\mathrm{Var}_1[p_2]}{(\mu + \tau_v\mathrm{Var}_1[p_2])\gamma\gamma_1}((1+a_1)(\mu\gamma+\gamma_1) + (\gamma\tau_v\Lambda_1 - 1)\gamma_1) = 0. \tag{A.27}$$

Solving the above for Λ_1 yields: $\Lambda_1 = (-\mu\gamma - a_1(\mu\gamma+\gamma_1))/\gamma_1\gamma\tau_v$. Since $\Lambda_1 > 0$, the previous equation implies that at equilibrium $a_1 < 0$. Furthermore, using (A.26), $1 + a_1 > 0$, which proves our result.

Taking the limit for $\tau_\eta \to \infty$ in $\psi(\Lambda_1)$ yields:

$$\lim_{\tau_\eta \to \infty} \psi(\Lambda_1) = \frac{1 - \Lambda_1(\gamma_2 + \mu\gamma)(\gamma_1(\gamma+\gamma_2) + \mu\gamma^2(1-\mu))\tau_{u_2}\tau_v^2}{\gamma\tau_v(1 + \mu(\gamma_2+\mu\gamma)^2\tau_{u_2}\tau_v)}. \tag{A.28}$$

Identifying Λ_1 yields

$$\Lambda_1 = \frac{1}{\tau_v(\gamma + (\mu\gamma+\gamma_1)(\mu\gamma+\gamma_2)(\gamma_2+\gamma)\tau_{u_2}\tau_v)}. \tag{A.29}$$

□

Proof of Corollary 6

The expressions for the strategies' coefficients are obtained in the proof of Proposition 2. Here we prove that lack of transparency reduces second period traders' hedging aggressiveness. With a transparent market, solving for a_2 yields $a_2 = -\mu\gamma/(\gamma_2 + \mu\gamma)$. When the market is not transparent, instead, replacing Λ_2^* in (A.5) and solving for a_2 yields $a_2 = -\mu\gamma/(\gamma_2 + \mu\gamma\kappa)$. Since $\kappa > 1$, the result follows. □

Proof of Corollary 7

From Proposition 2 it is immediate that $\text{Var}_2[v - p_2]$ is increasing in Λ_1. Differentiating $\text{Cov}[p_2, u_1]$ yields

$$\frac{\partial \text{Cov}[p_2, u_1]}{\partial \Lambda_1} = \frac{1-\mu}{\mu}\left((\mu\gamma\kappa + \alpha\gamma_2)\Lambda_2^* + \frac{\mu\gamma\gamma_2\kappa'(1-\alpha)}{(\mu\gamma\kappa + \gamma_2)^2}\Lambda_1\right) \geq 0,$$

for $\mu \leq 1$ (where $\kappa' \equiv \partial\kappa/\partial\Lambda_1$). Finally, from Proposition 2 we know that a_2 is increasing in κ, and thus in $\text{Var}_[v - p_2]$. This implies that a_2 is increasing in Λ_1. □

Proof of Proposition 3

The equilibrium quintic (A.24) 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:

$$\begin{aligned} f(\Lambda_1) = \Big[&\Lambda_1(\tau_{u_1}((\gamma_2^L/\mu\gamma\tau_v) + \text{Var}_2[v - p_2]) + \tau_\eta((\gamma_2/\gamma\tau_v) + \text{Var}_2[v - p_2]))(\mu\gamma + \gamma_1)\times \\ &(\tau_{u_1} + \tau_\eta)((\gamma_2/\mu\gamma\tau_v) + \text{Var}_2[v - p_2])\tau_\eta + (1/\mu\gamma\tau_v)^2((\gamma_2\tau_\eta/\mu\gamma\tau_v)(1-\mu)\gamma\tau_v\Lambda_1)^2(\gamma\tau_v\Lambda_1 - 1)\mu\Big]\tau_{u_2}\tau_v^2 \\ &+ (1/\mu\gamma\tau_v)^2(\tau_{u_1} + \tau_\eta)^2\mu\tau_\eta(\gamma\tau_v\Lambda_1 - 1). \end{aligned} \tag{A.30}$$

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.30) (i.e., the one that multiplies τ_{u_2}). We first concentrate on the analysis of this quintic:

$$\begin{aligned} \hat{f}(\Lambda_1) = &\Lambda_1(\tau_{u_1}((\gamma_2/\mu\gamma\tau_v) + \text{Var}_2[v - p_2]) + \tau_\eta((\gamma_2/\gamma\tau_v) + \text{Var}_2[v - p_2]))(\mu\gamma + \gamma_1)\times \\ &(\tau_{u_1} + \tau_\eta)((\gamma_2/\mu\gamma\tau_v) + \text{Var}_2[v - p_2])\tau_\eta + (1/\mu\gamma\tau_v)^2((\gamma_2\tau_\eta/\mu\gamma\tau_v)(1-\mu)\gamma\tau_v\Lambda_1)^2(\gamma\tau_v\Lambda_1 - 1)\mu. \end{aligned} \tag{A.31}$$

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:

$$h(0) = \frac{\tau_\eta(\mu\gamma+\gamma_1)(\mu\gamma+\gamma_2)(\tau_{u_1}+\tau_\eta)(\mu\gamma(\tau_{u_1}+\tau_\eta)+\gamma_2(\tau_{u_1}+\mu\tau_\eta))}{\gamma^2\mu^2} > 0. \tag{A.32}$$

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

$$\begin{aligned} h(\bar{\Lambda}_1^*) = \frac{\tau_\eta}{\gamma^2\tau_\eta^4}\times \\ \Big((\mu\gamma+\gamma_1)(\mu(\mu\gamma+\gamma_2)(\tau_{u_1}+\tau_\eta)+\gamma\tau_v)(\mu\gamma_2\mu(\tau_{u_1}+\mu\tau_\eta)+\gamma(\mu^2(\tau_\eta+\tau_{u_1})+\tau_v)) \\ -\mu(\gamma_2)^2\tau_\eta(1-\mu-\gamma\tau_v)\Big), \end{aligned} \tag{A.33}$$

which is negative when the following parameter restrictions are satisfied:

$$0 < \mu < \bar{\mu} \equiv \frac{\gamma_2(\sqrt{5\gamma^2+\gamma_2(2\gamma+\gamma_2)}-(\gamma+\gamma_2))}{2\gamma^2} \tag{A.34a}$$

$$0 < \tau_v < \bar{\tau}_v \equiv \frac{(1-\mu)(\gamma_2)^2-\mu\gamma(\mu\gamma+\gamma_2)}{\gamma(\gamma_2)^2} \tag{A.34b}$$

$$\tau_\eta > \underline{\tau}_\eta \equiv \frac{\gamma(\mu(\mu\gamma+\gamma_2)\tau_{u_1}+\gamma\tau_v)}{(1-\mu-\gamma\tau_v)(\gamma_2)^2-\mu\gamma(\mu\gamma+\gamma_2)} \tag{A.34c}$$

$$0 < \gamma_1 < \bar{\gamma}_1 \equiv \frac{\mu((\gamma_2)^2\tau_\eta(1-\mu-\gamma\tau_v)-\mu\gamma\gamma_2(\tau_{u_1}+\tau_\eta)-\gamma^2(\tau_v+\mu^2(\tau_{u_1}+\tau_\eta)))}{\gamma\tau_v+\mu(\mu\gamma+\gamma_2)(\tau_{u_1}+\tau_\eta)}. \tag{A.34d}$$

Therefore, when (A.34a)-(A.34d) 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.34a)-(A.34d) hold, three equilibria: $0 < (\Lambda_1^*)^I < \bar{\Lambda}_1^* < (\Lambda_1^*)^L < 1/\gamma\tau_v$, arise. Consider now the general quintic (A.30). First, note that

$$f(0) = -\frac{\tau_\eta(\tau_\eta+\tau_{u_1})^2}{\mu(\gamma\tau_v)^2} < 0. \tag{A.35}$$

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

$$\begin{aligned} f(\underline{\Lambda}_1^*) = \frac{\tau_\eta}{\mu\gamma^2}\Bigg\{\frac{(\tau_\eta+\tau_{u_1})^2(\mu\gamma\tau_v-(1-\mu))}{(1-\mu)\tau_v^2}+\frac{\tau_{u_2}}{1-\mu}\Bigg((\gamma_2)^2\tau_\eta(\mu\gamma\tau_v-(1-\mu))+ \\ +(\mu\gamma+\gamma_1)(\gamma_2(\tau_{u_1}+\tau_\eta)+\mu\gamma(\tau_{u_1}+\tau_\eta+\tau_v))(\gamma_2(\tau_{u_1}+\mu\tau_\eta)+\mu\gamma(\tau_{u_1}+\tau_\eta+\tau_v))\Bigg)\Bigg\}. \end{aligned} \tag{A.36}$$

The sign of (A.36) 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

$$\gamma > \underline{\gamma} \equiv \frac{(1-\mu)\tau_\eta-\gamma_1(\tau_{u_1}+\tau_\eta)(\tau_{u_1}+\mu\tau_\eta)}{\mu(\tau_v\tau_\eta+(\tau_{u_1}+\tau_\eta)(\tau_{u_1}+\mu\tau_\eta))}. \tag{A.37}$$

Hence, if (A.37) holds, and

$$\tau_{u_2} > \underline{\tau}_{u_2} \equiv \big((\tau_v^2(\gamma_2^2(\tau_\eta(\mu\gamma\tau_v - (1-\mu)) + (\mu\gamma+\gamma_1)(\tau_1+\tau_\eta)(\tau_{u_1}+\mu\tau_\eta)) + \mu\gamma\gamma_2 \times (\mu\gamma+\gamma_1)(\tau_1+\tau_v+\tau_\eta)(2\tau_{u_1}+\tau_\eta(1+\mu)) + (\mu\gamma+\gamma_1)(\mu\gamma)^2(\tau_{u_1}+\tau_\eta+\tau_v)^2))\big)^{-1} \times \big((\tau_{u_1}+\tau_\eta)^2(1-\mu-\mu\gamma\tau_v)\big), \tag{A.38}$$

expression (A.36) is positive. This establishes the existence of an equilibrium $0 < (\Lambda_1^*)^H < \underline{\Lambda}_1^*$. Finally, provided (A.34a)-(A.34d) 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.34a)-(A.34d), and (A.37), (A.38) hold, the model displays three equilibria:

$$0 < (\Lambda_1^*)^H < \underline{\Lambda}_1^* < (\Lambda_1^*)^I < \bar{\Lambda}_1^* < (\Lambda_1^*)^L < \frac{1}{\gamma\tau_v}. \tag{A.39}$$

□

Proof of Corollary 9

See the proof of Proposition 3. □

Proof of Corollary 10

Taking the limit for $\tau_{u_2} \to 0$ in $\Psi(\Lambda_1)$, and collecting terms, yields the following equilibrium condition:

$$f(\Lambda_1) = \mu^3(\tau_{u_1}+\tau_\eta)^2(\gamma\tau_v\Lambda_1 - 1) = 0. \tag{A.40}$$

The above equation has a unique real solution $\Lambda_1^* = 1/\gamma\tau_v$. Substituting this solution in (A.10) implies that at this equilibrium $a_1 = -1$. □

Proof of Corollary 11

If we let $\tau_\eta \to 0$, (A.24) reads as follows:

$$f(\Lambda_1) = -\mu^3\tau_{u_1}^2(1-\gamma\tau_v\Lambda_1) + \Lambda_1(\gamma_1+\mu\gamma)\tau_{u_2}\tau_v^2(\gamma\tau_v\Lambda_1^2(1-\mu)^2 + \mu\tau_{u_1}(\gamma_2+\mu\gamma))^2 = 0. \tag{A.41}$$

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.41) at (A.29) yields $f(\Lambda_1^*) > 0$ which implies that a strongly opaque market is more liquid than a fully transparent one. □

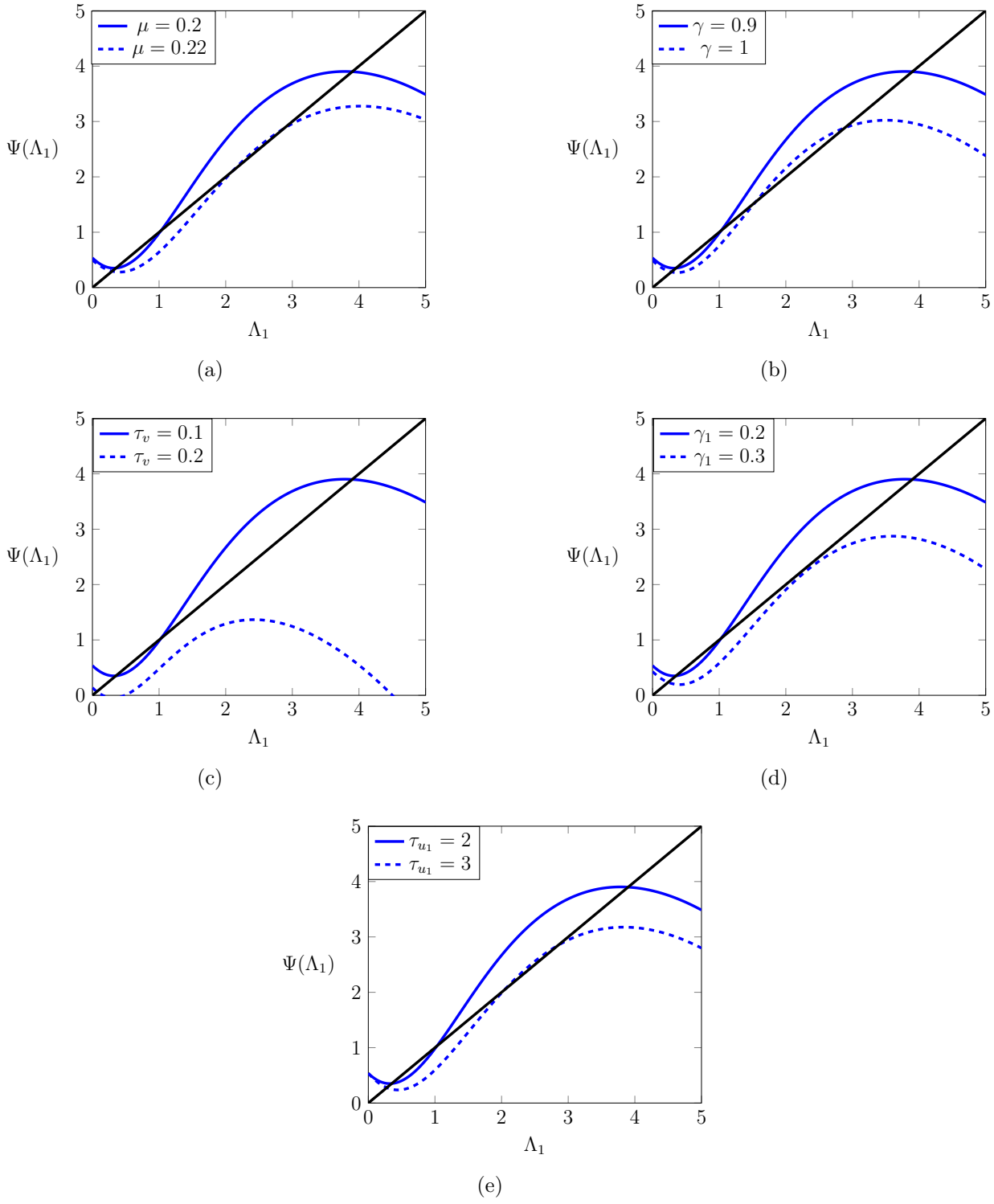


Figure 11: Comparative statics. The figure displays the effect of a change in the proportion of FDs (Panel (a)), the risk tolerance of SDs and FDs (Panel (b)), the volatility of the payoff (Panel (c)), the risk tolerance of first period traders (Panel (d)), and the volatility of the first period endowment shock (Panel (e)). The continuous blue curve represents the function $\psi(\Lambda_1)$ for the set of parameters' values used in Figure 6.

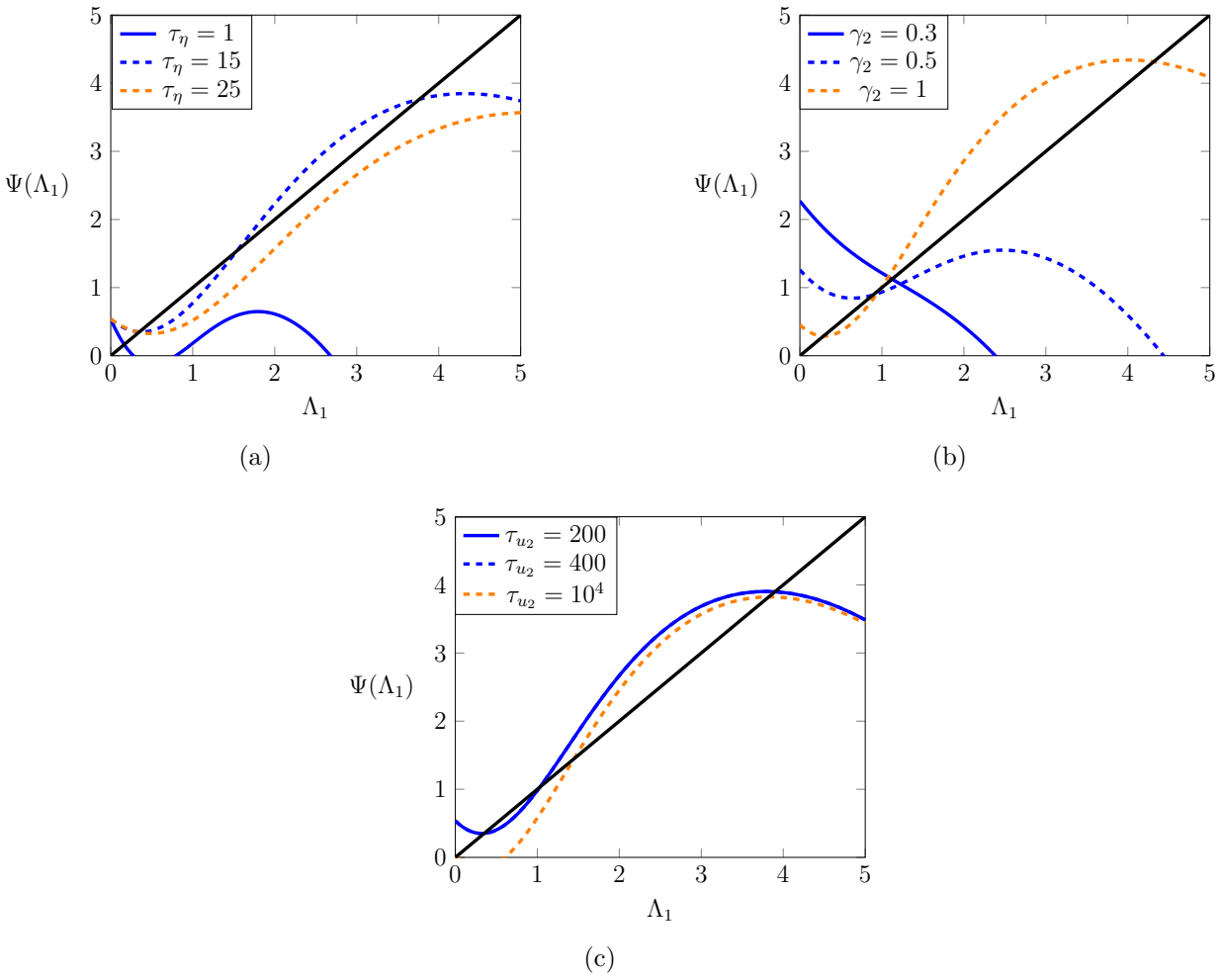


Figure 12: Comparative statics. The figure displays the effect of a change in the precision of second period liquidity traders' signal (Panel (a)), second period liquidity traders' risk tolerance (Panel (b)), and the precision of second period liquidity traders' endowment shock (Panel (c)). The other parameters' values are as in Figure 6.